Electron. J. Probab. **27** (2022), article no. 103, 1–29.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP821

Global lower mass-bound for critical configuration models in the heavy-tailed regime*

Shankar Bhamidi† Souvik Dhara‡ Remco van der Hofstad§
Sanchayan Sen¶

Abstract

We establish the global lower mass-bound property for the largest connected components in the critical window for the configuration model when the degree distribution has an infinite third moment. The scaling limit of the critical percolation clusters, viewed as measured metric spaces, was established in our prior work with respect to the Gromov-weak topology. Our result extends those scaling limit results to the stronger Gromov-Hausdorff-Prokhorov topology under slightly stronger assumptions on the degree distribution. This implies the distributional convergence of global functionals such as the diameter of the largest critical components. Further, our result gives a sufficient condition for compactness of the random metric spaces that arise as scaling limits of critical clusters in the heavy-tailed regime.

Keywords: global lower mass-bound; critical configuration model; heavy-tailed degrees.
MSC2020 subject classifications: 60C05; 05C80.
Submitted to EJP on May 8, 2020, final version accepted on July 3, 2022.

*The authors are grateful to two anonymous referees for their careful reading and many comments and suggestions on an earlier version of the paper. SB was partially supported by NSF grants DMS-1613072, DMS-1606839, and ARO grant W911NF-17-1-0010. SD and RvdH were supported by the Netherlands Organisation for Scientific Research (NWO) through Gravitation Networks grant 024.002.003. In addition, RvdH was supported by VICI grant 639.033.806. SD also thanks Elchanan Mossel for supporting his research through Vannevar Bush Faculty Fellowship (ONR-N00014-20-1-2826), Simons Investigator award (622132). SS has been supported in part by the Infosys foundation, Bangalore, and by MATRICS grant MTR/2019/000745 from SERB. SD thanks Eindhoven University of Technology, where a major part of this work was done.

†Department of Statistics and Operations Research, University of North Carolina.
E-mail: bhamidi@email.unc.edu
‡Simons Institute for the Theory of Computing, University of California, Berkeley.
E-mail: dharasouvik1991@gmail.com
§Department of Mathematics and Computer Science, Eindhoven University of Technology.
E-mail: bhamidi@email.unc.edu
¶Department of Mathematics, Indian Institute of Science.
E-mail: sanchayan.sen1@gmail.com

1 Introduction

Any finite, connected graph $\mathscr{C}$ can be viewed as a metric space with the distance between points given by $a\mathrm{d}(\cdot,\cdot)$ for some constant $a>0$, where $\mathrm{d}(\cdot,\cdot)$ is used as a generic notation to denote the graph-distance (i.e., number of edges in the shortest path between vertices). There is a natural probability measure μ associated to the metric space $(\mathscr{C}, a\mathrm{d})$ given by $\mu(A)=|A|/|\mathscr{C}|$ for any $A\subset\mathscr{C}$, where $|A|$ denotes the number of vertices in A. We denote this metric measure space by $(\mathscr{C},a)$. Fix any $\delta>0$ and define the δ-lower mass of $(\mathscr{C},a)$ by

$$\mathfrak{m}(\delta) := \frac{\inf_{u\in\mathscr{C}}\big|\{v\in\mathscr{C}: a\mathrm{d}(v,u)\le\delta\}\big|}{|\mathscr{C}|}. \tag{1.1}$$

Thus, $\mathfrak{m}(\delta)$ is the least *mass* in any δ-neighborhood of a vertex in $(\mathscr{C},a)$. For a sequence $(\mathscr{C}_n,a_n)_{n\ge1}$ of graphs viewed as metric measure spaces, the global lower mass-bound property is defined as follows:

Definition 1 (Global lower mass-bound property [5])**.** For $\delta>0$, let $\mathfrak{m}_n(\delta)$ denote the δ-lower mass of $(\mathscr{C}_n,a_n)$. Then $(\mathscr{C}_n,a_n)_{n\ge1}$ is said to satisfy the global lower mass-bound property if and only if $\sup_{n\ge1}\mathfrak{m}_n(\delta)^{-1}<\infty$ for any $\delta>0$. When $(\mathscr{C}_n)_{n\ge1}$ is a collection of random graphs, $(\mathscr{C}_n,a_n)_{n\ge1}$ is said to satisfy the global lower mass-bound property if and only if $(\mathfrak{m}_n(\delta)^{-1})_{n\ge1}$ is a tight sequence of random variables for any $\delta>0$.

The aim of this paper is to prove the global lower mass-bound property for largest connected components of random graphs with given degrees (configuration model) at criticality, when the third moment of the empirical degree distribution tends to infinity (Theorem 1.5). Informally speaking, the global lower mass-bound property ensures that all the small neighborhoods of vertices in the 'large' critical component have mass bounded away from zero, so that the component does not have any *light spots* and the total mass is well-distributed over the whole component. This has several interesting consequences in the theory of critical random graphs. Our main motivation comes from the work of Athreya, Löhr, and Winter [5], who have shown that the global lower mass-bound property can be used to prove Gromov-Hausdorff-Prokhorov (GHP) convergence of random metric spaces. In a previous paper [7], we have studied the critical percolation clusters for the configuration model in the heavy-tailed universality class. We have proved that the ordered vector of components converges in distribution to suitable random objects in the Gromov-weak topology. The global lower mass-bound in this paper shows that the result of [7] in fact holds with respect to the stronger GHP-topology. One motivating reason for proving the GHP-convergence is that it yields the scaling limit of global functionals like the diameter of large critical components. Finding the scaling limit for the diameter of critical components is a daunting task even for the Erdős-Rényi random graph. Nachmias and Peres [31] estimated the tail probabilities of the diameter, but showing a distributional convergence result was a difficult question, until the seminal paper by Addario-Berry, Broutin and Goldschmidt [2] that proved the GHP-convergence for critical Erdős-Rényi random graphs. As a corollary of Theorem 1.5, we also get distributional convergence of the suitably rescaled diameter of the critical percolation clusters in the heavy-tailed regime (Theorem 1.9), where the scaling limit and exponents turn out to be different than those for the Erdős-Rényi case.

We will further discuss the applications and the scope of this work as well as its technical contributions after stating our results in Section 1.3. We start by defining the configuration model and state the precise assumptions.

1.1 The configuration model

Consider a non-increasing sequence of degrees $\boldsymbol{d} = (d_i)_{i\in[n]}$ such that $\ell_n = \sum_{i\in[n]} d_i$ is even. For notational convenience, we suppress the dependence of the degree sequence on n. The configuration model on n vertices having degree sequence $\boldsymbol{d}$ is constructed as follows [6, 11]:

> Equip vertex j with d_j stubs, or *half-edges*. Two half-edges create an edge once they are paired. Therefore, initially we have $\ell_n = \sum_{i\in[n]} d_i$ half-edges. Pick any one half-edge and pair it with another uniformly chosen half-edge from the remaining unpaired half-edges and keep repeating the above procedure until all the unpaired half-edges are exhausted.

Let $\mathrm{CM}_n(\boldsymbol{d})$ denote the graph constructed by the above procedure. Note that $\mathrm{CM}_n(\boldsymbol{d})$ may contain self-loops or multiple edges. Given any degree sequence, let $\mathrm{UM}_n(\boldsymbol{d})$ denote the graph chosen uniformly at random from the collection of all simple graphs with degree sequence $\boldsymbol{d}$. It can be shown that the conditional law of $\mathrm{CM}_n(\boldsymbol{d})$, conditioned on it being simple, is the same as $\mathrm{UM}_n(\boldsymbol{d})$ (see e.g. [23, Proposition 7.13]).

1.2 Main results

Fix a constant $\tau \in (3, 4)$, which will denote the power-law exponent of the asymptotic degree distribution of $\mathrm{CM}_n(\boldsymbol{d})$. Throughout this paper we will use the shorthand notation

$$\alpha = 1/(\tau-1), \quad \rho = (\tau-2)/(\tau-1), \quad \eta = (\tau-3)/(\tau-1). \tag{1.2}$$

We use the standard notation of $\xrightarrow{\mathrm{P}}$ and $\xrightarrow{d}$ to denote convergence in probability and in distribution, respectively. Also, we use a generic notation C to denote a positive universal constant whose exact value may change from line to line. We use Bachmann–Landau asymptotic notation $o(\cdot)$, $O(\cdot)$, $\Theta(\cdot)$, $\omega(\cdot)$, $\Omega(\cdot)$. A sequence of events $(\mathcal{E}_n)_{n\geq 1}$ is said to occur with high probability (whp) with respect to the probability measures $(\mathbb{P}_n)_{n\geq 1}$ when $\mathbb{P}_n(\mathcal{E}_n) \to 1$. For (random) variables X_n and Y_n, define $X_n = O_{\mathrm{P}}(Y_n)$ when $(|X_n|/|Y_n|)_{n\geq 1}$ is a tight sequence; $X_n = o_{\mathrm{P}}(Y_n)$ when $X_n/Y_n \xrightarrow{\mathrm{P}} 0$; $X_n = \Theta_{\mathrm{P}}(Y_n)$ if both $X_n = O_{\mathrm{P}}(Y_n)$ and $Y_n = O_{\mathrm{P}}(X_n)$.

We first state the general assumptions that are used to prove scaling limits for critical configuration models with heavy-tailed degree distributions as identified previously in [7, 16]:

Assumption 1.1 (General assumptions)**.** For each $n \geq 1$, let $\boldsymbol{d} = \boldsymbol{d}_n = (d_1, \ldots, d_n)$ be a degree sequence satisfying $d_1 \geq d_2 \geq \ldots \geq d_n$. We assume the following about $(\boldsymbol{d}_n)_{n\geq 1}$ as $n \to \infty$:

(i) (*High-degree vertices*) For each fixed $i \geq 1$,

$$n^{-\alpha} d_i \to \theta_i, \tag{1.3}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots) \in \ell^3_{\downarrow} \setminus \ell^2_{\downarrow}$, where $\ell^p_{\downarrow} := \{(x_i)_{i\geq 1} : x_1 \geq x_2 \geq \ldots \text{ and } \sum_i x_i^p < \infty\}$.

(ii) (*Moment assumptions*) Let D_n denote the degree of a typical vertex, i.e., a vertex chosen uniformly at random from the vertex set $[n]$, independently of $\mathrm{CM}_n(\boldsymbol{d})$. Then, D_n converges in distribution to some discrete random variable D and

$$\mathbb{E}[D_n] = \frac{1}{n}\sum_{i\in[n]} d_i \to \mu := \mathbb{E}[D], \qquad \mathbb{E}[D_n^2] = \frac{1}{n}\sum_{i\in[n]} d_i^2 \to \mu_2 := \mathbb{E}[D^2], \tag{1.4}$$

$$\lim_{K\to\infty} \limsup_{n\to\infty} n^{-3\alpha} \sum_{i=K+1}^{n} d_i^3 = 0. \tag{1.5}$$

(iii) Let n_1 be the number of degree-one vertices. Then $n_1 = \Theta(n)$, which is equivalent to assuming that $\mathbb{P}(D = 1) > 0$.

Remark 1.2. As important examples, Assumption 1.1 was shown to hold when the degree distribution is power-law with exponent $\tau \in (3, 4)$ [16, Section 2]. More precisely, if F is a distribution function on the nonnegative integers satisfying $[1 - F](x) = (1 + o(1))Cx^{-(\tau-1)}$ as $x \to \infty$, then Assumptions 1.1(i), 1.1(ii) are satisfied when (a) $d_i = [1 - F]^{-1}(i/n)$, and when (b) d_i are the order statistics of an i.i.d. sample from F (we add a dummy half-edge to vertex 1 if $\sum_{i\in[n]} d_i$ is odd). Assumptions 1.1(iii) is also satisfied in these examples if F has non-zero mass at 1.

We further assume that the configuration model lies within the critical window of the phase transition, i.e., for some $\lambda \in \mathbb{R}$,

$$\nu_n = \frac{\sum_{i\in[n]} d_i(d_i - 1)}{\sum_{i\in[n]} d_i} = 1 + \lambda n^{-\eta} + o(n^{-\eta}). \tag{1.6}$$

Denote the i-th largest connected component of $\mathrm{CM}_n(\boldsymbol{d})$ by $\mathscr{C}_{(i)}$, breaking ties arbitrarily. For each $v \in [n]$ and $\delta > 0$, let $\mathcal{N}_v(\delta)$ denote the δn^{η} neighborhood of v in $\mathrm{CM}_n(\boldsymbol{d})$ in the graph distance. For each $i \geq 1$, define

$$\mathfrak{m}_i^n(\delta) = \inf_{v\in\mathscr{C}_{(i)}} n^{-\rho}|\mathcal{N}_v(\delta)|. \tag{1.7}$$

Our goal is to prove the global lower mass-bound property for the critical components $\mathscr{C}_{(i)}$. For $\mathrm{CM}_n(\boldsymbol{d})$ satisfying Assumption 1.1 and (1.6), it was shown in [16, Theorem 1] that

$$(n^{-\rho}|\mathscr{C}_{(i)}|)_{i\geq 1} \xrightarrow{d} (\xi_i)_{i\geq 1}, \tag{1.8}$$

with respect to the $\ell^2_{\downarrow}$-topology, where the ξ_i's are non-degenerate random variables with support $(0, \infty)$. Therefore, it is enough to rescale by n^{ρ} in (1.7) instead of the component sizes as given in (1.1). In order to prove tightness of $\mathfrak{m}_i^n(\delta)$, we will need a further technical assumption on the degrees.

Assumption 1.3. Let V_n^* be a vertex chosen in a size-biased manner with sizes being $(d_i/\ell_n)_{i\in[n]}$, i.e., $\mathbb{P}(V_n^* = i) = d_i/\ell_n$, and let D_n^* be the degree of V_n^*. There exist constants $c_0 > 0$ and $c_1 > 1$ such that for all $n \geq 1$,

$$\mathbb{P}(l < D_n^* \leq c_1 l) \geq \frac{c_0}{l^{\tau-2}} \quad \text{for} \quad 1 \leq l < d_1\,. \tag{1.9}$$

Remark 1.4. Assumption 1.3 says that the mass distribution in the tail of D_n^* is well-behaved in the sense that we have a uniform (over n) lower bound of the form (1.9). Such lower bounds can be used to obtain tail-bounds on the heights of branching processes; see Proposition 4.7 below. (See also [1, Theorem 1.3].) It can be easily shown that Assumption 1.3 holds in the examples discussed in Remark 1.2 by observing that the size-biased distribution is a power-law with exponent $\tau - 1$.

The following theorem is the main result of this paper:

Theorem 1.5 (Global lower mass-bound for $\mathrm{CM}_n(\boldsymbol{d})$)**.** *Suppose that* Assumptions 1.1, 1.3 *and the criticality condition* (1.6) *hold. Then, for each fixed $i \geq 1$, $(\mathscr{C}_{(i)}, n^{-\eta})_{n\geq 1}$ satisfies the global lower mass-bound, i.e., for any $\delta > 0$, the sequence $(\mathfrak{m}_i^n(\delta)^{-1})_{n\geq 1}$ is tight.*

By [24, Theorem 1.1], under the condition (1.4) in Assumption 1.1,

$$\liminf_{n\to\infty} \mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}) > 0. \tag{1.10}$$

This immediately implies the following:

Theorem 1.6 (Global lower mass-bound for $\mathrm{UM}_n(\boldsymbol{d})$)**.** *Under* Assumption 1.1, 1.3 *and* (1.6)*, the largest components of* $\mathrm{UM}_n(\boldsymbol{d})$ *also satisfy the global lower mass-bound property.*

Next we state another important corollary, which says that the global lower mass-bound property is also satisfied by critical percolation clusters in $\mathrm{CM}_n(\boldsymbol{d})$ and $\mathrm{UM}_n(\boldsymbol{d})$. To this end, let us assume that

$$\lim_{n\to\infty} \frac{\sum_{i\in[n]} d_i(d_i-1)}{\sum_{i\in[n]} d_i} = \nu > 1. \tag{1.11}$$

In this regime, $\mathrm{CM}_n(\boldsymbol{d})$ is supercritical in the sense that there exists a unique *giant* component whp for $\nu > 1$, and when $\nu < 1$, all the components have size $o_{\mathbb{P}}(n)$ [26, 30]. Percolation refers to deleting each edge of a graph independently with probability $1-p$. The critical window for percolation on $\mathrm{CM}_n(\boldsymbol{d})$ in the heavy-tailed setting was studied in [7, 16], and is defined by the values of p given by

$$p_c(\lambda) = \frac{1}{\nu_n} + \frac{\lambda}{n^\eta} + o(n^{-\eta}). \tag{1.12}$$

Let $\mathscr{C}_{(i)}(p_c(\lambda))$ denote the i-th largest component of the graph obtained by percolation with probability $p_c(\lambda)$ on the graph $\mathrm{CM}_n(\boldsymbol{d})$. Then the following result holds:

Theorem 1.7 (Global lower mass-bound for critical percolation)**.** *Under* Assumptions 1.1(i), 1.1(ii), 1.3, (1.11) *and* (1.12)*,* $(\mathscr{C}_{(i)}(p_c(\lambda)), n^{-\eta})_{n\geq 1}$ *satisfies the global lower mass-bound property, for each fixed* $i \geq 1$*. This result also holds for percolation on* $\mathrm{UM}_n(\boldsymbol{d})$*.*

Let $\mathcal{G}_n$ denote the graph obtained by doing percolation with edge retention probability $p_c(\lambda)$ (defined in (1.12)) on $\mathrm{CM}_n(\boldsymbol{d})$. Let $\boldsymbol{d}^p = (d_i^p)_{i\in[n]}$ denote the degree sequence of $\mathcal{G}_n$. By [19, Lemma 3.2], the conditional law of $\mathcal{G}_n$, conditionally on $\boldsymbol{d}^p$, is same as the law of $\mathrm{CM}_n(\boldsymbol{d}^p)$. Thus, Theorem 1.7 follows from Theorem 1.5 if we can show that the percolated degree sequence $\boldsymbol{d}^p$ satisfies (with possibly different parameters) Assumptions 1.1 and 1.3 with high probability when the original degree sequence $(d_i)_{i\in[n]}$ satisfies Assumptions 1.1(i), 1.1(ii), 1.3, and also (1.6) holds for $\boldsymbol{d}^p$ if further the percolation probability is given by (1.12). The verification of these assumptions are provided in Section 5.

