Novel scaling limits for critical inhomogeneous random graphs

Shankar Bhamidi *
Remco van der Hofstad †
Johan S.H. van Leeuwaarden †

September 8, 2009

Abstract

We find scaling limits for the sizes of the largest components at criticality for the rank-1 inhomogeneous random graphs with power-law degrees with exponent τ. We investigate the case where $\tau \in (3,4)$, so that the degrees have finite variance but infinite third moment. The sizes of the largest clusters, rescaled by $n^{-(\tau-2)/(\tau-1)}$, converge to hitting times of a 'thinned' Lévy process. This process is intimately connected to the general multiplicative coalescents studied in [1] and [3]. In particular, we use the results in [3] to show that, when interpreting the location λ inside the critical window as time, the limiting process is a multiplicative process with diffusion constant 0 and the entrance boundary describing the size of relative components in the $\lambda \to -\infty$ regime proportional to $\big(i^{-1/(\tau-1)}\big)_{i\geq 1}$. A crucial ingredient is the identification of the scaling of the largest connected components in the barely subcritical regime.

Our results should be contrasted to the case where the degree exponent τ satisfies $\tau > 4$, so that the third moment is finite. There, instead, we see that the sizes of the components rescaled by $n^{-2/3}$ converge to the excursion lengths of an inhomogeneous Brownian motion, as proved in [1] for the Erdős-Rényi random graph and extended to the present setting in [6, 26]. The limit again is a multiplicative coalescent, the only difference with the limit for $\tau \in (3,4)$ being the initial state, corresponding to the barely subcritical regime.

Key words: critical random graphs, phase transitions, inhomogeneous networks, thinned Lévy processes, multiplicative coalescent.
MSC2000 subject classification. 60C05, 05C80, 90B15.

1 Introduction

1.1 Model

We start by describing the model considered in this paper. In our random graph model, vertices have *weights*, and the edges are independent with the edge probability being approximately equal to the rescaled product of the weights of the two end vertices of the edge. While there are many different versions of such random graphs (see Section 1.5 for a discussion of these), it will be convenient for us to work with

*Department of Statistics and Operations Research, University of North Carolina, 304 Hanes Hall, Chapel Hill, NC 27510. Email: `bhamidi@email.unc.edu`

†Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: `rhofstad@win.tue.nl`, `j.s.h.v.leeuwaarden@tue.nl`

the so-called *Poissonian random graph* or Norros-Reittu model [23]. To define the model, we consider the vertex set $[n] := \{1, 2, \dots, n\}$, and attach an edge with probability p_{ij} between vertices i and j, where

$$p_{ij} = 1 - \exp\left(-\frac{w_i w_j}{l_n}\right), \tag{1.1}$$

with

$$l_n = \sum_{i=1}^{n} w_i. \tag{1.2}$$

Different edges are independent. In this model, the average degree of vertex i is close to w_i, thus incorporating *inhomogeneity* in the model. There are many adaptations of this model, for which equivalent results hold. We defer the discussion of these models to Section 1.5.

We let the weight sequence $\boldsymbol{w} = (w_i)_{i\in[n]}$ be defined by

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

where $F(x)$ is a distribution function for which we assume that there exists a $\tau \in (3, 4)$ and $0 < c_F < \infty$ such that, as $x \to \infty$,

$$1 - F(x) = c_F x^{-(\tau-1)}(1 + o(1)), \tag{1.4}$$

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

$$[1 - F]^{-1}(u) = \inf\{s : [1 - F](s) \le u\}. \tag{1.5}$$

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

$$F(x) = \begin{cases} 0 & \text{for } x < a, \\ 1 - (a/x)^{\tau-1} & \text{for } x \ge a, \end{cases} \tag{1.6}$$

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

We shall frequently make use of the fact that (1.4) implies that, as $u \downarrow 0$, (see e.g., [14, (B.9)])

$$[1 - F]^{-1}(u) = \big(c_F/u\big)^{1/(\tau-1)}(1 + o(1)). \tag{1.7}$$

Under the key assumption in (1.4), we have that the third moment of the degrees tends to infinity, i.e., with $W \sim F$, we have $\mathbb{E}[W^3] = \infty$. We write

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

Then, by [7], when $\nu > 1$ there is one giant component of size proportional to n, while all other components are of smaller size $o(n)$, and when $\nu \le 1$, the largest connected component contains a proportion of vertices that converges to zero in probability. Thus, the critical value of the model is $\nu = 1$. In the example in (1.6), we have that

$$\mathbb{E}[W] = \frac{a(\tau-1)}{\tau-2}, \qquad \mathbb{E}[W^2] = \frac{a^2(\tau-1)}{\tau-3}, \tag{1.9}$$

so that the critical case arises when

$$\nu = \frac{\mathbb{E}[W^2]}{\mathbb{E}[W]} = \frac{a(\tau-2)}{\tau-3} = 1, \tag{1.10}$$

i.e., when $a = (\tau-3)/(\tau-2)$.

With the above definition of the weights, we shall write $\mathcal{G}_n^0(\boldsymbol{w})$ to be the graph constructed with the probabilities in (1.1), while, for any fixed $\lambda \in \mathbb{R}$, we shall write $\mathcal{G}_n^\lambda(\boldsymbol{w})$ when we use the weight sequence $(1+\lambda n^{-(\tau-3)/(\tau-1)})\boldsymbol{w}$. This setting has been studied in some detail in [17], where, for the largest connected component $\mathcal{C}_{\max}$ and each $\lambda \in \mathbb{R}$, it is proved that $|\mathcal{C}_{\max}|n^{-(\tau-2)/(\tau-1)}$ is a *tight* random variable, and also that $n^{(\tau-2)/(\tau-1)}/|\mathcal{C}_{\max}|$ is tight. In this paper, we bring the discussion of the critical behavior of such inhomogeneous random graphs substantially further, by identifying the scaling limit of $(|\mathcal{C}_{(i)}|n^{-(\tau-2)/(\tau-1)})_{i\geq 1}$, where $(\mathcal{C}_{(i)})_{i\geq 1}$ denote the connected components ordered in size, i.e., $|\mathcal{C}_{\max}| = |\mathcal{C}_{(1)}| \geq |\mathcal{C}_{(2)}| \geq \ldots$. We now state our main results.

1.2 The scaling limit for $\tau \in (3,4)$

In this section, we investigate the scaling limit of the largest connected components ordered in size. Our first main result is as follows:

Theorem 1.1 (Weak convergence of the ordered critical clusters for $\tau \in (3,4)$)**.** *Fix the Norros-Reittu random graph with weights $\boldsymbol{w}(\lambda) = \{(1+\lambda n^{-(\tau-3)/(\tau-1)})w_i\}_{i=1}^n$, where $(w_i)_{i\in[n]}$ are as in (1.3). Assume that $\nu = 1$ and that (1.4) holds. Then, for all $\lambda \in \mathbb{R}$,*

$$\big(|\mathcal{C}_{(i)}|n^{-(\tau-2)/(\tau-1)}\big)_{i\geq 1} \xrightarrow{d} (\gamma_i(\lambda))_{i\geq 1}, \tag{1.11}$$

in the product topology, for some non-degenerate limit $(\gamma_i(\lambda))_{i\geq 1}$.

Theorem 1.1 proves [17, Conjecture 1.7]. In the course of the proof, we shall identify the limiting random variables explicitly in terms of certain hitting times of zero of the scaling limit of the exploration process. We next investigate some interesting properties of the limiting largest clusters:

Theorem 1.2 (High-weight vertices are part of $\mathcal{C}_{\max}$ with positive probability)**.** *Under the assumptions in Theorem 1.1, for every $i, j \geq 1$ fixed,*

$$\lim_{n\to\infty} \mathbb{P}(j \in \mathcal{C}(i)) = q_{ij}(\lambda) \in (0,1), \tag{1.12}$$

and

$$\lim_{n\to\infty} \mathbb{P}(i \in \mathcal{C}_{\max}) = q_i(\lambda) \in (0,1). \tag{1.13}$$

The following theorem, which comes out of the construction used in the proof, essentially says that, for each fixed λ, the maximal size components are those attached to the largest weight vertices. We shall also crucially need this theorem in the proof of Theorem 1.1. In order to state the result, we first introduce some notation. Let

$$\mathcal{C}_{\leq}(i) = \begin{cases} \mathcal{C}(i) & \text{if } i < j\ \forall j \in \mathcal{C}(i), \\ \varnothing & \text{otherwise.} \end{cases} \tag{1.14}$$

Then, clearly, $|\mathcal{C}_{\max}| = \max_{i\in[n]} |\mathcal{C}(i)| = \max_{i\in[n]} |\mathcal{C}_{\leq}(i)|$, and $(|\mathcal{C}_{(i)}|)_{i\geq 1}$ is equal to the sequence $(|\mathcal{C}_{\leq}(i)|)_{i\geq 1}$ ordered in size. Then, we have the following result on the cluster sizes $(|\mathcal{C}_{\leq}(i)|)_{i\in[n]}$:

Theorem 1.3 (Maximal cluster contains a high-weight vertex)**.** *Assume that the conditions in Theorem 1.1 hold. Then:*
(a) for any $\varepsilon \in (0,1)$, there exists a $K \geq 1$, such that, for all n,

$$\mathbb{P}\big(\max_{i\geq K} |\mathcal{C}_{\leq}(i)| \geq \varepsilon n^{(\tau-2)/(\tau-1)}\big) \leq \varepsilon; \tag{1.15}$$

(b) for any $m \geq 1$, for $\varepsilon > 0$ small, the probability that the vector $(|\mathcal{C}_{\leq}(i)|)_{i\in[n]}$ contains at least m components of size at least $\varepsilon n^{(\tau-2)/(\tau-1)}$ converges to 1 uniformly in n as $\varepsilon \downarrow 0$.

1.3 Relation to the standard multiplicative coalescent

In this section we shall give a quick overview of Aldous's standard multiplicative coalescent and how it relates to the limiting random variables in Theorem 1.1, seen as functions of the parameter λ. It shall not be possible to give a full description of the process and its many fascinating properties here and we refer the interested reader to the paper [3], the survey paper [2] and the book [5].

Write $l^2_\downarrow$ to be the metric space of infinite real-valued sequences $\mathbf{x} = (x_1, x_2 \ldots)$ with $x_1 \geq x_2 \geq \cdots \geq 0$. The standard multiplicative coalescent is described as the Markov process with states in $l^2_\downarrow$ whose dynamics is described as follows: for each pair of clusters (x, y), the pair merges at rate xy.

First in [1], Aldous showed that there is a Feller process on the space l^2 defined for all times $-\infty < t < \infty$ starting from infinitesimally small masses at time $-\infty$ following the above merging dynamics, and further the distribution of the coalescent process at any time t is the same as the limiting cluster sizes of an Erdős-Rényi random graph with edge probabilities $p_n = (1 + tn^{-1/3})/n$.

In [3], the entrance boundary at $-\infty$ of the above Markov process was explicitly characterized, in the sense that it is proved that every *extreme* version of the above Markov proces is characterized by a diffusion parameter κ, a translation parameter β, and a vector $\boldsymbol{c} = (c_1, c_2, \ldots)$ which describes the relative sizes of the large clusters at time $-\infty$. We refer the interested reader to [3] for the full description of the process. In this terminology, the multiplicative coalescent can be described as the ordered excursions beyond past minima of the process

$$W^{\kappa,\beta,\boldsymbol{c}}(s) = \kappa^{1/2} W(s) + \beta s - \frac{1}{2}\kappa s^2 + V^{\boldsymbol{c}}(s), \tag{1.16}$$

where $(W(s))_{s\geq 0}$ is a standard Brownian motion, while

$$V^{\boldsymbol{c}}(s) = \sum_{j=1}^{\infty} c_j (\mathbb{1}_{\{T_j \leq s\}} - c_j s), \tag{1.17}$$

with $(T_j)_{j\geq 1}$ independent exponential random variables with mean $1/c_j$. Then, the $(\kappa, \beta, \boldsymbol{c})$-multiplicative coalescent is the set of ordered excursions from zero of the reflected process

$$B^{\kappa,\beta,\boldsymbol{c}}(s) = W^{\kappa,\beta,\boldsymbol{c}}(s) - \min_{0\leq s' \leq s} W^{\kappa,\beta\boldsymbol{c}}(s'). \tag{1.18}$$

Part of the proof of [3] is that these ordered excursions can be defined properly.

The following theorem draws a connection between the components of the graph for a fixed λ and the sizes of clusters at the same value of time in a multiplicative coalescent with a particular entrance boundary and scale and translation parameters. For this, define the sequence $\mathbf{c} = c\boldsymbol{w}^*$ where

$$\boldsymbol{w}^* = (i^{-1/(\tau-1)})_{i\geq 1} \tag{1.19}$$

and $c = c_F^{1/(\tau-1)}$. Then, we have the following theorem, where we write $\xrightarrow{\mathbb{P}}$ to denote convergence in probability:

Theorem 1.4 (Relation to multiplicative coalescents)**.** *Assume that the conditions in Theorem 1.1 hold. Consider the sequence valued random variables* $\mathbf{X}^*(\lambda) = (\gamma_1(\mathbb{E}[W]\lambda), \gamma_2(\mathbb{E}[W]\lambda), \ldots)$ *obtained in Theorem 1.1. Then* $\mathbf{X}^*(\lambda)$ *has the same distribution as a multiplicative coalescent at time* λ *with entrance boundary* $\boldsymbol{c}$*, diffusion constant* $\kappa = 0$ *and centering constant* $\beta = -\zeta/\mathbb{E}[W]$*, where* ζ *is identified explicitly in* (A.2)*. In particular,*

$$|\lambda|\gamma_j(\lambda) \xrightarrow{\mathbb{P}} c_j/\mathbb{E}[W] \text{ as } \lambda \to -\infty \text{ for each } j \geq 1. \tag{1.20}$$

See Section 7 for a full proof of this result. The setting in this paper is the first example where the multiplicative coalescent with $\kappa = 0$ arises in random graph theory. Indeed, all random graph examples in [3] have largest component sizes of the order $n^{2/3}$, like for the Erdős-Rényi random graph studied in [1]. Our example links the multiplicative coalescent also to random graphs with largest critical connected components of the order $n^{(\tau-2)/(\tau-1)}$.

A crucial part of the proof is the analysis of the *subcritical* phase of our model. The asymptotics in the subcritical case acts as the entrance boundary of the multiplicative coalescent, as explained in more detail in [3, Proposition 7]. This entrance boundary is identified in the following theorem, which is of independent interest:

Theorem 1.5 (Subcritical phase)**.** *Assume that the conditions in Theorem 1.1 hold, but now take* $\lambda = \lambda_n \to -\infty$ *as* $n \to \infty$ *such that* $\lambda_n \geq -n^{-(\tau-3)/(\tau-1)}$*. Then, for each* $j \in \mathbb{N}$*,*

$$|\lambda_n| n^{-(\tau-2)/(\tau-1)} |\mathcal{C}_{(j)}| \xrightarrow{\mathbb{P}} c_j. \tag{1.21}$$

Theorem 1.5 is proved in Section 6. The lower bound on λ_n is only there to ensure that $\tilde{w}_i = (1+\lambda n^{-(\tau-3)/(\tau-1)})w_i \geq 0$ for every $i \in [n]$. A result similar to Theorem 1.5 is proved for the near-critical phase of the configuration model in [18], but the proof we give here is entirely different.

1.4 Overview of the proof

We first give an overview of the proof. We note that $\boldsymbol{w}$ is ordered in size, i.e., $w_1 \geq w_2 \geq w_3 \geq \ldots$. Let $\mathcal{C}(i)$ denote the cluster found by starting the exploration from vertex i. We start by exploring the clusters from the largest weight vertices onwards. The reason for exploring the clusters from the high-weight vertices onwards is that vertex 1 is part of $\mathcal{C}_{\max}$ with positive probability, which is strictly smaller than 1. This is rather different in the setting where $\tau > 4$, which we shall discuss in more detail in Section 1.5.

In Theorem 1.3 we have seen that for any $m \geq 1$ and for K sufficiently large, the vector

$$\big(|\mathcal{C}_{(i)}| n^{-(\tau-2)/(\tau-1)}\big)_{i\in[m]}$$

is equal to the first m elements of $(|\mathcal{C}_{\leq}(i)|)_{i\in[K]}$ ordered in size. This leads us to study the scaling limit of the connected components of the first K vertices. For this, in turn, it suffices to compute the scaling limit of $(|\mathcal{C}_{\leq}(i)|)_{i\in[K]}$, which is our next aim.

In Section 2, we shall start by identifying the scaling limit of $|\mathcal{C}_{\leq}(1)| n^{-(\tau-2)/(\tau-1)} = |\mathcal{C}(1)| n^{-(\tau-2)/(\tau-1)}$. The weak limit of $|\mathcal{C}(1)| n^{-(\tau-2)/(\tau-1)}$ is given in terms of the hitting time of 0 of an exploration process exploring the cluster of vertex 1. See Theorems 2.1 and 2.2. The scaling limit of the exploration process of a cluster exists (see Theorem 2.2), and is a novel kind of stochastic process, which we can think of as a 'thinned' Lévy process. Therefore, the convergence in distribution of $|\mathcal{C}(1)| n^{-(\tau-2)/(\tau-1)}$ as in Theorem 2.1 is equivalent to the convergence of the first hitting time of zero of the exploration process to the one of this thinned Lévy process. In proving this, we shall employ a careful analysis of hitting times of a spectrally positive Lévy process that stochastically dominates the thinned Lévy process.

Following the proof of convergence of $|\mathcal{C}(1)| n^{-(\tau-2)/(\tau-1)}$ in Theorem 2.1, we shall prove the convergence in distribution of $(|\mathcal{C}_{\leq}(i)| n^{-(\tau-2)/(\tau-1)})_{i\in[K]}$ in Theorem 4.1. This proof makes crucial use of the estimates in the proof of Theorem 2.1, and allows to extend the result in Theorem 2.1 to the (joint) convergence of several rescaled clusters by an inductive argument. By Proposition 1.3, with high probability, the largest m clusters are given by the largest m of the vector $(|\mathcal{C}_{\leq}(i)| n^{-(\tau-2)/(\tau-1)})_{i\in[K]}$, so that this completes the proof of Theorem 1.1. This conclusion shall be carried out in Section 5 below.

In Section 6, we prove Theorem 1.5 by a second moment argument. In Section 7, we use the results proved in Section 6, jointly with the results in [3], to prove Theorem 1.4.

Throughout this paper, we shall make use of the following standard notation. We write $f(n) = O(g(n))$ for functions $f, g \geq 0$ and $n \to \infty$ if there exists a constant $C > 0$ such that $f(n) \leq Cg(n)$ in the limit, and $f(n) = o(g(n))$ if $g(n) \neq O(f(n))$. Furthermore, we write $f = \Theta(g)$ if $f = O(g)$ and $g = O(f)$. We write $O_{\mathbb{P}}(b_n)$ for a sequence of random variables X_n for which $|X_n|/b_n$ is tight as $n \to \infty$, $o_{\mathbb{P}}(b_n)$ for a sequence of random variables X_n for which $|X_n|/b_n \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$. Finally, we write that a sequence of events $(E_n)_{n\geq 1}$ occurs *with high probability* (**whp**) when $\mathbb{P}(E_n) \to 1$.

1.5 Discussion

In this section, we discuss our results in relation to several other results in the literature.

Related rank-1 inhomogeneous random graphs. The model considered here is a special case of the so-called *rank-1 inhomogeneous random graph* defined in [7]. It is asymptotically equivalent with many related models, such as the *random graph with given prescribed degrees* or Chung-Lu model, where instead

$$p_{ij} = \max\{w_i w_j / l_n, 1\}, \tag{1.22}$$

and which has been studied intensively by Chung and Lu (see [9, 10, 11, 12, 13]). A further adaptation is the *generalized random graph* [8], for which

$$p_{ij} = \frac{w_i w_j}{l_n + w_i w_j}, \tag{1.23}$$

See [17, Section 2], which in turn is based on [20], for more details on the asymptotic equivalence of such inhomogeneous random graphs. Since these models are asymptotically equivalent, all results proved here also hold for these related rank-1 models.

Comparison to the case where $\tau > 4$. In [1, 6, 26], the scaling limit was considered in the case where $\tau > 4$ and thus $\mathbb{E}[W^3] < \infty$. In this case, the scaling limit turns out to be a constant multiple times the scaling limit for the Erdős-Rényi random graph as identified in [1]. Thus, the setting when $\tau \in (3, 4)$ is fundamentally different. When $\mathbb{E}[W^3] < \infty$, the probability that $1 \in \mathcal{C}_{\max}$ is negligible, while in our setting this is not true, as shown in Theorem 1.2.

Other weights. Our proof reveals that the precise limit of $w_i n^{-1/(\tau-1)}$, for fixed $i \geq 1$, arises in the scaling limit. We make crucial use of the fact that, in our setting, by (1.7) $\lim_{n\to\infty} w_i n^{-1/(\tau-1)} = (c_F/i)^{1/(\tau-1)}$. However, we believe that also when $\lim_{n\to\infty} w_i n^{-1/(\tau-1)}$ exists for every $i \geq 1$ and is *asymptotically* equal to $a i^{-1/(\tau-1)}$, our results remain valid. This suggests that, by varying the precise values of high weights, there are many possibly scaling limits to be found. It would be of interest to investigate this further. Also, we restrict to $1 - F(x)$ that are, for large $x \geq 0$, asymptotic to an inverse power of x (see (1.4)). It would be of interest to investigate the scaling behavior when (1.4) is replaced with the assumption that $1 - F(x)$ is regularly varying with exponent $\tau - 1$. In this case, we believe that the scaling of the largest critical clusters depend on the slowly varying function, and are given by $\ell(n) n^{(\tau-2)/(\tau-1)}$ for some suitable slowly varying function $n \mapsto \ell(n)$, instead.

I.i.d. weights. In our analysis, we make crucial use of the choice for w_i in (1.3). In the literature, also the setting where $(W_i)_{i\in[n]}$ are i.i.d. random variables with distribution function F, has been considered. We expect the behavior in this model to be quite different. Indeed, then take $w_i = W_{(i)}$, where $W_{(i)}$ are the order statistics of the i.i.d. sequence $(W_i)_{i\in[n]}$. It is well known that

$$n^{-1/(\tau-1)} W_{(i)} \xrightarrow{d} \xi_i \equiv (E_1 + \cdots + E_i)^{-1/(\tau-1)}, \tag{1.24}$$

where $\{E_i\}_{i=1}^{\infty}$ are i.i.d. exponential random variables with mean 1. In particular, when $\tau \in (3, 4)$, $\mathbb{E}[\xi_1^a] < \infty$ whenever $a < \tau - 1$. The extra randomness of the order statistics has an effect on the scaling limit, which is *different*. In most cases, the two settings have the *same* behavior (see, for example, [6], where this is shown to hold for weights for which $\mathbb{E}[W^3] < \infty$, where W has distribution function F).

