Universality for critical heavy-tailed network models: Metric structure of maximal components[*]

Shankar Bhamidi[†]   Souvik Dhara[‡]   Remco van der Hofstad[§]

Sanchayan Sen[¶]

Abstract

We study limits of the largest connected components (viewed as metric spaces) obtained by critical percolation on uniformly chosen graphs and configuration models with heavy-tailed degrees. For rank-one inhomogeneous random graphs, such results were derived by Bhamidi, van der Hofstad, Sen (2018) [15]. We develop general principles under which the identical scaling limits as in [15] can be obtained. Of independent interest, we derive refined asymptotics for various susceptibility functions and the maximal diameter in the barely subcritical regime.

Contents

[†]Department of Statistics and Operations Research, University of North Carolina Chapel Hill, USA.
   E-mail: bhamidi@email.unc.edu
[‡]Department of Mathematics, Massachusetts Institute of Technology & Microsoft Research, USA.
   E-mail: sdhara@mit.edu
[§]Department of Mathematics and Computer Science, Eindhoven University of Technology, Netherlands.
   E-mail: r.w.v.d.hofstad@tue.nl
[¶]Department of Mathematics, Indian Institute of Science, India.
   E-mail: sanchayan.sen1@gmail.com

D Limit of exploration process: Proof sketch for Proposition 7.8   **57**

1 Introduction

Over the last decades, applications arising from complex systems in different fields have inspired a host of models for networks as well as models of dynamically evolving networks. One of the major themes in the study of these models has been in the nature of the emergence of the giant component. A classical example is the percolation process, where each edge of the network is independently kept with probability p, and deleted otherwise. As p increases from 0 to 1, the graph experiences a transition in the connectivity structure, i.e., there exists a "critical percolation value" p_c such that for any $\varepsilon > 0$ and $p < p_c(1 - \varepsilon)$, the proportion of vertices in the largest component is asymptotically negligible, while for $p > p_c(1 + \varepsilon)$, a unique giant component emerges containing an asymptotically positive proportion of vertices [8, 35, 43, 46, 51].

Understanding the behavior at criticality is one of the key questions in statistical physics because the components exhibit unique and key features in the critical regime. In the physics literature, the critical behavior of percolation relates to studying optimal paths in networks in the so-called strong disorder regime. A wide array of conjectures and heuristic deductions of the associated critical exponents can be found in [21, 22, 28, 41]. In a nutshell, these conjectures can be described as follows:

> *The intrinsic nature of the critical behavior does not depend on the exact description of the model, but only on moment conditions on the degree distribution. There are two major universality classes corresponding to the critical regime and the nature of emergence of the giant depending on whether the degree distribution has asymptotically finite third moment or infinite third moment. For example, in case of power-law degree distributions (i.e., $\mathbb{P}(D \geq x) \approx x^{-(\tau-1)}$ the precise nature of the approximation left implicit), the nature of the critical behavior depends only on the power-law degree exponent τ: (a) For $\tau > 4$, the maximal component sizes are of the order $n^{2/3}$ in the critical regime, whilst typical distances in these maximal connected components scale like $n^{1/3}$; (b) For $\tau \in (3,4)$, the maximal component sizes are of the order $n^{(\tau-2)/(\tau-1)}$, whilst distances scale like $n^{(\tau-3)/(\tau-1)}$.*

The above conjectures have inspired a large and beautiful collection of works in probability theory. In a seminal work, Aldous [3] provided a detailed understanding for the vector of rescaled component sizes at criticality for Erdős-Rényi random graphs, and the scaling limits for component sizes are now well understood under quite general setups in both finite third-moment [16, 32, 42, 52, 57, 58, 60] and infinite third-moment [17, 31, 42, 52] settings. We refer the reader to [30, Chapter 1], [44, Chapter 4] for detailed discussions about this topic. A recent and emerging direction in this literature aims at understanding the critical component structures, and distances within these components from a very general perspective. This line of work was pioneered by Addario-Berry, Broutin and Goldschmidt [1], where the largest connected components were shown to converge when viewed as metric spaces (see below for exact definitions). Subsequently, [10, 11, 13] have explored the *universality class* corresponding to [1], showing that the universality in the finite third-moment setting holds not only with respect to functionals like component sizes, but also the entire metric structure. On the other hand, in the infinite third-moment setting, a recent result [15] shows that the metric structure turns out to be fundamentally different. The results in [15] was obtained for one fundamental random graph model (rank-one model, closely related to the Chung-Lu [26, 27] and Norros-Reittu model [23]) under the assumption that the weights follow a power-law distribution. In

this paper, we explore the universality class corresponding to the candidate limit law established in [15]. Informally, the main contributions of this paper are as follows:

▷ **Universality theorem:** We establish sufficient conditions that imply convergence to the limits established in [15]. This is described later in Theorem 5.4. Since we need to set up a number of constructs, a formal statement is deferred until all of these objects have been defined. We refer to Theorem 5.4 as a *universality theorem* because it identifies the domain of attraction of the limit laws in [15]. Informally, the theorem implies that if a sequence of dynamic networks satisfies some entrance boundary conditions in the barely subcritical regime, and evolves approximately according to the multiplicative coalescent dynamics over the critical window, then the metric structure of the critical components are close to those for rank-one inhomogeneous random graphs. Theorem 5.4 is similar in spirit to [10, Theorem 3.4], but our result holds for the infinite third-moment degrees. Technically, we do not need additional restrictions as in [10, Assumption 3.3], since we compare the metric structures in the Gromov-weak topology, instead of the Gromov-Hausdorff-Prokhorov topology. The universality theorem holds under arguably optimal assumptions (see Remark 5.5).

▷ **Critical percolation on graphs with given degrees:** Our primary motivation was to analyze the critical regime for percolation on the uniform random graph model (and the closely associated configuration model) with a prescribed degree distribution that converges to a heavy-tailed degree distribution. Limit laws for the metric structure of maximal components in the critical regime are described in Theorems 2.3 and 2.4. These results are proved under Assumption 2.1, which is the most general set of assumptions under which the component sizes were shown to converge in [31] (see [31, Section 2 and 3] for the applicability and necessity of these assumptions).

▷ **Barely subcritical regime:** In order to carry out the above analysis and in particular to apply the universality theorem for percolation on configuration models, we establish refined bounds for component sizes, various susceptibility functionals, and diameters of connected components in the barely subcritical regime of the configuration model which are of independent interest; these are described in Theorems 2.10 and 2.11.

1.1 Organization of the paper

In Section 2, we describe the configuration model and critical behavior of percolation, which is the main motivation of this paper, and then describe the main results relevant to this model. Section 3 has a detailed discussion about the relevance of the results in this paper, some open problems, and an informal description of the proof ideas. We provide a full description of the limit objects and various notions of convergence of metric-space-valued random variables in Section 4. Section 5 describes and proves the general universality result. Section 6 proves results about the configuration model in the barely subcritical regime. Finally, Section 7 combines the above estimates with a coupling of the evolution of the configuration model through the critical percolation scaling window to finish the proof of Theorem 2.3.

2 Critical percolation on the configuration model

In this section, we state our main results. In Section 2.1, we state the results about the metric structure of the largest critical percolation clusters of the configuration model. We defer full definitions of the limit objects as well as notions of convergence of measured metric spaces to Section 4. In Section 2.2, we state the results about the barely subcritical regime, and we conclude this section with an overview of the proofs in

Section 2.3.

2.1 Metric structure of the critical components

The configuration model

Consider n vertices labeled by $[n] := \{1, 2, ..., n\}$ and 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 [18, 56]:

> 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 a 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. 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 $\mathrm{CM}_n(\boldsymbol{d})$, conditioned being simple, is same as $\mathrm{UM}_n(\boldsymbol{d})$ (see [43, Proposition 7.15]). It was further shown in [47] that, if the degree distribution satisfies a finite second-moment condition (a condition which will hold in the context of this paper), then the asymptotic probability of the graph being simple converges to a positive limit.

Let us now describe the assumptions on the degree sequences. For $p > 0$, define the metric space

$$\ell^p_{\downarrow} = \Big\{ (x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}}_+ : x_1 \geq x_2 \geq \dots, \sum_i x_i^p < \infty \Big\}, \tag{2.1}$$

with metric $d(\boldsymbol{x}, \boldsymbol{y}) = \big(\sum_i |x_i - y_i|^p \big)^{1/p}$. Fix $\tau \in (3, 4)$. Throughout this paper we use the following functionals of τ:

$$\alpha = 1/(\tau - 1), \qquad \rho = (\tau - 2)/(\tau - 1), \qquad \eta = (\tau - 3)/(\tau - 1). \tag{2.2}$$

Assumption 2.1 (Degree sequence). For each $n \geq 1$, let $\boldsymbol{d} = \boldsymbol{d}_n = (d_1, \dots, d_n)$ be a degree sequence (d_i's may depend on n, but we suppress n in the notation for clarity). We assume the following about $(\boldsymbol{d}_n)_{n \geq 1}$ as $n \to \infty$:

(i) (*High-degree vertices*) For each fixed $i \geq 1$, $n^{-\alpha} d_i \to \theta_i$, where $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots) \in \ell^3_{\downarrow} \setminus \ell^2_{\downarrow}$.