Remark 1.8. It is worthwhile to point out that Theorem 1.5 can be proved when the $\mathscr{C}_{(i)}$'s are endowed with a more general measure rather than the counting measure. To be precise, for any sequence of vertex weights $(w_v)_{v\in[n]}$, the component $\mathscr{C}_{(i)}$ can be equipped with the measure $\mu_{(i)}(A) = \sum_{v\in A} w_v / \sum_{v\in\mathscr{C}_{(i)}} w_v$, for any $A \subset \mathscr{C}_{(i)}$. Then Theorem 1.5 can also be proved using identical methods as in this paper, with the additional assumptions that

$$\lim_{n\to\infty} \frac{1}{\ell_n} \sum_{i\in[n]} d_i w_i = \mu_w, \quad \max\left\{ \sum_{i\in[n]} d_i w_i^2, \sum_{i\in[n]} d_i^2 w_i \right\} = O(n^{3\alpha}).$$

These additional assumptions are required when we apply the results from [16] (see [16, Theorem 21]). We adopted the simpler version of the counting measure here because it relates directly to [7, Theorem 2.1].

1.3 Discussion

Scaling limit of critical percolation clusters We write $n^{-\eta}\mathscr{C}_{(i)}(p_c(\lambda))$ to denote the i-th largest component of $\mathrm{CM}_n(\boldsymbol{d}, p_c(\lambda))$, viewed as a measured metric space with the metric being the graph distance re-scaled by n^η, and the measure being proportional to the counting measure. Athreya, Löhr, and Winter [5] showed that the global lower

mass-bound property forms a crucial ingredient to prove convergence of random metric spaces such as $n^{-\eta}\mathscr{C}_{(i)}(p_c(\lambda))$ with respect to the Gromov-Hausdorff-Prokhorov (GHP) topology on the space of compact metric spaces. The other key ingredient is the scaling limit for $n^{-\eta}\mathscr{C}_{(i)}(p_c(\lambda))$ with respect to the Gromov-weak topology, which was established in [7, Theorem 2.1]. The Gromov-weak topology is an analogue of finite-dimensional convergence, since it considers distances between a finite number of sampled points from the underlying metric space. Thus, global functionals such as the diameter are not continuous with respect to this topology. Indeed, it may be the case that there is a long path of growing length, that has asymptotically negligible mass. In our context, the problem could arise due to paths of length much larger than n^{η}. The global lower mass-bound property ensures that the components have sufficient mass everywhere. This forbids the existence of long thin paths, when the total mass of the component converges. For this reason, Gromov-weak convergence and global lower mass-bound together imply GHP-convergence when the support of the limiting measure is the entire limiting space [5, Theorem 6.1]. For formal definitions of the Gromov-weak topology, and the GHP-topology on the space of compact measured metric spcaes, we refer the reader to [5, 8, 22].

Following the above discussion, the next theorem is a direct consequence of Theorem 1.7, [7, Theorem 2.3] and [5, Theorem 6.1]: Let $\mathbb{M}$ denote the space of measured compact metric spaces equipped with the GHP-topology, and let $\mathbb{M}^{\mathbb{N}}$ denote the product space with the associated product topology.

Theorem 1.9 (GHP convergence of critical percolation clusters)**.** *There exists a sequence of measured metric spaces* $(\mathscr{M}_i)_{i\geq 1} = ((M_i, \mathrm{d}_i, \mu_i))_{i\geq 1} \in \mathbb{M}^{\mathbb{N}}$ *such that, under* Assumptions 1.1(i), 1.1(ii), 1.3, (1.11) *and* (1.12)*, as* $n \to \infty$*,*

$$(n^{-\eta}\mathscr{C}_{(i)}(p_c(\lambda)))_{i\geq 1} \xrightarrow{d} (\mathscr{M}_i)_{i\geq 1} \quad \textit{in } \mathbb{M}^{\mathbb{N}}. \tag{1.13}$$

Moreover, the results also hold for $\mathrm{UM}_n(\boldsymbol{d}, p_c(\lambda))$*.*

The exact description of the space $\mathscr{M}_i$ can be found in [7]. It is worthwhile mentioning a recent work by Conchon-Kerjan and Goldschmidt [15] which is closely related to Theorem 1.9. Conchon-Kerjan and Goldschmidt [15] deduce scaling limits for the vector of components in GHP-topology for critical configuration models having i.i.d power law degrees with exponent $\tau \in (3, 4)$. In Remarks 1.2 and 1.4, we noted that Assumptions 1.1(i), 1.1(ii), and 1.3 hold when the degrees are i.i.d samples from a power-law distribution with exponent $\tau \in (3, 4)$. Therefore, Theorem 1.9 implies that the conditional law of $(n^{-\eta}\mathscr{C}_{(i)}(p_c(\lambda)))_{i\geq 1}$, conditioned on the i.i.d degree sequence, converges to the law of $(\mathscr{M}_i)_{i\geq 1}$ in $\mathbb{M}^{\mathbb{N}}$ for almost every realization of the i.i.d degree sequence. Hence, Theorem 1.9 gives a quenched result whereas [15] proves an annealed result. The method of [15] relies on an alternative approach showing convergence of the height processes corresponding to the components. The associated limiting object was studied in [21], which interestingly turns out to have a quite different description than those in [7, 8].

Scaling limit of maximal distances For any metric space (X, d) and a point $x \in X$, define the radius of x in X and the diameter of X by

$$\mathrm{Rad}(x, X) = \sup_{y\in X} \mathrm{d}(x, y) \quad \text{and} \quad \mathrm{diam}(X) = \sup_{x\in X} \mathrm{Rad}(x, X) = \sup_{x,y\in X} \mathrm{d}(x, y). \tag{1.14}$$

An important corollary of Theorem 1.9 is the convergence of the radius and the diameter of the critical components: Let $V_{n,i}$ be a uniformly chosen vertex in $\mathscr{C}_{(i)}(p_c(\lambda))$, where $(V_{n,i})_{i\geq 1}$ is an independent collection conditionally on $(\mathscr{C}_{(i)}(p_c(\lambda)))_{i\geq 1}$. Similarly, using

the notation of the scaling limits in Theorem 1.9, let V_i be chosen from M_i according to the measure μ_i and let $(V_i)_{i\geq 1}$ be an independent collection conditionally on $(\mathscr{M}_i)_{i\geq 1}$.

Corollary 1.10 (Convergence of radius and diameter)**.** *Under* Assumptions 1.1(i), 1.1(ii), *1.3*, (1.11) *and* (1.12)*, as* $n\to\infty$*,*

$$\begin{aligned}\big(n^{-\eta}\mathrm{Rad}(V_{n,i},\mathscr{C}_{(i)}(p_c(\lambda)))\big)_{i\geq 1} &\xrightarrow{d} (\mathrm{Rad}(V_i,\mathscr{M}_i))_{i\geq 1},\\ \big(n^{-\eta}\mathrm{diam}(\mathscr{C}_{(i)}(p_c(\lambda)))\big)_{i\geq 1} &\xrightarrow{d} (\mathrm{diam}(\mathscr{M}_i))_{i\geq 1},\end{aligned} \tag{1.15}$$

with respect to the product topology, where $(\mathscr{M}_i)_{i\geq 1}$ *is given by* Theorem 1.9*. Moreover, the result also holds for* $\mathrm{UM}_n(\boldsymbol{d})$*.*

Proving scaling limits for the diameter of the critical tree-like objects is often a difficult task. In [32], Szekeres proved that, for the uniform random rooted labelled tree on m vertices, the diameter, rescaled by $\sqrt{m}$, converges in distribution. Szekeres also provided an explicit formula for the density of the limiting distribution in [32, Page 395, (12)]. Szekeres' method was based on generating functions. Łuczak [28] also considered enumeration of trees with diameter $\gg\sqrt{m}$. On the other hand, Aldous [3] (see [3, Section 3.4]) noted that the GHP-convergence can be used as an effective tool to prove scaling limit results for the diameter. This is the motivating idea behind Corollary 1.10. Aldous [3] also raised a natural question whether it is possible to obtain an explicit formula from a result such as Corollary 1.10. In a recent paper, Wang [33] showed that it is indeed possible to get such a formula for the Brownian tree. In the context of Corollary 1.10, the difficulty is two-fold: First, the critical components have surplus edges. For the scaling limits of critical Erdős-Rényi random graphs, Miermont and Sen [29] recently gave a breadth-first construction, which yields an alternative description of the scaling limit of the radius function from a fixed point (rescaled by $n^{1/3}$). However, the description for the diameter and an explicit formula such as the one by Wang [33] is still an open question. Second, the scaling limit in Corollary 1.10 is in the heavy-tailed universality class. Even for **p**-trees (see [14]) that satisfy $p_i/(\sum_i p_i^2)^{1/2}\to\beta_i>0$, with $(\beta_i)_{i\geq 1}\in\ell^2_\downarrow\setminus\ell^1_\downarrow$, obtaining an explicit description for the limiting distribution of the diameter is an interesting question.

Compactness of the limiting metric space The limiting spaces $\mathscr{M}_i$ are constructed by tilting the distribution of an inhomogeneous continuum random tree (ICRT), and then identifying a Poisson number of vertices to create cycles. This object is well-defined as a metric measure space for $\boldsymbol{\theta}\in\ell^3_\downarrow\setminus\ell^2_\downarrow$. However, it may not be compact for all $\boldsymbol{\theta}\in\ell^3_\downarrow\setminus\ell^2_\downarrow$. It is interesting to find an explicit criterion for the compactness of the limiting objects $\mathscr{M}_i$ in terms of the underlying parameters.

Indeed, in the context of compactness of ICRTs, Aldous, Miermont, and Pitman [4, Section 7] state an additional condition, which was conjectured to be necessary and sufficient for the compactness of ICRTs. This conjectured was recently proved in [10]. In the context of critical random graphs, a recent paper by Broutin, Duquesne, and Wang [13] shows that the following criterion, analogous to [4], is sufficient for the almost sure compactness of $\mathscr{M}_i$ [1]:

[1] **Note:** The condition (1.17) does not hold for all $\boldsymbol{\theta}\in\ell^3_\downarrow\setminus\ell^2_\downarrow$. Indeed, take $\theta_i=i^{-1/2}$. For $u\in(\theta_2^{-1},\infty)$, let $i_0=i_0(u)$ be such that $\theta_{i_0}^{-1}<u\leq\theta_{i_0+1}^{-1}$, i.e., $\sqrt{i_0}<u\leq\sqrt{i_0+1}$. Then,

$$\Psi_\theta(u)\leq C\Big[\sum_{i\leq i_0}\theta_i(u\theta_i)+\sum_{i>i_0}\theta_i(u\theta_i)^2\Big]=C\Big[u\sum_{i<u^2}\frac{1}{i}+u^2\sum_{i\geq u^2}\frac{1}{i^{3/2}}\Big]\leq C[u\log u+u], \tag{1.16}$$

and thus (1.17) cannot hold.

$$\int_1^\infty \frac{\mathrm{d}u}{\Psi_\theta(u)} < \infty, \quad \text{where} \quad \Psi_\theta(u) = \sum_{i\geq 1} \theta_i(\mathrm{e}^{-u\theta_i} - 1 + u\theta_i). \tag{1.17}$$

Our GHP convergence from Theorem 1.9 indirectly yields a sufficient condition for the compactness of the limiting metric space almost surely, by considering an asymptotic version of Assumption 1.3: Suppose $\boldsymbol{\theta} \in \ell^3_\downarrow \setminus \ell^2_\downarrow$ and there exist constants $c_0 > 0$ and $c_1 > 1$ such that

$$x^{\tau-2} \times \sum_{i=1}^\infty \theta_i \mathbb{1}\{x < \theta_i \leq c_1 x\} \geq c_0 \quad \text{for all } x \in (0, \theta_1). \tag{1.18}$$

The fact that (1.18) is a sufficient condition for the compactness of $\mathscr{M}_i$ follows immediately from Theorem 1.9 and the following proposition:

Proposition 1.11. *Consider any $\boldsymbol{\theta} \in \ell^3_\downarrow \setminus \ell^2_\downarrow$ such that* (1.18) *holds. Then there exists a sequence of degree sequences satisfying* Assumptions 1.1(i), 1.1(ii), 1.3, *and* (1.11).

We will prove Proposition 1.11 in Appendix B. A natural question is how the conditions in (1.17) and (1.18) compare. We argue below that, in fact, (1.18) is strictly stronger than (1.17).

Recall that $C > 0$ is a generic notation for a constant whose value can be different in different expressions. We first show that (1.18) implies (1.17). Suppose $\theta_i > \theta_{i+1}$. Then

$$\theta_{i+1}^{\tau-2} \sum_{j=1}^i \theta_j \geq \theta_{i+1}^{\tau-2} \sum_{j=1}^\infty \theta_j \cdot \mathbb{1}\{\theta_{i+1} < \theta_j \leq c_1\theta_{i+1}\} \geq c_0, \tag{1.19}$$

where the last step uses (1.18). Now, for $u \in (\frac{1}{\theta_i}, \frac{1}{\theta_{i+1}}]$,

$$\begin{aligned} \Psi_\theta(u) &\geq C\Big[\sum_{k=1}^i u\theta_k^2 + \sum_{k=i+1}^\infty u^2\theta_k^3\Big] \geq Cu\theta_{i+1}\sum_{k=1}^i \theta_k \\ &= \frac{Cu\theta_{i+1}^{\tau-2}\sum_{k=1}^i \theta_k}{\theta_{i+1}^{\tau-3}} \geq \frac{Cuc_0}{\theta_{i+1}^{\tau-3}} \geq Cc_0u^{\tau-2}. \end{aligned} \tag{1.20}$$

Thus, $\int_{\theta_1^{-1}}^\infty \frac{\mathrm{d}u}{\Psi_\theta(u)} \leq C\int_{\theta_1^{-1}}^\infty u^{-(\tau-2)}\mathrm{d}u < \infty$, since $\tau > 3$. This yields (1.17).