High-weight vertices. The fact that the vertex i is in the largest connected component with non-vanishing probability as $n \to \infty$ (see Theorem 1.2) is remarkable and invites some further discussion. In our setting, a *uniform* vertex is an element of $\mathcal{C}_{\max}$ with negligible probability. The point is that vertex i has weight w_i, which, for i fixed, is close to $\big(c_F/i\big)^{1/(\tau-1)} n^{1/(\tau-1)}$, while a uniform vertex has a bounded weight. Thus, Theorem 1.2 can be interpreted by saying that the highest weight vertices characterize the largest components. In the subcritical case (see e.g., [19]) the largest connected component is the one of the vertex with the highest weight, and the critical situation arises when the highest weight vertices start connecting to each other. Proving that this picture is correct all the way to the critical window is part of the aim of [18].

Connection to the multiplicative coalescent. In this paper we give two completely separate proofs of the convergence of the maximal components as $n \to \infty$. The first proof, comprising Sections 2 to 5, gives an inductive proof of the convergence of the maximal components, while Section 7 gives a second proof via modifications of the results in [3]. The advantage of the first proof is that the ideas developed in this are crucial for showing Theorems 1.2 and 1.3 that in this regime, unlike the regime when $\tau > 4$, there is a non-trivial asymptotic probability for the high end vertices to be in the maximal components, as well as being in the same component (in the $\tau > 4$ regime both these probabilities are asymptotically zero). The mental picture associated with the entrance boundary of the coalescent here seems to be different from [3], where in spirit many of the component sizes are of order $n^{2/3}$. Here they describe the sizes of the maximal components rescaled by $n^{-(\tau-2)/(\tau-1)}$ in the $\lambda \to -\infty$ regime, whilst in [3] they arise as limits of random graphs similar to critical Erdős-Rényi random graphs where in addition to the random edges, there are initially a number of large "planted" components of sizes $\lfloor c_i n^{2/3} \rfloor$, see [3, Section 1.3]. However, the results of [3] are crucial in identifying the distribution of the limiting component sizes for fixed λ. It would be interesting to see if the stochastic calculus techniques developed in [3] can be further modified to give useful information about the surplus of edges in the maximal components (the surplus of a component $\mathcal{C}$ with $E(\mathcal{C})$ edges and $V(\mathcal{C})$ vertices is equal to $E(\mathcal{C}) - (V(\mathcal{C}) - 1)$ and denotes the number of edges that must be removed from the component to make it a tree).

2 The scaling limit of the cluster of vertex 1

In this section, we identify the scaling limit of $|\mathcal{C}(1)|$. We note from (1.3) that the weight of vertex 1 is *maximal*, i.e., $w_1 \geq w_2 \geq \ldots \geq w_n$. When $\tau > 4$, the probability that vertex 1 belongs to $\mathcal{C}_{\max}$ is negligible. When $\tau \in (3,4)$, vertex 1 is in $\mathcal{C}_{\max}$ with *positive* probability, so that it is quite reasonable to start exploring the cluster of vertex 1 first. Theorem 2.1 below states that $|\mathcal{C}(1)|$ is of order $n^{(\tau-2)/(\tau-1)}$. By [17, Theorems 1.2 and 1.4], the same is valid for $|\mathcal{C}_{\max}|$.

Theorem 2.1 (Weak convergence of the cluster of vertex 1 for $\tau \in (3,4)$)**.** *Fix the Norros-Reittu random graph with weights $\boldsymbol{w}(\lambda) = \{(1+\lambda n^{-(\tau-3)/(\tau-1)}) w_i\}_{i=1}^n$, where $(w_i)_{i\in[n]}$ are as in (1.3). Assume that $\nu = 1$ and that (1.4) holds. Then, for all $\lambda \in \mathbb{R}$,*

$$|\mathcal{C}(1)| n^{-(\tau-2)/(\tau-1)} \xrightarrow{d} H_1(0), \tag{2.1}$$

for some non-degenerate limit $H_1(0)$.

Throughout the proof, we shall write $\tilde{w}_i = (1+\lambda n^{-(\tau-3)/(\tau-1)}) w_i$, and we denote

$$\nu_n = \frac{1}{l_n} \sum_{j\in[n]} w_j^2. \tag{2.2}$$

We also write $\tilde{\nu}_n = (1+\lambda n^{-(\tau-3)/(\tau-1)}) \nu_n$, and note that this is ν_n computed for the weights $\boldsymbol{w}(\lambda) = (\tilde{w}_i)_{i\in[n]}$.

We shall frequently make use of the asymptotics

$$\nu_n = \nu + \zeta n^{-(\tau-3)/(\tau-1)} + o(n^{-(\tau-3)/(\tau-1)}), \qquad l_n = \sum_{i=1}^{n} w_i = n\mu + O(n^{1/(\tau-1)}), \tag{2.3}$$

where μ is the mean of the distribution F and $\zeta \leq 0$. The asymptotics for l_n and the fact that $\nu_n - \nu = O(n^{-(\tau-3)/(\tau-1)})$ follow from [17, Cor. 4.2]. The sharper asymptotics for ν_n in (2.3) is obtained by a more careful analysis of the arising sum, which is deferred to Lemma A.1 in the appendix. Thus, we shall use that in the critical regime where $\nu = 1$,

$$\tilde{\nu}_n = 1 + \theta n^{-(\tau-3)/(\tau-1)} + o(n^{-(\tau-3)/(\tau-1)}), \tag{2.4}$$

where $\theta = \lambda + \zeta$. The parameter $\theta \in \mathbb{R}$ indicates the location inside the critical window formed by the weights $\boldsymbol{w}(\lambda)$. Indeed, in the asymptotics for $\tilde{\nu}_n$ in (2.4), the fact that $\theta = \zeta + \lambda$ arises from $\tilde{\nu}_n = (1 + \lambda n^{-(\tau-3)/(\tau-1)})\nu_n$, together with the sharp asymptotics of ν_n in (2.3). The value of ζ is constant and does not depend λ, while the value of λ indicates the location inside the scaling window, so we can, alternatively, measure the location inside the scaling window by $\theta \in \mathbb{R}$.

In order to prove Theorem 2.1, we make heavy use of the *cluster exploration*, which is described in detail in [23] and [17]. The model in [23] is a *random multigraph*, i.e., a random graph potentially having self-loops and multiple edges. Indeed, for each $i, j \in [n]$, we let the number of edges between vertex i and j be $\mathrm{Poi}(w_i w_j / l_n)$, where, for $\lambda \geq 0$, we let $\mathrm{Poi}(\lambda)$ denote a Poisson random variable with mean λ. The number of edges between different pairs of vertices are *independent*. To retrieve our random graph model, we merge multiple edges and erase self-loops. Then, the probability that an edge exists between two vertices $i, j \in [n]$ is equal to

$$p_{ij} = \mathbb{P}\big(\mathrm{Poi}(w_i w_j / l_n) \geq 1\big) = 1 - \mathrm{e}^{-w_i w_j / l_n}, \tag{2.5}$$

as required. Further, the number of edges from a vertex i has a Poisson distribution with mean w_i. We shall work with the above Poisson random graph instead, and we shall refer to the Poisson random variable $\mathrm{Poi}(w_i)$ as the number of *potential neighbors*. When we find what the vertices are that correspond to these $\mathrm{Poi}(w_i)$ potential neighbors, i.e., when we determine their *marks*, then we can see how many real neighbors there are.

We denote by $(Z_l)_{l \geq 0}$ the exploration process in the breadth-first search, where $Z_0 = 1$ and Z_l denotes the number of potential elements in the cluster of the initial vertex, which is in the case of Theorem 2.1 equal to vertex 1, of which we have not yet explored their neighbors. Thus, we set $Z_0 = 1, Z_1 = \mathrm{Poi}(\tilde{w}_1)$, and note that, for $l \geq 2$, Z_l satisfies the recursion relation

$$Z_l = Z_{l-1} + X_l - 1, \tag{2.6}$$

where X_l denotes the number of potential neighbors of the l^{th} vertex which is explored. In [23, Proposition 3.1] (see also [17, Section 4.2]), the cluster exploration was described in terms of a thinned marked mixed Poisson branching process. This implies that the distribution of X_l (for $2 \leq l \leq n$) is equal to $\mathrm{Poi}(\tilde{w}_{M_l}) J_l$, where the marks $\{M_l\}_{l=1}^{\infty}$ are i.i.d. random variables with distribution

$$\mathbb{P}(M = m) = \frac{w_m}{l_n}, \quad 1 \leq m \leq n, \tag{2.7}$$

and $J_l = \mathbb{1}_{\{M_l \notin \{1\} \cup \{M_1, \ldots, M_{l-1}\}\}}$ is the indicator that the mark M_l has not been found before and is not equal to 1. Here, the mark M_l is the label of the potential element of the cluster that we are exploring, and, clearly, if a vertex has already been observed to be part of $\mathcal{C}(1)$ and its neighbors have been explored, then we should not do so again. We call the drawing of the random mark a *vertex check*. We conclude that we arrive at, for $l \geq 2$,

$$Z_l = Z_{l-1} + \mathrm{Poi}(\tilde{w}_{M_l}) J_l - 1. \tag{2.8}$$

Then, the number of vertex checks that have been performed when exploring the cluster of vertex 1 equals $V(1)$, which is given by

$$V(1) = \min\{l : Z_l = 0\}, \tag{2.9}$$

since the first time at which there are no more vertices to be checked, all vertices in the cluster have been checked.

Further, the number of real vertices found to be part of $\mathcal{C}(1)$ after l vertex checks equals

$$|\mathcal{C}(1; l)| = 1 + \sum_{j=2}^{l} J_j. \tag{2.10}$$

Therefore, we conclude that

$$|\mathcal{C}(1)| = |\mathcal{C}_{\leq}(1)| = 1 + \sum_{j=2}^{V(1)} J_j = V(1) - \sum_{j=2}^{V(1)} (1 - J_j). \tag{2.11}$$

It turns out that the second contribution is an error term (see Lemma 3.8 below), so that the cluster size of 1 asymptotically corresponds to the first hitting time of 0 of $l \mapsto Z_l$. Theorem 2.1 shall follow from the fact that we can identify the scaling limit of the process $(Z_l)_{l\geq 0}$. To identify this scaling limit, we let

$$\mathcal{Z}_t^{(n)} = n^{-1/(\tau-1)} Z_{tn^{(\tau-2)/(\tau-1)}}. \tag{2.12}$$

The intuition behind (2.12) is as follows. First, since the largest connected components are of order $n^{(\tau-2)/(\tau-1)}$ as proved in [17, Theorems 1.2 and 1.4], and the successive elapsed time between hits of zero of the process $(Z_l)_{l\geq 0}$ correspond to the cluster sizes, the relevant time scale is $tn^{(\tau-2)/(\tau-1)}$. Further, by Proposition 1.3, we see that the large clusters correspond to the clusters of the high-weight vertices. The maximal weight is of the order $n^{1/(\tau-1)}$, so that this needs to be the relevant scale on which the process Z_l runs. The proof below shall make this intuition precise.

In order to define the scaling limit, we define the non-negative continuous-time process $(\mathcal{S}_t)_{t\geq 0}$. For some $a > 0$, we let $\{\mathcal{I}_i(t)\}_{i=1}^{\infty}$ denote an independent increasing indicator processes with

$$\mathbb{P}(\mathcal{I}_i(s) = 0 \ \forall s \in [0, t]) = \mathrm{e}^{-ati^{-1/(\tau-1)}}. \tag{2.13}$$

We further let, for some $b > 0$ and $c \in \mathbb{R}$, and a as in (2.13),

$$\mathcal{S}_t = b - abt + ct + \sum_{i=2}^{\infty} \frac{b}{i^{1/(\tau-1)}} [\mathcal{I}_i(t) - \frac{at}{i^{1/(\tau-1)}}], \tag{2.14}$$

for all $t \geq 0$. We call $(\mathcal{S}_t)_{t\geq 0}$ a *thinned Lévy process*, a name we shall explain in more detail after the theorem. The main result is the following theorem:

Theorem 2.2 (The scaling limit of Z_l)**.** *As $n \to \infty$, under the conditions of Theorem 1.1,*

$$(\mathcal{Z}_t^{(n)})_{t\geq 0} \xrightarrow{d} (\mathcal{S}_t)_{t\geq 0}, \tag{2.15}$$

where $a = c_F^{1/(\tau-1)}/\mu$, $b = c_F^{1/(\tau-1)}$, $c = \theta$, in the sense of convergence in the J_1 Skorokhod topology on the space of right-continuous left-limited functions on $\mathbb{R}^+$. Consequently, $H_1(0)$ is the hitting time of 0 of $(\mathcal{S}_t)_{t\geq 0}$.

It is worthwhile to note that while the convergence in Theorem 2.2 only has implications for our random graph for $t \leq H_1(0)$, which is the hitting time of zero of the process $(\mathcal{S}_t)_{t\geq 0}$, the processes

$(\mathcal{S}_t^{(n)})_{t\geq 0}$ and $(\mathcal{S}_t)_{t\geq 0}$ are well defined also for larger t, and convergence holds for *all* t. This shall, in fact, be useful in the proof.

The proof of Theorem 2.2 shall be given in Section 3 below. We now first discuss the limiting process $(\mathcal{S}_t)_{t\geq 0}$ and its connection to Lévy processes. To do this, we denote by $(\mathcal{R}_t)_{t\geq 0}$ the process given by

$$\mathcal{R}_t = b - abt + ct + \sum_{i=2}^{\infty} \frac{b}{i^{1/(\tau-1)}}[N_i(t) - \frac{at}{i^{1/(\tau-1)}}], \tag{2.16}$$

where N_i are independent Poisson processes with rates $\frac{at}{i^{1/(\tau-1)}}$. Clearly, the process $(\mathcal{R}_t)_{t\geq 0}$ is a spectrally positive Lévy process (see e.g., [4, 22] for more information on Lévy processes), with exponent $\psi(\theta)$ (for which $\mathbb{E}(\mathrm{e}^{-\theta(\mathcal{R}_t-\mathcal{R}_0)}) = \mathrm{e}^{-t\psi(\theta)}$) given by

$$\psi(\theta) = (c-ab)\theta + \sum_{i=2}^{\infty} \frac{a}{i^{1/(\tau-1)}}\Big[1 - \mathrm{e}^{-\theta\frac{b}{i^{1/(\tau-1)}}} - \frac{bt\theta}{i^{1/(\tau-1)}}\Big]. \tag{2.17}$$

Alternatively, the exponent $\psi(\theta)$ can be expressed in terms of the integral

$$\psi(\theta) = (c-ab)\theta - \theta\int_1^{\infty} \Pi(dx) + \int_0^{\infty} (\mathrm{e}^{\theta x} - 1 - \theta\mathbb{1}_{\{x<1\}})\Pi(dx), \tag{2.18}$$

where the Lévy measure Π is defined as

$$\Pi(dx) = \frac{a}{b}\sum_{i=2}^{\infty} x\delta_{x,\frac{b}{i^{1/(\tau-1)}}}. \tag{2.19}$$

Thus, since $\Pi(-\infty,0) = 0$, the Lévy process is spectrally positive, so that the process $(\mathcal{R}_t)_{t\geq 0}$ has only positive jumps. Also, $\Pi(b,\infty) = 0$, so that the jumps of $(\mathcal{R}_t)_{t\geq 0}$ are bounded by b. Finally,

$$\int_0^{\infty} (1\wedge x^2)\Pi(dx) \leq \int_0^{\infty} x^2\Pi(dx) = \frac{a}{b}\sum_{i=2}^{\infty}\Big(\frac{b}{i^{1/(\tau-1)}}\Big)^3 = ab^2\sum_{i=2}^{\infty} i^{-3/(\tau-1)} < \infty, \tag{2.20}$$

since $\tau \in (3,4)$ so that $3/(\tau-1) > 1$. Therefore, the process $(\mathcal{R}_t)_{t\geq 0}$ is a well-defined Lévy process.

We may reformulate (2.14) as

$$\mathcal{S}_t = b - abt + ct + \sum_{i=2}^{\infty} \frac{b}{i^{1/(\tau-1)}}[\mathbb{1}_{\{N_i(t)\geq 1\}} - \frac{at}{i^{1/(\tau-1)}}], \tag{2.21}$$

so that the process $(\mathcal{S}_t)_{t\geq 0}$ does not include multiple counts of the independent processes $(N_i(t))_{t\geq 0}$. This is the reason that we call the process $(\mathcal{S}_t)_{t\geq 0}$ a *thinned* Lévy process. In [3], this process is called a Lévy process without repetitions. Naturally, we have that the descriptions in (2.16) and (2.21) satisfy that, a.s., for all $t \geq 0$,

$$\mathcal{S}_t \leq \mathcal{R}_t. \tag{2.22}$$

This allows us to make use of Lévy process methodology at various places in our proofs. We do note that $\mathcal{R}_t$ is a rather poor approximation for $\mathcal{S}_t$, particularly on large time scales, because the *thinning* becomes more important as time progresses.

3 Proof of Theorems 2.1 and 2.2

In this section, we prove Theorems 2.1 and 2.2. We start by proving Theorem 2.2 in Section 3.1, and make use of Theorem 2.2 to prove Theorem 2.1 in Section 3.2.

3.1 Proof of Theorem 2.2

Instead of $(Z_l)_{l\geq 0}$, it shall be convenient to work with a related process $(S_l)_{l\geq 0}$, which is defined as $S_0 = 1, S_1 = \tilde{w}_1$ and satisfies the recursion relation, for $l \geq 2$,

$$S_l = S_{l-1} + \tilde{w}_{M_l} J_l - 1, \tag{3.1}$$

i.e., the Poisson random variables $\mathrm{Poi}(\tilde{w}_{M_l})$ appearing in the recursion for Z_l in (2.8) are replaced with their (random) weights $\tilde{w}_{M_l}$. We shall first show that S_l and Z_l are quite close:

Lemma 3.1 (The difference between S_l and Z_l is small)**.** *For any $m \geq 0$,*

$$\sup_{l\leq m} |Z_l - S_l| = O_{\mathbb{P}}(m^{1/2}). \tag{3.2}$$

Proof. We have that $\{Z_l - S_l\}_{l\geq 0}$ is a martingale w.r.t. the filtration $\mathcal{F}_l = \sigma(\{M_i\}_{i=1}^l)$. Therefore, by the Doob-Kolmogorov inequality [16, Theorem (7.8.2), p. 338],

$$\mathbb{P}\big(\sup_{l\leq m} |Z_l - S_l| > \varepsilon\sqrt{m}\big) \leq \frac{1}{m\varepsilon^2}\mathbb{E}\Big[|Z_m - S_m|^2\Big]. \tag{3.3}$$

Now,

$$\mathbb{E}\Big[|Z_m - S_m|^2\Big] = \mathbb{E}\Big[\mathbb{E}\big[|Z_m - S_m|^2 \mid \{M_i\}_{i=1}^m\big]\Big] = \mathbb{E}\big[\sum_{l=1}^m \tilde{w}_{M_l} J_l\big] \leq \mathbb{E}\big[\sum_{l=1}^m \tilde{w}_{M_l}\big] = m\tilde{\nu}_n = m(1+o(1)), \tag{3.4}$$

by (2.4). This proves the claim. □

We proceed by investigating the scaling limit of $(S_l)_{l\geq 1}$. For this, we define

$$\mathcal{S}_t^{(n)} = n^{-1/(\tau-1)} S_{tn^{(\tau-2)/(\tau-1)}}. \tag{3.5}$$

We shall prove that

$$(\mathcal{S}_t^{(n)})_{t\geq 0} \xrightarrow{d} (\mathcal{S}_t)_{t\geq 0}, \tag{3.6}$$

which shall be enough to prove Theorem 2.2. Indeed, to see that (3.6) proves Theorem 2.2, we note that by Lemma 3.1, for every $t = o(n^{(4-\tau)/(\tau-1)})$,

$$\sup_{s\leq t} |\mathcal{Z}_t^{(n)} - \mathcal{S}_t^{(n)}| = O_{\mathbb{P}}\big(\sqrt{t} n^{\frac{4-\tau}{2(\tau-1)}}\big) = o_{\mathbb{P}}(1). \tag{3.7}$$

We continue with the proof of (3.6). We investigate the process $\mathcal{S}_t^{(n)}$ up to the first time it hits 0. We shall prove that, due to (2.11) and Lemma 3.1, this hitting time is close to $n^{-(\tau-2)/(\tau-1)}|\mathcal{C}_{\leq}(1)|$. We note that, by (3.1),

$$S_l = \tilde{w}_1 + \sum_{i\in\mathcal{V}_l^{(n)}} \tilde{w}_i - l = \tilde{w}_1 + \sum_{i=2}^n \tilde{w}_i \mathcal{I}_i^{(n)}(l) - l, \tag{3.8}$$

where

$$\mathcal{I}_i^{(n)}(l) = \mathbb{1}_{\{i\in\mathcal{V}_l^{(n)}\}}, \tag{3.9}$$

and

$$\mathcal{V}_l^{(n)} = \bigcup_{j=2}^l \{M_j\}. \tag{3.10}$$

Using that

$$\tilde{\nu}_n = \sum_{i=1}^{n} \frac{\tilde{w}_i w_i}{l_n}, \tag{3.11}$$

we can rewrite this as

$$S_l = \tilde{w}_1 - \frac{\tilde{w}_1 w_1 l}{l_n} + \sum_{i=2}^{n} \tilde{w}_i \big[\mathcal{I}_i^{(n)}(l) - \frac{w_i l}{l_n}\big] + (\tilde{\nu}_n - 1)l. \tag{3.12}$$

Now we recall (see (2.4))

$$l = tn^{(\tau-2)/(\tau-1)}, \qquad \tilde{\nu}_n - 1 = \theta n^{-(\tau-3)/(\tau-1)}\nu_n, \tag{3.13}$$

and we recall from (1.4) that, for i such that $i/n \to \infty$,

$$w_i = [1-F]^{-1}(i/n) = b(n/i)^{1/(\tau-1)}(1+o(1)), \tag{3.14}$$

where $b = c_F^{1/(\tau-1)}$ and c_F is defined in (1.4). As a result,

$$\begin{aligned} \mathcal{S}_t^{(n)} &= n^{-1/(\tau-1)} S_{tn^{(\tau-2)/(\tau-1)}} \\ &= b - \frac{b^2}{\mu}t + \sum_{i=2}^{n} n^{-1/(\tau-1)} \tilde{w}_i \big[\mathcal{I}_i^{(n)}(tn^{(\tau-2)/(\tau-1)}) - n^{-1/(\tau-1)}\frac{w_i t}{\mu_n}\big] + \theta t + o(1), \end{aligned} \tag{3.15}$$

where we write $\mu_n = l_n/n = \mu + o(1)$.