To see that the implication is strict, take $\theta_i = (i^\alpha \log(i+2))^{-1}$. Then

$$\theta_{i+1}^{\tau-2}\sum_{j=1}^i \theta_j \leq \big((i+1)^\alpha \log(i+3)\big)^{-(\tau-2)} \sum_{j=1}^i j^{-\alpha} \leq \frac{C}{\log^{\tau-2} i}, \tag{1.21}$$

which tends to zero as $i \to \infty$. However, as we have seen in (1.19), (1.18) would imply that the left side of (1.21) is bounded away from zero. Thus, (1.18) does not hold in this case. To see that (1.17) does hold, note that $\theta_i \geq \theta_i' := i^{-\alpha'}$ for all large enough i, where $\alpha' = \frac{1}{\tau'-1}$ and $3 < \tau' < \tau$. Then $(\theta_i')_{i\geq 1}$ satisfies (1.18). Therefore, a computation similar to (1.20) yields, for $u \in (\frac{1}{\theta_i}, \frac{1}{\theta_{i+1}}]$,

$$\Psi_\theta(u) \geq Cu\theta_{i+1}\sum_{k=1}^i \theta_k \geq Cu\theta_{i+1}'\sum_{k=1}^i \theta_k' \geq \frac{Cu}{(\theta_i')^{\tau-3}}. \tag{1.22}$$

Since $u \leq (i+1)^\alpha \log(i+3)$, we can choose $\delta > 0$ such that $u^{1+\delta} \leq Ci^{\alpha'} = C/\theta_i'$. Therefore, $\Psi_\theta(u) \geq Cu^{1+(\tau-3)(1+\delta)}$. Thus, (1.17) follows.

Proof ideas and technical motivation for this work The proof of Theorem 1.5 consists of two main steps, that form the key ideas in the argument. The first step is to show that the neighborhoods of the high-degree vertices, called *hubs*, have mass $\Theta_{\mathbb{P}}(n^{\rho})$. Secondly, any small εn^{η} neighborhood contains a hub with high probability. These two facts, summarized in Propositions 2.1 and 2.2 below, together ensure that the total mass of any neighborhood of $\mathscr{C}_{(i)}$ of radius εn^{η} is bounded away from zero. These two facts were proved in [8] in the context of rank-one inhomogeneous random graphs. However, the proof techniques are completely different here. The main advantage in [8] was that the breadth-first exploration of components could be dominated by a branching process with *mixed Poisson* progeny distribution that is *independent of* n. This allows one to use existing literature to estimate the probabilities that a long path exists in the branching process. However, such a technique is specific to rank-one inhomogeneous random graphs and does not work in the cases where the above stochastic domination does not hold. This was one of the technical motivations for this work. Moreover, the final section contains results about exponential tail-bounds for the number of edges in large critical components (Proposition 4.1), as well as a coupling of the neighborhood exploration with a branching process with stochastically larger progeny distribution (Section 4.2), which are both interesting in their own right.

Organization of this paper The rest of this paper is organized as follows: In Section 2, we state two key propositions, the first involving the total mass of small neighborhoods, and the second involving a bound on the diameter of a *slightly subcritical* $\mathrm{CM}_n(\boldsymbol{d})$. The proof of Theorem 1.5 is completed in Section 2. In Section 3, we derive the required bounds on the total mass of small neighborhoods. In Section 4, we obtain bounds on the diameter of the connected components after removing the high-degree vertices. In Section 5, we prove Assumptions 1.1, 1.3 for the percolated degree sequence, which allows us to conclude Theorem 1.7.

2 Proof of the global lower mass-bound

In this section, we first state the two key ingredients in Propositions 2.1 and 2.2, and then complete the proof of Theorem 1.5. The proofs of Propositions 2.1 and 2.2 are given in the subsequent sections. The first ingredient shows that hub i has sufficient mass close to it with high probability:

Proposition 2.1. *Assume that* Assumptions 1.1 *and* (1.6) *hold. Recall that* $\mathcal{N}_v(\delta)$ *denotes the* δn^{η} *neighborhood of* v. *For each fixed* $i \geq 1$ *and* $\varepsilon_2 > 0$, *there exists* $\delta_{i,\varepsilon_2} > 0$ *and* $n_{i,\varepsilon_2} \geq 1$ *such that, for any* $\delta \in (0, \delta_{i,\varepsilon_2}]$ *and* $n \geq n_{i,\varepsilon_2}$,

$$\mathbb{P}\big(|\mathcal{N}_i(\delta)| \leq \theta_i \delta n^{\rho}\big) \leq \frac{\varepsilon_2}{2^{i+1}}. \tag{2.1}$$

Next, we need some control on the diameter of the graph after removing the hubs. Denote by $\mathcal{G}_n^{>K}$ the graph obtained by removing the vertices $[K] = \{1, \ldots, K\}$ having the largest degrees and the edges incident to them from $\mathrm{CM}_n(\boldsymbol{d})$. Note that $\mathcal{G}_n^{>K}$ is a configuration model conditionally on its degree sequence. Let $\Delta^{>K}$ denote the maximum of the diameters of the connected components of $\mathcal{G}_n^{>K}$. The following proposition shows that, for large K, $\Delta^{>K}$ is small with high probability:

Proposition 2.2. *Assume that* Assumptions 1.1, 1.3 *and* (1.6) *hold. Then, for any* $\varepsilon_1, \varepsilon_2 > 0$, *there exists* $K = K(\varepsilon_1, \varepsilon_2)$ *and* $n_0 = n_0(\varepsilon_1, \varepsilon_2)$ *such that for all* $n \geq n_0$,

$$\mathbb{P}\left(\Delta^{>K} > \varepsilon_1 n^{\eta}\right) \leq \frac{\varepsilon_2}{4}. \tag{2.2}$$

We now prove Theorem 1.5 assuming Propositions 2.1 and 2.2:

Proof of Theorem 1.5. Fix $i \geq 1$ and $\varepsilon_1, \varepsilon_2 > 0$. For a component $\mathscr{C} \subset \mathrm{CM}_n(\boldsymbol{d})$, we write $\Delta(\mathscr{C})$ to denote its diameter. Let us choose K and n_0 so that (2.2) holds for all $n \geq n_0$. In view of Proposition 2.1, let $\delta_0 = \min\{\varepsilon_1, \delta_{1,\varepsilon_2}, \ldots, \delta_{K,\varepsilon_2}\}/2$, and $n_0' = \max\{n_0, n_{1,\varepsilon_2}, \ldots, n_{K,\varepsilon_2}\}$. Thus, for all $n \geq n_0'$, (2.1) is satisfied for all $i \in [K]$. Define

$$F_1 := \{\Delta^{>K} < \varepsilon_1 n^\eta/2\}, \quad F_2 := \{\Delta(\mathscr{C}_{(i)}) > \varepsilon_1 n^\eta/2\}. \tag{2.3}$$

Notice that, on the event $F_1 \cap F_2$, it must be the case that one of the vertices in $[K]$ belongs to $\mathscr{C}_{(i)}$, and that the union of the neighborhoods of $[K]$ of radius $\lceil \varepsilon_1 n^\eta/2 \rceil + 1 \approx \varepsilon_1 n^\eta/2$ covers $\mathscr{C}_{(i)}$. Therefore, given any vertex $v \in \mathscr{C}_{(i)}$, $\mathcal{N}_v(\varepsilon_1)$ contains at least one of the neighborhoods $(\mathcal{N}_j(\varepsilon_1/2))_{j\in[K]}$. This observation yields that

$$\inf_{v\in\mathscr{C}_{(i)}} n^{-\rho}|\mathcal{N}_v(\varepsilon_1)| \geq \min_{j\in[K]} n^{-\rho}|\mathcal{N}_j(\varepsilon_1/2)| \geq \min_{j\in[K]} n^{-\rho}|\mathcal{N}_j(\delta_0)|. \tag{2.4}$$

Thus, for all $n \geq n_0'$,

$$\begin{aligned} &\mathbb{P}\Big(F_1 \cap F_2 \cap \Big\{\inf_{v\in\mathscr{C}_{(i)}} n^{-\rho}|\mathcal{N}_v(\varepsilon_1)| \leq \theta_K\delta_0\Big\}\Big) \\ &\qquad \leq \sum_{j\in[K]} \mathbb{P}\big(|\mathcal{N}_j(\delta)| \leq \theta_j\delta_0 n^\rho\big) \leq \sum_{j=1}^{K} \frac{\varepsilon_2}{2^{j+1}} \leq \frac{\varepsilon_2}{2}, \end{aligned} \tag{2.5}$$

where the one-but-last step follows from Proposition 2.1. Further, on the event F_2^c, $|\mathcal{N}_v(\varepsilon_1)| = |\mathscr{C}_{(i)}|$ for all $v \in \mathscr{C}_{(i)}$. Moreover, using (1.8), it follows that $n^{-\rho}|\mathscr{C}_{(i)}|$ converges in distribution to a random variable with strictly positive support. Using the Portmanteau theorem, the above implies that for any $\delta_0' > 0$, there exists $\tilde{n}_0 = \tilde{n}_0(\varepsilon_2, \delta_0')$ such that, for all $n \geq \tilde{n}_0$,

$$\mathbb{P}\big(n^{-\rho}|\mathscr{C}_{(i)}| \leq \delta_0'\big) \leq \frac{\varepsilon_2}{4}. \tag{2.6}$$

Therefore,

$$\mathbb{P}\left(F_2^c \cap \left\{\inf_{v\in\mathscr{C}_{(i)}} n^{-\rho}|\mathcal{N}_v(\varepsilon_1)| \leq \delta_0'\right\}\right) \leq \frac{\varepsilon_2}{4}. \tag{2.7}$$

Now, using (2.5) and (2.7), together with Proposition 2.2, it follows that, for any $n \geq \max\{n_0', \tilde{n}_0\}$, and K chosen as above,

$$\mathbb{P}\left(\inf_{v\in\mathscr{C}_{(i)}} n^{-\rho}|\mathcal{N}_v(\varepsilon_1)| \leq \min\{\delta_0', \theta_K\delta_0\}\right) \leq \varepsilon_2. \tag{2.8}$$

This completes the proof of Theorem 1.5. □

3 Lower bound on the total mass of neighborhoods of hubs

In this section, we prove Proposition 2.1.

Proof of Proposition 2.1. Let us denote the component of $\mathrm{CM}_n(\boldsymbol{d})$ containing vertex i by $\mathscr{C}(i)$. Consider the breadth-first exploration of $\mathscr{C}(i)$ starting from vertex i, given by the following exploration algorithm [16]:

Algorithm 1 (Exploring the graph)**.** The algorithm carries along vertices that can be alive, active, exploring and killed, and half-edges that can be alive, active or killed. We sequentially explore the graph as follows:

(S0) At stage $l = 0$, all the vertices and the half-edges are *alive*, and only the half-edges associated to vertex i are *active*. Also, there are no *exploring* vertices except i.

(S1) At each stage l, if there is an exploring vertex, take an active half-edge e of an exploring vertex v and pair it uniformly to another alive half-edge f. Kill e, f. If f is incident to a vertex v' that has not been discovered before, then declare all the half-edges incident to v' (if any) active, except f. If $\mathrm{degree}(v') = 1$ (i.e. the only half-edge incident to v' is f) then kill v'. Otherwise, declare v' to be active and larger than all other vertices that are alive. After killing e, if v does not have another active half-edge, then kill v also. If there is no exploring vertex at the beginning of stage l, we pick the oldest active half-edge, declare the corresponding vertex to be exploring, and then execute the same process as above.

(S2) Repeat (S1) until there is no active half-edges left.

Call a vertex *discovered* if it is either active or killed. Let $\mathscr{V}_l$ denote the set of vertices discovered up to time l and $\mathcal{I}_j^n(l) := \mathbb{1}\{j \in \mathscr{V}_l\}$. Define the exploration process by

$$S_n(l) = d_i + \sum_{j\neq i} d_j \mathcal{I}_j^n(l) - 2l = d_i + \sum_{j\neq i} d_j \left(\mathcal{I}_j^n(l) - \frac{d_j}{\ell_n} l\right) + \left(\frac{1}{\ell_n}\sum_{j\neq i} d_j^2 - 2\right) l. \tag{3.1}$$

Note that the exploration process keeps track of the number of active half-edges. Thus, $\mathscr{C}(i)$ is explored when $\boldsymbol{S}_n$ hits zero. Moreover, since one edge is explored at each step, the hitting time of zero is the total number of edges in $\mathscr{C}(i)$. Define the re-scaled version $\bar{\boldsymbol{S}}_n$ of $\boldsymbol{S}_n$ by $\bar{S}_n(t) = n^{-\alpha} S_n(\lfloor t n^{\rho} \rfloor)$. Then, by Assumption 1.1 and (1.6),

$$\bar{S}_n(t) = \theta_i - \frac{\theta_i^2 t}{\mu} + n^{-\alpha} \sum_{j\neq i} d_j \left(\mathcal{I}_j^n(tn^{\rho}) - \frac{d_j}{\ell_n} tn^{\rho}\right) + \lambda t + o(1). \tag{3.2}$$

The convergence of this exploration process was considered in [16, Theorem 8] except for the fact that the exploration process started at zero in [16]. However, using identical arguments to [16, Theorem 8], it can be shown that

$$\bar{\boldsymbol{S}}_n \xrightarrow{d} \boldsymbol{S}_\infty, \tag{3.3}$$

with respect to the Skorohod J_1-topology, where

$$S_\infty(t) = \theta_i - \frac{\theta_i^2 t}{\mu} + \sum_{j\neq i} \theta_j \Big(\mathcal{I}_j(t) - \frac{\theta_j t}{\mu}\Big) + \lambda t, \tag{3.4}$$

with $\mathcal{I}_j(s) := \mathbb{1}\{\xi_j \le s\}$ and $\xi_j \sim \mathrm{Exponential}(\theta_j/\mu)$ independently of each other.

Let $h_n(u)$ (respectively $h_\infty(u)$) denote the first hitting time of $\bar{\boldsymbol{S}}_n$ (respectively $\boldsymbol{S}_\infty$) of u. More precisely,

$$h_n(u) := \inf\Big\{t : \bar{S}_n(t) \le u \text{ or } \lim_{t'\nearrow t} \bar{S}_n(t') \le u\Big\}, \tag{3.5}$$

and define $h_\infty(u)$ similarly by replacing $\bar{S}_n(t)$ by $\bar{S}_\infty(t)$. Note that, by [16, Lemma 36], the distribution of $h_\infty(u)$ does not have any atoms and therefore, for any $\varepsilon_2 > 0$, there exists $\beta_{\varepsilon_2,i} > 0$ such that

$$\mathbb{P}\big(h_\infty(\theta_i/2) \le \beta_{\varepsilon_2,i}\big) \le \frac{\varepsilon_2}{2^{i+1}}.$$

Now we use the following fact:

Fact 3.1. *Let $(X_n(t))_{t\ge 0} \xrightarrow{d} (X(t))_{t\ge 0}$ in Skorohod J_1-topology and let $h(X_n)$ (respectively $h(X)$) denote the hitting time to zero of X_n (respectively X). Then,*

$$\liminf_{n\to\infty} \mathbb{P}(h(X_n) > a) \ge \mathbb{P}(h(X) > a) \quad \textit{for all } a > 0. \tag{3.6}$$

Proof. Let $(f_n)_{n\geq 1}$ be such that $h(f_n) \leq a$ for all $n \geq 1$ and $f_n \to f$ in the Skorohod J_1-topology as $n \to \infty$. Now, $h(f_n) \leq a$ implies that $\inf_{t\in[0,a]} f_n(t) \leq 0$. Using [34, Theorem 13.4.1], it follows that $\inf_{t\in[0,a]} f(t) \leq 0$ and thus $h(f) \leq a$. Therefore, we have shown that $\{f\colon h(f) \leq a\}$ is a closed set in the Skorohod J_1-topology, and therefore $\{f : h(f) > a\}$ is an open set. The proof follows using the Portmanteau theorem [9, Theorem 2.1 (iv)]. □

Using (3.3) and Fact 3.1, there exists $n_{i,\varepsilon_2} \geq 1$ such that, for all $n \geq n_{i,\varepsilon_2}$,

$$\mathbb{P}(h_n(\theta_i/2) \leq \beta_{\varepsilon_2,i}) \leq \frac{\varepsilon_2}{2^i}. \tag{3.7}$$

Our first goal is to show that there exists a $\delta_{i,\varepsilon}$ such that for any $\delta \in (0, \delta_{i,\varepsilon_2}]$,

$$\sum_{k\in\mathcal{N}_i(\delta)} d_k \leq \theta_i \delta n^\rho \quad \Longrightarrow \quad h_n(\theta_i/2) \leq \beta_{\varepsilon_2,i}. \tag{3.8}$$