We proceed by showing that the sum in (3.15) is predominantly carried by the first few terms. For this, we compute the variance of the sum over $i \geq K$ for K large. We start by noting that $\mathcal{I}_i^{(n)}(l)$ is the indicator that $i \in \mathcal{V}_l^{(n)}$, and $\mathcal{V}_l^{(n)}$ contains the first l marks drawn, where the marks $\{M_i\}_{i=1}^{l}$ are i.i.d. with distribution given by (2.7). Therefore, $\mathcal{I}_i^{(n)}(l)$ and $\mathcal{I}_j^{(n)}(l)$ are, for different i, j, negatively correlated, so that

$$\begin{aligned} \mathrm{Var}\Big(\sum_{i=K}^{n} n^{-1/(\tau-1)} \tilde{w}_i \big[\mathcal{I}_i^{(n)}(l) - \frac{w_i l}{l_n}\big]\Big) &= \mathrm{Var}\Big(\sum_{i=K}^{n} n^{-1/(\tau-1)} \tilde{w}_i \mathcal{I}_i^{(n)}(l)\Big) \\ &\leq \sum_{i=K}^{n} \big(n^{-1/(\tau-1)}\tilde{w}_i\big)^2 \mathrm{Var}\big(\mathcal{I}_i^{(n)}(l)\big). \end{aligned} \tag{3.16}$$

Now,

$$\mathrm{Var}\big(\mathcal{I}_i^{(n)}(l)\big) = \mathbb{P}\big(\mathcal{I}_i^{(n)}(l) = 0\big)\big(1 - \mathbb{P}\big(\mathcal{I}_i^{(n)}(l) = 0\big)\big), \tag{3.17}$$

and

$$1 - \mathbb{P}\big(\mathcal{I}_i^{(n)}(l) = 0\big) = 1 - (1 - \frac{w_i}{l_n})^{l-1} \leq \frac{w_i l}{l_n}, \tag{3.18}$$

so that

$$\mathrm{Var}\big(\mathcal{I}_i^{(n)}(l)\big) \leq \frac{w_i l}{l_n}. \tag{3.19}$$

Therefore, when $l = tn^{(\tau-2)/(\tau-1)}$,

$$\begin{aligned} \mathrm{Var}\Big(\sum_{i=K}^{n} n^{-1/(\tau-1)} \tilde{w}_i \big[\mathcal{I}_i^{(n)}(l) - \frac{w_i l}{l_n}\big]\Big) &\leq \sum_{i=K}^{n} \big(n^{-1/(\tau-1)}\tilde{w}_i\big)^2 \frac{w_i tn^{(\tau-2)/(\tau-1)}}{l_n} \\ &\leq Ct \sum_{i=K}^{n} i^{-3/(\tau-1)} \leq CtK^{1-3/(\tau-1)} = o(1), \end{aligned} \tag{3.20}$$

when $K \to \infty$, since $\tau \in (3,4)$. Moreover,

$$M_l^{(n,K)} = \sum_{i=K}^{n} n^{-1/(\tau-1)} \tilde{w}_i \big[\mathcal{I}_i^{(n)}(l) - n^{-1/(\tau-1)} \frac{w_i l}{l_n}\big] \tag{3.21}$$

is a supermartingale since

$$\begin{aligned} &\mathbb{E}[M_{l+1}^{(n,K)} - M_l^{(n,K)} \mid \{\mathcal{I}_i^{(n)}(l)\}_{i=1}^n] \\ &\quad = \mathbb{E}\Big[\sum_{i=K}^{n} n^{-1/(\tau-1)} \tilde{w}_i \big[\mathcal{I}_i^{(n)}(l+1) - \mathcal{I}_i^{(n)}(l) - n^{-1/(\tau-1)} \frac{w_i}{l_n}\big] \mid \{\mathcal{I}_i^{(n)}(l)\}_{i=1}^n\Big] \\ &\quad \le \sum_{i=K}^{n} n^{-1/(\tau-1)} \tilde{w}_i (1 - \mathcal{I}_i^{(n)}(l)) \Big(\mathbb{E}\big[\mathcal{I}_i^{(n)}(l+1) \mid \{\mathcal{I}_i^{(n)}(l)\}_{i=1}^n\big] - n^{-1/(\tau-1)} \frac{w_i}{l_n}\Big) = 0. \end{aligned} \tag{3.22}$$

Therefore, by the maximal inequality [16, Theorem (12.6.1), p. 496] and the Cauchy-Schwarz inequality, and writing $m = tn^{(\tau-2)/(\tau-1)}$,

$$\mathbb{P}\big(\max_{l \le m} M_l^{(n,K)} \ge x\big) \le \frac{\mathbb{E}[M_0^{(n,K)}] + \mathbb{E}[|M_m^{(n,K)}|]}{x} \le c\sqrt{t} K^{(\tau-4)/2(\tau-1)}. \tag{3.23}$$

Since $\tau < 4$, we obtain that, for $K \ge 1$ large and uniformly in n, $\mathbb{P}(\max_{l \le m} M_l^{(n,K)} \ge \varepsilon) \le \varepsilon$.

We now summarize the statements in Lemma 3.1 and (3.23). For this, we denote

$$\mathcal{S}_t^{(n,K)} = b - \frac{b^2}{\mu} t + \sum_{i=2}^{K} n^{-1/(\tau-1)} \tilde{w}_i \big[\mathcal{I}_i^{(n)}(tn^{(\tau-2)/(\tau-1)}) - n^{-1/(\tau-1)} \frac{w_i t}{\mu_n}\big] + \theta t, \tag{3.24}$$

$$\mathcal{S}_t^{(\infty,K)} = b - \frac{b^2}{\mu} t + \sum_{i=2}^{K} b i^{-1/(\tau-1)} \big[\mathcal{I}_i(t) - a i^{-1/(\tau-1)}\big] + \theta t. \tag{3.25}$$

Then we obtain the following corollary:

Corollary 3.2 (Finite sum approximation of Z)**.** *For every $\varepsilon, \delta > 0$, there exists $K > 0$ and $N \ge 1$ such that for all $n \ge N$, for all $u \le \infty$,*

$$\mathbb{P}(\sup_{t \le u} \big|\mathcal{Z}_t^{(n)} - \mathcal{S}_t^{(n,K)}\big| \ge \delta) \le \varepsilon. \tag{3.26}$$

Consequently, with

$$H^{(n)}(x) = \inf\{t : \mathcal{Z}_t^{(n)} = x\}, \qquad H^{(n,K)}(x) = \inf\{t : \mathcal{S}_t^{(n,K)} = x\}, \tag{3.27}$$

it follows that, for every $\varepsilon, \delta > 0$, there exists $K > 0$ and $N \ge 1$ such that, for all $n \ge N$,

$$\mathbb{P}\big(H^{(n,K)}(x+\delta) \le H^{(n)}(x) \le H^{(n,K)}(x-\delta)\big) \ge 1 - \varepsilon. \tag{3.28}$$

The above suggests that it suffices to investigate $(\mathcal{I}_i^{(n)}(tn^{(\tau-2)/(\tau-1)}))_{i\in[K]}$. That is the content of the following lemma:

Lemma 3.3 (Convergence of indicators)**.** *As $n \to \infty$, for all $K \ge 1$ and $T > 0$,*

$$(\mathcal{I}_i^{(n)}(tn^{(\tau-2)/(\tau-1)}))_{i\in[K], t\in[0,T]} \xrightarrow{d} (\mathcal{I}_i(t))_{i\in[K], t\in[0,T]}. \tag{3.29}$$

As a consequence, for all $K \ge 1$ and $T > 0$,

$$(\mathcal{S}_t^{(n,K)})_{t\in[0,T]} \xrightarrow{d} (\mathcal{S}_t^{(\infty,K)})_{t\in[0,T]}. \tag{3.30}$$

In both statements, $\xrightarrow{d}$ denotes convergence in the J_1 Skorokhod topology on the space of right-continuous left-limited functions on $\mathbb{R}^+$.

Proof. Since $\mathcal{I}_i^{(n)}(tn^{(\tau-2)/(\tau-1)})$ are all indicator processes of the form

$$\mathcal{I}_i^{(n)}(tn^{(\tau-2)/(\tau-1)}) = \mathbb{1}_{\{T_i \le tn^{(\tau-2)/(\tau-1)}\}}, \tag{3.31}$$

where T_i is the first time that mark i is chosen, it suffices to prove that

$$(n^{(\tau-2)/(\tau-1)}T_i)_{i\in[K]} \xrightarrow{d} (E_i)_{i\in[K]}, \tag{3.32}$$

where E_i are independent exponentials with mean $ai^{-1/(\tau-1)}$. For this, in turn, it suffices to prove that, for every sequence $t_1, \dots, t_K$,

$$\mathbb{P}(n^{(\tau-2)/(\tau-1)}T_i > t_i \ \forall i \in [K]) \to \mathrm{e}^{-a\sum_{i=1}^K i^{-1/(\tau-1)}t_i}. \tag{3.33}$$

The latter is equivalent to

$$\mathbb{P}\big(\mathcal{I}_i^{(n)}(t_i n^{(\tau-2)/(\tau-1)}) = 0 \ \forall i \in [K]\big) \to \mathbb{P}(\mathcal{I}_i(t_i) = 0 \ \forall i \in [K]) = \mathrm{e}^{-a\sum_{i=1}^K i^{-1/(\tau-1)}t_i}. \tag{3.34}$$

Now, since the marks are i.i.d., we obtain that

$$\mathbb{P}\big(\mathcal{I}_i^{(n)}(m_i) = 0 \ \forall i \in [K]\big) = \prod_{l=1}^{\infty} \mathbb{P}(M_l \notin \{i \in [K] : l \le m_i\}) = \prod_{l=1}^{\infty} \big(1 - \sum_{i:l\le m_i} \frac{w_i}{l_n}\big). \tag{3.35}$$

A Taylor expansion gives that

$$\mathbb{P}\big(\mathcal{I}_i^{(n)}(m_i) = 0 \ \forall i \in [K]\big) = \mathrm{e}^{-\sum_{l=1}^n \sum_{i:m_i\le l} \frac{w_i}{l_n} + o(1)} = \mathrm{e}^{-\sum_{i\in[K]} \frac{w_i m_i}{l_n} + o(1)}. \tag{3.36}$$

Applying this to $m_i = t_i n^{(\tau-2)/(\tau-1)}$, and noting that for this choice

$$\frac{m_i w_i}{l_n} = \frac{b i^{-1/(\tau-1)} t_i}{\mu}, \tag{3.37}$$

we arrive at the claim in (3.29) with $a = b/\mu$. The claim in (3.30) follows from the fact that, by (3.24), $\mathcal{S}_t^{(n,K)}$ is a weighted sum of the $(\mathcal{I}_i^{(n)}(tn^{(\tau-2)/(\tau-1)}))_{i\in[K]}$, and the (deterministic) weights converge. Thus, the continuous mapping theorem gives the claim. □

Proof of Theorem 2.2. By (3.23), **whp** we can restrict the sum in the definition of S_l in (3.8) to the first $[K]$ terms. Then, we have a sum of stochastic processes that by Lemma 3.3 converge weakly to the independent indicator processes in (2.21). This proves the claim. □

3.2 Proof of Theorem 2.1

In this section, we give a proof of Theorem 2.1. We start by looking at the first hitting time of zero of the process $l \mapsto Z_l$, and use the fact that by (2.11), $V(1) = \min\{l : Z_l = 0\}$, where $V(1)$ denotes the number of vertex checks performed in exploring the cluster of vertex 1. The proof shall proceed as follows. We shall first use Theorem 2.2 and Lemma 3.1 to prove that $V(1)n^{-(\tau-2)/(\tau-1)}$ converges in distribution to $H_{\mathcal{S}}(0)$, where $H_{\mathcal{S}}(0)$ denotes the first hitting time of 0 of the process $(\mathcal{S}_t)_{t\ge0}$ (see Proposition 3.4 below). We then prove that $V(1)n^{-(\tau-2)/(\tau-1)}$ and $\mathcal{C}(1)n^{-(\tau-2)/(\tau-1)}$ have identical scaling limits, by looking at the second term in (2.11) (see Lemma 3.8 below). We then complete the proof of Theorem 2.1. Finally, we state and prove an auxiliary result useful in the proof of Theorems 1.1, 1.2 and 4.1, and which will play a crucial role in the next section, where we investigate the scaling limit of several clusters simultaneously.

By Theorem 2.2 and Lemma 3.1, the process $(\mathcal{Z}_t^{(n)})_{t\ge0}$, where $\mathcal{Z}_t^{(n)} = n^{-1/(\tau-1)}Z_{tn^{(\tau-2)/(\tau-1)}}$ converges in distribution to the process $(\mathcal{S}_t)_{t\ge0}$. Note that

$$n^{-(\tau-2)/(\tau-1)}V(1) = \min\{t : \mathcal{Z}_t^{(n)} = 0\} = H^{(n)}(0). \tag{3.38}$$

By Corollary 3.2, we have that

$$\mathbb{P}(H^{(n,K)}(\delta) > u) - \varepsilon \le \mathbb{P}(H^{(n)}(0) > u) \le \mathbb{P}(H^{(n,K)}(-\delta) > u) + \varepsilon. \tag{3.39}$$

Therefore, to prove convergence in distribution of $n^{-(\tau-2)/(\tau-1)}V(1)$, it suffices to prove the following proposition:

Proposition 3.4 (Convergence of hitting times)**.** *For every* $u \in [0,\infty)$*,*

$$\lim_{\delta\to 0}\lim_{K\to\infty}\lim_{n\to\infty}\mathbb{P}(H^{(n,K)}(\delta) > u) = \mathbb{P}(H_{\mathcal{S}}(0) > u), \tag{3.40}$$

where

$$H_{\mathcal{S}}(x) = \inf\{t\colon \mathcal{S}_t \le x\} \tag{3.41}$$

is the first hitting time of level x *of* $(\mathcal{S}_t)_{t\ge 0}$*.*

Proof. We shall prove the upper bound only, the proof for the lower bound being equivalent. We use that the process $(\mathcal{S}_t^{(n,K)})_{t\in[0,T]}$ converges in distribution to $(\mathcal{S}_t^{(\infty,K)})_{t\in[0,T]}$ in the sense of convergence in the J_1 Skorokhod topology on the space of right-continuous left-limited functions on $\mathbb{R}^+$. This implies that, uniformly in $n \ge 1$ sufficiently large, and for any $\eta > 0$ sufficiently small,

$$\mathbb{P}(H^{(n,K)}(-\delta) > u) \le \mathbb{P}(H^{(\infty,K)}(-\delta-\eta) > u-\eta). \tag{3.42}$$

Further, **whp**, we have that, as $K \to \infty$,

$$\sup_{u\le t}|\mathcal{S}_t - \mathcal{S}_t^{(\infty,K)}| \le \eta. \tag{3.43}$$

Therefore, for sufficiently large $K \ge 1$,

$$\mathbb{P}(H^{(n,K)}(-\delta) > u) \le \mathbb{P}(H_{\mathcal{S}}(-\delta-2\eta) > u-\eta) \le \mathbb{P}(H_{\mathcal{S}}(-2\delta) > u-\eta), \tag{3.44}$$

when we take $\eta \le \delta/2$. We conclude that it suffices to show that

$$\lim_{\delta\downarrow 0}\lim_{\eta\downarrow 0}\mathbb{P}(H_{\mathcal{S}}(-\delta) > u-\eta) = \mathbb{P}(H_{\mathcal{S}}(0) > u). \tag{3.45}$$

This is a kind of *joint continuity* of the distribution of the hitting time, both in space and time.

We shall continue to explore the proof of (3.45). We let $\mathbb{P}_x$ denote the distribution of the process $(\mathcal{S}_t)_{t\ge 0}$ when $\mathcal{S}_0 = x$. Then,

$$\mathbb{P}(H_{\mathcal{S}}(-\delta) > u-\eta) = \mathbb{P}_b(H_{\mathcal{S}}(-\delta) > u-\eta) = \mathbb{E}_b\big[\mathbb{1}_{\{H_{\mathcal{S}}(\delta)>u-\eta\}}\mathbb{P}_b(H_{\mathcal{S}}(-\delta) > u-\eta \mid \mathcal{F}_{H_{\mathcal{S}}(\delta)})\big]. \tag{3.46}$$

The process $(\mathcal{S}_{t+H_{\mathcal{S}}(\delta)} - \delta)_{t\ge 0}$ is not quite a Markov process, but it is, conditionally on $\mathcal{F}_{H_{\mathcal{S}}(\delta)}$, in law equal to the process $(\mathcal{S}_t^{(\delta)})_{t\ge 0}$ given by

$$\mathcal{S}_t^{(\delta)} = -abt + ct - t\sum_{i\in\mathcal{V}_{H_{\mathcal{S}}(\delta)}} abi^{-2/(\tau-1)} + \sum_{i\in\mathcal{U}_t^c\setminus\mathcal{V}_{H_{\mathcal{S}}(\delta)}} bi^{-1/(\tau-1)}[\mathcal{I}_i(t) - \frac{at}{i^{1/(\tau-1)}}]. \tag{3.47}$$

In particular, $\mathcal{S}_t^{(\delta)}$ is stochastically dominated by $\mathcal{S}_t$. Thus,

$$\mathbb{P}_b(H_{\mathcal{S}}(-\delta) - H_{\mathcal{S}}(\delta) > a \mid \mathcal{F}_{H_{\mathcal{S}}(\delta)}) \le \mathbb{P}_\delta(H_{\mathcal{S}}(-\delta) > a). \tag{3.48}$$

Therefore, in order to prove (3.45), we split

$$\begin{aligned}&\mathbb{P}(H_{\mathcal{S}}(-\delta) > u-\eta) - \mathbb{P}(H_{\mathcal{S}}(0) > u)\\ &\quad= [\mathbb{P}(H_{\mathcal{S}}(-\delta) > u-\eta) - \mathbb{P}(H_{\mathcal{S}}(0) > u-2\eta)] + [\mathbb{P}(H_{\mathcal{S}}(0) > u-2\eta) - \mathbb{P}(H_{\mathcal{S}}(0) > u)],\end{aligned} \tag{3.49}$$

and note that the second term converges to zero when $\eta \to 0$, since we shall show that $H_{\mathcal{S}}(0)$ has a continuous distribution, while the first term is bounded by

$$\begin{aligned}\mathbb{P}(H_{\mathcal{S}}(-\delta) > u - \eta) - \mathbb{P}(H_{\mathcal{S}}(0) > u - 2\eta) &= \mathbb{P}_b(H_{\mathcal{S}}(-\delta) > u - \eta, H_{\mathcal{S}}(0) \le u - 2\eta) \\ &\le \mathbb{P}(H_{\mathcal{S}}(-\delta) - H_{\mathcal{S}}(0) > \eta) \le \mathbb{P}_\delta(H_{\mathcal{S}}(-\delta) > \eta). \qquad (3.50)\end{aligned}$$

Now, we observe that by (2.22), $\mathcal{S}_t \le \mathcal{R}_t$ a.s. for all $t \ge 0$, where $(\mathcal{R}_t)_{t\ge0}$ is the Lévy process defined in (2.16), and hence

$$\mathbb{P}_\delta(H_{\mathcal{S}}(-\delta) > \eta) \le \mathbb{P}_\delta(H_{\mathcal{R}}(-\delta) > \eta), \qquad (3.51)$$

where, for a stochastic process $\mathcal{X}_t$, we let

$$H_{\mathcal{X}}(x) = \inf\{t : \mathcal{X}_t \le x\} \qquad (3.52)$$

be its first hitting time of x. We conclude that, to prove Proposition 3.4, it suffices to show that

$$\lim_{\eta\downarrow0}\lim_{\delta\downarrow0} \mathbb{P}_\delta(H_{\mathcal{R}}(-\delta) > \eta) = 0. \qquad (3.53)$$

Therefore, we are lead to study the probability that the Lévy process $(\mathcal{R}_t)_{t\ge0}$ started from $\delta > 0$, which is small, has a large hitting time of 0.

Because $(\mathcal{R}_t)_{t\ge0}$ has independent increments, and only non-negative jumps, we have the following result (see e.g. [24, (1.5)]):

Lemma 3.5 (Hitting times for spectrally positive Lévy processes)**.** *For $\psi(\theta)$ defined in (2.17) we have that $\psi(\theta) \to -\infty$ as $\theta \to \infty$. Suppose $b \ge 0$.*
(a) For each $s > 0$ there is a unique positive solution $\theta = \eta(s)$ of $\psi(\theta) = -s$, and

$$\mathbb{E}_b(\mathrm{e}^{-sH_{\mathcal{R}}(0)}) = \mathrm{e}^{-b\eta(s)}. \qquad (3.54)$$

(b) Define $F_{H_{\mathcal{R}}(0)}(t|b) = \mathbb{P}_b(H_{\mathcal{R}}(0) \le t)$ and $F_{\mathcal{R}_t}(x|b) = \mathbb{P}_b(\mathcal{R}_t \le x)$, and let $f_{H_{\mathcal{R}}(0)}(t|b)$ and $f_{\mathcal{R}_t}(x|b)$ be the respective densities (if they exist). Then,

$$f_{H_{\mathcal{R}}(0)}(t|b) = \frac{b}{t} f_{\mathcal{R}_t}(-b|0). \qquad (3.55)$$

We shall show below that indeed $\mathcal{R}_t$ has a density for every $t > 0$, so that we are allowed to use Lemma 3.5. Lemma 3.5(a) yields

$$\mathbb{P}_\delta(H_{\mathcal{R}}(-\delta) = \infty) = \mathbb{P}_{2\delta}(H_{\mathcal{R}}(0) = \infty) = 1 - \mathrm{e}^{-2\delta\eta(0)}, \qquad (3.56)$$

for which it can be shown that $\eta(0) = 0$ if $\mathbb{E}(\mathcal{R}_1 - \mathcal{R}_0) \le 0$ and $\eta(0) \ge 0$ if $\mathbb{E}(\mathcal{R}_1 - \mathcal{R}_0) > 0$. From Lemma 3.5(b) we then conclude that

$$\begin{aligned}\mathbb{P}_\delta(H_{\mathcal{R}}(-\delta) > a) &= \int_a^\infty f_{H_{\mathcal{R}}(0)}(t|2\delta)dt + \mathbb{P}_\delta(H_{\mathcal{R}}(-\delta) = \infty) \\ &= \int_a^\infty \frac{2\delta}{t} f_{\mathcal{R}_t}(-2\delta|0)dt + 1 - \mathrm{e}^{-2\delta\eta(0)}. \qquad (3.57)\end{aligned}$$

Lemma 3.6 (Bound on the density of $\mathcal{R}_t$)**.** *The density of $\mathcal{R}_t$ can be bounded by*

$$\sup_x f_{\mathcal{R}_t}(x|b) \le a_2 t^{-1/(\tau-2)}. \qquad (3.58)$$

Proof. In order to bound $f_{\mathcal{R}_u}(x)$, we use the Fourier inversion theorem to obtain

$$f_{\mathcal{R}_t}(x|b) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\theta x} \hat{f}_{\mathcal{R}_t}(\theta|b) \frac{d\theta}{2\pi} \tag{3.59}$$

with the characteristic function $\hat{f}_{\mathcal{R}_t}(\theta|b) = \mathbb{E}_0(\mathrm{e}^{\mathrm{i}\theta\mathcal{R}_t}) = \mathrm{e}^{-t\Psi(\theta)}$ and $\Psi(\theta) = \psi(-\mathrm{i}\theta)$. The existence of the density $f_{\mathcal{R}_t}(x|b)$ follows from the fact that $|\hat{f}_{\mathcal{R}_t}(\theta|b)|$ is integrable, which we shall prove below. This also establishes the fact that $\mathcal{R}_t$ has a density, as stated below Lemma 3.5.