Recall that $\mathcal{N}_v(\delta)$ denotes the δn^η neighborhood of v in $\mathrm{CM}_n(\boldsymbol{d})$. To prove (3.8), let $\partial(j)$ denote the set of vertices at distance j from i. Let E_{j1} denote the total number of edges between vertices in $\partial(j)$ and $\partial(j-1)$, and let E_{j2} denote the number of edges within the vertices in $\partial(j-1)$. Define $E_j = E_{j1} + E_{j2}$. Fix any $\delta < 2\beta_{\varepsilon_2,i}/\theta_i$. Note that if $\sum_{k\in\mathcal{N}_i(\delta)} d_k \leq \theta_i \delta n^\rho$, then the total number of edges in $\mathcal{N}_i(\delta)$ is at most $\theta_i \delta n^\rho/2$. Thus there exists $j \leq \delta n^\eta$ such that $E_j \leq \theta_i \delta n^\rho / 2\delta n^\eta = \theta_i n^\alpha/2$. This implies that $\boldsymbol{S}_n$ must go below $\theta_i n^\alpha/2$ before exploring all the vertices in $\mathcal{N}_i(\delta)$. This is because we are exploring the components in a breadth-first manner and $\bar{\boldsymbol{S}}_n$ keeps track of the number of active half-edges, which in turn are the potential connections to vertices at the next level. Since one edge is explored in each time step, and we rescale time by n^ρ, this implies that

$$h_n(\theta_i/2) \leq \frac{1}{2} n^{-\rho} \sum_{k\in\mathcal{N}_i(\delta)} d_k \leq \theta_i \delta/2 \leq \beta_{\varepsilon_2,i}. \tag{3.9}$$

Therefore, for all $n \geq n_{i,\varepsilon_2}$,

$$\mathbb{P}\Big(\sum_{k\in\mathcal{N}_i(\delta)} d_k \leq \theta_i \delta n^\rho \Big) \leq \mathbb{P}(h_n(\theta_i/2) \leq \beta_{\varepsilon_2,i}) \leq \frac{\varepsilon_2}{2^i}. \tag{3.10}$$

Finally, to conclude Proposition 2.1 from (3.10), we use the result from [16, Lemma 22] that, for any $T > 0$,

$$\sup_{u\leq T} \Big| \sum_{i\in[n]} \mathcal{I}_i^n(un^\rho) - un^\rho \Big| = o_{\mathbb{P}}(n^\rho). \tag{3.11}$$

This implies that the difference between the number of edges and the number of vertices explored up to time un^ρ is $o_{\mathbb{P}}(n^\rho)$ uniformly over $u \leq T$. The proof of Proposition 2.1 now follows. □

4 Diameter after removing hubs

Throughout the remainder of the paper, we fix the convention that $C, C', C'' > 0$ etc. denote constants whose value can change from line to line. Recall the definition of the graph $\mathcal{G}_n^{>K}$ from Proposition 2.2. If we keep on exploring $\mathcal{G}_n^{>K}$ in a breadth-first manner using Algorithm 1 and ignore the cycles created, then we get a random tree. The idea is to couple neighborhoods in $\mathcal{G}_n^{>K}$ with a suitable branching process such that the progeny distribution of the branching process dominates the number of children of each vertex in the breadth-first tree. Therefore, when there is a long path in $\mathcal{G}_n^{>K}$ that makes the diameter large, that long path must be present in the branching process

as well under the above coupling. In this way, the question about the diameter of $\mathcal{G}_n^{>K}$ reduces to the question about the height of a branching process. To estimate the height suitably, we use a recent beautiful proof technique by Addario-Berry [1] which allows one to relate the height of a branching process to the sum of inverses of the associated breadth-first random walk.

In Section 4.1, we establish tail bounds for the number of edges within components. This allows us to formulate the desired coupling in Section 4.2. In Section 4.3, we analyze the breadth-first random walk to show that it is unlikely that the height of the branching process is larger than εn^{η}. These bounds are different from those derived in [1] since our branching process depends on n and there is a joint scaling involved between the distances and the law of the branching process.

4.1 Asymptotics for the number of edges

For a graph G, let $\mathrm{E}(G)$ denote the number of edges in G.

Proposition 4.1. *Suppose that* Assumption 1.1 *and* (1.6) *hold. For all* $\varepsilon \in (0, \frac{4-\tau}{\tau-1})$*, and sufficiently large* n*,*

$$\mathbb{P}(\mathrm{E}(\mathscr{C}(i)) > n^{\rho+\varepsilon}) \leq C\mathrm{e}^{-C' n^{\varepsilon/2}}, \tag{4.1}$$

for some absolute constants $C, C' > 0$ *and all* $i \in [n]$*.*

The proof of Proposition 4.1 relies on concentration techniques for martingales. We start by defining the relevant notation. Consider exploring $\mathrm{CM}_n(\boldsymbol{d})$ with Algorithm 1, and let the associated exploration process be defined in (3.1). Let us denote the degree of the vertex found at step l by $d_{(l)}$. If no new vertex is found at step l, then $d_{(l)} = 0$. Also, let $\mathscr{F}_l$ denote the sigma-algebra containing all the information revealed by the exploration process up to time l. Thus,

$$S_n(0) = d_i, \quad \text{and} \quad S_n(l) = S_n(l-1) + (d_{(l)} - 2). \tag{4.2}$$

Using the Doob-Meyer decomposition, one can write

$$S_n(l) = S_n(0) + M_n(l) + A_n(l), \tag{4.3}$$

where M_n is a martingale with respect to $(\mathscr{F}_l)_{l\geq 1}$. The drift A_n and the quadratic variation $\langle M_n \rangle$ of M_n are given by

$$A_n(l) = \sum_{j=1}^{l} \mathbb{E}\big[d_{(j)} - 2 | \mathscr{F}_{j-1}\big], \quad \langle M_n \rangle(l) = \sum_{j=1}^{l} \mathrm{Var}\left(d_{(j)} | \mathscr{F}_{j-1}\right). \tag{4.4}$$

We will show that for any $\varepsilon \in (0, \varepsilon_0)$, the following two lemmas hold:

Lemma 4.2. *Suppose that* Assumption 1.1 *and* (1.6) *hold. For all* $\varepsilon \in (0, \frac{4-\tau}{\tau-1})$*, and sufficiently large* n*,*

$$\mathbb{P}(n^{-(\alpha+\varepsilon)} M_n(n^{\rho+\varepsilon}) > 1) \leq C\mathrm{e}^{-C' n^{\varepsilon}}, \tag{4.5}$$

for some absolute constants $C, C' > 0$*.*

Lemma 4.3. *Suppose that* Assumption 1.1 *and* (1.6) *hold. For all fixed* $K \geq 1$*,* $\varepsilon \in (0, \frac{4-\tau}{\tau-1})$*, and sufficiently large* n*,*

$$\mathbb{P}\left(n^{-(\alpha+\varepsilon)} A_n(n^{\rho+\varepsilon}) \geq -C \sum_{i=1}^{K} \theta_i^2 \right) \leq C\mathrm{e}^{-C' n^{\varepsilon/2}}, \tag{4.6}$$

for some absolute constants $C, C' > 0$*.*

Proof of Proposition 4.1 subject to Lemmas 4.2, 4.3. Throughout, we write $t_n := n^{\rho+\varepsilon}$. Note that we can choose $K \geq 1$ such that $\sum_{i=1}^{K} \theta_i^2$ is arbitrarily large since $\boldsymbol{\theta} \notin \ell^2_{\downarrow}$. Thus, if $n^{-(\alpha+\varepsilon)} M_n(t_n) \leq 1$ and $n^{-(\alpha+\varepsilon)} A_n(t_n) \leq -C \sum_{i=1}^{K} \theta_i^2$, then $n^{-(\alpha+\varepsilon)} S_n(t_n) < 0$, and therefore $\mathscr{C}(i)$ must be explored before time t_n, and thus $\mathrm{E}(\mathscr{C}(i)) \leq t_n$. As a result, Lemmas 4.2 and 4.3 together complete the proof of Proposition 4.1. □

Proof of Lemma 4.2. First note that $\frac{4-\tau}{\tau-1} + \rho = 2\alpha < 1$ and therefore $t_n = o(n)$. Thus, uniformly over $j \leq t_n$,

$$\mathrm{Var}\,(d_{(j)} | \mathscr{F}_{j-1}) \leq \mathbb{E}[d^2_{(j)} | \mathscr{F}_{j-1}] = \frac{\sum_{j \notin \mathscr{V}_{j-1}} d_j^3}{\ell_n - 2j + 2} \leq \frac{\sum_{j \in [n]} d_j^3}{\ell_n - 2t_n + 2} \leq C n^{3\alpha - 1}, \tag{4.7}$$

so that, almost surely,

$$\langle M_n \rangle(t_n) \leq t_n C n^{3\alpha-1} = C n^{2\alpha+\varepsilon}. \tag{4.8}$$

Also, $d_{(j)} \leq C n^{\alpha}$ almost surely. We can now use Freedman's inequality [20, Proposition 2.1] which says that if $Y(k) = \sum_{j \leq k} X_j$ with $\mathbb{E}[X_j | \mathcal{F}_{j-1}] = 0$ (for some filtration $(\mathcal{F}_j)_{j \geq 1}$) and $\mathbb{P}(|X_j| \leq R,\ \forall j \geq 1) = 1$, then, for any $a, b > 0$,

$$\mathbb{P}(Y(k) \geq a, \text{ and } \langle Y \rangle(k) \leq b) \leq \exp\left(\frac{-a^2}{2(Ra + b)} \right). \tag{4.9}$$

We apply (4.9) with $a = n^{\alpha+\varepsilon}$, $b = C n^{2\alpha+\varepsilon}$ and $R = C n^{\alpha}$. Note that $\langle M_n \rangle(t_n) \leq b$ almost surely by (4.8). It follows that

$$\mathbb{P}(M_n(t_n) > n^{\alpha+\varepsilon}) \leq \exp\left(- \frac{n^{2\alpha+2\varepsilon}}{2C(n^{\alpha} n^{\alpha+\varepsilon} + n^{2\alpha+\varepsilon})} \right) \leq C \mathrm{e}^{-C' n^{\varepsilon}}, \tag{4.10}$$

and the proof follows. □

Proof of Lemma 4.3. Note that

$$\begin{aligned} \mathbb{E}\big[d_{(i)} - 2 | \mathscr{F}_{i-1}\big] &= \frac{\sum_{j \notin \mathscr{V}_{i-1}} d_j^2}{\ell_n - 2i + 1} - 2 \\ &= \frac{1}{\ell_n} \sum_{j \in [n]} d_j (d_j - 2) - \frac{1}{\ell_n} \sum_{j \in \mathscr{V}_{i-1}} d_j^2 + \frac{(2i-1) \sum_{j \notin \mathscr{V}_{i-1}} d_j^2}{\ell_n (\ell_n - 2i + 1)} \\ &\leq \lambda n^{-\eta} - \frac{1}{\ell_n} \sum_{j \in \mathscr{V}_{i-1}} d_j^2 + \frac{(2i-1)}{(\ell_n - 2i + 1)^2} \sum_{j \in [n]} d_j^2 + o(n^{-\eta}) \end{aligned} \tag{4.11}$$

uniformly over $i \leq t_n$. Therefore, for all sufficiently large n,

$$\begin{aligned} A_n(t_n) = \sum_{j=1}^{t_n} \mathbb{E}\big[d_{(j)} - 2 | \mathscr{F}_{j-1}\big] &\leq \lambda t_n n^{-\eta} - \frac{1}{\ell_n} \sum_{i=1}^{t_n} \sum_{j \in \mathscr{V}_{i-1}} d_j^2 + \frac{C t_n^2}{\ell_n} + o(n^{\alpha+\varepsilon}) \\ &= \lambda n^{\alpha+\varepsilon} - \frac{1}{\ell_n} \sum_{i=1}^{t_n} \sum_{j \in \mathscr{V}_{i-1}} d_j^2 + o(n^{\alpha+\varepsilon}), \end{aligned} \tag{4.12}$$

where in the second step we have used $\sum_{i \in [n]} d_i^2 / \ell_n = O(1)$, and in the last step we have used that $t_n^2 / \ell_n = O(n^{2\rho+2\varepsilon-1}) = o(n^{\alpha+\varepsilon})$ for $\varepsilon < 1 + \alpha - 2\rho = \frac{4-\tau}{\tau-1}$. Let us denote the second term in (4.12) by (A). To analyze (A), define the event

$$\mathcal{A}_n := \big\{ \exists j : d_j > n^{\alpha - \varepsilon/2}, j \notin \mathscr{V}_{t_n/2} \big\}. \tag{4.13}$$

Then, for all sufficiently large n,

$$\mathbb{P}(\mathcal{A}_n) \le \sum_{j:d_j>n^{\alpha-\varepsilon/2}} \left(1-\frac{d_j}{\ell_n-2t_n}\right)^{t_n} \le n\mathrm{e}^{-Cn^{\varepsilon/2}}. \tag{4.14}$$

On the event $\mathcal{A}_n^c$,

$$(\mathrm{A}) = \frac{1}{\ell_n}\sum_{i=1}^{t_n}\sum_{j\in[n]} d_j^2 \mathbb{1}\{j\in\mathscr{V}_{i-1}\} \ge \frac{1}{\ell_n}\sum_{i=\frac{t_n}{2}+1}^{t_n}\sum_{j=1}^{K} d_j^2 \ge Cn^{\alpha+\varepsilon}\sum_{j=1}^{K}\theta_j^2. \tag{4.15}$$

Combining (4.12), (4.14) and (4.15) now completes the proof. □

4.2 Coupling with branching processes

Recall that $\mathscr{C}(i)$ is the connected component in $\mathrm{CM}_n(\boldsymbol{d})$ containing vertex i. Define the event $\mathcal{K}_n := \{\mathrm{E}(\mathscr{C}(i)) > n^{\rho+\varepsilon}\}$. Proposition 4.1 implies that the probability of $\mathcal{K}_n$ happening is exponentially small in n. On the event $\mathcal{K}_n^c$, we can couple the breadth-first exploration starting from vertex i with a suitable branching process. Let n_k denote the number of vertices of degree k and consider the branching process $\mathcal{X}_n(i)$ starting with d_i individuals, and the progeny distribution $\bar{\xi}_n$ given by

$$\mathbb{P}\left(\bar{\xi}_n = k\right) = \bar{p}_k = \begin{cases} \frac{(k+1)n_{k+1}}{\underline{\ell}_n} & \text{for } k\ge 1,\\ \frac{n_1-2n^{\rho+\varepsilon}}{\underline{\ell}_n} & \text{for } k=0, \end{cases} \tag{4.16}$$

where $\underline{\ell}_n = \ell_n - 2n^{\rho+\varepsilon}$. Note that, at each step of the exploration, we have at most $(k+1)n_{k+1}$ half-edges that are incident to vertices having k further unpaired half-edges. Further, on the event $\mathcal{K}_n^c$, we have at least $\underline{\ell}_n$ choices for pairing. Therefore, the number of active half-edges discovered at each step in the breadth-first exploration of the neighborhoods of i is stochastically dominated by $\bar{\xi}_n$. This proves the next proposition, which we state after setting up some further notation. Recall that $\mathcal{G}_n^{>i-1}$ denotes the graph obtained by deleting vertices in $[i-1]$ and the associated edges from $\mathrm{CM}_n(\boldsymbol{d})$. Let $\partial_i(r)$ denote the number of vertices at distance r from vertex i in the graph $\mathcal{G}_n^{>i-1}$. Let $\bar{\xi}_n(i)$ denote the random variable with the distribution in (4.16) truncated in such a way that $\{d_1,\dots,d_{i-1}\}$ are excluded from the support. More precisely,

$$\mathbb{P}(\bar{\xi}_n(i)=k) = \begin{cases} 0 & \text{for } k>d_i,\\ \frac{(k+1)}{L}\#\{j\ge i:\ d_j=k+1\} & \text{for } 1\le k\le d_i,\\ \frac{n_1-2n^{\rho+\varepsilon}}{L} & \text{for } k=0, \end{cases} \tag{4.17}$$

where $L = \underline{\ell}_n - \sum_{j=1}^{i-1} d_j$ is the appropriate normalizing constant. Let $\mathcal{X}_{n,\mathrm{res}}(i)$ denote the branching process starting with d_i individuals and progeny distribution $\bar{\xi}_n(i)$ and let $\bar{\partial}_i(r)$ denote the number of individuals in generation r of $\mathcal{X}_{n,\mathrm{res}}(i)$. Then the above stochastic domination argument immediately yields the next proposition:

Proposition 4.4. *Suppose that* Assumption 1.1 *and* (1.6) *hold. Let K_n be as described in* Lemma A.1 *below. For all $r\ge 1$, $1\le i\le K_n$, $\varepsilon\in(0,\frac{4-\tau}{\tau-1})$ and $n\ge 1$,*

$$\mathbb{P}(\partial_i(r)\neq\varnothing) \le \mathbb{P}(\bar{\partial}_i(r)\neq\varnothing) + \mathbb{P}(\mathrm{E}(\mathscr{C}(i))>n^{\rho+\varepsilon}). \tag{4.18}$$

Before proceeding with the next section in which we investigate $\mathbb{P}(\bar{\partial}_i(r)\neq\varnothing)$, we estimate the expectation and variance of the progeny distribution in the branching

process $\mathcal{X}_{n,\mathrm{res}}(i)$ using Assumptions 1.1, 1.3, and (1.6). Using $\sum_i \theta_i^2 = \infty$, we can choose $i_2(\lambda)$ (depending only on λ) such that

$$\frac{1}{\mu}\sum_{i=1}^{i_2(\lambda)} \theta_i^2 \geq 5\lambda. \tag{4.19}$$

Also the normalizing constant in (4.17) satisfies

$$L = \ell_n(1 + o(n^{-\eta})) \tag{4.20}$$

uniformly over $1 \leq i \leq K_n$. To see this, first observe that $\underline{\ell}_n = \ell_n - 2n^{\rho+\varepsilon} = o(n^{-\eta})$ since $\varepsilon < 1 - \rho - \eta = \frac{4-\tau}{\tau-1}$. Also, $\frac{1}{\ell_n}\sum_{j\leq i} d_j = O(d_1 K_n n^{-1}) = o(n^{2\alpha-1}) = o(n^{-\eta})$, as $K_n = o(n^{\alpha})$ by Assumption 1.3 and Lemma A.1 and $2\alpha - 1 = -\eta$. Hence, (4.20) follows. Now using Assumption 1.1 and (1.6), note that there exists $N_\lambda \geq 1$ such that for all $n \geq N_\lambda$ and $i_2(\lambda) \leq i \leq K_n$,

$$\begin{aligned}
\mathbb{E}\left[\bar{\xi}_n(i)\right] &= \frac{1}{L}\sum_{j\geq i} d_j(d_j-1) = \frac{1}{\ell_n}\sum_{j\geq i} d_j(d_j-1) + o(n^{-\eta}) \\
&= 1 + \lambda n^{-\eta} - \frac{1}{\ell_n}\sum_{j<i} d_j(d_j-1) + o(n^{-\eta}) \\
&\leq 1 + \lambda n^{-\eta} - \frac{1}{2\ell_n}\sum_{j<i} d_j^2 + o(n^{-\eta}) \\
&\leq 1 - \left(Cn^{-2\alpha}\sum_{j<i} d_j^2\right) n^{-\eta} + o(n^{-\eta}),
\end{aligned} \tag{4.21}$$

where the third step uses (1.6), the penultimate step uses the fact that $d_i \geq 2$ so that $\sum_{j<i} d_j \leq \sum_{j<i} d_j^2/2$ for $i \leq K_n$, and the last step uses (4.19). Thus, for $n \geq N_\lambda$ and $i_2(\lambda) \leq i \leq K_n$,

$$\mathbb{E}\left[\bar{\xi}_n(i)\right] \leq 1 - \beta_i^n n^{-\eta} \quad \text{where} \quad \beta_i^n = Cn^{-2\alpha}\sum_{j<i} d_j^2. \tag{4.22}$$

The estimate in (4.22) will be crucial in the next section.

4.3 Estimating heights of trees via random walks

We will prove the following theorem in this section:

Theorem 4.5. *Suppose that* Assumptions 1.1, 1.3, *and* (1.6) *hold. Fix* $\varepsilon > 0$. *Then, for all* $i_2(\lambda) \leq i \leq K_n$ *(where* $i_2(\lambda)$ *and* K_n *are given by* (4.19) *and* Lemma A.1 *respectively) and* $n \geq N_\lambda$,

$$\mathbb{P}(\bar{\partial}_i(\varepsilon n^\eta) \neq \varnothing) \leq \frac{Cd_i}{n^\alpha} \mathrm{e}^{-\frac{\varepsilon\beta_i^n}{2}}, \tag{4.23}$$

for some constant $C = C(\varepsilon, \lambda) > 0$.

Define $\mathcal{X}_n^1(i)$ to be the Galton-Watson tree starting with one offspring and progeny distribution $\bar{\xi}_n(i)$ and let $\bar{\partial}_i^1(r)$ denote the number of individuals in generation r of $\mathcal{X}_n^1(i)$. The crucial ingredient for the proof of Theorem 4.5 is the following:

Proposition 4.6. *Under identical conditions as in* Theorem 4.5, *for all* $n \geq N_\lambda$,

$$\mathbb{P}(\bar{\partial}_{i_2(\lambda)}^1(\varepsilon n^\eta) \neq \varnothing) \leq \frac{C}{n^\alpha}, \tag{4.24}$$

for some constant $C = C(\varepsilon, \lambda) > 0$.

Proof of Theorem 4.5 using Proposition 4.6. Let M_r denote the number of children at generation r of $\mathcal{X}_{n,\mathrm{res}}(i)$, and note that

$$\mathbb{P}(\bar{\partial}_i(\varepsilon n^\eta) \neq \varnothing) \leq \mathbb{E}[M_{\varepsilon n^\eta/2}] \times \mathbb{P}(\bar{\partial}_i^1(\varepsilon n^\eta/2) \neq \varnothing). \tag{4.25}$$

Now, using (4.22),

$$\mathbb{E}[M_{\varepsilon n^\eta/2}] \leq d_i(1-\beta_i^n n^{-\eta})^{\varepsilon n^\eta/2} \leq d_i \mathrm{e}^{-\frac{\varepsilon \beta_i^n}{2}}, \tag{4.26}$$

and $\bar{\xi}_n(i) \preceq \bar{\xi}_n(i-1) \preceq \cdots \preceq \bar{\xi}_n(i_2(\lambda))$, where $\preceq$ denotes stochastic domination. Thus,

$$\mathbb{P}(\bar{\partial}_i(\varepsilon n^\eta) \neq \varnothing) \leq d_i \mathrm{e}^{-\frac{\varepsilon \beta_i^n}{2}} \times \mathbb{P}(\bar{\partial}_{i_2(\lambda)}^1(\varepsilon n^\eta/2) \neq \varnothing), \tag{4.27}$$

and the proof of Theorem 4.5 follows using Proposition 4.6. □

The rest of this section is devoted to the proof of Proposition 4.6. We leverage some key ideas from [1]. Define the breadth-first random walk $\boldsymbol{s}_n$ by $s_n(0) = 1$ and

$$s_n(u) = s_n(u-1) + \zeta_u - 1, \tag{4.28}$$

where $(\zeta_u)_{u\geq 1}$ are i.i.d. observations from the distribution of $\bar{\xi}_n(i_2(\lambda))$. Let $\sigma := \inf\{u : s_n(u) = 0\}$ and for $t = 0, 1, \ldots, \sigma$, define the function

$$H_n(t) := \sum_{u=0}^{t-1} \frac{1}{s_n(u)}. \tag{4.29}$$

A remarkable fact observed in [1, Proposition 1.7] states that the height of a tree with breadth-first exploration process $\boldsymbol{s}_n$ is at most $3H_n(\sigma)$. Thus Proposition 4.6 can be concluded directly from the following estimate:

Proposition 4.7. *Under identical conditions as in* Theorem 4.5*, for all* $n \geq N_\lambda$*,*

$$\mathbb{P}(H_n(\sigma) > \varepsilon n^\eta) \leq \frac{C}{n^\alpha}, \tag{4.30}$$

for some constant $C = C(\varepsilon, \lambda) > 0$.

In what follows, we fix $\delta > 0$ such that $\delta n^\alpha + 2 < d_{i_2(\lambda)}/100$ for all $n \geq N_\lambda$. Define $I_l := [2^{l-1}, 2^{l+1})$ for $l \geq 1$. Let $\mathbb{P}_x$ denote the law of the random walk $\boldsymbol{s}_n$, starting from x and satisfying the recurrence relation in (4.28). Let $\sigma_{nl} := \min\{t \geq 1 : s_n(t) \notin I_l\}$ and $r_{nl} := \min\{t \geq 1 : \sup_{x\in I_l} \mathbb{P}_x(\sigma_{nl} > t) \leq 1/2\}$. We first obtain the following bound on r_{nl}:

Lemma 4.8. *Under identical conditions as in* Theorem 4.5*, there exists* $n_\star \geq 1$ *depending only on* $(d_i\,;\, i \in [n], n \geq 1)$ *such that for all* $n \geq n_\star$ *and all* $l \geq 1$ *satisfying* $2^{l+1} \leq \delta n^\alpha$*, we have* $r_{nl} \leq C2^{(\tau-2)l}$ *for some (sufficiently large) constant* $C > 0$.

Proof. By (4.17), $\mathbb{P}(\bar{\xi}_n(i_2(\lambda)) = j) = (1+o(1))\mathbb{P}(D_n^* = j+1)$ uniformly over $1 \leq j \leq d_{i_2(\lambda)}$. Thus, by Assumption 1.3,

$$\mathbb{P}\Big(\frac{u}{c_1} < \bar{\xi}_n(i_2(\lambda)) \leq u\Big) \geq Cu^{-(\tau-2)}, \tag{4.31}$$

for all $c_1 \leq u \leq \delta n^\alpha$.

Next, in order to estimate σ_{nl}, we bound $\sup_{x\in I_l} \mathbb{P}_x(s_n(t) \in I_l)$ using an upper bound on Lévy's concentration function due to Esseen [18], that we describe now. For a random variable Z, define Lévy's concentration function

$$Q(Z, L) := \sup_{x\in\mathbb{R}} \mathbb{P}(Z \in [x, x+L)). \tag{4.32}$$

By [18, Theorem 3.1], for any $u > 0$,

$$Q(s_n(t), u) \le \frac{Cu}{\left(t \times \mathbb{E}[|\zeta_1 - \zeta_2|^2 \mathbb{1}\{|\zeta_1 - \zeta_2| \le u\}]\right)^{1/2}}, \tag{4.33}$$

where ζ_1 and ζ_2 are i.i.d. realizations from the distribution of $\bar{\xi}_n(i_2(\lambda))$. To get an upper bound on the right side of (4.33), we first observe that for any random variable Y supported on $\mathbb{Z}_{\ge 0}$,

$$\begin{aligned}\mathbb{E}[Y^2 \mathbb{1}\{Y \le u\}] &= \sum_{1 \le y \le u} y^2 \mathbb{P}(Y = y) = \sum_{1 \le y \le u} \sum_{1 \le x \le y} y \mathbb{P}(Y = y) \\ &= \sum_{1 \le x \le u} \sum_{x \le y \le u} y \mathbb{P}(Y = y) \ge \sum_{1 \le x \le u} x \mathbb{P}(x \le Y \le u).\end{aligned} \tag{4.34}$$

Now, it follows from (4.17) and Assumption 1.1 (iii) that $\liminf_{n\to\infty} \mathbb{P}(\bar{\xi}_n(i_2(\lambda)) = 0) > 0$. Similarly, using (4.17) and Assumption 1.1 (ii), we can choose an integer $j_\star > c_1$ such that $\liminf_{n\to\infty} \mathbb{P}\big(\bar{\xi}_n(i_2(\lambda)) - 1 = j_\star\big) > 0$. Let $n_\star$ be such that

$$\inf_{n \ge n_\star} \mathbb{P}\big(\bar{\xi}_n(i_2(\lambda)) = 0\big) > 0 \quad \text{and} \quad \inf_{n \ge n_\star} \mathbb{P}\big(\bar{\xi}_n(i_2(\lambda)) - 1 = j_\star\big) > 0\,. \tag{4.35}$$

Then for any $n \ge n_\star$ and $c_1 \le u \le \delta n^\alpha$,

$$\begin{aligned}\mathbb{E}[|\zeta_1 - \zeta_2|^2 \mathbb{1}\{|\zeta_1 - \zeta_2| \le u\}] &\ge \sum_{1 \le x \le u} x \mathbb{P}(x \le |\zeta_1 - \zeta_2| \le u) \\ &\ge \sum_{1 \le x \le u/c_1} x \mathbb{P}(x \le \zeta_1 \le u) \mathbb{P}(\zeta_2 = 0) \ge C u^{4-\tau},\end{aligned} \tag{4.36}$$

where the penultimate step uses the fact that if $x \le \zeta_1 \le u$ and $\zeta_2 = 0$, then $x \le |\zeta_1 - \zeta_2| \le u$, and the final step follows using (4.31) and the first inequality in (4.35). Thus, (4.33) yields, for $n \ge n_\star$ and any $l \ge 1$ satisfying $c_1 \le 2^{l+1} \le \delta n^\alpha$,

$$\sup_{x \in I_l} \mathbb{P}_x(\sigma_{nl} > t) \le \sup_{x \in I_l} \mathbb{P}_x(s_n(t) \in I_l) \le Q(s_n(t), 2^{l+1}) \le \frac{C 2^l}{(t 2^{l(4-\tau)})^{1/2}}, \tag{4.37}$$

which is at most $1/2$ by choosing $t = C 2^{l(\tau-2)}$ for some large constant $C > 0$.

Finally, for all $n \ge n_\star$ and $l \ge 1$ satisfying $2^{l+1} < c_1$,

$$\sup_{x \in I_l} \mathbb{P}_x(\sigma_{nl} > t) \le \mathbb{P}\big(\bar{\xi}_n(i_2(\lambda)) - 1 \ne j_\star\big)^t \le \exp(-Ct)\,,$$

where the last step uses the second inequality in (4.35). This in particular implies that $r_{nl} \le C$ for all $n \ge n_\star$ and $l \ge 1$ satisfying $2^{l+1} < c_1$. This completes the proof. □

We now decompose the possible values of the random walk (4.28) starting from $s_n(0) = 1$ into different scales. Recall that $I_l := [2^{l-1}, 2^{l+1})$. At each time t, the scale of $s_n(t)$, denoted by $\mathrm{scl}(s_n(t))$, is an integer. Let $\mathrm{scl}(s_n(0)) = 1$. Suppose that $\mathrm{scl}(s_n(u)) = l$ for some $u > 0$. A change of scale occurs when $\boldsymbol{s}_n$ leaves I_l, i.e., at time $T := \inf\{t > u : s_n(t) \notin I_l\}$, and the new scale is given by $\mathrm{scl}(s_n(T)) = l'$, where l' is such that $s_n(T) \in [2^{l'-1}, 2^{l'})$. Now, the next change of scale occurs at time $T' := \inf\{t > T : s_n(t) \notin I_{l'}\}$, and the scale remains the same until T', i.e., $\mathrm{scl}(s_n(t)) = l'$ for all $T \le t < T'$. Define

$$H_{nl}(t) := \sum_{u \in [0,t),\ \mathrm{scl}(s_n(u)) = l} \frac{1}{s_n(u)}, \quad \text{so that} \quad H_n(t) = \sum_{l \ge 1} H_{nl}(t). \tag{4.38}$$

Let $T_{nl}(t) := \#\{u \in [0,t) : \mathrm{scl}(s_n(u)) = l\}$, and note that

$$2^{l-1} H_{nl}(t) \le T_{nl}(t) \le 2^{l+1} H_{nl}(t). \tag{4.39}$$

Therefore, for any $x > 0$,

$$\mathbb{P}\Big(H_{nl}(\sigma) \ge \frac{x r_{nl}}{2^{l-1}}\Big) \le \mathbb{P}(T_{nl}(\sigma) \ge x r_{nl}). \tag{4.40}$$

The next lemma estimates $\mathbb{P}(T_{nl}(\sigma) \ge x r_{nl})$:

Lemma 4.9. *For all $n \ge 1$, $l \ge 1$, and $x > 0$,*

$$\mathbb{P}(T_{nl}(\sigma) \ge x r_{nl}) \le C 2^{-l - C' x}, \tag{4.41}$$

for some absolute constants $C, C' > 0$.

Proof. Let us first show that for any $l \ge 2$,

$$\mathbb{P}(T_{nl}(\sigma) \ne 0) \le 2^{-(l-1)}. \tag{4.42}$$

For any $t \ge 0$, let $\mathcal{F}_t$ denote the sigma-field generated by $(\zeta_u)_{u=0}^t$, where we take $\zeta_0 = 1$. Note that if $T_{nl}(\sigma) \ne 0$, then $s_n(u)$ hits 2^{l-1} before hitting zero. For $H > 1$, let $\gamma_H := \min\{t : s_n(t) \ge H, \text{ or } s_n(t) = 0\}$. Since $\mathbb{E}[\zeta_u - 1] < 0$ by (4.22), $(s_n(t))_{t \ge 0}$ is a supermartingale with respect to the filtration $(\mathcal{F}_t)_{t \ge 0}$. Consequently, an application of the optional stopping theorem yields

$$H \mathbb{P}(s_n(\gamma_H) \ge H) \le \mathbb{E}[s_n(\gamma_H)] \le \mathbb{E}[s_n(0)] = 1, \tag{4.43}$$

and therefore,

$$\mathbb{P}(s_n(\gamma_H) \ge H) \le \frac{1}{H}. \tag{4.44}$$

Thus, (4.42) follows by taking $H = 2^{l-1}$ together with the fact that $T_{nl}(\sigma) \ne 0$ implies that $s_n(\gamma_H) \ge H$.