We obtain that, uniformly in $x \in \mathbb{R}$,

$$f_{\mathcal{R}_t}(x|b) \leq \int_{-\infty}^{\infty} |\hat{f}_{\mathcal{R}_t}(\theta|b)| \frac{d\theta}{2\pi} = \int_{-\infty}^{\infty} \mathrm{e}^{-t\mathrm{Re}(\Psi(\theta))} \frac{d\theta}{2\pi}, \tag{3.60}$$

where, for $z \in \mathbb{C}$, $\mathrm{Re}(z)$ denotes the real part of z.

From (2.17) we see that (with $\alpha = 1/(\tau - 1)$)

$$\mathrm{Re}(\Psi(\theta)) = \sum_{j=2}^{\infty} \frac{a}{j^{\alpha}} [1 - \cos(b\theta j^{-\alpha})]. \tag{3.61}$$

To prove a lower bound on $\mathrm{Re}(\Psi(\theta))$, we let

$$j_\theta = \min\{j \geq 2 : b\theta j^{-\alpha} \leq \pi/2\}, \tag{3.62}$$

so that

$$j_\theta = (2b\theta/\pi)^{1/\alpha} \wedge 2 = (2b\theta/\pi)^{\tau-1} \wedge 2. \tag{3.63}$$

Then we bound

$$\mathrm{Re}(\Psi(\theta)) \geq \sum_{j=j_\theta}^{\infty} \frac{a}{j^{\alpha}} [1 - \cos(b\theta j^{-\alpha})]. \tag{3.64}$$

Next, we use that

$$1 - \cos(x) \geq \frac{2}{\pi} x^2, \quad x \in [-\tfrac{1}{2}\pi, \tfrac{1}{2}\pi], \tag{3.65}$$

to arrive at

$$\mathrm{Re}(\Psi(\theta)) \geq c\theta^2 \sum_{j=j_\theta}^{\infty} j^{-3\alpha}, \tag{3.66}$$

where $c > 0$ denotes a positive constant that possibly changes from line to line. We arrive at the fact that

$$\mathrm{Re}(\Psi(\theta)) \geq c\theta^2 j_\theta^{1-3\alpha} \geq c\theta^2 \vee \theta^{\tau-2}. \tag{3.67}$$

In particular,

$$f_{\mathcal{R}_t}(x|b) \leq \int_{-\infty}^{\infty} \mathrm{e}^{-t\mathrm{Re}(\Psi(\theta))} \frac{d\theta}{2\pi} \leq \int_{-\infty}^{\infty} \mathrm{e}^{-ct\theta^2 \vee \theta^{\tau-2}} \frac{d\theta}{2\pi} \leq Ct^{1/(\tau-2)}, \tag{3.68}$$

as claimed in (3.58). This completes the proof of Lemma 3.6. □

Lemma 3.7 ($\mathcal{S}_t$ has a density)**.** *For all $t > 0$, $\mathcal{S}_t$ has a density. As a result, the distribution of $H_{\mathcal{S}}(0)$ is continuous.*

Proof. We use a similar method as in the proof of Lemma 3.6. We note that $\mathcal{S}_t$ has a density if and only if $\mathcal{S}_t'$ has, where

$$\mathcal{S}_t' = \sum_{j=2}^{\infty} j^{-1/(\tau-1)} \big[\mathcal{I}_j(t) - j^{-1/(\tau-1)}\big], \tag{3.69}$$

and $(\mathcal{I}_j(t))_{j\geq 2}$ are independent indicator processes with rate $j^{-1/(\tau-1)}$. This, in turn, follows when the characteristic function of $\mathcal{S}'_t$ is integrable. The characteristic function of $\mathcal{S}'_t$ is given by

$$\hat{f}_{\mathcal{S}'_t}(k) = \mathbb{E}[\mathrm{e}^{ik\mathcal{S}'_t}] = \prod_{j=2}^{\infty} \mathrm{e}^{-j^{-2/(\tau-1)}ik}\Big(1 + (\mathrm{e}^{-j^{-1/(\tau-1)}ik} - 1)\mathrm{e}^{-j^{-1/(\tau-1)}t}\Big). \tag{3.70}$$

Thus, for every $j_k \geq 2$,

$$|\hat{f}_{\mathcal{S}'_t}(k)| \leq \prod_{j\geq j_k}^{\infty} \big|1 + (\mathrm{e}^{-j^{-1/(\tau-1)}ik} - 1)\mathrm{e}^{-j^{-1/(\tau-1)}t}\big|. \tag{3.71}$$

Next, note that

$$\begin{aligned}
&\big|1 + (\mathrm{e}^{-j^{-1/(\tau-1)}ik} - 1)\mathrm{e}^{-j^{-1/(\tau-1)}t}\big|^2 \\
&\quad= \mathrm{e}^{-2j^{-1/(\tau-1)}t}\sin(j^{-1/(\tau-1)}k)^2 + \big(1 - \mathrm{e}^{-j^{-1/(\tau-1)}t} + \cos(j^{-1/(\tau-1)}k)\mathrm{e}^{-j^{-1/(\tau-1)}t}\big)^2 \\
&\quad= 1 - 2(1 - \mathrm{e}^{-j^{-1/(\tau-1)}t})\mathrm{e}^{-j^{-1/(\tau-1)}t}[1 - \cos(j^{-1/(\tau-1)}k)] \\
&\quad\leq \mathrm{e}^{-2(1-\mathrm{e}^{-j^{-1/(\tau-1)}t})\mathrm{e}^{-j^{-1/(\tau-1)}t}[1-\cos(j^{-1/(\tau-1)}k)]}.
\end{aligned} \tag{3.72}$$

From here on, we can follow the proof of the fact that the characteristic function of $\mathcal{R}_t$ is integrable. To prove that $H_{\mathcal{S}}(0)$ has a continuous distribution, note that when $\mathbb{P}(H_{\mathcal{S}}(0) = u) > 0$ for some $u \geq 0$, then, in particular, $\mathbb{P}(\mathcal{S}_u = 0) > 0$, which is in contradiction to the fact that $\mathcal{S}_u$ has a density. □

Now we are ready to complete the proof of Proposition 3.4. By (3.58) and (3.57), we arrive at the statement that, for all $\delta > 0$ and $\eta > 0$,

$$\mathbb{P}_\delta(H_{\mathcal{R}}(-\delta) > a) \leq \int_\eta^\infty \frac{2a_2\delta}{u^{(\tau-1)/(\tau-2)}}du + \mathbb{P}_\delta(H_{\mathcal{R}}(-\delta) = \infty) \leq c\delta\eta^{-1/(\tau-2)} + 1 - \mathrm{e}^{-2\delta\eta(0)}. \tag{3.73}$$

When first $\delta \downarrow 0$ followed by $\eta \downarrow 0$, this converges to 0, and we have proved (3.53). This completes the proof of Proposition 3.4. □

We proceed by showing that the scaling limits of the number of vertex checks of a cluster and the cluster size are identical. For this, we shall make use of the following lemma:

Lemma 3.8 (Number of multiple hits is small)**.** *As $n \to \infty$, for any $m \geq 1$,*

$$\mathbb{E}\Big[\sum_{j=2}^{m}[1 - J_j]\Big] \leq \frac{mw_1}{l_n} + \frac{m(m-1)\nu_n}{2l_n}. \tag{3.74}$$

Consequently, there exists $T_n \to \infty$, such that

$$n^{-(\tau-2)/(\tau-1)} \sum_{j=2}^{T_n n^{(\tau-2)/(\tau-1)}} [1 - J_j] \xrightarrow{\mathbb{P}} 0. \tag{3.75}$$

Proof. We note that $J_j = 0$ precisely when $M_l = 1$ or when there exists an $l < j$ such that $M_l = M_j$. Therefore,

$$\mathbb{E}[1 - J_j] \leq \frac{w_1}{l_n} + \sum_{l=2}^{j-1}\sum_{i=2}^{n} \frac{w_i^2}{l_n^2} \leq \frac{w_1}{l_n} + (j-1)\frac{\nu_n}{l_n}. \tag{3.76}$$

Summing the above inequality over $2 \leq j \leq m$ proves the claim in (3.74). For (3.75), we use the Markov inequality to bound

$$\mathbb{P}\Big(n^{-(\tau-2)/(\tau-1)} \sum_{j=2}^{T_n n^{(\tau-2)/(\tau-1)}} [1-J_j] \geq \varepsilon_n\Big) \leq \varepsilon_n^{-1} n^{-(\tau-2)/(\tau-1)} \mathbb{E}\Big[\sum_{j=2}^{T_n n^{(\tau-2)/(\tau-1)}} [1-J_j]\Big]$$

$$\leq \frac{T_n w_1}{\varepsilon_n l_n} + \frac{T_n^2 n^{(\tau-2)/(\tau-1)} \nu_n}{2 l_n \varepsilon_n} = o(1),$$

whenever $T_n^2 n^{-1/(\tau-1)}/\varepsilon_n = o(1)$. Choosing, for example, $T_n = \log n$ and $\varepsilon_n = 1/\log n$ does the trick. □

Now we are ready to complete the proof of Theorem 2.1:

Proof of Theorem 2.1. By Corollary 3.2 and Proposition 3.4, as well as (2.9), we obtain that

$$n^{-(\tau-2)/(\tau-1)} V(1) \xrightarrow{d} H_{\mathcal{S}}(0). \tag{3.77}$$

In particular, this implies that $|\mathcal{C}(1)| \leq n^{(\tau-2)/(\tau-1)} T_n$ for any $T_n \to \infty$. Therefore, by (2.11), and **whp**,

$$n^{-(\tau-2)/(\tau-1)} V(1) - n^{-(\tau-2)/(\tau-1)} \sum_{j=2}^{T_n n^{(\tau-2)/(\tau-1)}} [1-J_j] \leq n^{-(\tau-2)/(\tau-1)} |\mathcal{C}(1)| \leq n^{-(\tau-2)/(\tau-1)} V(1). \tag{3.78}$$

Now, by Lemma 3.8, the difference between the left-hand and right-hand side of (3.78) converges to zero in probability, so that also

$$n^{-(\tau-2)/(\tau-1)} |\mathcal{C}(1)| \xrightarrow{d} H_{\mathcal{S}}(0). \tag{3.79}$$

This completes the proof of Theorem 2.1, and identifies $H_1(0) = H_{\mathcal{S}}(0)$. □

In the next section, where we study the joint convergence of various clusters simultaneously, we shall also need the following joint convergence result:

Proposition 3.9 (Weak convergence of functionals)**.** *As $n \to \infty$,*

$$\Big(n^{-(\tau-2)/(\tau-1)} |\mathcal{C}(1)|, (\mathbb{1}_{\{q \in \mathcal{C}(1)\}})_{q \geq 1}\Big) \xrightarrow{d} \Big(H_1(0), (\mathcal{I}_q(H_1(0)))_{q \geq 1}\Big), \tag{3.80}$$

where $\mathcal{I}_q(H_1(0))$ denotes the indicator that $\mathcal{I}_q(t) = 1$ at the hitting time of 0 of $(\mathcal{S}_t)_{t \geq 0}$, in the product topology. Moreover, (i) *the random variable $H_1(0)$ is non-degenerate; and* (ii) *the indicators $(\mathcal{I}_q(H_1(0)))_{q \geq 2}$ are non-trivial in the sense that they take the values 0 and 1 each with positive probability.*

We note that, while the indicator processes $(\mathcal{I}_q(t))_{t \geq 0}$ are independent for different q, the random variables $(\mathbb{1}_{\{\mathcal{I}_q(H_1(0))\}})_{q \geq 1}$ are *not independent* as $H_1(0)$, the hitting time of 0 of the process $(\mathcal{S}_t)_{t \geq 0}$, depends sensitively on all of the indicator processes.

Proof. We shall use a randomization trick. Indeed, let $N_j^{(n)}(t)$ be a sequence of independent Poisson processes with rate w_j/l_n. Let

$$T_j = \min\{t : N(t) = j\}, \qquad \text{where} \qquad N(t) = \sum_{j=1}^{n} N_j^{(n)}(t). \tag{3.81}$$

Then, $t \mapsto N(t)$ is a rate 1 Poisson process, and we have that

$$S_l = S'_{T_l}, \tag{3.82}$$

where the continuous-time process $(S'_t)_{t \geq 0}$ is defined by

$$S'_t = w_1 - \frac{w_1^2 N(t)}{l_n} + \sum_{i=2}^{n} w_i \big[\mathbb{1}_{\{N_i^{(n)}(t) \geq 1\}} - \frac{w_i N(t)}{l_n}\big] + (\nu_n - 1) N(t). \tag{3.83}$$

By construction, the processes $(\mathbb{1}_{\{N_q^{(n)}(n^{(\tau-2)/(\tau-1)}t)\geq 1\}})_{t\geq 0}$ are *independent*, and are characterized by the birth times

$$E_q^{(n)} = \inf\{t : N_q^{(n)}(n^{(\tau-2)/(\tau-1)}t) \geq 1\}. \tag{3.84}$$

Again by construction, these birth times are independent for different $q \geq 2$, and

$$E_q^{(n)} = \mathrm{Exp}(n^{(\tau-2)/(\tau-1)}w_q/l_n). \tag{3.85}$$

The parameters of these exponential random variables converge to

$$n^{(\tau-2)/(\tau-1)}w_q/l_n \to aq^{-1/(\tau-1)}, \tag{3.86}$$

which are the parameters of the limiting exponential random variables $\mathcal{I}_q(t) = \mathbb{1}_{\{N_q(t)\geq 0\}}$ in (2.21). By the convergence of the parameters, we can *couple* $E_q^{(n)}$ with $E_q = \mathrm{Exp}(aq^{-1/(\tau-1)})$ in such a way that, for every $q \geq 2$ fixed,

$$\mathbb{P}\big(E_q^{(n)} \neq \mathrm{Exp}(aq^{-1/(\tau-1)})\big) = o(1). \tag{3.87}$$

Indeed, (3.87) follows by noting that, by (3.86), the density of $E_q^{(n)}$ converges pointwise to the one of E_q, which, by [25, (7.3)] implies that we can couple $(E_q^{(n)})_{n\geq 1}$ in such a way that (3.87) holds.

Equation (3.87) immediately implies that, for each $K \geq 1$,

$$\mathbb{P}(\mathbb{1}_{\{N_q^{(n)}(n^{(\tau-2)/(\tau-1)}t)\geq 1\}} = \mathcal{I}_q(t)\ \forall t \geq 0, q \in [K]) = 1 - o(1), \tag{3.88}$$

so that we have also, **whp**, perfectly coupled the entire processes $(\mathbb{1}_{\{N_q^{(n)}(n^{(\tau-2)/(\tau-1)}t)\geq 1\}})_{t\geq 0}$ and $(\mathcal{I}_q(t))_{t\geq 0}$. In particular, this implies that, for every $q \geq 2$,

$$\mathbb{P}\big(\mathbb{1}_{\{N_q^{(n)}(T_l)\geq 1\}} = \mathcal{I}_q(T_l)\ \forall l \geq 1, q \in [K]\big) = 1 - o(1), \tag{3.89}$$

and, by construction, $\mathbb{1}_{\{N_q^{(n)}(T_l)\geq 1\}} = \mathcal{I}_q^{(n)}(l)$. In particular, this applies for $l = V(1)$, for which $\mathbb{1}_{\{N_q^{(n)}(T_l)\geq 1\}} = \mathbb{1}_{\{q\in\mathcal{C}(1)\}}$. This provides a perfect coupling between $\mathbb{1}_{\{q\in\mathcal{C}(1)\}}$ and $\mathcal{I}_q(T_{V(1)})$. We then note that

$$n^{-(\tau-2)/(\tau-1)}|\mathcal{C}(1)| \xrightarrow{d} H_{\mathcal{S}}(0), \qquad \sup_{t\leq u} |n^{-(\tau-2)/(\tau-1)}T_{tn^{(\tau-2)/(\tau-1)}} - t| \xrightarrow{\mathbb{P}} 0. \tag{3.90}$$

Moreover, we next show that, for all $m \geq 1$,

$$\lim_{\eta\downarrow 0} \mathbb{P}\Big(\exists q \leq m : \mathcal{I}_q(H_{\mathcal{S}}(0) - \eta) \neq \mathcal{I}_q(H_{\mathcal{S}}(0) + \eta)\Big) = 0, \tag{3.91}$$

which completes the proof of convergence, for all $m \geq 1$, of

$$\Big(n^{-(\tau-2)/(\tau-1)}|\mathcal{C}(1)|, (\mathbb{1}_{\{q\in\mathcal{C}(1)\}})_{q\in[m]}\Big) \xrightarrow{d} \Big(H(0), (\mathcal{I}_q(H(0)))_{q\in[m]}\Big). \tag{3.92}$$

To prove (3.91), we make heavy use of the techniques in the proof of Proposition 3.4. Indeed, we shall prove that

$$\lim_{\eta\downarrow 0} \mathbb{P}\Big(\exists q \leq m : \mathcal{I}_q(H_{\mathcal{S}}(0) - \eta) \neq \mathcal{I}_q(H_{\mathcal{S}}(0))\Big) = 0. \tag{3.93}$$

A similar argument then shows

$$\lim_{\eta\downarrow 0} \mathbb{P}\Big(\exists q \leq m : \mathcal{I}_q(H_{\mathcal{S}}(0)) \neq \mathcal{I}_q(H_{\mathcal{S}}(0) + \eta)\Big) = 0, \tag{3.94}$$

which combines to (3.92). Now, we have that

$$\lim_{\delta\downarrow 0}\lim_{\eta\downarrow 0} \mathbb{P}(H_{\mathcal{S}}(0) - \eta \leq H_{\mathcal{S}}(\delta)) = 0, \tag{3.95}$$

so that (3.93) follows when

$$\lim_{\delta\downarrow 0}\mathbb{P}\Big(\exists q\leq m:\mathcal{I}_q(H_{\mathcal{S}}(\delta))\neq\mathcal{I}_q(H_{\mathcal{S}}(0))\Big)=0. \tag{3.96}$$

In a similar spirit,

$$\lim_{\varepsilon\downarrow 0}\lim_{\delta\downarrow 0}\mathbb{P}(H_{\mathcal{S}}(0)\geq H_{\mathcal{S}}(\delta)+\varepsilon)=0, \tag{3.97}$$

so that (3.94) follows when

$$\lim_{\varepsilon\downarrow 0}\lim_{\delta\downarrow 0}\mathbb{P}\Big(\exists q\leq m:\mathcal{I}_q(H_{\mathcal{S}}(\delta))\neq\mathcal{I}_q(H_{\mathcal{S}}(\delta)+\varepsilon)\Big)=0. \tag{3.98}$$

Now, by the strong Markov property for $t\mapsto(\mathcal{I}_q(t))_{q\geq 1}$ and the fact that $H_{\mathcal{S}}(\delta)$ is a stopping time w.r.t. the natural filtration $\mathcal{F}_t=\sigma\big(((\mathcal{I}_q(t))_{q\geq 1})_{s\in[0,t]}\big)$, we have

$$\mathbb{P}\Big(\exists q\leq m:\mathcal{I}_q(H_{\mathcal{S}}(\delta))\neq\mathcal{I}_q(H_{\mathcal{S}}(\delta)+\varepsilon)\mid\mathcal{F}_{H_{\mathcal{S}}(\delta)}\Big)\leq\mathbb{P}\Big(\exists q\leq m:\mathcal{I}_q(0)\neq\mathcal{I}_q(\varepsilon)\Big), \tag{3.99}$$

the inequality arising since, when we condition on $\mathcal{F}_{H_{\mathcal{S}}(\delta)}$, $\mathcal{I}_q(H_{\mathcal{S}}(\delta))=1$ for some $q\leq m$, and then, clearly, $\mathcal{I}_q(H_{\mathcal{S}}(\delta)+t)=\mathcal{I}_q(H_{\mathcal{S}}(\delta))=1$ for all $t\geq 0$.

For each $m\geq 1$ fixed, the right-hand side of (3.99) converges to 0 when $\varepsilon\downarrow 0$. This completes the proof of (3.93), as required.

Weak convergence of $(\mathbb{1}_{\{q\in\mathcal{C}(1)\}})_{q\geq 1}$ in the product topology is equivalent to the weak convergence of $(\mathbb{1}_{\{q\in\mathcal{C}(1)\}})_{q\in[m]}$ for any $m\geq 1$ (see [21, Theorem 4.29]), which implies the claimed weak convergence.