Next, we define $U_n(t, [a,b))$–the number of upcrossings of an interval $[a,b)$ by $\boldsymbol{s}_n$ up to time t–to be the supremum of all integers k such that there exist times $(u_j, t_j)_{j=1}^k$ satisfying $0 \le u_1 < t_1 < u_2 < \cdots < t_k \le t$, and $s_n(u_j) < a < b \le s_n(t_j)$ for all $j \in [k]$. We will use the following simple fact (see [1, Proposition 3.2]): for any positive integers k, z, a, b with $0 < z < a < b$,

$$\mathbb{P}_z\big(U_n(\sigma, [a,b)) \ge k\big) \le \Big(\frac{a-1}{b}\Big)^k. \tag{4.45}$$

Next define $\mathrm{visit}(l,t)$ to be the number of visits to scale l upto time t, i.e., this is the supremum over $k \in \mathbb{N}$ such that one can find $(u_j, t_j)_{j=1}^k$ with $u_1 < t_1 < \cdots < u_k < t_k \le t$ satisfying $\mathrm{scl}(s_n(u_j)) \ne l$ but $\mathrm{scl}(s_n(t_j)) = l$. For the random walk $\boldsymbol{s}_n$ started at $s_n(0) = 1$, we set $\mathrm{visit}(1,0) = 1$ and $\mathrm{visit}(l,t) = 0$ if $\boldsymbol{s}_n$ does not enter scale l before time t. Further, define $M_{nl} = \mathrm{visit}(l, \sigma)$ (the total number of visits to scale l) and $t_{jl} = \#\{t < \sigma : \mathrm{scl}(s_n(t)) = l, \mathrm{visit}(l,t) = j\}$ (the time spent at scale l during the j-th visit). Note that, if $T_{nl}(\sigma) \ne 0$ occurs, then $M_{nl} \ge 1$, and $T_{nl}(\sigma) = \sum_{j=1}^{M_{nl}} t_{jl}$. Thus, for any $m \ge 2$ and $x \in \mathbb{Z}_{\ge 2}$,

$$\mathbb{P}\big(T_{nl}(\sigma) > 5 x r_{nl}\big) = \mathbb{P}\Big(\sum_{j=1}^{M_{nl}} t_{jl} > 5 x r_{nl}\Big) \le \mathbb{P}(M_{nl} > m) + \mathbb{P}\Big(\sum_{j=1}^{m} t_{jl} > 5 x r_{nl}\Big). \tag{4.46}$$

Now, $M_{nl} > m$ implies that $T_{nl}(\sigma) \neq 0$, and after the first visit to scale l, the walk comes back to scale l at least m times before hitting zero. In any of the subsequent visits, if $\boldsymbol{s}_n$ enters scale l from below (this can only happen for $l \geq 3$), then that would imply an upcrossing of the interval $[2^{l-2}, 2^{l-1})$ has taken place. Otherwise, if $\boldsymbol{s}_n$ enters scale l from above in any of the subsequent visits, then it must be the case that while leaving the scale l during the previous visit, the walk went from scale l to a higher scale. This yields an upcrossing of $[2^l, 2^{l+1})$. Therefore, for any $l \geq 3$, $M_{nl} > m$ implies that $T_{nl}(\sigma) \neq 0$, and after the first visit to scale l and before hitting zero, either at least $m/2$ many upcrossings of $[2^{l-2}, 2^{l-1})$ have taken place, or at least $m/2$ many upcrossings of $[2^l, 2^{l+1})$ have taken place. Thus, using (4.42), (4.45), and the strong Markov property, for any $l \geq 3$,

$$\mathbb{P}(M_{nl} > m) \leq \frac{C}{2^{l+m/2}}. \tag{4.47}$$

Next, by the definition of r_{nl} given right above Lemma 4.8, $\mathbb{P}_z(t_{jl} > kr_{nl}) \leq 2^{-k}$ for any $z > 0$, which implies that $\lfloor t_{jl}/r_{nl} \rfloor$ can be stochastically dominated by a Geometric$(1/2)$ random variable. Using the strong Markov property, it follows that for any $z > 0$, under $\mathbb{P}_z$, $\sum_{j=1}^m \lfloor t_{jl}/r_{nl} \rfloor$ is stochastically dominated by $\sum_{i=1}^m g_i$, where $(g_i)_{i \geq 1}$ is an i.i.d. collection of Geometric$(1/2)$ random variables. Thus, for any $z > 0$,

$$\begin{aligned}\mathbb{P}_z\Big(\sum_{j=1}^m t_{jl} \geq (k+m)r_{nl}\Big) \leq \mathbb{P}_z\Big(\sum_{j=1}^m \Big\lfloor \frac{t_{jl}}{r_{nl}} \Big\rfloor > k\Big) \leq \mathbb{P}\Big(\sum_{i=1}^m g_i > k\Big) \\ = \mathbb{P}(\mathrm{Bin}(k, 1/2) < m) \leq \mathrm{e}^{-(k-2m)^2/2k},\end{aligned}$$

for $2m \leq k$, where the last step follows using standard concentration inequalities such as [27, Theorem 2.1]. Consequently, using (4.42) and the strong Markov property, $\mathbb{P}\big(\sum_{j=1}^m t_{jl} \geq (k+m)r_{nl}\big) \leq 2^{-(l-1)} \cdot \mathrm{e}^{-(k-2m)^2/2k}$ for $2m \leq k$. Combining this with (4.46) and (4.47), and taking $k = 4x$ and $m = x$ yields

$$\mathbb{P}(T_{nl}(\sigma) > 5xr_{nl}) \leq C2^{-l-C'x} \tag{4.48}$$

for any $l \geq 3$. The proofs for $l = 1$ and $l = 2$ follow similar steps. This completes the proof of Lemma 4.9. □

We are now ready to prove Proposition 4.7.

Proof of Proposition 4.7. Recall the definition of $\boldsymbol{s}_n$ from (4.28) starting from one, so that $s_n(0) = 1$. Fix $\delta > 0$. We first estimate the probability of the event $\mathcal{B}_n$ that $\boldsymbol{s}_n$ hits $\delta n^\alpha/2$ before hitting zero. Let $\gamma := \min\{t\colon s_n(t) \geq \delta n^\alpha/2, \text{ or } s_n(t) = 0\}$. By (4.44),

$$\mathbb{P}(\mathcal{B}_n) = \mathbb{P}\Big(s_n(\gamma) \geq \frac{\delta n^\alpha}{2}\Big) \leq \frac{2}{\delta n^\alpha}. \tag{4.49}$$

Let $m := \max\{l \geq 1 : 2^{l+1} \leq \delta n^\alpha\}$. On $\mathcal{B}_n^c$, $H_{nl}(\sigma) = 0$ for $l > m$. Thus, for any sequence of positive numbers $(b_l)_{l \geq 1}$,

$$\begin{aligned}\mathbb{P}\Big(H_n(\sigma) \geq \sum_{l=1}^m \frac{b_l r_{nl}}{2^{l-1}}\Big) &\leq \frac{2}{\delta n^\alpha} + \mathbb{P}\Big(H_n(\sigma) \geq \sum_{l=1}^m \frac{b_l r_{nl}}{2^{l-1}}, \text{ and } \mathcal{B}_n^c \text{ occurs}\Big) \\ &\leq \frac{2}{\delta n^\alpha} + \mathbb{P}\Big(H_{nl}(\sigma) \geq \frac{b_l r_{nl}}{2^{l-1}} \text{ for some } 1 \leq l \leq m\Big).\end{aligned} \tag{4.50}$$

Using (4.40) and Lemma 4.9, (4.50) yields

$$\mathbb{P}\Big(H_n(\sigma) \geq \sum_{l=1}^m \frac{b_l r_{nl}}{2^l}\Big) \leq \frac{2}{\delta n^\alpha} + \sum_{l=1}^m \mathbb{P}(T_{nl}(\sigma) \geq b_l r_{nl}) \leq \frac{2}{\delta n^\alpha} + C\sum_{l=1}^m 2^{-l-C'b_l}. \tag{4.51}$$

Letting $b_l = \frac{1}{C'}(m-l+1+2\log_2(m-l+1))$ for $1 \le l \le m$, and using Lemma 4.8,

$$\begin{aligned}\sum_{l=1}^{m} \frac{b_l r_{nl}}{2^{l-1}} &\le C \sum_{l=1}^{m} \big(m-l+1+2\log_2(m-l+1)\big)2^{l(\tau-3)} \\ &= C \sum_{j=1}^{m} \frac{(j+2\log_2 j)2^{(m+1)(\tau-3)}}{2^{j(\tau-3)}} \le C(\delta n^{\alpha})^{\tau-3} \sum_{j=1}^{m} \frac{(j+2\log_2 j)}{2^{j(\tau-3)}} \\ &\le C n^{\eta} \delta^{\tau-3},\end{aligned} \tag{4.52}$$

where we have used $\sum_{j=1}^{\infty} \frac{(j+2\log_2 j)}{2^{j(\tau-3)}} < \infty$ in the last step; the bound in (4.52) holds for all $n \ge n_\star$, where $n_\star$ is as in Lemma 4.8. Also,

$$\sum_{l=1}^{m} 2^{-l-C'b_l} = \sum_{l=1}^{m} 2^{-(m+1)}(m-l+1)^{-2} \le \frac{4}{\delta n^{\alpha}} \sum_{l=1}^{\infty} \frac{1}{l^2}. \tag{4.53}$$

Thus, the claim in Proposition 4.7 follows for $n \ge n_\star$ by combining (4.52) and (4.53) with (4.51). We conclude that the claimed bound holds for $n \ge N_\lambda$ by choosing a larger constant C on the right side of (4.30). □

4.4 Proof of Proposition 2.2

Let us now complete the proof of Proposition 2.2 using Proposition 4.4 and Theorem 4.5. We take K_n as in Lemma A.1 so that the results in Section 4.3 hold. Note that these bounds work for $i_2(\lambda) \le i \le K_n$, and we will use path counting arguments from [7, 25] to bound the diameter for $i > K_n$. Define $\mathscr{C}_{\mathrm{res}}(i)$ to be the connected component containing vertex i in the graph $\mathcal{G}_n^{>i-1} = \mathrm{CM}_n(\boldsymbol{d}) \setminus [i-1]$. Note that if $\Delta^{>K} > \varepsilon n^{\eta}$, then there exists a path of length εn^{η} in $\mathrm{CM}_n(\boldsymbol{d})$ avoiding all the vertices in $[K]$. Suppose that the minimum index among vertices on that path is i_0. Then $\mathrm{diam}(\mathscr{C}_{\mathrm{res}}(i_0)) > \varepsilon n^{\eta}$. Therefore, $\Delta^{>K} > \varepsilon n^{\eta}$ implies that either there exists $i \in (K, K_n)$ satisfying $\mathrm{diam}(\mathscr{C}_{\mathrm{res}}(i)) > \varepsilon n^{\eta}$, or $\mathrm{diam}(\mathrm{CM}_n(\boldsymbol{d}) \setminus [K_n]) > \varepsilon n^{\eta}$. We will use the following lemma first to complete the proof of Proposition 2.2 and prove the lemma subsequently:

Lemma 4.10. *Under* Assumptions 1.1 and 1.3, for any $\varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}(\mathrm{diam}(\mathrm{CM}_n(\boldsymbol{d}) \setminus [K_n]) > \varepsilon n^{\eta}) = 0$, where K_n as in Lemma A.1.

Proof of Proposition 2.2. As defined earlier around (4.17), $\partial_i(r)$ denotes the number of vertices at distance r from the vertex i in the graph $\mathcal{G}_n^{>i-1}$. Recall the definition of $\bar{\partial}$ in Proposition 4.4. Thus, Proposition 4.4 and Theorem 4.5 together with Lemma 4.10, yield that

$$\begin{aligned}\mathbb{P}\left(\Delta^{>K} > \varepsilon n^{\eta}\right) &\le \sum_{i\in(K,K_n)} \mathbb{P}(\bar{\partial}_i(\varepsilon n^{\eta}/2) \ne \varnothing) + \mathbb{P}(\mathrm{diam}(\mathrm{CM}_n(\boldsymbol{d}) \setminus [K_n]) > \varepsilon n^{\eta}) \\ &\le C \sum_{i\in(K,K_n)} \Big(\frac{d_i}{n^{\alpha}}\Big) \mathrm{e}^{-\varepsilon\beta_i^n/4} + o(1),\end{aligned} \tag{4.54}$$

where the last line tends to zero if we first take $n \to \infty$ and then take $K \to \infty$ using Assumption 1.3 and Lemma A.1 below. Thus the proof of Proposition 2.2 follows. □

Proof of Lemma 4.10. Let $\boldsymbol{d}' := (d_i'\,; i \in [n] \setminus [K_n])$, where d_i' denotes the degree of i in $\mathrm{CM}_n(\boldsymbol{d}) \setminus [K_n]$. Note that $\mathrm{CM}_n(\boldsymbol{d}) \setminus [K_n]$ is again distributed as a configuration model conditionally on its degree sequence $\boldsymbol{d}'$, with the criticality parameter

$$\nu_n' = \frac{\sum_{i>K_n} d_i'(d_i'-1)}{\sum_{i>K_n} d_i'} \le \frac{\sum_{i>K_n} d_i(d_i-1)}{\ell_n - 2\sum_{i=1}^{K_n} d_i} \le 1 - R_n n^{-\eta}, \quad R_n = \omega(\log n), \tag{4.55}$$

where the penultimate step follows using $d'_i \le d_i$ and $\ell'_n := \sum_{i>K_n} d'_i = \ell_n - 2\sum_{i=1}^{K_n} d_i$, and the last bound follows from the definition of K_n given in Lemma A.1 (ii) and an argument identical to that in (4.21). Let $\mathbb{P}'(\cdot)$ denote the probability measure conditionally on $\boldsymbol{d}'$. We will use path-counting techniques for subcritical configuration models. An argument similar to the one given in [25, Lemma 6.1] shows that for any $l \ge 1$, conditional on $\boldsymbol{d}'$, the expected number of paths of length l starting from vertex i is at most

$$\frac{\ell'_n d'_i (\nu'_n)^{l-1}}{\ell'_n - 2l + 3} \le \ell'^2_n (\nu'_n)^{l-1}.$$

Thus, for any $i > K_n$,

$$\mathbb{P}'(\exists \text{ path of length at least } \varepsilon n^\eta \text{ from } i \text{ in } \mathrm{CM}_n(\boldsymbol{d}) \setminus [K_n]) \le C\ell'^2_n \sum_{l>\varepsilon n^\eta} (\nu'_n)^l, \tag{4.56}$$

Thus, for $i > K_n$, the probability in (4.56) is at most

$$Cn^2(1 - R_n n^{-\eta})^{\varepsilon n^\eta}/(R_n n^{-\eta}) \le Cn^{2+\eta} \mathrm{e}^{-\varepsilon R_n} = o(1/n), \tag{4.57}$$

since $R_n \gg \log n$. Therefore,

$$\mathbb{P}'(\exists i > K_n : \exists \text{ path of length at least } \varepsilon n^\eta \text{ from } i \text{ in } \mathrm{CM}_n(\boldsymbol{d}) \setminus [K_n]) = o(1), \tag{4.58}$$

and the proof of Lemma 4.10 follows. □

5 Verification of the assumptions for percolated degrees: Proof of Theorem 1.7

Let $\mathcal{G}_n$ denote the graph obtained by performing percolation with edge retention probability $p_c(\lambda)$ (defined in (1.12)) on $\mathrm{CM}_n(\boldsymbol{d})$. Let $\boldsymbol{d}^p = (d_i^p)_{i\in[n]}$ denote the degree sequence of $\mathcal{G}_n$. By [19, Lemma 3.2], the law of $\mathcal{G}_n$, conditionally on $\boldsymbol{d}^p$, is the same as the law of $\mathrm{CM}_n(\boldsymbol{d}^p)$. Thus, it is enough to show that if the original degree sequence $(\boldsymbol{d}_n\,,\, n \ge 1)$ satisfies Assumptions 1.1(i), 1.1(ii) and 1.3, then we can construct $(\boldsymbol{d}^p\,,\, n \ge 1)$ on the same probability space so that Assumption 1.1, (1.6), and Assumption 1.3 are satisfied almost surely (with possibly different parameters), since then the claim in Theorem 1.7 will follow from Theorem 1.5.

First, note that $\mathbb{E}[d_i^p] = d_i p_c(\lambda)(1 + o(1))$. Also, given $\mathrm{CM}_n(\boldsymbol{d})$, changing the status of an edge (deleted or retained) can change d_i^p by at most 2 when the edge is incident to i. There are at most d_i choices for such an edge. Thus, the bounded difference inequality [27, Corollary 2.27] implies, for each fixed $i \ge 1$, and for any $\varepsilon > 0$,

$$\mathbb{P}\big(|d_i^p - d_i p_c(\lambda)| > \varepsilon d_i p_c(\lambda)\big) \le 2\mathrm{e}^{-\frac{\varepsilon^2}{4} d_i p_c^2(\lambda)}\,. \tag{5.1}$$

In particular, for each $i \ge 1$, $n^{-\alpha} d_i^p \xrightarrow{d} \theta_i/\nu$ as $n \to \infty$, which verifies Assumption 1.1 (i).

Next, let $M_r^p = \sum_{i\in[n]} (d_i^p)_r$ and $M_r = \sum_{i\in[n]} (d_i)_r$, where $(x)_r := x(x-1)\cdots(x-r+1)$. To verify the moment conditions in Assumption 1.1 (ii), note that (1.5) holds for $\boldsymbol{d}^p$ since $\sum_{i>K} (d_i^p)_3 \le \sum_{i>K} (d_i)_3$. We will show that

$$M_1^p = (1 + O_{\mathrm{P}}(n^{-1/2}))p_c(\lambda)M_1 \quad \text{and} \quad M_2^p = (1 + O_{\mathrm{P}}(n^{\frac{3\alpha}{2}-1}))p_c(\lambda)^2 M_2. \tag{5.2}$$

Using (5.2), the first and second moment assumptions in Assumption 1.1 (ii) holds for the percolated degree sequence. The estimate (5.2) also shows that (1.6) holds. Indeed,

$$\frac{M_2^p}{M_1^p} = \frac{p_c(\lambda)M_2}{M_1}(1 + O_{\mathrm{P}}(n^{\frac{3\alpha}{2}-1})) = 1 + \nu\lambda n^{-\eta} + o(n^{-\eta}), \tag{5.3}$$

where the last step follows using (1.11), (1.12), and the fact that $-1+3\alpha/2 < -1+2\alpha = -\eta$.

It remains to prove (5.2). Since $\frac{1}{2}\sum_{i\in[n]} d_i^p$ has a binomial distribution with parameter $\ell_n/2$ and $p_c(\lambda)$, the first asymptotics follows from Chebyshev's inequality. For the asymptotics of M_2^p, we use the following construction of $\mathcal{G}_n$ from [19].

Algorithm 2. $\boldsymbol{d}^p = (d_i^p)_{i\in[n]}$ can be generated as follows:

(S0) Sample $R_n \sim \mathrm{Bin}(\ell_n/2, p_c(\lambda))$.

(S1) Conditionally on R_n, sample a uniform subset of $2R_n$ half-edges from the set of ℓ_n half-edges. Let $I_j^{(i)}$ denote the indicator that j-th half-edge of i is selected. Then $d_i^p = \sum_{j=1}^{d_i} I_j^{(i)}$ for all $i \in [n]$.

Using the above construction, note that

$$M_2^p = \sum_{i\in[n]} \sum_{1\le j_1\ne j_2\le d_i} I_{j_1}^{(i)} I_{j_2}^{(i)}. \tag{5.4}$$

Let $\mathbb{P}_1(\cdot) = \mathbb{P}(\cdot|R_n)$ and similarly define $\mathbb{E}_1[\cdot]$, $\mathrm{Var}_1(\cdot)$, and $\mathrm{Cov}_1(\cdot,\cdot)$. Then,

$$\begin{aligned}\mathbb{E}_1[M_2^p] &= \sum_{i\in[n]} \sum_{1\le j_1\ne j_2\le d_i} \mathbb{P}_1(I_{j_1}^{(i)}=1, I_{j_2}^{(i)}=1) = \sum_{i\in[n]} \sum_{1\le j_1\ne j_2\le d_i} \frac{\binom{\ell_n-2}{2R_n-2}}{\binom{\ell_n}{2R_n}} \\ &= \sum_{i\in[n]} \sum_{1\le j_1\ne j_2\le d_i} \frac{2R_n(2R_n-1)}{\ell_n(\ell_n-1)} = (1+O_{\mathbb{P}}(n^{-1/2}))p_c(\lambda)^2 M_2,\end{aligned} \tag{5.5}$$

where the last step follows using $R_n = (1+O_{\mathbb{P}}(n^{-1/2}))p_c(\lambda)\ell_n/2$.

Next, recall that a collection of random variables $(X_1,\dots,X_t)$ is called negatively associated if for every index set $I \subset [k]$,

$$\mathrm{Cov}\big(f(X_i, i\in I), g(X_i, i\in I^c)\big) \le 0, \tag{5.6}$$

for all functions $f : \mathbb{R}^{|I|} \mapsto \mathbb{R}$ and $g : \mathbb{R}^{t-|I|} \mapsto \mathbb{R}$ that are component-wise non-decreasing ([17, Definition 3]). Then, conditionally on R_n, $I_j^{(i)}$, $j = 1,\dots,d_i$, $i \in [n]$ are negatively associated (cf. [17, Theorem 10]), which yields the almost sure bound

$$\mathrm{Var}_1(M_2^p) \le \sum_{i\in[n]} \sum_{1\le j_1\ne j_2\le d_i} \mathrm{Var}_1(I_{j_1}^{(i)} I_{j_2}^{(i)}) + \sum_{i\in[n]} \sum_{\substack{1\le j_1\ne j_2\le d_i\\ 1\le j_3\ne j_4\le d_i\\ |\{j_1,j_2\}\cap\{j_3,j_4\}|=1}} \mathrm{Cov}_1(I_{j_1}^{(i)} I_{j_2}^{(i)}, I_{j_3}^{(i)} I_{j_4}^{(i)}), \tag{5.7}$$

since the contribution of $|\{j_1,j_2\}\cap\{j_3,j_4\}| = 0$ can be ignored due to negative association. Also, $\mathrm{Var}_1(I_{j_1}^{(i)} I_{j_2}^{(i)}) \le 1$ and $|\mathrm{Cov}_1(I_{j_1}^{(i)} I_{j_2}^{(i)}, I_{j_3}^{(i)} I_{j_4}^{(i)})| \le (\mathrm{Var}_1(I_{j_1}^{(i)} I_{j_2}^{(i)})\mathrm{Var}_1(I_{j_3}^{(i)} I_{j_4}^{(i)}))^{1/2} \le 1$. Therefore,

$$\mathrm{Var}_1(M_2^p) \le \sum_{i\in[n]} d_i^2 + 4\sum_{i\in[n]} d_i^3 = O(n^{3\alpha}). \tag{5.8}$$

Thus, for any $A > 0$,

$$\mathbb{P}\big(|M_2^p - \mathbb{E}_1[M_2^p]| > An^{3\alpha/2}\big) = \mathbb{E}\big[\mathbb{P}_1\big(|M_2^p - \mathbb{E}_1[M_2^p]| > An^{3\alpha/2}\big)\big] \le \frac{\mathbb{E}\big[\mathrm{Var}_1(M_2^p)\big]}{A^2 n^{3\alpha}}, \tag{5.9}$$

which can be made arbitrarily small by choosing $A > 0$ large. Thus, we conclude the asymptotics of M_2^p in (5.2) by using (5.5) and (5.9).

Finally, we need to show convergence in distribution of empirical measure of $\boldsymbol{d}^p$ to finish verifying Assumptions 1.1 (ii) and (iii). Let $n_k^p = \#\{i : d_i^p = k\}$, and $n_{\geq k}^p = \sum_{r\geq k} n_r^p$. It suffices to show that

$$\frac{n_{\geq k}^p}{n} \xrightarrow{d} \mathbb{P}(D^p \geq k) \quad \text{for all } k \geq 1, \tag{5.10}$$

where D^p satisfies $\big(D^p \mid D = l\big) \sim \mathrm{Bin}(l, 1/\nu)$, for all $l \geq 1$. Let V_n be a uniformly chosen vertex and $D_n^p = d_{V_n}^p$ and $D_n = d_{V_n}$. By the construction in Algorithm 2,

$$\begin{aligned}\mathbb{P}(D_n^p = k \mid D_n = l, R_n) &= \frac{\binom{l}{k}\binom{\ell_n - l}{2R_n - k}}{\binom{\ell_n}{2R_n}} = (1+o(1))\binom{l}{k}\left(\frac{2R_n}{\ell_n}\right)^k\left(1-\frac{2R_n}{\ell_n}\right)^{l-k} \\ &= (1+o_{\mathbb{P}}(1))\binom{l}{k}\left(\frac{1}{\nu}\right)^k\left(1-\frac{1}{\nu}\right)^{l-k},\end{aligned} \tag{5.11}$$

where in the final step we have used that $R_n = p_c(\lambda)\ell_n/2(1+o_{\mathbb{P}}(1))$ and $p_c(\lambda) = \nu^{-1}(1+o(1))$. Thus,

$$\mathbb{E}\left[\frac{n_k^p}{n} \;\middle|\; R_n\right] = \mathbb{P}(D_n^p = k \mid R_n) \xrightarrow{d} \mathbb{P}(D^p = k). \tag{5.12}$$

Moreover, note that $d_i^p = \sum_{j=1}^{d_i} I_j^{(i)}$, and the definition of negative association in (5.6) allows us to conclude negative correlation between increasing functions of $I_j^{(i)}$ that depend on disjoint sets of indices. Therefore,

$$\mathrm{Var}\big(n_{\geq k}^p \mid R_n\big) = \mathrm{Var}\Big(\sum_{i:\, d_i \geq k} \mathbb{1}_{\{d_i^p \geq k\}} \;\Big|\; R_n\Big) \leq \sum_{i:\, d_i \geq k} \mathrm{Var}\big(\mathbb{1}_{\{d_i^p \geq k\}} \mid R_n\big) \leq n, \tag{5.13}$$

and thus $\mathbb{E}[\mathrm{Var}(n_{\geq k}^p/n \mid R_n)] = O(1/n)$. This together with (5.12) yields (5.10).

We finally verify that Assumption 1.3 holds with high probability. Let the constants c_1 and c_0 be as in Assumption 1.3. Let $c_1' := 9c_1\nu$. It is enough to show that there exists a deterministic constant $c_0' > 0$ such that

$$\mathbb{P}\Big(\frac{1}{n}\sum_{i\in[n]} d_i^p \mathbb{1}\{l < d_i^p \leq c_1' l\} \geq \frac{c_0'}{l^{\tau-2}} \quad \text{for all} \quad 1 \leq l \leq d_1/c_1'\Big) \to 1 \quad \text{as} \quad n \to \infty. \tag{5.14}$$

Write $\sum_*$ for $\sum_{i:8l<d_i p_c(\lambda)\leq 8c_1 l}$. Then, for $1 \leq l \leq d_1/c_1'$ and all large n,

$$\begin{aligned}\frac{1}{n}\sum_{i\in[n]} \mathbb{E}\big[d_i^p \mathbb{1}\{l < d_i^p \leq c_1' l\}\big] &\geq \frac{1}{n}\sum\nolimits_* \mathbb{E}\big[d_i^p \mathbb{1}\{l < d_i^p \leq c_1' l\}\big] \\ &= \frac{1}{n}\sum\nolimits_* \mathbb{E}\big[d_i^p \mathbb{1}\{l < d_i^p\}\big],\end{aligned} \tag{5.15}$$

where the last step uses the fact that when $d_i p_c(\lambda) \leq 8c_1 l$, we have $d_i^p \leq d_i \leq 8c_1 l/p_c(\lambda) \leq c_1' l$ for all large n. Let $X_i \sim \mathrm{Bin}(\lfloor d_i/2 \rfloor, p_c(\lambda))$. Then d_i^p is stochastically larger than X_i. Using (5.15), we see that for $1 \leq l \leq d_1/c_1'$ and all large n,

$$\begin{aligned}\frac{1}{n}\sum_{i\in[n]} \mathbb{E}\big[d_i^p \mathbb{1}\{l < d_i^p \leq c_1' l\}\big] &\geq \frac{1}{n}\sum\nolimits_* \mathbb{E}\big[X_i \cdot \mathbb{1}\{l < X_i\}\big] \\ &\geq \frac{1}{n}\sum\nolimits_* \mathbb{E}[X_i]\mathbb{P}(X_i > l) \geq \frac{1}{n}\sum\nolimits_* \mathbb{E}[X_i]\mathbb{P}\big(X_i \geq \lfloor \lfloor d_i/2 \rfloor p_c(\lambda) \rfloor\big) \\ &\geq \frac{1}{n}\sum\nolimits_* \mathbb{E}[X_i] \cdot \frac{1}{2} \geq \frac{C}{n}\sum\nolimits_* d_i \geq \frac{C'}{l^{\tau-2}},\end{aligned} \tag{5.16}$$

where the third step uses $l < d_i p_c(\lambda)/8 \le \lfloor \lfloor d_i/2 \rfloor p_c(\lambda) \rfloor$, and the final step follows using Assumption 1.3.

Now let $F_1 := \sum_{i\in[n]} d_i^p \mathbb{1}\{l < d_i^p \le c_1' l\}$ and $F_2 := \mathbb{E}[F_1|\mathrm{CM}_n(\boldsymbol{d})]$. We will apply the bounded difference inequality from [27, Corollary 2.27]. Given the graph $\mathrm{CM}_n(\boldsymbol{d})$, if we keep one extra edge in the percolated graph, then F_1 can change by at most $2c_1' l$. Thus, for any $\varepsilon > 0$,

$$\mathbb{P}\Big(|F_1 - F_2| > \frac{n\varepsilon}{l^{\tau-2}} \,\Big|\, \mathrm{CM}_n(\boldsymbol{d})\Big) \le 2\exp\Big(-\frac{n^2\varepsilon^2}{l^{2(\tau-2)}(2c_1' l)^2 \cdot \frac{\ell_n}{2}}\Big) \le 2\mathrm{e}^{-C\varepsilon^2 n l^{-2(\tau-1)}}. \quad (5.17)$$

Also, we can apply concentration inequalities such as [12, Lemma 2.5] to conclude that

$$\mathbb{P}\Big(|F_2 - \mathbb{E}[F_2]| > \frac{n\varepsilon}{l^{\tau-2}}\Big) \le 2\mathrm{e}^{-C\varepsilon^2 n l^{-2(\tau-1)}}. \quad (5.18)$$

Combining (5.17) and (5.18) together with (5.16) shows that there exists an $\varepsilon_0 > 0$ such that (5.14) holds if we replace "for all $1 \le l \le d_1/c_1'$" by "for all $1 \le l \le n^{\varepsilon_0}$." For $l \ge n^{\varepsilon_0}$, we use (5.1) together with a union bound to complete the proof of (5.14).

A A technical lemma

Lemma A.1. *Let $\beta_i^n = n^{-2\alpha}\sum_{j=1}^{i-1} d_j^2$. Then* Assumption 1.1(i), (1.5), and Assumption 1.3 *imply the following:*

(i) *For all $\varepsilon > 0$,*

$$\lim_{K\to\infty} \limsup_{n\to\infty} \sum_{i>K} \Big(\frac{d_i}{n^\alpha}\Big) \times \mathrm{e}^{-\varepsilon\beta_i^n} = 0. \quad (A.1)$$

(ii) *There exists a sequence $(K_n)_{n\ge1}$ with $K_n \to \infty$, and $K_n = o(n^\alpha)$ such that $\beta_{K_n}^n > \log^3 n$ for all large n.*

Proof. We will use $C_0, C_1, \ldots$ etc. as generic notation for positive constants that do not depend on n. Recall Assumption 1.3. Let $\theta_{i,n} := n^{-\alpha} d_i$, $i \in [n]$. We first claim that Assumption 1.3 implies

$$\min_{2\le i\le n} \theta_{i,n}^{\tau-2} \sum_{j=1}^{i-1} \theta_{j,n} \ge C_0. \quad (A.2)$$

To see this, let $1 = i_1 < i_2 < i_3 < \ldots$ be the indices such that $d_{i_{k-1}} = d_{i_{k-1}+1} = \ldots = d_{i_k-1} > d_{i_k}$ for $k \ge 2$. Then for $k \ge 2$,

$$\frac{1}{\ell_n}\sum_{j=1}^{i_k-1} d_j = \mathbb{P}(D_n^* > d_{i_k}) \ge \mathbb{P}\big(d_{i_k} < D_n^* \le c_1 d_{i_k}\big) \ge c_0 (d_{i_k})^{-(\tau-2)},$$

and consequently,

$$\min_k \theta_{i_k,n}^{\tau-2} \sum_{j=1}^{i_k-1} \theta_{j,n} \ge C_0. \quad (A.3)$$

If $i_k > i \ge i_{k-1}$, then $\theta_{i,n}^{\tau-2}\sum_{j=1}^{i-1}\theta_{j,n} = \theta_{i_{k-1},n}^{\tau-2}\sum_{j=1}^{i-1}\theta_{j,n} \ge \theta_{i_{k-1},n}^{\tau-2}\sum_{j=1}^{i_{k-1}-1}\theta_{j,n}$. Thus we conclude (A.2) from (A.3).