We continue to show the properties of the limiting variables. The random variable $H_1(0)$ is non-degenerate, since it has a continuous distribution. We shall next show that $\mathbb{1}_{\{q\in\mathcal{C}(1)\}}$ is non-trivial. We shall show this only for $q=2$. For this, we use the fact that $\mathbb{P}(H_1(0)>K)$ can be made arbitrarily small, so that

$$\mathbb{P}(2\notin\mathcal{C}(1))\leq\mathbb{P}(H_1(0)>K)+\mathbb{P}(\mathcal{I}_2(K)=0)=\mathbb{P}(H_1(0)>K)+\mathrm{e}^{-a2^{-1/(\tau-1)}K}<1, \tag{3.100}$$

when $K\geq 1$ is large enough. Now, further,

$$\mathbb{P}(2\notin\mathcal{C}(1))\geq\mathbb{P}(H_1(0)\geq\varepsilon,\mathcal{I}_2(\varepsilon)=1)=\mathbb{P}(\mathcal{I}_2(\varepsilon)=1)-\mathbb{P}(H_1(0)<\varepsilon,\mathcal{I}_2(\varepsilon)=1). \tag{3.101}$$

The first probability is of $\Theta(\varepsilon)$, whereas

$$\mathbb{P}(H_1(0)<\varepsilon,\mathcal{I}_2(\varepsilon)=1)\leq\mathbb{P}(H_1(0)<\varepsilon)\mathbb{P}(\mathcal{I}_2(\varepsilon)=1), \tag{3.102}$$

by the Fortuin-Kasteleyn-Ginibre-inequality (see [15, Thm. 2.4]) and the fact that both random variables are monotone in the independent exponential random variables that describe the first hit of q for all $q\geq 1$. Thus,

$$\mathbb{P}(2\notin\mathcal{C}(1))\geq\mathbb{P}(H_1(0)\geq\varepsilon)\mathbb{P}(\mathcal{I}_2(\varepsilon)=1)>0, \tag{3.103}$$

which proves the claim. □

4 Convergence of multiple clusters

In this section, we extend the analysis of one cluster in Section 2 to multiple clusters. The main result is as follows:

Theorem 4.1 (Weak convergence of the cluster of first vertices for $\tau\in(3,4)$). *Fix the Norros-Reittu random graph with weights $\boldsymbol{w}(\lambda)=\{(1+\lambda n^{-(\tau-3)/(\tau-1)})w_i\}_{i=1}^n$, where $(w_i)_{i\in[n]}$ are as in (1.3). Assume that $\nu=1$ and that (1.4) holds. Then, for all $\lambda\in\mathbb{R}$,*

$$\big(|\mathcal{C}_{\leq}(i)|n^{-(\tau-2)/(\tau-1)}\big)_{i\in[K]}\xrightarrow{d}\big(\kappa_i(\lambda)\big)_{i\in[K]}, \tag{4.1}$$

for some non-degenerate limit $\big(\kappa_i(\lambda)\big)_{i\geq 1}$.

We let $I_1^{(n)} = 1$, and

$$I_2^{(n)} = \min[n] \setminus \mathcal{C}(1) \tag{4.2}$$

be the minimal element that is not part of $\mathcal{C}(1)$, where, for a set of indices A, we let $\min A$ denote the minimal element of A. To extend the above definitions further, we define, recursively,

$$\mathcal{D}_i^{(n)} = \mathcal{C}_{\leq}(I_i^{(n)}), \tag{4.3}$$

and we let

$$\mathcal{D}_{\leq i}^{(n)} = \bigcup_{j \leq i} \mathcal{D}_j^{(n)}. \tag{4.4}$$

Then, we let $I_{i+1}^{(n)}$ be given by

$$I_{i+1}^{(n)} = \min[n] \setminus \mathcal{D}_{\leq i}^{(n)}, \tag{4.5}$$

which is the smallest index of which we have not yet explored the cluster.

Obviously, $|\mathcal{C}_{\leq}(i)| = 0$ unless $i = I_j^{(n)}$ for some j, so we proceed to investigate the weak convergence of $n^{-(\tau-2)/(\tau-1)}|\mathcal{D}_i^{(n)}|$. This will be done by induction on i. The induction hypothesis is that

$$\Big(n^{-(\tau-2)/(\tau-1)}|\mathcal{D}_j^{(n)}|, (\mathbb{1}_{\{q \in \mathcal{D}_{\leq j}^{(n)}\}})_{q \geq 1}\Big)_{1 \leq j \leq i} \xrightarrow{d} \Big(H_j(0), (\mathbb{1}_{\{q \in \mathcal{D}_{\leq j}\}})_{q \geq 1}\Big)_{1 \leq j \leq i}, \tag{4.6}$$

in the product topology, for some limiting random variables. Part of the induction hypothesis is that these limiting random variables satisfy the following facts: (1) the limiting random variables $(H_j(0))_{j \in [i]}$ are *non-degenerate*; and (2) the random indicators $(\mathbb{1}_{\{q \in \mathcal{D}_{\leq j}\}})_{j \in [i], q > i}$ are all *non-trivial*, in the sense that they take the values zero and one, each with positive probability. Note that, by construction, $\mathbb{1}_{\{q \in \mathcal{D}_{\leq j}\}} = 1$ for $j \leq i$, so condition (2) is the most we can hope for.

We shall start by initializing the induction hypothesis for $j = 1$, which relies on Proposition 3.9. Indeed, we have that $\mathcal{D}_1^{(n)} = \mathcal{D}_{\leq 1}^{(n)} = \mathcal{C}(1)$, so that (4.6) is identical to the statement in Proposition 3.9.

We next advance the induction hypothesis by verifying that (4.6) also holds for $j = i+1$. In order to simplify the notation, we let $H_j(0)$ be the weak limit of $n^{-(\tau-2)/(\tau-1)}|\mathcal{D}_j^{(n)}|$. We shall show that $H_j(0)$ is the hitting time of zero of an appropriate process alike $(\mathcal{S}_t)_{t \geq 0}$ in Section 2. We let $\mathcal{D}_j$ be the (random) set of indices for which

$$(\mathbb{1}_{\{q \in \mathcal{D}_{\leq j}^{(n)}\}})_{q \geq 1} \xrightarrow{d} (\mathbb{1}_{\{q \in \mathcal{D}_{\leq j}\}})_{q \geq 1}. \tag{4.7}$$

Then, we note that, by (4.6), we have that

$$I_{i+1}^{(n)} \xrightarrow{d} I_{i+1} \equiv \min\{q : \mathbb{1}_{\{q \in \mathcal{D}_{\leq i}\}} = 0\}, \tag{4.8}$$

and we see that $I_{i+1}^{(n)}$ and I_{i+1}, respectively, are *deterministic* functions of the sets $\mathcal{D}_{\leq i}^{(n)}$ and $\mathcal{D}_{\leq i}$, respectively. The latter random variable is *finite*, since, for $K, Q \geq 1$ large,

$$\mathbb{P}(I_{i+1}^{(n)} > K) \leq \mathbb{P}(|\mathcal{D}_{\leq j}^{(n)}| \geq Qn^{(\tau-2)/(\tau-1)}) + \mathbb{P}(I_{i+1}^{(n)} \geq K, |\mathcal{D}_{\leq j}^{(n)}| < Qn^{(\tau-2)/(\tau-1)}). \tag{4.9}$$

The first probability is small for $Q \geq 1$, while for the second and for $i \leq K/2$, we can bound

$$\begin{aligned} &\mathbb{P}(I_{i+1}^{(n)} > K, |\mathcal{D}_{\leq j}^{(n)}| < Qn^{(\tau-2)/(\tau-1)}) \\ &\quad \leq \mathbb{P}(\text{vertex } K \text{ drawn in } Qn^{(\tau-2)/(\tau-1)} \text{ vertex checks}). \end{aligned} \tag{4.10}$$

The latter probability is bounded above by

$$\mathbb{P}(\exists l \leq Qn^{(\tau-2)/(\tau-1)} : M_l = K) \leq \sum_{l=1}^{Qn^{(\tau-2)/(\tau-1)}} \frac{w_K}{l_n} \leq CQK^{-1/(\tau-1)}, \tag{4.11}$$

which converges to zero when $Q = K^\beta$ when $\beta < 1/(\tau-1)$.

We conclude that, from the induction hypothesis, we get the joint convergence

$$\Big(n^{-(\tau-2)/(\tau-1)}|\mathcal{D}_j^{(n)}|, I_{i+1}^{(n)}, (\mathbb{1}_{\{q\in\mathcal{D}_{\leq j}^{(n)}\}})_{q\geq 1}\Big)_{j\in[i]} \xrightarrow{d} \Big(H_j(0), I_{i+1}, (\mathbb{1}_{\{q\in\mathcal{D}_{\leq j}\}})_{q\geq 1}\Big)_{j\in[i]}. \tag{4.12}$$

We now start exploring the cluster of $I_{i+1}^{(n)}$, and we need to show that this cluster size, as well as the indices in it, converge in distribution. More precisely, to obtain the joint convergence in (4.6) for $i+1$ (and thus advance the induction hypothesis), when we prove that, conditionally on $\mathcal{D}_{\leq j}^{(n)}$,

$$\Big(n^{-(\tau-2)/(\tau-1)}|\mathcal{D}_{i+1}^{(n)}|, I_{i+1}^{(n)}, (\mathbb{1}_{\{q\in\mathcal{D}_{\leq i+1}^{(n)}\}})_{q\geq 1}\Big) \xrightarrow{d} \Big(H_{i+1}(0), I_{i+1}, (\mathbb{1}_{\{q\in\mathcal{D}_{\leq i+1}\}})_{q\geq 1}\Big). \tag{4.13}$$

An important step in the proof of (4.13) will consist of showing that the limiting distribution above

For this, we follow the approach in Section 2 as closely as possible. We note that after the exploration of $\mathcal{D}_{\leq i}^{(n)}$ and conditionally on it, the remaining graph is again a rank-1 inhomogeneous random graph, with (a) vertex set $[n]\setminus\mathcal{D}_{\leq i}^{(n)}$, and (b) edge probabilities, for $u,v\in[n]\setminus\mathcal{D}_{\leq i}^{(n)}$, given by

$$p_{uv} = 1-\mathrm{e}^{-w_u w_v/l_n}. \tag{4.14}$$

Thus, similarly to the setting in Section 2, we set $Z_0(i)=1$ and let $Z_1(i)$ denote the number of neighbors of vertex $I_{i+1}^{(n)}$, which is equal to

$$Z_1(i) = \sum_{j\notin\mathcal{D}_{\leq i}^{(n)}} \mathrm{Poi}(\tilde{w}_{I_{i+1}^{(n)}} w_j/l_n) = \mathrm{Poi}(\tilde{w}_{I_{i+1}^{(n)}} l_n(i)/l_n), \tag{4.15}$$

where

$$l_n(i) = \sum_{j\notin\mathcal{D}_{\leq i}^{(n)}} w_j. \tag{4.16}$$

We further note that, for $l\geq 2$, $Z_l(i)$ satisfies the recursion relation

$$Z_l(i) = Z_{l-1}(i) + X_l(i) - 1, \tag{4.17}$$

where $X_l(i)$ denotes the number of neighbors outside of $\mathcal{D}_{\leq i}^{(n)}$ of the l^{th} vertex which is explored. As explained in more detail in Section 2, the distribution of $X_l(i)$ (for $2\leq l\leq n$) is equal to $\mathrm{Poi}(w_{M_l(i)}l_n(i)/l_n)J_l(i)$, where the marks $\{M_l(i)\}_{l=1}^\infty$ are i.i.d. random variables with distribution $M(i)$ given by

$$\mathbb{P}(M(i)=m) = \frac{w_m}{l_n(i)}, \quad m\in[n]\setminus\mathcal{D}_{\leq i}^{(n)}, \tag{4.18}$$

and

$$J_l(i) = \mathbb{1}_{\{M_l(i)\notin\{I_{i+1}^{(n)}\}\cup\{M_1(i),\ldots,M_{l-1}(i)\}\}} \tag{4.19}$$

is the indicator that the mark $M_l(i)$ has not been found up to time l and is not equal to vertex $I_{i+1}^{(n)}$.

We conclude that we arrive at, for $l\geq 2$,

$$Z_l(i) = Z_{l-1}(i) + \mathrm{Poi}(\tilde{w}_{M_l(i)})J_l(i). \tag{4.20}$$

Then, the number of vertex checks $V(I_{i+1}^{(n)})$ in the exploration of $\mathcal{D}_{i+1}^{(n)} = \mathcal{C}_{\leq}(I_{i+1}^{(n)})$ equals

$$V(I_{i+1}^{(n)}) = \min\{l : Z_l(i)=0\}, \tag{4.21}$$

and

$$(\mathbb{1}_{\{a\in\mathcal{D}_{i+1}^{(n)}\}})_{a\neq I_{i+1}^{(n)}} = (\mathbb{1}_{\{\exists l\leq|\mathcal{D}_{i+1}^{(n)}|:M_l=a\}})_{a\neq I_{i+1}^{(n)}}, \tag{4.22}$$

while $\mathbb{1}_{\{I_{i+1}^{(n)}\in\mathcal{D}_{i+1}^{(n)}\}}=1$. We again note that

$$|\mathcal{D}_{i+1}^{(n)}|=|\mathcal{C}_{\leq}(I_{i+1}^{(n)})|\leq V(I_{i+1}^{(n)}), \tag{4.23}$$

while

$$n^{-(\tau-2)/(\tau-1)}\big[V(I_{i+1}^{(n)})-|\mathcal{D}_{i+1}^{(n)}|\big]\xrightarrow{\mathbb{P}}0, \tag{4.24}$$

as in the proof of Lemma 3.8. This gives us a convenient description of the random variables involved.

In order to prove the weak convergence of $V(I_{i+1}^{(n)})$, we again investigate the scaling limit of the process $(Z_l(i))_{l\geq 0}$. For this, we define $S_0(i)=1$, $S_1(i)=\tilde{w}_{I_{i+1}^{(n)}}l_n(i)/l_n$ and, for $l\geq 2$,

$$S_l(i)=S_{l-1}(i)+\tilde{w}_{M_l(i)}J_l(i)-1. \tag{4.25}$$

Then, as in Lemma 3.1, it is easy to show that, conditionally on $\mathcal{D}_{\leq i}^{(n)}$, the processes $(S_l(i))_{l\geq 0}$ and $(Z_l(i))_{l\geq 0}$ are uniformly close. Denote by $\mathcal{B}_i^{(n)}=\mathcal{D}_{\leq i}^{(n)}\cup\{I_{i+1}^{(n)}\}$ the union of all vertices explored in the first i clusters and the minimal element not in the first i clusters. Now, we rewrite

$$S_l(i)=\tilde{w}_{I_{i+1}^{(n)}}\frac{l_n(i)}{l_n}+\sum_{j=2}^{l}w_{M_j(i)}J_j(i)-(l-1)=\tilde{w}_{I_{i+1}^{(n)}}\frac{l_n(i)}{l_n}+\sum_{q\in[n]\setminus\mathcal{B}_i^{(n)}}\tilde{w}_q\mathcal{I}_q^{(n)}(l)-(l-1), \tag{4.26}$$

where

$$\mathcal{I}_q^{(n)}(l)=\mathbb{1}_{\{\exists j\leq l:M_j(i)=q\}}. \tag{4.27}$$

We further rewrite the above as

$$S_l(i)=\tilde{w}_{I_{i+1}^{(n)}}\frac{l_n(i)}{l_n}+\sum_{q\in[n]\setminus\mathcal{B}_i^{(n)}}\tilde{w}_q\big(\mathcal{I}_q^{(n)}(l)-\frac{w_ql}{l_n}\big)+l\Big(\sum_{q\in[n]\setminus\mathcal{B}_i^{(n)}}\frac{\tilde{w}_qw_q}{l_n}-1\Big)+1.$$

We note that we can rewrite the last sum, using (2.4), as

$$\begin{aligned}(1+\lambda n^{-(\tau-3)/(\tau-1)})\sum_{q\in[n]\setminus\mathcal{B}_i^{(n)}}\frac{w_q^2}{l_n}-1&=(\tilde{\nu}_n-1)-(1+\lambda n^{-(\tau-2)/(\tau-1)})\sum_{q\in\mathcal{B}_i^{(n)}}\frac{w_q^2}{l_n}\\&=\theta n^{-(\tau-3)/(\tau-1)}-\sum_{q\in\mathcal{B}_i^{(n)}}\frac{w_q^2}{l_n}+o_{\mathbb{P}}(n^{-(\tau-3)/(\tau-1)}),\end{aligned}$$

In turn, the sum can be approximated by

$$\sum_{q\in\mathcal{B}_i^{(n)}}\frac{w_q^2}{l_n}=dn^{-(\tau-3)/(\tau-1)}\sum_{q\in\mathcal{B}_i^{(n)}}q^{-2/(\tau-1)}(1+o_{\mathbb{P}}(1)), \tag{4.28}$$

where $d=c_F^{2/(\tau-1)}/\mu$. Denoting

$$D_i^{(n)}=d\sum_{q\in\mathcal{B}_i^{(n)}}q^{-2/(\tau-1)}, \tag{4.29}$$

we therefore have that

$$S_l(i)=w_{I_{i+1}^{(n)}}\frac{l_n(i)}{l_n}+\sum_{q\in[n]\setminus\mathcal{B}_i^{(n)}}w_q\big(\mathcal{I}_q^{(n)}(l)-\frac{w_ql}{l_n(i)}\big)+l(\theta-D_i^{(n)})n^{(\tau-3)/(\tau-1)}+o_{\mathbb{P}}(ln^{-(\tau-3)/(\tau-1)}). \tag{4.30}$$

We conclude that we arrive at a similar process as when exploring $\mathcal{C}(1)$, apart from the fact that: (i) fewer vertices are allowed to participate, and (ii) a negative drift $-D_i^{(n)}$ is introduced.

We proceed by investigating the convergence of $D_i^{(n)}$:

Lemma 4.2 (Weak convergence of random drift)**.** *As $n \to \infty$, and assuming (4.12),*

$$D_i^{(n)} \xrightarrow{d} D_i \equiv \sum_{q \in \mathcal{D}_{\leq i} \cup \{I_{i+1}\}} q^{-2/(\tau-1)}, \tag{4.31}$$

where $(\mathcal{D}_{\leq i}, I_{i+1})$ is the weak limit of $(\mathcal{D}_{\leq i}^{(n)}, I_{i+1}^{(n)})$.

Proof. We start by bounding $\mathbb{P}(q \in \mathcal{D}_{\leq i}^{(n)})$, for $q > 0$ large. We shall first prove that the probability that $|\mathcal{D}_{\leq i}^{(n)}| \leq n^{(\tau-2)/(\tau-1)} K$ is $1 - o(1)$ when $K > 0$ grows large. Indeed, by [17, Theorem 1.4], we have that, $|\mathcal{C}_{\max}| = \max_i |\mathcal{C}_{\leq i}| \leq \omega n^{(\tau-2)/(\tau-1)}$ with probability $1 - o(1)$, as $\omega \to \infty$. Thus, $|\mathcal{D}_{\leq i}^{(n)}| \leq n^{(\tau-2)/(\tau-1)}(i\omega) = n^{(\tau-2)/(\tau-1)} K$, with probability $1 - o(1)$ as $K \to \infty$, when we take $K = \omega i$. Thus, denoting

$$\mathcal{B}_{i,K}^{(n)} = \{|\mathcal{D}_{\leq i}^{(n)}| \leq n^{(\tau-2)/(\tau-1)} K\}, \tag{4.32}$$

we have that

$$\mathbb{P}(\{q \in \mathcal{D}_{\leq i}^{(n)} \setminus \{I_j^{(n)}\}_{j=1}^{i}\} \cap \mathcal{B}_{i,K}^{(n)}) \leq n^{(\tau-2)/(\tau-1)} K \frac{w_q}{\sum_{j > K n^{(\tau-2)/(\tau-1)}} w_j}, \tag{4.33}$$

since, independently of the choices before, the probability of drawing q is at most $w_q / \sum_{j > K n^{(\tau-2)/(\tau-1)}} w_j$. Now,

$$\sum_{j > K n^{(\tau-2)/(\tau-1)}} w_j = l_n(1 + o(1)) = \mu n(1 + o(1)). \tag{4.34}$$

Thus, for some $C > 0$,

$$\mathbb{P}(\{q \in \mathcal{D}_{\leq i}^{(n)} \setminus \{I_j^{(n)}\}_{j=1}^{i}\} \cap \mathcal{B}_{i,K}^{(n)}) \leq C K q^{-1/(\tau-1)}, \tag{4.35}$$

so that

$$\mathbb{E}\Big[\sum_{q \in \mathcal{B}_i^{(n)} : q > Q} q^{-2/(\tau-1)} \mathbb{1}_{\{\mathcal{B}_{i,K}^{(n)}\}}\Big] \leq i Q^{-2/(\tau-1)} + C K Q^{(\tau-4)/(\tau-1)}, \tag{4.36}$$

where the first contribution arises from the (at most i) values of $j = 1, \ldots, i+1$ for which $I_{i+1}^{(n)} > Q$.

Equation (4.36) implies that the weak convergence of $D_i^{(n)}$ follows from the weak convergence of

$$\sum_{q \in \mathcal{B}_i^{(n)} : q \leq Q} q^{-2/(\tau-1)}, \tag{4.37}$$

which, in turn, follows from (4.12) and the continuous mapping theorem. □

Define

$$\mathcal{Z}_t^{(n)}(i) = n^{-1/(\tau-1)} Z_{tn^{(\tau-2)/(\tau-1)}}(i), \qquad \mathcal{S}_t^{(n)}(i) = n^{-1/(\tau-1)} S_{tn^{(\tau-2)/(\tau-1)}}(i), \tag{4.38}$$

and

$$\mathcal{S}_t(i) = b I_{i+1}^{1/(\tau-1)} + \sum_{q \in \mathcal{D}_{\leq i} \cup \{I_{i+1}\}} a q^{-1/(\tau-1)} \big(\mathcal{I}_q(t) - b t q^{-1/(\tau-1)}\big) + t(c - D_i). \tag{4.39}$$

Then, using Lemma 4.2, the proof of Theorem 2.1 can easily be adapted to prove that $n^{-(\tau-2)/(\tau-1)} |\mathcal{D}_{i+1}^{(n)}| \xrightarrow{d} H_{i+1}(0)$, where $H_{i+1}(0)$ is the hitting time of 0 of $(\mathcal{S}_t(i))_{t \geq 0}$, and where a, b, c are given by $a = c_F^{1/(\tau-1)}/\mu$, $b = c_F^{1/(\tau-1)}$ and $c = \theta$.

Indeed, in more detail, we shall work *conditionally* on $(\mathbb{1}_{\{a \in \mathcal{D}_{\leq i+1}^{(n)}\}})_{a \in [n]}$. The proof of Theorem 2.1 reveals that the main contribution to $(\mathcal{S}_t(i))_{t \geq 0}$ and $(\mathcal{S}_t^{(n)}(i))_{t \geq 0}$ arises from the vertices $q \in [K]$. Now, since $(\mathbb{1}_{\{a \in \mathcal{D}_{\leq i+1}^{(n)}\}})_{a \in [K]}$ is a sequence of *discrete* random variables taking a finite number of outcomes, that converge in distribution, we have that its probability mass function converges pointwise. By [25], this implies that we can *couple* $(\mathbb{1}_{\{a\in\mathcal{D}^{(n)}_{\leq i+1}\}})_{a\in[K]}$ to $(\mathbb{1}_{\{a\in\mathcal{D}_{\leq i+1}\}})_{a\in[K]}$ in such a way that

$$\mathbb{P}\Big((\mathbb{1}_{\{a\in\mathcal{D}^{(n)}_{\leq i+1}\}})_{a\in[K]}\neq(\mathbb{1}_{\{a\in\mathcal{D}_{\leq i+1}\}})_{a\in[K]}\Big)=o(1). \tag{4.40}$$

Therefore, **whp**, there is a perfect coupling between the elements in $[K]$ of $\mathcal{D}^{(n)}_{\leq i+1}$ and $\mathcal{D}_{\leq i+1}$. When this is the case, we can basically think of the set of summands in (4.28) as being deterministic and follow the proof of Theorem 2.1 verbatim.

Further, the proof of Proposition 3.9 can be adapted to prove the joint convergence of

$$\Big(n^{-(\tau-2)/(\tau-1)}|\mathcal{D}^{(n)}_{i+1}|,(\mathbb{1}_{\{q\in\mathcal{D}^{(n)}_{i+1}\}})_{q\geq 1}\Big)\xrightarrow{d}\Big(H_{i+1}(0),(\mathcal{I}_q(H_{i+1}(0)))_{q\geq 1}\Big). \tag{4.41}$$

Together with the induction hypothesis, this proves that (4.6) also holds for all $j\leq i+1$, and, thus, we have advanced the induction hypothesis. This, in particular, proves Theorem 4.1. □

5 Proof of Theorems 1.1, 1.2 and 1.3

In this section, we prove Theorems 1.1 and 1.2 and 1.3, using the results in Theorems 2.1 and 4.1, as well as Proposition 3.9. We shall first start with a proof of Theorem 1.3. Note that, combining parts (a) and (b) in Theorem 1.3, we obtain that, with high probability as K becomes large, the largest m clusters are all among the first $(|\mathcal{C}_{\leq}(i)|)_{i\in[K]}$. This explains why we start the cluster exploration from the vertices with the highest weights.

Proof of Theorem 1.3. (a) For $\max_{i\geq K}|\mathcal{C}_{\leq}(i)|\geq\varepsilon n^{(\tau-2)/(\tau-1)}$ to occur, we must have that there exists a cluster using the vertices in $[n]\setminus[K]$ such that (1) $|\mathcal{C}_{\leq}(i)|\geq\varepsilon n^{(\tau-2)/(\tau-1)}$, and (2) the cluster is not connected to any of the vertices in $[K]$.
Now, by construction, when we restrict ourselves to the vertices in $[n]\setminus[K]$, we again have a Norros-Reittu model, with edge probabilities $p_{ij}=1-\mathrm{e}^{-w_iw_j/l_n}$, $i,j\in[n]\setminus[K]$. However, no vertex in $[n]\setminus[K]$ used in this cluster is allowed to have an edge to any of the vertices in $[K]$. Therefore, with

$$Z^{[K]}_{\geq k}=\sum_{v=1}^{n}\mathbb{1}_{\{|\mathcal{C}(v)|\geq k,\mathcal{C}(v)\cap[K]=\varnothing\}}, \tag{5.1}$$

we obtain that

$$\{\max_{i\geq K}|\mathcal{C}_{\leq}(i)|\geq k\}=\{Z^{[K]}_{\geq k}\geq k\}. \tag{5.2}$$

Therefore,

$$\mathbb{P}\big(\max_{i\geq K}|\mathcal{C}_{\leq}(i)|\geq k\big)=\mathbb{P}(Z^{[K]}_{\geq k}\geq k)\leq\frac{\mathbb{E}[Z^{[K]}_{\geq k}]}{k}=\frac{1}{k}\sum_{v=K+1}^{n}\mathbb{P}(|\mathcal{C}(v)|\geq k,\mathcal{C}(v)\cap[K]=\varnothing). \tag{5.3}$$

Now, denote by $\mathcal{C}^{[K]}(v)$ the cluster of v restricted to $[n]\setminus[K]$. Then, due to the independence of disjoint sets of edges, and the fact that $\mathcal{C}(v)\cap[K]=\varnothing$ only depends on edges between $[K]$ and $[n]\setminus[K]$, while $|\mathcal{C}^{[K]}(v)|\geq k$ only on edges between pairs of vertices in $[n]\setminus[K]$, we obtain

$$\mathbb{P}(|\mathcal{C}(v)|\geq k,\mathcal{C}(v)\cap[K]=\varnothing)=\mathbb{E}[\mathrm{e}^{-W_{[K]}W_{\mathcal{C}^{[K]}(v)}/l_n}\mathbb{1}_{\{|\mathcal{C}^{[K]}(v)|\geq k\}}], \tag{5.4}$$

where, for a set of vertices $A\subseteq[n]$, we define

$$W_A=\sum_{a\in A}w_a. \tag{5.5}$$

We split depending on whether $W_{\mathcal{C}^{[K]}(v)} \geq k/2$ or not, to obtain

$$\mathbb{P}\big(\max_{i\geq K} |\mathcal{C}_{\leq}(i)| \geq k\big) \leq \frac{1}{k} \sum_{v=K+1}^{n} \mathrm{e}^{-W_{[K]}k/(2l_n)} \mathbb{P}(|\mathcal{C}^{[K]}(v)| \geq k) \tag{5.6}$$

$$+ \frac{1}{k} \sum_{v=K+1}^{n} \mathbb{P}(|\mathcal{C}^{[K]}(v)| \geq k, W_{\mathcal{C}^{[K]}(v)} \leq k/2). \tag{5.7}$$

For the first term we compute that

$$W_{[K]} = \sum_{j=1}^{K} w_j \geq c \sum_{j=1}^{K} (n/j)^{1/(\tau-1)} \geq cn^{1/(\tau-1)} K^{(\tau-2)/(\tau-1)}. \tag{5.8}$$

Thus, when $k = k_n = \varepsilon n^{(\tau-2)/(\tau-1)}$, we obtain, for some $a > 0$,

$$\begin{aligned} \mathrm{e}^{-W_{[K]}k_n/(2l_n)} \sum_{v\in[n]} \mathbb{P}(|\mathcal{C}^{[K]}(v)| \geq k_n) &\leq \mathrm{e}^{-a\varepsilon K^{(\tau-2)/(\tau-1)}} \mathbb{P}(|\mathcal{C}^{[K]}(v)| \geq k_n) \\ &\leq \mathrm{e}^{-a\varepsilon K^{(\tau-2)/(\tau-1)}} \sum_{v\in[n]} \mathbb{P}(|\mathcal{C}(v)| \geq k_n) \\ &= \mathrm{e}^{-a\varepsilon K^{(\tau-2)/(\tau-1)}} n\mathbb{P}(|\mathcal{C}(V)| \geq k_n), \end{aligned} \tag{5.9}$$

where $V \in [n]$ is a uniform vertex. By [17, Proposition 5.1] and

$$\mathbb{P}(|\mathcal{C}(V)| \geq k_n) \leq a_1\big(k_n^{-1/(\tau-2)} + \big(\varepsilon_n \vee n^{-(\tau-3)/(\tau-1)}\big)^{1/(\tau-3)}\big) \leq a_1\big(k_n^{-1/(\tau-2)} + n^{-1/(\tau-1)}\big), \tag{5.10}$$

we get, for $k_n = \varepsilon n^{(\tau-2)/(\tau-1)}$,

$$n\mathbb{P}(|\mathcal{C}(V)| \geq k_n) \leq a_1' \varepsilon^{-1/(\tau-2)} n^{-(\tau-2)/(\tau-1)}. \tag{5.11}$$

Therefore, the term in (5.6) is bounded by

$$\mathrm{e}^{-a\varepsilon K^{(\tau-2)/(\tau-1)}} a_1' \varepsilon^{-(\tau-1)/(\tau-2)}. \tag{5.12}$$

When we pick $K = K_\varepsilon$ sufficiently large, we can make this as small as we wish.

We continue with the term in (5.7), for which we use a large deviation argument. We formulate this result in the following lemma:

Lemma 5.1 (Large deviations for cluster weight of large clusters)**.** *For every $k = o(n)$, there exists a $J > 0$ such that*

$$\mathbb{P}(\exists v : |\mathcal{C}(v)| \geq k, w_{\mathcal{C}(j)} \leq k/2) \leq n\mathrm{e}^{-Jk}. \tag{5.13}$$

Proof. Indeed, when $|\mathcal{C}(v)| \geq k$, then $w_{\mathcal{C}(v)}$ is stochastically bounded from below by the sum $\sum_{i=1}^{k} w_{v(i)}$, where $\{v(i)\}_{i=1}^{k}$ are the sized-biased ordering of $\{1,\ldots,n\}$, i.e., for every $j \notin \{v(s)\}_{s=1}^{i-1}$,

$$\mathbb{P}(v(i) = j \mid \{v(s)\}_{s=1}^{i-1}) = \frac{w_j}{\sum_{l\notin\{v(s)\}_{s=1}^{i-1}} w_s}. \tag{5.14}$$

In particular, for each i and conditionally on $\{v(j)\}_{j=1}^{i-1}$, $w_{v(i)}$ is stochastically bounded from above by W_i', which is equal to w_j for $j \in [n] \setminus [i-1]$ with probability

$$\frac{w_j}{l_n - \sum_{s=1}^{i-1+K} w_s}, \tag{5.15}$$

i.e., we have removed the vertices with the largest $i-1$ weights. Now take $\eta > 0$ very small, and note that, whenever $k \leq \eta n$, in turn, W_i' is stochastically bounded from above by an i.i.d. sequence of random variables $W_i^{(n)}(\eta)$ which equals to w_j for $j \in [n] \setminus [\eta n]$ with probability

$$\frac{w_j}{l_n - \sum_{s=1}^{\eta n} w_s}. \tag{5.16}$$

Now take $\eta > 0$ so small that

$$\mathbb{E}[W^{(n)}(\eta)] = \sum_{j=\eta n}^{n} \frac{w_j^2}{l_n - \sum_{s=1}^{\eta n} w_s} \geq 3/4. \tag{5.17}$$

Then, the term in (5.7) is bounded above by

$$\frac{1}{k} \sum_{v=K+1}^{n} \mathbb{P}(|\mathcal{C}^{[K]}(v)| \geq k, W_{\mathcal{C}^{[K]}(v)} \leq k/2) \leq \frac{n}{k} \mathbb{P}(\sum_{i=1}^{k} W_i^{(n)}(\eta) \leq k/2). \tag{5.18}$$

The Chernoff bound proves that $\mathbb{P}(\sum_{i=1}^{k} W_i^{(n)}(\eta) \leq k/2)$ is exponentially small in k, so that the term in (5.7) is exponentially small. We now make this intuition precise. By the Chernoff bound, for each $\theta \geq 0$, and by the fact that $(W_i^{(n)}(\eta))_{i \in [k]}$ are i.i.d. random variables, we have

$$\mathbb{P}(\sum_{i=1}^{k} W_i^{(n)}(\eta) \leq k/2) \leq \mathrm{e}^{\theta k/2} \mathbb{E}\big[\mathrm{e}^{-\theta \sum_{i=1}^{k} W_i^{(n)}(\eta)}\big] = \Big(\mathrm{e}^{\theta/2} \phi_{n,\eta}(\theta)\Big)^k, \tag{5.19}$$

where

$$\phi_{n,\eta}(\theta) = \mathbb{E}[\mathrm{e}^{-\theta W_1^{(n)}(\eta)}] \tag{5.20}$$

denotes the moment generating function of $W_1(\eta)$. By (5.19), it suffices to prove that there exists a $\theta > 0$ such that, for n sufficiently large,

$$\theta/2 + \log \phi_{n,\eta}(\theta) < 0.$$

This is what we shall show now. By dominated convergence, for each fixed $\theta > 0$,

$$\log \phi_{n,\eta}(\theta) \to \log \phi_{\eta}(\theta) = \log \mathbb{E}[\mathrm{e}^{-\theta W(\eta)}], \tag{5.21}$$

where

$$\mathbb{P}(W(\eta) \leq x) = \mathbb{E}[(1-F)^{-1}(U) \mid U \geq \eta], \tag{5.22}$$

and U is a uniform random variable. As a result, the distribution of U *conditionally on* $U \geq \eta$ is uniform on $[\eta, 1]$. Let U_η denote a uniform random variable on $[\eta, 1]$, so that $W(\eta) \stackrel{d}{=} (1-F)^{-1}(U_\eta)$. Then, $W(\eta)$ has mean $\mathbb{E}[W(\eta)] \geq 3/4$ and bounded variance σ_η^2 (since $W(\eta) \leq (1-F)^{-1}(\eta) < \infty$ a.s.). Therefore, a Taylor expansion yields that, for *fixed* $\eta > 0$,

$$\log \phi_\eta(\theta) \leq -3\theta/4 + \sigma_\eta^2 \theta^2 + o(\theta^2). \tag{5.23}$$

Now, fix a $\theta > 0$ so small that

$$\theta/2 - 3\theta/4 + \sigma_\eta^2 \theta^2 \leq -\theta/6, \tag{5.24}$$

and then N so large that, for all $n \geq N$,

$$\log \phi_{n,\eta}(\theta) \leq \log \phi_\eta(\theta) + \frac{\theta}{12}. \tag{5.25}$$

Then, indeed, for $n \geq N$, since $\theta > 0$,

$$\theta/2 + \log \phi_{n,\eta}(\theta) \leq -\frac{\theta}{6} + \frac{\theta}{12} = -\frac{\theta}{12} < 0, \tag{5.26}$$

so that

$$e^{\theta/2}\phi_{n,\eta}(\theta) \leq e^{-\frac{\theta}{12}}, \tag{5.27}$$

which, in turn, implies that

$$\sum_{v=1}^{n} \mathbb{P}(|\mathcal{C}(v)| \geq k, W_{\mathcal{C}(v)} \leq k/2) \leq n e^{-\frac{k\theta}{12}}. \tag{5.28}$$

When $n \to \infty$, this proves the claim for $J = \theta/12$. □

We apply Lemma 5.1 to the term in (5.7), which is then extremely small when we take $k = \varepsilon n^{(\tau-2)/(\tau-1)}$.
(b) We denote by

$$Z_{\geq k} = \sum_{v=1}^{n} \mathbb{1}_{\{|\mathcal{C}(v)|\geq k\}} \tag{5.29}$$

the number of vertices that are contained in connected components of size at least k. In [17], the random variable $Z_{\geq k}$ has been used in a crucial way to prove the asymptotics of $|\mathcal{C}_{\max}|$. We now slightly extend these results.

We shall prove that, for all $\varepsilon > 0$ sufficiently small, there exist constants a_2, C such that

$$\mathbb{P}\big(Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}} \leq n \frac{a_2}{\varepsilon^{1/(\tau-2)} n^{1/(\tau-1)}}\big) \leq C\varepsilon^{2/(\tau-2)}. \tag{5.30}$$

We first note that it suffices to prove (5.30) when $\nu_n \leq 1 - Kn^{(\tau-3)/(\tau-1)}$. Indeed, the random variable $Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}}$ is increasing in the edge occupation statuses, and, therefore, we may take $\lambda < 0$ so that $-\lambda > K$ to achieve the claim.

We shall use a second moment method. By [17, Proposition 6.1],

$$\mathbb{E}[Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}}] \geq n \frac{a_2}{\varepsilon^{1/(\tau-2)} n^{1/(\tau-1)}}, \tag{5.31}$$

so that

$$\mathbb{P}\big(Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}} \leq n \frac{a_2}{\varepsilon^{1/(\tau-2)} n^{1/(\tau-1)}}\big) \leq \mathbb{P}\big(Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}} \leq \mathbb{E}[Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}}]/2\big). \tag{5.32}$$

We take $\varepsilon > 0$ small, and bound, by the Chebychev inequality,

$$\mathbb{P}\big(Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}} \leq \mathbb{E}[Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}}]/2\big) \leq \frac{4\mathrm{Var}(Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}})}{\mathbb{E}[Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}}]^2}. \tag{5.33}$$

By [17, Proposition 3.2],

$$\mathrm{Var}(Z_{\geq k}) \leq n\mathbb{E}[|\mathcal{C}(V)|], \tag{5.34}$$

where $V \in [n]$ is uniform, and $\mathcal{C}(v)$ is the cluster of $v \in [n]$. By [17, Proposition 5.2], in turn,

$$\mathbb{E}[|\mathcal{C}(V)|] \leq Kn^{(\tau-3)/(\tau-1)}. \tag{5.35}$$

As a result, we obtain

$$\mathbb{P}\big(Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}} \leq \mathbb{E}[Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}}]/2\big) \leq \frac{4Kn^{2(\tau-2)/(\tau-1)}}{a_2^2 \varepsilon^{-2/(\tau-2)} n^{2(\tau-2)/(\tau-1)}} = C\varepsilon^{2/(\tau-2)}, \tag{5.36}$$

which is small when $\varepsilon > 0$ is small. We conclude that, **whp**,

$$Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}} \geq \mathbb{E}[Z_{\geq \varepsilon n^{(\tau-2)/(\tau-1)}}]/2 \geq \frac{a_2}{2\varepsilon^{1/(\tau-2)}} n^{(\tau-2)/(\tau-1)}. \tag{5.37}$$

Since, by [17, Theorem 1.2], $|\mathcal{C}_{\max}| \leq \varepsilon^{-1/2} n^{(\tau-2)/(\tau-1)}$ **whp**, there are, again **whp**, at least

$$\frac{a_2}{2\varepsilon^{1/(\tau-2)}} n^{(\tau-2)/(\tau-1)} / \varepsilon^{-1/2} n^{(\tau-2)/(\tau-1)} = C\varepsilon^{1/(\tau-2)-1/2} \tag{5.38}$$

clusters of size at least $\varepsilon n^{(\tau-2)/(\tau-1)}$. By part (a), with high probability for $K \geq 1$ large, these clusters will be part of $(|\mathcal{C}_{\leq}(i)|)_{i\in[K]}$ when $\varepsilon > 0$ is small, and $K \geq 1$ is large. □

We now complete the proof of Theorems 1.1 and 1.2:

Proof of Theorems 1.1 and 1.2. Weak convergence of $(|\mathcal{C}_{(i)}|n^{-(\tau-2)/(\tau-1)})_{i\geq 1}$ in the product topology is equivalent to the weak convergence of $(|\mathcal{C}_{(i)}|n^{-(\tau-2)/(\tau-1)})_{i\in[m]}$ for any $m\geq 1$ (see [21, Theorem 4.29]). In turn, by Proposition 1.3, this follows from the convergence in distribution of $(|\mathcal{C}_{\leq}(i)|n^{-(\tau-2)/(\tau-1)})_{i\in[m]}$ for all m. The latter follows from Theorem 4.1. Since, **whp**, again by Proposition 1.3, $(|\mathcal{C}_{(i)}|n^{-(\tau-2)/(\tau-1)})_{i\in[m]}$ is equal to the largest m components of $(|\mathcal{C}_{(i)}|n^{-(\tau-2)/(\tau-1)})_{i\in[m]}$, we have identified

$$(\gamma_i(\lambda))_{i\geq 1} \stackrel{d}{=} (H_{(i)}(0))_{i\geq 1}, \tag{5.39}$$

where $(H_{(i)}(0))_{i\geq 1}$ is $(H_i(0))_{i\geq 1}$ ordered in size. This completes the proof of Theorem 1.1, and identifies the limiting random variables. To prove Theorem 1.2, we use Proposition 3.9, and note that the limiting variables are all non-trivial (i.e., they are equal to 0 or 1 each with positive probability). This proves (1.12). The proof of (1.13) is similar, noting that the limit of $n^{-(\tau-2)/(\tau-1)}|\mathcal{C}_{\leq}(i)|$ is maximal with strictly positive probability. □

6 Proof of Theorem 1.5

In this section, we shall prove Theorem 1.5 on the largest subcritical clusters. We shall extend the result also to the ordered weights of subcritical clusters, which shall be a crucial ingredient in the proof of Theorem 1.4. For this, we let $W_{(j)}$ denote the j^{th} element of the ordered version of the vector

$$w_{\mathcal{C}_{\leq}(j)} = \sum_{i\in\mathcal{C}_{\leq}(j)} w_i. \tag{6.1}$$

Then, we shall prove that Theorem 1.5 holds both for $W_{(j)}$ as well as for $|\mathcal{C}_{(j)}|$. Indeed, it shall also follow from the result that **whp**, $W_{(j)} = w_{\mathcal{C}_{(j)}}$, i.e., the j^{th} largest cluster weight is the weight of the j^{th} largest cluster.

To prove this scaling, we shall prove that

$$|\lambda_n|n^{-(\tau-2)/(\tau-1)}|\mathcal{C}(j)| \xrightarrow{\mathbb{P}} c_j, \qquad |\lambda_n|n^{-(\tau-2)/(\tau-1)}W_{(j)} \xrightarrow{\mathbb{P}} c_j, \tag{6.2}$$

where we recall that $c_j = c_F^{1/(\tau-1)} j^{-1/(\tau-1)}$. Since $j\mapsto c_j$ is strictly decreasing, this means that, **whp**, $\mathcal{C}(j) = \mathcal{C}_{\leq}(j)$. Thus, this also implies that **whp**, for all $j\leq m$, $\mathcal{C}(j) = \mathcal{C}_{(j)}$. Then (6.2) proves the result for the ordered cluster weights.