Next, define

$$f_n(x) := \begin{cases} \frac{1}{\theta_{i+1,n}}, & \text{if } \sum_{j=1}^{i-1}\theta_{j,n} \le x < \sum_{j=1}^{i}\theta_{j,n} \text{ for some } i \in [n-1], \\ 0, & \text{if } x \ge \sum_{j=1}^{n-1}\theta_{j,n}, \end{cases} \quad (A.4)$$

and

$$g_n(x) := \begin{cases} \sum_{j=1}^{i} \theta_{j,n}^2\,, & \text{if } \sum_{j=1}^{i-1} \theta_{j,n} \le x < \sum_{j=1}^{i} \theta_{j,n} \text{ for some } i \in [n], \\ 0\,, & \text{if } x \ge \sum_{j=1}^{n} \theta_{j,n}\,. \end{cases} \tag{A.5}$$

Since $\sum_{j=1}^{i} \theta_{j,n} \le 2\sum_{j=1}^{i-1} \theta_{j,n}$ for $2 \le i \le n$, we have, using (A.2), $\theta_{i+1,n}^{\tau-2} \sum_{j=1}^{i-1} \theta_{j,n} \ge C_0/2$. Therefore, $f_n(x)^{-(\tau-2)} \times 2x \ge C_0$ for any $\theta_{1,n} \le x < \sum_{j=1}^{n-1} \theta_{j,n}$, and consequently,

$$f_n(x) \le C_1 x^{\frac{1}{\tau-2}} \quad \text{for} \quad \theta_{1,n} \le x < \sum_{j=1}^{n-1} \theta_{j,n}. \tag{A.6}$$

Next, for $i \in [n-1]$,

$$\begin{aligned} \sum_{j=1}^{i} \theta_{j,n}^2 \ge \sum_{j=1}^{i} \theta_{j,n}\theta_{j+1,n} = \theta_{1,n}\theta_{2,n} + \int_{\theta_{1,n}}^{\sum_{j=1}^{i} \theta_{j,n}} \frac{\mathrm{d}x}{f_n(x)} \\ \ge C_2 \int_{\theta_{1,n}}^{\sum_{j=1}^{i} \theta_{j,n}} \frac{\mathrm{d}x}{x^{1/(\tau-2)}} \ge C_3 \Big(\sum_{j=1}^{i} \theta_{j,n} \Big)^{\frac{\tau-3}{\tau-2}} - C_4. \end{aligned} \tag{A.7}$$

Therefore,

$$g_n(x) \ge C_3 x^{\frac{\tau-3}{\tau-2}} - C_4 \quad \text{for} \quad 0 \le x < \sum_{j=1}^{n-1} \theta_{j,n}. \tag{A.8}$$

Now,

$$\sum_{i=K}^{n-1} \theta_{i,n} \mathrm{e}^{-\varepsilon \sum_{j=1}^{i} \theta_{j,n}^2} = \int_{\sum_{j=1}^{K-1} \theta_{j,n}}^{\sum_{j=1}^{n-1} \theta_{j,n}} \mathrm{e}^{-\varepsilon g_n(x)} \mathrm{d}x \le C_5 \int_{\sum_{j=1}^{K-1} \theta_{j,n}}^{\infty} \mathrm{e}^{-\varepsilon C_3 x^{\frac{\tau-3}{\tau-2}}} \mathrm{d}x, \tag{A.9}$$

and the above integral is finite for each fixed $K \ge 1$. By Assumption 1.1(i), $\sum_{j=1}^{K-1} \theta_{j,n} \to \sum_{j=1}^{K-1} \theta_j$ as $n \to \infty$, which diverges if we take $K \to \infty$. Thus, the proof of (A.1) follows.

We next prove Lemma A.1(ii). Let $K_n := \lceil n^{\alpha/2} \rceil$. Suppose that $\beta_{K_n}^n \le \log^3 n$. Using (A.8), it follows that

$$\log^3 n \ge \beta_{K_n}^n \ge C_3 \Big(\sum_{j=1}^{K_n} \theta_{j,n} \Big)^{\frac{\tau-3}{\tau-2}} - C_4, \tag{A.10}$$

and an application of (A.2) yields

$$C_4 + \log^3 n \ge C(\theta_{K_n+1,n})^{-(\tau-3)} \implies \theta_{K_n,n} \ge \frac{C'}{(\log n)^{\frac{3}{\tau-3}}}\,. \tag{A.11}$$

Therefore, $\sum_{i=1}^{K_n} \theta_{i,n}^3 \ge C'^3 K_n (\log n)^{-9/(\tau-3)}$. Thus, if $\beta_{K_n}^n \le \log^3 n$ for infinitely many n, then

$$\liminf_{n\to\infty} n^{-3\alpha} \sum_{i\in[n]} d_i^3 \ge \liminf_{n\to\infty} \sum_{i=1}^{K_n} \theta_{i,n}^3 = \infty. \tag{A.12}$$

On the other hand, Assumption 1.1(i) and (1.5) imply that $\sup_n n^{-3\alpha} \sum_{i\in[n]} d_i^3 < \infty$. This leads to a contradiction. Thus the claim in Lemma A.1 (ii) also follows. □

B Degree sequence satisfying compactness criterion

In this section, we prove Proposition 1.11.

Proof of Proposition 1.11. Define $\boldsymbol{d}^{(1,n)} := (d_i^{(1,n)})_{i\in[n]}$ with $d_i^{(1,n)} := \lceil n^\alpha \theta_i \rceil$ for $i \in [n]$. Let $\boldsymbol{d}^{(2,n)} = (d_i^{(2,n)})_{i\in[n]}$ be such that, for some $0 < K_1 < K_2 < \infty$,

$$K_1\Big(\frac{n}{i}\Big)^\alpha \le d_i^{(2,n)} \le K_2\Big(\frac{n}{i}\Big)^\alpha, \quad \text{for } i \in [n], \tag{B.1}$$

and Assumption 1.1(ii) and (1.11) are satisfied. The idea is to change the high-degree vertices of $\boldsymbol{d}^{(2,n)}$ by those of $\boldsymbol{d}^{(1,n)}$. To this end, let

$$i_{(1,n)} := \max\big\{i \ge 1\colon d_i^{(1,n)} \ge (\frac{n}{\log n})^\alpha\big\} \quad \text{and} \quad i_{(2,n)} := \max\big\{i \ge 1\colon d_i^{(2,n)} \ge (\frac{n}{\log n})^\alpha\big\}.$$

For two finite sequences (x_i) and (y_j), we write $\texttt{Sort-Merge}((x_i),(y_j))$ as the sequence obtained by concatenating (x_i) and (y_j) and then sorting the sequence in a nonincreasing order. We define

$$\boldsymbol{d}^{(n)} = (d_i^{(n)}) := \texttt{Sort-Merge}\Big((d_i^{(1,n)})_{i=1}^{i_{(1,n)}}, (d_i^{(2,n)})_{i=i_{(2,n)}+1}^{n}\Big). \tag{B.2}$$

Note that $i_{(1,n)} \to \infty$. Also,

$$\infty > \sum_{i=1}^{\infty}\theta_i^3 \ge \sum_{i=1}^{i_{(1,n)}}\theta_i^3 \ge i_{(1,n)}\theta_{i_{(1,n)}}^3 \ge i_{(1,n)}\Big(\frac{1}{2\log n}\Big)^{3\alpha}, \tag{B.3}$$

and therefore $i_{(1,n)} \le C(\log n)^{3\alpha}$. Further, it follows from (B.1) that $i_{(2,n)} \le K_2^{1/\alpha}(\log n)$. Therefore, the degree sequence in (B.2) has length $n(1+o(1))$.

Since $i_{(1,n)} \to \infty$, Assumption 1.1 (i) is satisfied by $(\boldsymbol{d}^{(n)})_{n\ge1}$. Also, for each fixed $K \ge 1$,

$$n^{-3\alpha}\sum_{i>K}(d_i^{(n)})^3 \le \sum_{i>K} 8\theta_i^3 + n^{-3\alpha}\sum_{i>K}(d_i^{(2,n)})^3, \tag{B.4}$$

and thus (1.5) holds. Next, it can be easily checked that the remaining conditions in Assumption 1.1(ii) and (1.11) hold for $(\boldsymbol{d}^{(n)})_{n\ge1}$ by making use of the fact that $(\boldsymbol{d}^{(2,n)})_{n\ge1}$ satisfies Assumption 1.1(ii) and (1.11).

Finally we have to verify that $(\boldsymbol{d}^{(n)})_{n\ge1}$ satisfies Assumption 1.3. It suffices to show that there exist $C > 1$ and $C' > 0$ such that for all $n \ge 1$,

$$\sum_i d_i^{(n)}\mathbb{1}\big\{l < d_i^{(n)} \le Cl\big\} \ge C'n/l^{\tau-2} \quad \text{for} \quad 1 \le l < d_1^{(n)}.$$

This can be proved in a straightforward way by using (1.18), (B.1), and the definition of $\boldsymbol{d}^{(n)}$ given in (B.2). We omit the details. □

References

[1] Louigi Addario-Berry, *Most trees are short and fat*, Probab. Theory Relat. Fields **173** (2019), no. 1-2, 1–26. MR3916103

[2] Louigi Addario-Berry, Nicolas Broutin, and Christina Goldschmidt, *The continuum limit of critical random graphs*, Probab. Theory Relat. Fields **152** (2012), no. 3, 367–406. MR2892951

[3] David Aldous, *The continuum random tree II: an overview*, Stochastic Analysis: Proc. Durham Symp. Stochastic Analysis, 1990 (M T Barlow and N H Bingham, eds.), London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, 1991, pp. 23–70. MR1166406

[4] David Aldous, Gregory Miermont, and Jim Pitman, *The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin's local time identity*, Probab. Theory Relat. Fields **129** (2004), no. 2, 182–218. MR2063375

[5] Siva Athreya, Wolfgang Löhr, and Anita Winter, *The gap between Gromov-vague and Gromov–Hausdorff-vague topology*, Stoch. Proc. Appl. **126** (2016), no. 9, 2527–2553. MR3522292

[6] Edward A Bender and E. Rodney Canfield, *The asymptotic number of labeled graphs with given degree sequences*, J. Combin. Theory Ser. A **24** (1978), no. 3, 296–307. MR0505796

[7] Shankar Bhamidi, Souvik Dhara, Remco van der Hofstad, and Sanchayan Sen, *Universality for critical heavy-tailed random graphs: Metric structure of maximal components*, Electron. J. Probab. **25** (2020), no. 47, 1–57. MR4092766

[8] Shankar Bhamidi, Remco van der Hofstad, and Sanchayan Sen, *The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs*, Probab. Theory Relat. Fields **170** (2018), no. 1, 387–474. MR3748328

[9] Patrick Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, Inc., 1999. MR1700749

[10] Arthur Blanc-Renaudie, *Compactness and fractal dimensions of inhomogeneous continuum random trees*, Probab. Theory Relat. Fields (2022), 1–31.

[11] Béla Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European J. Combin. **1** (1980), no. 4, 311–316. MR0595929

[12] Christian Borgs, Jennifer T. Chayes, Souvik Dhara, and Subhabrata Sen, *Limits of sparse configuration models and beyond: Graphexes and Multi-Graphexes*, Ann. Probab. **49** (2021), no. 6, 2830–2873. MR4348680

[13] Nicolas Broutin, Thomas Duquesne, and Minmin Wang, *Limits of multiplicative inhomogeneous random graphs and lévy trees: limit theorems*, Probab. Theory Relat. Fields **181** (2021), no. 4, 865–973. MR4344135

[14] Michael Camarri and Jim Pitman, *Limit distributions and random trees derived from the birthday problem with unequal probabilities*, Electron. J. Probab. **5** (2000), 1–18. MR1741774

[15] Guillaume Conchon-Kerjan and Christina Goldschmidt, *The stable graph: the metric space scaling limit of a critical random graph with i.i.d. power-law degrees*, arXiv:2002.04954 (2020). MR4043223

[16] Souvik Dhara, Remco van der Hofstad, Johan S. H. van Leeuwaarden, and Sanchayan Sen, *Heavy-tailed configuration models at criticality*, Ann. Inst. H. Poincaré (B) Probab. Statist. **56** (2020), no. 3, 1515–1558. MR4116701

[17] Devdatt Dubhashi, Volker Priebe, and Desh Ranjan, *Negative dependence through the FKG inequality*, Basic Research in Computer Science (BRICS) (1996).

[18] C. G Esseen, *On the concentration function of a sum of independent random variables*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **9** (1968), no. 4, 290–308. MR0231419

[19] Nikolaos Fountoulakis, *Percolation on sparse random graphs with given degree sequence*, Internet Math. **4** (2007), no. 1, 329–356. MR2522948

[20] David A Freedman, *On tail probabilities for martingales*, Ann. Probab. **3** (1975), no. 1, 100–118. MR0380971

[21] Christina Goldschmidt, Bénédicte Haas, and Delphin Sénizergues, *Stable graphs: distributions and line-breaking construction*, arXiv:1811.06940 (2018).

[22] Andreas Greven, Peter Pfaffelhuber, and Anita Winter, *Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees)*, Probab. Theory Relat. Fields **145** (2009), no. 1, 285–322. MR2520129

[23] Remco van der Hofstad, *Random Graphs and Complex Networks*, vol. 1, Cambridge University Press, Cambridge, 2017. MR3617364

[24] Svante Janson, *The probability that a random multigraph is simple*, Comb. Probab. Comp. **18** (2009), no. 1-2, 205–225. MR2497380

[25] Svante Janson, *Susceptibility of random graphs with given vertex degrees*, J. Combin. **1** (2010), no. 3-4, 357–387. MR2799217

[26] Svante Janson and Malwina J. Luczak, *A new approach to the giant component problem*, Random Struct. Algor. **34** (2009), no. 2, 197–216. MR2490288

[27] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski, *Random Graphs.*, Wiley, New York, 2000. MR1782847

[28] Tomasz Łuczak, *The number of trees with large diameter*, J. Aust. Math. Soc. Series A. Pure Mathematics and Statistics **58** (1995), no. 3, 298–311. MR1329866

[29] Grégory Miermont and Sanchayan Sen, *On breadth-first constructions of scaling limits of random graphs and random unicellular maps*, Random Struct. Algorithms.

[30] Michael Molloy and Bruce Reed, *A critical-point for random graphs with a given degree sequence*, Random Struct. Algor. **6** (1995), no. 2-3, 161–179. MR1370952

[31] Asaf Nachmias and Yuval Peres, *Critical random graphs: Diameter and mixing time*, Ann. Probab. **36** (2008), no. 4, 1267–1286. MR2435849

[32] G Szekeres, *Distribution of labelled trees by diameter*, Combinatorial Mathematics X (Berlin, Heidelberg) (Louis Reynolds Antoine Casse, ed.), Springer Berlin Heidelberg, 1983, pp. 392–397.

[33] Minmin Wang, *Height and diameter of Brownian tree*, Electron. Commun. Probab. **20** (2015), no. 88, 1–15. MR3434205

[34] Ward Whitt, *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*, Springer-Verlag, New York, 2002. MR1876437

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[2]EJMS: Electronic Journal Management System: `https://vtex.lt/services/ejms-peer-review/`
[3]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[4]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[5]Project Euclid: `https://projecteuclid.org/`
[6]IMS Open Access Fund: `https://imstat.org/shop/donation/`