Before starting with the proofs of (6.2), we introduce some branching process notation. We write T for the total progeny of a branching process with mixed Poisson offspring distribution $\text{Poi}(M)$, where the distribution of M is given in (2.7). Below, we shall often let $(T_j)_{j\geq 1}$ be an i.i.d. sequence of such total progenies. Further, we let $T(j)$ be the total progeny of a two-stage branching process, given by

$$T(j) = 1 + \sum_{l=1}^{\text{Poi}(w_j)} T_l, \tag{6.3}$$

i.e., the first generation has distribution $\text{Poi}(w_j)$ and later generations have offspring distribution $\text{Poi}(M)$. Then, by the results in [23, Proposition 3.1] (see also [17, Section 4.2]), we have that we can couple $|\mathcal{C}(j)|$ and $T(j)$ such that $|\mathcal{C}(j)|\leq T(j)$ a.s. We also define

$$w_T = \sum_{l=1}^{T} w_{M_l}, \tag{6.4}$$

be the total weight of the branching process total progeny, where $(M_j)_{j\geq 1}$ are the i.i.d. marks used in the mixed-Poisson branching process. Similarly, we let

$$w_{T(j)} = \sum_{l=1}^{T(j)} w_{M_l} \tag{6.5}$$

be the total weight of the total progeny $T(i)$. Moments of $T, w_T, T(j)$ and $w_{T(j)}$ are proved in Lemma A.2 in the appendix. We shall frequently make use of these computations. The proof of Theorem 1.5 now consists of three key steps, which we shall prove one by one.

Asymptotics of mean cluster size and weight of high-weight vertices. In the following lemma we investigate the means of $|\mathcal{C}(j)|$ and $w_{\mathcal{C}(j)}$:

Lemma 6.1 (Mean cluster size and weights). *As $n \to \infty$, for every $j \in \mathbb{N}$ fixed, and when $\lambda_n \to -\infty$,*

$$\mathbb{E}[|\mathcal{C}(j)|] = \frac{w_j}{1-\nu_n}(1+o(1)), \qquad \mathbb{E}[w_{\mathcal{C}(j)}] = \frac{w_j \nu_n}{1-\nu_n}(1+o(1)). \tag{6.6}$$

Proof. By the fact that $|\mathcal{C}(j)|$ and $T(j)$ can be coupled so that $|\mathcal{C}(j)| \leq T(j)$ a.s., we obtain that

$$\mathbb{E}[|\mathcal{C}(j)|] \leq \mathbb{E}[T(j)] = \frac{w_j}{1-\nu_n}, \tag{6.7}$$

the latter equality following from Lemma A.2(a). A similar upper bound follows for $\mathbb{E}[w_{\mathcal{C}(j)}]$ again using Lemma A.2(a).

For the lower bound, we note that

$$\mathbb{E}[|\mathcal{C}(j)|] = \mathbb{E}[T(j)] - \mathbb{E}[T(j) - |\mathcal{C}(j)|]. \tag{6.8}$$

Now, for $a_n = n^{(\tau-2)/(\tau-1)} \gg \mathbb{E}[T(j)]$, we bound

$$\mathbb{E}[T(j) - |\mathcal{C}(j)|] \leq \mathbb{E}[T(j)\mathbb{1}_{\{T(j)>a_n\}}] + \mathbb{E}[[T(j) - |\mathcal{C}(j)|]\mathbb{1}_{\{T(j)\leq a_n\}}]. \tag{6.9}$$

The first term in (6.9) is bounded by

$$\mathbb{E}[T(j)\mathbb{1}_{\{T(j)>a_n\}}] \leq \frac{1}{a_n}\mathbb{E}[T(j)^2] = \frac{1}{a_n}\Big((1+\frac{w_j}{1-\nu_n})^2 + \frac{w_j}{(1-\nu_n)^3}\Big) = o(\frac{w_j}{1-\nu_n}), \tag{6.10}$$

since $a_n \gg \mathbb{E}[T(j)] = w_j/(1-\nu_n)$. Since $a_n = n^{(\tau-1)/(\tau-1)}$, we have (and have used) that

$$(1-\nu_n)^2 a_n = \lambda_n^2 n^{-2(\tau-3)/(\tau-1)+(\tau-2)/(\tau-1)} = \lambda_n^2 n^{-(4-\tau)/(\tau-1)} = o(1). \tag{6.11}$$

For the second term in (6.9), we note that differences between $T(j)$ and $|\mathcal{C}(j)|$ arise due to vertices which have been used *at least twice* in $T(j)$. In fact, the law of $|\mathcal{C}(j)|$ can be obtained from the branching process by removing vertices (and their complete offspring) of which the mark has already been used. Thus,

$$\begin{aligned}\mathbb{E}[[T(j) - |\mathcal{C}(j)|]\mathbb{1}_{\{T(j)\leq a_n\}}] &\leq \sum_{i\in[n]} \mathbb{E}\Big[[T(j) - |\mathcal{C}(j)|]\mathbb{1}_{\{T(j)\leq a_n\}}\mathbb{1}_{\{i \text{ used twice}\}}\Big] \\ &\leq \sum_{i\in[n]} \mathbb{E}[T(i)] \sum_{s_1,s_2=1}^{a_n} \mathbb{P}\big(i \text{ chosen at times } s_1, s_2\big),\end{aligned} \tag{6.12}$$

where $\mathbb{E}[T(i)]$ arises from removing the tree rooted at the vertex at time s_2 of which the root has mark i. Now, i can only be chosen at time s_1 when $T(j) \geq s_1 - 1$, which, as an event, is independent from the event that the mark i is chosen at times s_1, s_2. Therefore,

$$\begin{aligned}\mathbb{E}[[T(j) - |\mathcal{C}(j)|]\mathbb{1}_{\{T(j)\leq a_n\}}] &\leq \sum_{i\in[n]} \mathbb{E}[T(i)] \sum_{s_1,s_2=1}^{a_n} \mathbb{P}(T(j) \geq s_1 - 1)\frac{w_i^2}{l_n^2} \\ &\leq a_n \sum_{s_1=1}^{a_n} \mathbb{P}(T(j) \geq s_1 - 1) \sum_{i\in[n]} \mathbb{E}[T(i)]\frac{w_i^2}{l_n^2} \\ &\leq a_n \mathbb{E}[T(j)] \sum_{i\in[n]} \mathbb{E}[T(i)]\frac{w_i^2}{l_n^2}.\end{aligned} \tag{6.13}$$

This is $o(\mathbb{E}[T(j)])$ when $\lambda_n \to -\infty$ since

$$a_n \sum_{i\in[n]} \mathbb{E}[T(i)]\frac{w_i^2}{l_n^2} = a_n \sum_{i\in[n]} \frac{w_i^3}{l_n^2(1-\nu_n)} \leq \frac{C}{|\lambda_n|} n^{(\tau-2)/(\tau-1)-2} n^{3/(\tau-1)} n^{(\tau-3)/(\tau-1)} = \frac{C}{|\lambda_n|} = o(1). \tag{6.14}$$

This completes the proof for $\mathbb{E}[|\mathcal{C}(j)|]$. The proof for $w_{T(j)}$ is similar. Indeed, we split

$$\mathbb{E}[w_{T(j)} - w_{\mathcal{C}(j)}] \leq \mathbb{E}[w_{T(j)}\mathbb{1}_{\{T(j)>a_n\}}] + \mathbb{E}[[w_{T(j)} - w_{\mathcal{C}(j)}]\mathbb{1}_{\{T(j)\leq a_n\}}]. \tag{6.15}$$

The first term is now bounded by

$$\mathbb{E}[w_{T(j)}\mathbb{1}_{\{T(j)>a_n\}}] \leq \frac{1}{a_n}\mathbb{E}[w_{T(j)}T(j)], \tag{6.16}$$

which we can again bound using Lemma A.2(f). Further, in (6.13)–(6.14), $\mathbb{E}[T(i)]$ needs to be replaced with $\mathbb{E}[w_{T(i)}]$. □

Cluster size and weight of high weight vertices are concentrated. We note that, by the stochastic domination and the fact that $\mathbb{E}[|\mathcal{C}(j)|] = \frac{w_j}{1-\nu_n}(1+o(1))$, we have

$$\mathrm{Var}(|\mathcal{C}(j)|) \leq \mathrm{Var}(T(j)) + o(\mathbb{E}[T(j)]^2). \tag{6.17}$$

Now, by Lemma A.2(b),

$$\mathrm{Var}(T(j)) = 1 + \frac{2w_j}{1-\nu_n} + \frac{w_j}{(1-\nu_n)^3} = o(\frac{w_j^2}{(1-\nu_n)^2}) \tag{6.18}$$

precisely when $1/(\nu_n - 1) = o(w_j)$, which is the case since

$$\frac{1}{\nu_n - 1} = \frac{1}{|\lambda_n|} n^{(\tau-3)/(\tau-1)} = o(n^{1/(\tau-1)}), \tag{6.19}$$

since $|\lambda_n| \to \infty$ and $\tau < 4$. For $w_{T(j)}$ the argument is similar, apart from the fact that

$$\mathrm{Var}(T(j)) = \frac{w_j}{(1-\nu_n)^3}\Big(\frac{1}{l_n}\sum_{i=1}^n w_i^3\Big) = o(\frac{w_j^2}{(1-\nu_n)^2}), \tag{6.20}$$

since, for j fixed,

$$\frac{1}{1-\nu_n}\Big(\frac{1}{l_n}\sum_{i=1}^n w_i^3\Big) \leq \frac{C}{|\lambda_n|} n^{(\tau-3)/(\tau-1)+3/(\tau-1)-1} = \frac{C}{|\lambda_n|} n^{1/(\tau-1)} = o(w_j). \tag{6.21}$$

We conclude that, for j fixed, $\mathrm{Var}(|\mathcal{C}(j)|) = o(\mathbb{E}[|\mathcal{C}(j)|]^2)$, so that

$$\frac{|\mathcal{C}(j)|}{\mathbb{E}[|\mathcal{C}(j)|]} \xrightarrow{\mathbb{P}} 1, \qquad \frac{w_{\mathcal{C}(j)}}{\mathbb{E}[w_{\mathcal{C}(j)}]} \xrightarrow{\mathbb{P}} 1 \tag{6.22}$$

and then Lemma 6.1 completes the proof of (6.2).

Only high-weight vertices matter. We start by proving that the probability that, for $K \geq 1$, there exists a $j > K$ such that $w_{\mathcal{C}_{\leq}(j)} \geq \varepsilon n^{(\tau-2)/(\tau-1)}/|\lambda_n|$ is small. Since, for all $j \leq K$, we have that $|\lambda_n| n^{-(\tau-2)/(\tau-1)} w_{\mathcal{C}(j)} \xrightarrow{\mathbb{P}} c_j$, we have that, for all $i \leq m$ and m such that $c_m > \varepsilon$, $w_{\mathcal{C}(j)} = W_{(j)}$. Note that, if there exists a $j > K$ such that $w_{\mathcal{C}_{\leq}(j)} \geq \varepsilon n^{(\tau-2)/(\tau-1)}/|\lambda_n|$, then

$$\sum_{j>K} w_j w^2_{\mathcal{C}^{[K]}(j)} \geq \frac{\varepsilon^3}{|\lambda_n|^3} n^{3(\tau-2)/(\tau-1)}, \tag{6.23}$$

where we recall that $\mathcal{C}^{[K]}(j)$ is the cluster of j in the random graph only making use of the vertices in $[n] \setminus [K]$. Since

$$l_n \geq \sum_{j>K} w_j, \tag{6.24}$$

we see that this random graph is stochastically bounded by the random graph having weights $\boldsymbol{w}^{[K]}$, where $w_j^{[K]} = 0$ when $j \leq K$ and $w_j^{[K]} = w_j$ otherwise. By the Markov inequality, the probability that (6.23) happens is bounded above by

$$\frac{|\lambda_n|^3}{\varepsilon^3} n^{-3(\tau-2)/(\tau-1)} \sum_{j>K} w_j \mathbb{E}[w^2_{\mathcal{C}^{[K]}(j)}]. \tag{6.25}$$

By Lemma A.2(f), we obtain that

$$\mathbb{E}[w^2_{\mathcal{C}^{[K]}(j)}] = \Big(\frac{w_j^{[K]}}{1-\nu_n^{[K]}}\Big)^2 + \frac{w_j^{[K]}}{(1-\nu_n^{[K]})^3}\Big(\frac{1}{l_n^{[K]}} \sum_{i=1}^n (w_i^{[K]})^3\Big), \tag{6.26}$$

where all the superscripts $[K]$ refer to the fact that we should compute quantities for the weights $\boldsymbol{w}^{[K]}$. It is not hard to see that the r.h.s. of (6.26) is equal to

$$\Big(\frac{w_j}{1-\nu_n}\Big)^2 + (1+o(1))\frac{w_j}{(1-\nu_n)^3}\Big(\frac{1}{l_n} \sum_{i>K}^n w_i^3\Big), \tag{6.27}$$

which gives that

$$\sum_{j>K} w_j \mathbb{E}[w^2_{\mathcal{C}^{[K]}(j)}] \leq \frac{1}{(1-\nu_n)^2}\Big(1+\frac{1}{1-\nu_n}\Big) \sum_{j>K} w_j^3 \leq C K^{-(\tau-4)/(\tau-1)} \Big(n^{1/(\tau-1)}/|\lambda_n|\Big)^3, \tag{6.28}$$

which can be made arbitrarily small by taking K large.

We complete this section by proving that the probability that there exists a $j > K$ such that $|\mathcal{C}_{\leq}(j)| \geq \varepsilon n^{(\tau-2)/(\tau-1)}/|\lambda_n|$ is small. For this, we use Lemma 5.1, which proves that, with very high probability, if $|\mathcal{C}_{\leq}(j)| \geq \varepsilon n^{(\tau-2)/(\tau-1)}/|\lambda_n|$, then also $w_{\mathcal{C}_{\leq}(j)} \geq \varepsilon n^{(\tau-2)/(\tau-1)}/(2|\lambda_n|)$. Thus, the result for cluster sizes follows from the proof for cluster weights. This completes the proof of Theorem 1.5. □

7 Proof of Theorem 1.4

We break up the proof into three parts. In the first part, we identify the ordered sets of weights of clusters to be a multiplicative coalescent as we vary the parameter λ indicating the location inside the critical window. In the second part, we prove the convergence of this multiplicative coalescent, for which we identify the parameters of the limiting coalescent. Finally, we show that the cluster sizes are close to the cluster weights, which completes the proof of Theorem 1.4.

The random graph multiplicative coalescent. We make crucial use of [3, Proposition 7], whose application we now explain. Fix a sequence $\lambda_n \to -\infty$. For each fixed t, consider the construction of the inhomogeneous random graph as in (1.1) but with the weight sequence $\boldsymbol{w}(t) = (\bar{w}_j(t))_{1\leq j\leq n}$ given by

$$\bar{w}_j(t) = w_j(1 + (t+\lambda_n)l_n n^{-2(\tau-2)/(\tau-1)}). \tag{7.1}$$

Let $X^{(n)}(t)$ denote the ordered version of weighted component sizes, i.e., the ordered version of the vector

$$n^{-(\tau-2)/(\tau-1)} w_{\mathcal{C}_{\leq}(j)}(t), \tag{7.2}$$

where $w_{\mathcal{C}} = \sum_{v\in\mathcal{C}} w_v$ is the weight of a cluster $\mathcal{C}$.

Note that the above process, when taking $t = -\lambda_n + \lambda/\mu$, is closely related to the ordered clusters of our random graph with weights $\tilde{w}_j = w_j(1 + \lambda n^{-(\tau-3)/(\tau-1)})$, since $l_n = \mu n(1+o(1))$. We start by proving that $X^{(n)}$ can be constructed so that viewed as a function in t it is a multiplicative coalescent:

Lemma 7.1 (Discrete multiplicative coalescent)**.** *We can construct the process $X^{(n)} = (X^{(n)}(t))_{t\geq 0}$ such that for each fixed t, $X^{(n)}(t)$ has the distribution of the ordered rescaled weighted component sizes of the random graph with weight sequence given by (7.1) defined above and such that, for each fixed n, the process viewed as a process in t is a multiplicative coalescent. The initial state denoted by $\mathbf{x}^{(n)}(0)$ has the same distribution as the ordered weighted component sizes of a random graph with edge probabilities as in (1.1) and weight sequence*

$$\bar{w}_j(0) = w_j(1 + \lambda_n l_n n^{-2(\tau-2)/(\tau-1)}). \tag{7.3}$$

Proof. This is quite easy. For each unordered pair (i,j), let ξ_{ij} be exponential random variables with rate $w_i w_j / l_n$. For fixed t define the graph $\mathcal{G}_n^t$ to consist of all those edges (i,j) for which

$$\xi_{ij} \leq \left(1 + \frac{(\lambda_n + t)l_n}{n^{2(\tau-2)/(\tau-1)}}\right)$$

Then by construction, for all $t \geq 0$, the rescaled weighted component sizes of $\mathcal{G}_n^t$ have the same distribution as $X^{(n)}(t)$. Further for any time t we note that two *distinct* clusters $\mathcal{C}_{\leq}(i)(t)$ and $\mathcal{C}_{\leq}(j)(t)$ coalesce at rate

$$l_n n^{-2(\tau-2)/(\tau-1)} \sum_{s_1\in\mathcal{C}_{\leq(i)}(t), s_2\in\mathcal{C}_{\leq(j)}(t)} \frac{w_{s_1} w_{s_2}}{l_n} = \left(n^{-(\tau-2)/(\tau-1)} w_{\mathcal{C}_{\leq(i)}}(t)\right)\left(n^{-(\tau-2)/(\tau-1)} w_{\mathcal{C}_{\leq(j)}}(t)\right), \tag{7.4}$$

as required. □

Convergence of the random graph multiplicative coalescent. We now apply [3, Proposition 7], which shows that $\left(X^{(n)}(\lambda_n + \lambda)\right)_\lambda$ converges to a $(0, \tau, \boldsymbol{c})$-multiplicative coalescent when three conditions are satisfied about the initial state $\mathbf{x}^{(n)}(0)$. To state these conditions, we define, for $r = 2, 3$,

$$\sigma_r(\mathbf{x}^{(n)}) = \sum_j (x_j^{(n)})^r. \tag{7.5}$$

Then, the conditions in [3, Proposition 7] are that:
(a)

$$|\lambda_n|(|\lambda_n|\sigma_2(\mathbf{x}^{(n)}) - 1) \xrightarrow{\mathbb{P}} -\frac{\zeta}{\mu}; \tag{7.6}$$

(b)

$$\frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \xrightarrow{\mathbb{P}} c_j; \tag{7.7}$$

(c)

$$|\lambda_n|^3\sigma_3(\mathbf{x}^{(n)}) \xrightarrow{\mathbb{P}} \sum_{j=1}^{\infty} c_j^3. \tag{7.8}$$

The conditions (a)-(c) above are not precisely what is in [3, Proposition 7], and we start by explaining how (a)-(c) imply our result. Indeed, in [3, Proposition 7], the condition in (a) is replaced by $\sigma_2(\mathbf{x}^{(n)}) \to 0$, and the process

$$\Big(X^{(n)}\big(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+\lambda\big)\Big)_\lambda \tag{7.9}$$

is proved to converge to the multiplicative coalescent. Under condition (a) (and the fact that $\lambda_n \to -\infty$), (a) implies that $1/\sigma_2(\mathbf{x}^{(n)}) = |\lambda_n| - \zeta/\mu + o(1)$. Also, condition (c) is replaced by the condition that

$$\frac{\sigma_3(\mathbf{x}^{(n)})}{\sigma_2(\mathbf{x}^{(n)})^3} \xrightarrow{\mathbb{P}} \sum_{j=1}^{\infty} c_j^3, \tag{7.10}$$

which, combining (a) and (c), is equivalent. Further, in (a)-(c), we work with convergence in probability (as the initial state is a random variable), while in [3, Proposition 7], the initial state is considered to be deterministic. This is a minor change. We finally explain how the above conditions complete the proof of Theorem 1.4.

Cluster weights versus cluster size. We note that $X^{(n)}\big(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+\lambda\big)$ are the ordered cluster weights

$$n^{-(\tau-2)/(\tau-1)} w_{\mathcal{C}_{\leq}(j)}, \tag{7.11}$$

where the weights are

$$\bar{w}_j(|\lambda_n| - \zeta + \lambda + o(1)) = w_j(1+\lambda l_n n^{-2(\tau-2)/(\tau-1)}) = w_j(1+\lambda \mathbb{E}[W] n^{-(\tau-3)/(\tau-1)}(1+o(1))). \tag{7.12}$$

Thus, apart from a multiplication of λ by $\mathbb{E}[W]$, and the fact that we deal with *cluster weights* rather than with *cluster sizes*, this proves the claim. Now, we claim that $w_{\mathcal{C}_{\leq}(j)} = |\mathcal{C}_{\leq}(j)|(1+o_{\mathbb{P}}(1))$. Indeed, note that, by the analysis in Section 3 (see in particular (3.1)), we have that

$$S_{V(1)} = w_{\mathcal{C}(1)} - V(1), \tag{7.13}$$

where we recall that $V(1)$ is the number of vertex checks of the cluster of vertex 1. Further, by the analysis performed there, the random variable $n^{-(\tau-2)/(\tau-1)}V(1) = H_1(0)$ has the same limit in distribution as the first hitting time of $(-\infty, 0]$ of $(S_l)_{l\geq 0}$. Thus,

$$n^{-1/(\tau-1)} S_{V(1)} \xrightarrow{\mathbb{P}} 0. \tag{7.14}$$

Since $(\tau-2)/(\tau-1) > 1/(\tau-1)$, this immediately implies that $n^{-(\tau-2)/(\tau-1)} w_{\mathcal{C}(1)} \xrightarrow{d} H_1(0)$. The same applies to all other rescaled clusters $n^{-(\tau-2)/(\tau-1)} w_{\mathcal{C}_{\leq}(j)}$, so that the limit of $\Big(X^{(n)}\big(\frac{1}{\sigma_2(\mathbf{x}^{(n)})}+\lambda\big)\Big)_\lambda$ equals $\big((\gamma_j(\mathbb{E}[W]\lambda - \zeta))_{j\geq 1}\big)_\lambda$, as required.

Check of convergence conditions. We conclude that we are left to prove that conditions (a), (b) and (c) hold. We shall prove these conditions in the order (b), (c) and (a), condition (a) being the most difficult one.

Condition (b) follows immediately from (6.2). For condition (c), we apply similar ideas, and note that

$$|\lambda_n|^3\sigma_3(\mathbf{x}^{(n)}) = \sum_{j=1}^{n}\Big(|\lambda_n| n^{-(\tau-2)/(\tau-1)} w_{\mathcal{C}_{\leq}(j)}\Big)^3 = \big(|\lambda_n| n^{-(\tau-2)/(\tau-1)}\big)^3 \sum_{j=1}^{n} w_j w_{\mathcal{C}(j)}\Big)^3. \tag{7.15}$$

The latter equality follows, since

$$\sum_{j=1}^{n} w_{\mathcal{C}_{\leq}(j)}^3 = \sum_{j=1}^{n}\sum_{i_1,i_2,i_3} w_{i_1}w_{i_2}w_{i_3}\mathbb{1}_{\{i_1\longrightarrow i_2,i_3,\min\mathcal{C}(i_1)=j\}} = \sum_{i_1,i_2} w_{i_1}w_{i_2}w_{i_3}\mathbb{1}_{\{i_1\longrightarrow i_2,i_3\}} = \sum_{i_1} w_{i_1}w_{\mathcal{C}(i_1)}^2.$$

Now, the summands for $j \geq K$ are small in probability by (6.28), while the summand for $j \leq K$ converge in probability by (6.2). Thus, also condition (c) follows from (6.2)-(6.28).

We continue with condition (a), which is equivalent to the statement that

$$\sigma_2(\mathbf{x}^{(n)}) = \frac{1}{|\lambda_n|} - \frac{\zeta}{\mu\lambda_n^2} + o_{\mathbb{P}}(\lambda_n^{-2}). \tag{7.16}$$

We shall prove (7.16) by a second moment method. We first identify

$$\sigma_2(\mathbf{x}^{(n)}) = n^{-2(\tau-2)/(\tau-1)}\sum_{j=1}^{n} w_{\mathcal{C}_{\leq}(j)}^2 = n^{-2(\tau-2)/(\tau-1)}\sum_{i=1}^{n} w_i w_{\mathcal{C}(i)}. \tag{7.17}$$

The latter equality follows, since

$$\sum_{j=1}^{n} w_{\mathcal{C}_{\leq}(j)}^2 = \sum_{j=1}^{n}\sum_{i_1,i_2} w_{i_1}w_{i_2}\mathbb{1}_{\{i_1\longrightarrow i_2,\min\mathcal{C}(i_1)=j\}} = \sum_{i_1,i_2} w_{i_1}w_{i_2}\mathbb{1}_{\{i_1\longrightarrow i_2\}} = \sum_{i_1} w_{i_1}w_{\mathcal{C}(i_1)}. \tag{7.18}$$

Thus, in order to prove (7.16), it suffices to show that

$$\mathbb{E}[\sum_{i=1}^{n} w_i w_{\mathcal{C}(i)}] = n^{2(\tau-2)/(\tau-1)}(|\lambda_n|^{-1} - \frac{\zeta}{\mu}\lambda_n^{-2} + o(\lambda_n^{-2})), \tag{7.19}$$

and

$$\mathrm{Var}(\sum_{i=1}^{n} w_i w_{\mathcal{C}(i)}) = o(n^{4(\tau-2)/(\tau-1)}\lambda_n^{-4}). \tag{7.20}$$

Indeed, by (7.19), we have that

$$\mathbb{P}\Big(\big|\sigma_2(\mathbf{x}^{(n)}) - \lambda_n^{-1} - \frac{\zeta}{\mu}\lambda_n^{-2}\big| \geq \varepsilon\lambda_n^{-2}\Big) \leq \mathbb{P}\Big(\big|\sigma_2(\mathbf{x}^{(n)}) - \mathbb{E}[\sigma_2(\mathbf{x}^{(n)})]\big| \geq \varepsilon\lambda_n^{-2}/2\Big), \tag{7.21}$$

which, by the Chebychev inequality is bounded by

$$\mathbb{P}\Big(\big|\sigma_2(\mathbf{x}^{(n)}) - \lambda_n^{-1} - \zeta\lambda_n^{-2}\big| \geq \varepsilon\lambda_n^{-2}\Big) \leq \frac{4\lambda_n^4}{\varepsilon^2}\mathrm{Var}(\sigma_2(\mathbf{x}^{(n)})) = o(1). \tag{7.22}$$

Thus, (7.16) follows from (7.19) and (7.20). For this, we shall apply Lemma A.3 from the appendix, in the setting that $\tilde{\nu}_n = \nu_n(1+\lambda_n l_n n^{-2(\tau-2)/(\tau-1)}) = 1+\lambda_n l_n n^{-2(\tau-2)/(\tau-1)} + \zeta n^{-(\tau-3)/(\tau-1)} + o(n^{-(\tau-3)/(\tau-1)})$, so that, by Lemma A.3(a),

$$\begin{aligned}
\mathbb{E}[\sum_{i=1}^{n} w_i w_{\mathcal{C}(i)}] &= \frac{\sum_{i=1}^n w_i^2}{1-\nu_n} + o(n^{2(\tau-2)/(\tau-1)}|\lambda_n|^{-2}) \\
&= \nu_n l_n(|\lambda_n| l_n n^{-2(\tau-2)/(\tau-1)} + \zeta n^{-(\tau-3)/(\tau-1)} + o(n^{-(\tau-3)/(\tau-1)}))^{-1} \\
&= |\lambda_n|^{-1}n^{2(\tau-2)/(\tau-1)} - \frac{\zeta}{\mu}n^{-(\tau-3)/(\tau-1)}\lambda_n^{-2} + o(\lambda_n^{-2}n^{2(\tau-2)/(\tau-1)}),
\end{aligned} \tag{7.23}$$

which proves (7.19).

By Lemma A.3(b), we have that

$$\mathrm{Var}(\sum_{i=1}^{n} w_i w_{\mathcal{C}(i)}) \leq C\mathbb{E}[w_M^3]\mathbb{E}[w_T]^4 + C\mathbb{E}[w_M^2]\mathbb{E}[w_T]^2\mathbb{E}[w_T^2] = o(n^{4(\tau-2)/(\tau-1)}\lambda_n^{-4}), \tag{7.24}$$

precisely when both terms in the middle inequality satisfy this bound. We complete the proof by checking these estimates. For the first contribution, we note that

$$\mathbb{E}[w_M^3]\mathbb{E}[w_T]^4 = \frac{1}{l_n(1-\nu_n)^4}\sum_{j=1}^{n} w_j^4 \leq \frac{C}{\lambda_n^4} n^{4/(\tau-1)+3(\tau-3)/(\tau-1)-1} = \frac{C}{\lambda_n^4} n^{(3\tau-7)/(\tau-1)} = o(\lambda_n^{-4} n^{4(\tau-2)/(\tau-1)}). \tag{7.25}$$

For the second contribution, instead,

$$\begin{aligned}\mathbb{E}[w_M^2]\mathbb{E}[w_T]^2\mathbb{E}[w_T^2] &= \frac{1}{l_n^2(1-\nu_n)^5}\Big(\sum_{j=1}^{n} w_j^3\Big)^2 \leq \frac{C}{|\lambda_n|^5} n^{6/(\tau-1)+5(\tau-3)/(\tau-1)-2} \\ &= \frac{C}{|\lambda_n|^5} n^{(3\tau-7)/(\tau-1)} = o(\lambda_n^{-4} n^{4(\tau-2)/(\tau-1)}).\end{aligned} \tag{7.26}$$

This proves the required concentration for $\sigma_2(\mathbf{x}^{(n)})$ and hence completes the proof of Theorem 1.4. □

A Appendix: auxiliary results

In this section, we prove an auxiliary result on the asymptotics of ν_n in (2.3).

Lemma A.1 (Sharp asymptotics of ν_n)**.** *Let the distribution function F satisfy (1.4), and let ν_n be given by (2.2) and ν by (1.8). Then,*

$$\nu_n = \nu + \zeta n^{-(\tau-3)/(\tau-1)} + o(n^{-(\tau-3)/(\tau-1)}), \tag{A.1}$$

where

$$\zeta = -\frac{c_F^{2/(\tau-1)}}{\mu}\sum_{i=1}^{\infty}\Big[\int_{i-1}^{i} u^{-2/(\tau-1)}du - i^{-2/(\tau-1)}\Big] < \infty. \tag{A.2}$$

Proof. We recall that

$$\nu_n = \frac{\sum_{i=1}^{n} w_i^2}{\sum_{i=1}^{n} w_i}. \tag{A.3}$$

By the asymptotics of l_n in (2.3), we have that

$$\nu_n = \frac{\sum_{i=1}^{n} w_i^2}{n\mathbb{E}[W]} + o(n^{-(\tau-3)/(\tau-1)}). \tag{A.4}$$

We shall make use of the fact that, when f is non-increasing,

$$f(i) \leq \int_{i-1}^{i} f(u)du \leq f(i-1). \tag{A.5}$$

Applying this to $f(u) = [1-F]^{-1}(u)^2$, which is non-increasing, we obtain in particular that, for any $K \geq 1$,

$$\int_{K/n}^{1} [1-F]^{-1}(u)^2 du - \frac{1}{n} w_{K/n}^2 \leq \frac{1}{n}\sum_{i=K+1}^{n} w_i^2 \leq \int_{K/n}^{1} [1-F]^{-1}(u)^2 du. \tag{A.6}$$

Now,

$$\frac{1}{n}w_{K/n}^2 = \Theta(K^{-2/(\tau-1)}n^{-(\tau-3)/(\tau-1)}). \tag{A.7}$$

Thus, we conclude that

$$\nu-\nu_n = \frac{1}{\mu n}\sum_{i=1}^{K}\int_{(i-1)/n}^{i/n}[1-F]^{-1}(u)^2du - \frac{1}{\mu n}\sum_{i=1}^{K}w_i^2 + \Theta(K^{-2/(\tau-1)}n^{-(\tau-3)/(\tau-1)}) + o(n^{-(\tau-3)/(\tau-1)}). \tag{A.8}$$

Next, by (1.4), for every $K \geq 1$ fixed,

$$\frac{1}{n}\sum_{i=1}^{K}w_i^2 = n^{-(\tau-3)/(\tau-1)}\sum_{i=1}^{K}\big(\frac{c_F}{i}\big)^{2/(\tau-1)} + o(n^{-(\tau-3)/(\tau-1)}), \tag{A.9}$$

$$\frac{1}{n}\sum_{i=1}^{K}\int_{(i-1)/n}^{i/n}[1-F]^{-1}(u)^2du = n^{-(\tau-3)/(\tau-1)}\sum_{i=1}^{K}\int_{i-1}^{i}\big(\frac{c_F}{u}\big)^{2/(\tau-1)}du. \tag{A.10}$$

Combining these two estimates yields that

$$n^{(\tau-3)/(\tau-1)}[\nu-\nu_n] = \frac{c_F^{2/(\tau-1)}}{\mu}\sum_{i=1}^{K}\Big[\int_{i-1}^{i}u^{-2/(\tau-1)}du - i^{-2/(\tau-1)}\Big] + \Theta(K^{-2/(\tau-1)}) + o(1). \tag{A.11}$$

Letting first $n\to\infty$ followed by $K\to\infty$, we conclude that

$$\lim_{n\to\infty} n^{(\tau-3)/(\tau-1)}[\nu_n-\nu] = \zeta, \tag{A.12}$$

where

$$\zeta = -\frac{c_F^{2/(\tau-1)}}{\mu}\sum_{i=1}^{\infty}\Big[\int_{i-1}^{i}u^{-2/(\tau-1)}du - i^{-2/(\tau-1)}\Big], \tag{A.13}$$

as required. The fact that $\zeta < \infty$ follows from the fact that, for $i \geq 2$,

$$0 \leq \int_{i-1}^{i}u^{-2/(\tau-1)}du - i^{-2/(\tau-1)} \leq (i-1)^{-2/(\tau-1)} - i^{-2/(\tau-1)}, \tag{A.14}$$

which is a summable sequence. □

Lemma A.2 (Branching process computations)**.** *The following formulas hold:*
(a)

$$\mathbb{E}[T] = \frac{1}{1-\nu_n}, \qquad \mathbb{E}[w_T] = \frac{\nu_n}{1-\nu_n}, \tag{A.15}$$

where

$$\nu_n = \mathbb{E}[\mathrm{Poi}(w_M)] = \frac{1}{l_n}\sum_{j=1}^{n}w_j^2. \tag{A.16}$$

(b)

$$\mathbb{E}[T^r] = \frac{(2r-1)!!}{(1-\nu_n)^{2r-1}}. \tag{A.17}$$

(c)

$$\mathbb{E}[w_T^2] = \Big(\frac{1}{l_n}\sum_{j=1}^{n}w_j^3\Big)\frac{1}{(1-\nu_n)^3}. \tag{A.18}$$

(d)

$$\mathbb{E}[T(i)] = 1 + \frac{w_i}{1-\nu_n}, \qquad \mathbb{E}[w_{T(i)}] = \frac{w_i}{1-\nu_n}. \tag{A.19}$$

(e)

$$\mathbb{E}[T(i)^2] = (1 + \frac{w_i}{1-\nu_n})^2 + \frac{w_i}{(1-\nu_n)^3}, \qquad \mathbb{E}[w_{T(i)}^2] = \Big(\frac{w_i}{1-\nu_n}\Big)^2 + \frac{w_i}{(1-\nu_n)^3}\Big(\frac{1}{l_n}\sum_{j=1}^n w_j^3\Big). \tag{A.20}$$

(f)

$$\mathbb{E}[w_{T(i)}T(i)] = w_i\big(1 + \frac{1}{\nu_n - 1}\big) + w_i^2\frac{\nu_n}{(1-\nu_n)^2} + w_i\Big(\frac{1}{(1-\nu_n)^2} + \frac{1}{1-\nu_n}\frac{1}{l_n}\sum_{j=1}^n w_j^3\Big). \tag{A.21}$$

Proof. (a) and (b) for T are standard, and their proofs shall be omitted. For w_T, we use that

$$\mathbb{E}[w_T] = \sum_{j=1}^n \mathbb{E}[w_T \mid M_1 = j]\frac{w_j}{l_n} = \sum_{j=1}^n \frac{w_j}{l_n}(w_j + w_j\mathbb{E}[w_T]). \tag{A.22}$$

Solving for $\mathbb{E}[w_T]$ proves the result.
For (c), we work out the square in

$$\mathbb{E}[w_T^2] = \mathbb{E}\Big[\Big(w_M + \sum_{j=1}^{\mathrm{Poi}(w_M)} w_{T_j}\Big)^2\Big], \tag{A.23}$$

are rearrange terms.
For (d), we note that

$$\mathbb{E}[T(i)] = \mathbb{E}\Big[1 + \sum_{j=1}^{\mathrm{Poi}(w_i)} T_j\Big] = 1 + w_i\mathbb{E}[T] = 1 + \frac{w_i}{1-\nu_n}, \tag{A.24}$$

by (a). Similarly,

$$\mathbb{E}[w_{T(i)}] = \mathbb{E}\Big[w_i + \sum_{j=1}^{\mathrm{Poi}(w_i)} w_{T_j}\Big] = w_i(1 + \mathbb{E}[w_T]) = w_i(1 + \frac{\nu_n}{1-\nu_n}) = \frac{w_i}{1-\nu_n}. \tag{A.25}$$

All other computations are similar, and we refrain from giving their proofs. □

Lemma A.3 (Mean and variance of $\sigma_2(\mathbf{x}^{(n)})$)**.** *When the weights $\boldsymbol{w}$ satisfy that $\nu_n < 1 - \lambda_n n^{-(\tau-3)/(\tau-1)}$, then*
(a)

$$\mathbb{E}[\sum_{i=1}^n w_i w_{\mathcal{C}(i)}] = \frac{l_n\nu_n}{1-\nu_n} + o(n^{2(\tau-2)/(\tau-1)}\lambda_n^{-2}); \tag{A.26}$$

(b)

$$\mathrm{Var}\Big(\sum_{i=1}^n w_i w_{\mathcal{C}(i)}\Big) \leq l_n\mathbb{E}[w_T^3] \leq C\big(\mathbb{E}[w_M^3]\mathbb{E}[w_T]^4 + \mathbb{E}[w_M^2]\mathbb{E}[w_T]^2\mathbb{E}[w_T^2]\big). \tag{A.27}$$

Proof. (a) We make use of the analysis in Section 6. We bound

$$\mathbb{E}[\sum_{i=1}^n w_i w_{\mathcal{C}(i)}] \leq \mathbb{E}[\sum_{i=1}^n w_i w_{T(i)}] = l_n\mathbb{E}[w_T] = \frac{l_n\nu_n}{1-\nu_n}. \tag{A.28}$$

For the lower bound, we make use of the bound, for any a_n and K_n (see (6.13)),

$$\mathbb{E}[\sum_{i=1}^n w_i w_{\mathcal{C}(i)}] \geq \mathbb{E}[w_T] - \sum_{i=1}^n w_i \mathbb{E}[w_{T(i)}] \mathbb{1}_{\{w_i > K_n\}} - \frac{1}{a_n} \sum_{i=1}^n w_i \mathbb{E}[w_{T(i)} T(i)] \mathbb{1}_{\{w_i \leq K_n\}} - a_n \sum_{i=1}^n w_i \mathbb{E}[T(i)] \sum_{j\in[n]} \mathbb{E}[w_{T(j)}] \frac{w_j^2}{l_n^2}. \tag{A.29}$$

Now,

$$\sum_{i=1}^n w_i \mathbb{E}[w_{T(i)}] \mathbb{1}_{\{w_i > K_n\}} = \sum_{i=1}^n \frac{w_i^2 \mathbb{1}_{\{w_i > K_n\}}}{1-\nu_n} = o(n^{2(\tau-2)/(\tau-1)} \lambda_n^{-2}) \tag{A.30}$$

whenever $K_n \gg |\lambda_n|^{1/(\tau-3)}$, which occurs surely when $K_n \gg n^{1/(\tau-1)}$, since $\lambda_n \geq -l_n^{-1} n^{2(\tau-2)/(\tau-1)} = \Theta(n^{(\tau-3)/(\tau-1)})$, and

$$\frac{1}{a_n} \sum_{i=1}^n w_i \mathbb{E}[w_{T(i)} T(i)] \mathbb{1}_{\{w_i \leq K_n\}} \leq \frac{C \sum_{i=1}^n w_i^3 \mathbb{1}_{\{w_i \leq K_n\}}}{a_n l_n (1-\nu_n)^2} \leq \frac{C K_n}{a_n (1-\nu_n)^2} = o(n^{2(\tau-2)/(\tau-1)} |\lambda_n|^{-2}) \tag{A.31}$$

precisely when $a_n \gg K_n n^{-2/(\tau-1)}$. Further,

$$a_n \sum_{i=1}^n w_i \mathbb{E}[T(i)] \sum_{j\in[n]} \mathbb{E}[w_{T(j)}] \frac{w_j^2}{l_n^2} \leq a_n \frac{C}{l_n(1-\nu_n)^2} \sum_{j\in[n]} w_j^3 = o(n^{2(\tau-2)/(\tau-1)} \lambda_n^{-2}), \tag{A.32}$$

precisely when $a_n = o(n^{(\tau-2)/(\tau-1)})$. Thus, for every $\lambda_n \geq -l_n n^{-2(\tau-2)/(\tau-1)}$, we can find an appropriate K_n and a_n so that both terms are $o(n^{2(\tau-2)/(\tau-1)} \lambda_n^2)$.
(b) We shall start by bounding the second moment. For this, we rewrite

$$\mathbb{E}\Big[\Big(\sum_{i=1}^n w_i w_{\mathcal{C}(i)}\Big)^2\Big] = \sum_{i_1,i_2} w_{i_1} w_{i_2} \mathbb{E}[w_{\mathcal{C}(i_1)} w_{\mathcal{C}(i_2)}]. \tag{A.33}$$

Now we split

$$\mathbb{E}[w_{\mathcal{C}(i_1)} w_{\mathcal{C}(i_2)}] = \mathbb{E}[w_{\mathcal{C}(i_1)} w_{\mathcal{C}(i_2)} \mathbb{1}_{\{i_1 \longrightarrow i_2\}}] + \mathbb{E}[w_{\mathcal{C}(i_1)} w_{\mathcal{C}(i_2)} \mathbb{1}_{\{i_1 \not\longrightarrow i_2\}}]. \tag{A.34}$$

The second term is bounded from above by

$$\mathbb{E}[w_{\mathcal{C}(i_1)}] \mathbb{E}[w_{\mathcal{C}(i_2)}], \tag{A.35}$$

since the clusters have to make use of different sets of vertices. Therefore, summing over i_1, i_2, we obtain that

$$\mathrm{Var}\Big(\sum_{i=1}^n w_i w_{\mathcal{C}(i)}\Big) \leq \sum_{i_1,i_2} \mathbb{E}[w_{\mathcal{C}(i_1)} w_{\mathcal{C}(i_2)} \mathbb{1}_{\{i_1 \longrightarrow i_2\}}] = \sum_i w_i \mathbb{E}[w_{\mathcal{C}(i)}^3] \leq l_n \sum_i \frac{w_i}{\mathbb{E}} [w_{T(i)}^3] = l_n \mathbb{E}[w_T^3]. \tag{A.36}$$

The upper bound on $\mathbb{E}[w_T^3]$ follows as in the proof of Lemma A.2. □

Acknowledgements. The work of SB was supported by PIMS and NSERC, Canada. The work of RvdH and JvL was supported in part by the Netherlands Organisation for Scientific Research (NWO).

References

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

[2] D. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, pages 3–48, 1999.

[3] D. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab*, 3:1–59, 1998.

[4] J. Bertoin. *Lévy processes*, volume **121** of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, (1996).

[5] J. Bertoin. *Random fragmentation and coagulation processes.* Cambridge University Press, 2006.

[6] S. Bhamidi, R. van der Hofstad, and J.S.H. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. Preprint (2009).

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

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

[9] 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).

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

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

[12] 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).

[13] 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).

[14] 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).

[15] G.R. Grimmett. *Percolation.* Springer-Verlag, Berlin, second edition, (1999).

[16] G.R. Grimmett and D.R. Stirzaker. *Probability and random processes.* Oxford University Press, New York, third edition, (2001).

[17] R. van der Hofstad. Critical behavior in inhomogeneous random graphs. Preprint (2009).

[18] R. van der Hofstad, S. Janson, and M. Luczak. The near-critical behavior for the configuration model with finite-variance degrees. In preparation.

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

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

[21] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, (2002).

[22] A. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications.* Universitext. Springer-Verlag, Berlin, (2006).

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

[24] A. Pakes. A hitting time for Lévy processes, with applications to dams and branching processes. *Ann. Fac. Sci. Toulouse Math. (6)*, **5**(3):521–544, (1996).

[25] H. Thorisson. *Coupling, stationarity, and regeneration.* Probability and its Applications (New York). Springer-Verlag, New York, (2000).

[26] T.S. Turova. Diffusion approximation for the components in critical inhomogeneous random graphs of rank 1. (2009).