(ii) (*Moment assumptions*) Let V_n be chosen uniformly from $[n]$ (independently of $\mathrm{CM}_n(\boldsymbol{d})$), and $D_n = d_{V_n}$. Then D_n converges in distribution to some positive integer-valued random variable D, and

$$\frac{1}{n} \sum_{i \in [n]} d_i \to \mu := \mathbb{E}[D], \quad \frac{1}{n} \sum_{i \in [n]} d_i^2 \to \mu_2 := \mathbb{E}[D^2], \quad \lim_{K \to \infty} \limsup_{n \to \infty} n^{-3\alpha} \sum_{i=K+1}^{n} d_i^3 = 0. \tag{2.3}$$

Remark 2.2. Assumption 2.1 is identical to [31, Assumption 1]. We refer the reader to [31, Sections 2 and 3] for discussions about the relevance and necessity of these assumptions. It was shown in [31, Section 2] that Assumption 2.1 is satisfied in two key settings, when (i) the degrees are taken to be an i.i.d. sample from a power-law distribution, and (ii) the degrees are chosen according to the quantiles of a power-law distribution. The first setting has been considered in [52], and the latter setting has been considered for the rank-one inhomogeneous random graphs in [15, 17].

The component sizes of $\mathrm{CM}_n(\boldsymbol{d})$ are known to undergo a phase transition [50, 56] depending on the parameter

$$\nu_n = \frac{\sum_{i \in [n]} d_i(d_i - 1)}{\sum_{i \in [n]} d_i} \to \nu = \frac{\mathbb{E}\left[D(D-1)\right]}{\mathbb{E}\left[D\right]}. \tag{2.4}$$

When $\nu > 1$, $\mathrm{CM}_n(\boldsymbol{d})$ is supercritical in the sense that there exists a unique *giant* component with high probability, and when $\nu < 1$, all the components have size $o(n)$ with high probability and $\mathrm{CM}_n(\boldsymbol{d})$ is subcritical. In this paper, when considering percolation on $\mathrm{CM}_n(\boldsymbol{d})$, we will always assume that

$$\nu > 1, \text{ i.e. } \mathrm{CM}_n(\boldsymbol{d}) \text{ is supercritical.} \tag{2.5}$$

Percolation refers to deleting each edge of a graph independently with probability $1 - p$. In the case of percolation on random graphs, the deletion of edges is also independent from the underlying graph. Let $\mathrm{CM}_n(\boldsymbol{d}, p_n)$ and $\mathrm{UM}_n(\boldsymbol{d}, p_n)$ denote the graphs obtained from percolation with probability p_n on the graphs $\mathrm{CM}_n(\boldsymbol{d})$ and $\mathrm{UM}_n(\boldsymbol{d})$, respectively. For $p_n \to p$, it was shown in [46] that the critical point for the phase transition of the component sizes is $p = 1/\nu$. The critical window for percolation was studied in [31, 32] to obtain the asymptotics of the largest component sizes and their surplus edges. In the infinite third-moment setting, $\mathrm{CM}_n(\boldsymbol{d}, p_n)$ lies in the critical window when, for some $\lambda \in \mathbb{R}$,

$$p_n = p_n(\lambda) = \frac{1}{\nu_n} + \frac{\lambda}{n^\eta} + o(n^{-\eta}). \tag{2.6}$$

We now explain the precise meaning of convergence of components as metric spaces. Let $\mathscr{C}^p_{(i)}(\lambda)$ denote the i-th largest component of $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$. A measured metric space is a metric space equipped with a measure on the associated Borel sigma-algebra. Each component $\mathscr{C}$ can be viewed as a measured metric space with (i) the metric being the graph distance where each edge has length one; (ii) the measure being proportional to the counting measure, i.e., for any $A \subset \mathscr{C}$, the measure of A is given by $\mu_{\mathrm{ct},i}(A) = |A|/|\mathscr{C}^p_{(i)}(\lambda)|$, where $|A|$ denotes the cardinality of A. For a generic measured metric space $M = (M, \mathrm{d}, \mu)$ and $a > 0$, aM denotes the measured metric space $(M, a\mathrm{d}, \mu)$. We write $\mathscr{S}_*$ for the space of all measured metric spaces equipped with the Gromov-weak topology (see Section 4.1) and let $\mathscr{S}_*^{\mathbb{N}}$ denote the corresponding product space with the accompanying product topology. For each $n \geq 1$, view $\left(n^{-\eta}\mathscr{C}^p_{(i)}(\lambda)\right)_{i \geq 1}$ as an object in $\mathscr{S}_*^{\mathbb{N}}$ by appending an infinite sequence of empty metric spaces after enumerating the components in $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$. The main results for critical percolation on the configuration model are as follows:

Theorem 2.3. *Consider* $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$ *satisfying* Assumption 2.1, (2.5) *and* (2.6) *for some* $\lambda \in \mathbb{R}$. *There exists a sequence of random measured metric spaces* $(\mathscr{M}_i(\lambda))_{i \geq 1}$ *such that on* $\mathscr{S}_*^{\mathbb{N}}$, *as* $n \to \infty$,

$$\left(n^{-\eta}\mathscr{C}^p_{(i)}(\lambda)\right)_{i \geq 1} \xrightarrow{d} \left(\mathscr{M}_i(\lambda)\right)_{i \geq 1}. \tag{2.7}$$

Theorem 2.4. *Under* Assumption 2.1, (2.5) *and* (2.6) *for some* $\lambda \in \mathbb{R}$, *the convergence in* (2.7) *also holds for the components of* $\mathrm{UM}_n(\boldsymbol{d}, p_n(\lambda))$, *with the identical limiting object.*

Remark 2.5. The limiting objects are precisely described in Section 4.6.

Remark 2.6. The notion of convergence in Theorems 2.3 and 2.4 implies weak convergence of a wide array of continuous functionals with respect to the Gromov-weak topology. For example, it implies the joint convergence of the distances between an arbitrary (but fixed) number of uniformly (and independently) chosen vertices in the i-th largest component of $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$ or $\mathrm{UM}_n(\boldsymbol{d}, p_n(\lambda))$.

Remark 2.7. The conclusion of Theorem 2.3 holds if the measure $\mu_{ct,i}$ on $\mathscr{C}^p_{(i)}$ is replaced by more general measures. Indeed, define the probability measure $\mu_{w,i} := \sum_{k \in A} w_k / \sum_{k \in \mathscr{C}^p_{(i)}} w_k$ for $A \subset \mathscr{C}^p_{(i)}$. To prove analogous results as Theorems 2.3 and 2.4 with $\mu_{w,i}$'s, we require w_i's to satisfy some regularity conditions (see Assumption 2.9 below). The reason will be discussed in Remark 7.18.

Remark 2.8. The results above can be extended to the case $\mathbb{P}(D_n \geq x) \sim L(x)x^{-(\tau-1)}$, where $L(\cdot)$ is a slowly-varying function. The scaling limits would be the same, however the scaling exponents will be different as observed in [31]. In particular, the width of the scaling window now turns out to be $n^{-\eta}L_1(n)^2$ (for some slowly varying $L_1(\cdot)$) instead of $n^{-\eta}$, and results identical to Theorem 2.3 can be obtained by scaling the distances by $n^\eta L_1(n)^{-2}$.

2.2 Mesoscopic properties of the critical clusters: barely subcritical regime

One of the main ingredients in the proof of Theorem 2.3 is a refined analysis of various susceptibility functions in the *barely subcritical* regime (see (2.9) below for a definition) for the percolation process. The barely subcritical and supercritical regimes correspond to regimes that are just below or above the critical window. For the percolation process under Assumption 2.1, barely subcritical (supercritical) behavior is observed for p satisfying $n^\eta(p - p_n(0)) \to -\infty$ ($n^\eta(p - p_n(0)) \to \infty$), where $p_n(0)$ is defined in (2.6) for $\lambda = 0$. These behaviors are well understood for Erdős-Rényi random graphs [49, Section 23], [19, 51] and configuration models in the Erdős-Rényi universality class [40, 53, 60]. For barely supercritical configuration models in the heavy-tailed setting, the size of the emerging giant component was obtained in [45]. We provide a detailed picture about the component sizes and susceptibility functions in the subcritical regime below.

We will prove general statements about the susceptibility functions applicable not just to percolation on the configuration model, but rather to any barely subcritical configuration model. Since percolation on a configuration model yields a configuration model [35, 46], the above yields susceptibility functions for percolation on configuration model as a special case. To set this up we need a little more notation, where each vertex in the network is associated with both degree and weight, satisfying the following assumptions:

Assumption 2.9 (Barely subcritical degree sequence). Let $\boldsymbol{d'} = (d'_1, \ldots, d'_n)$ be a degree sequence and let $w_{(\cdot)} : [n] \mapsto \mathbb{R}$ be a non-negative weight function such that the following conditions hold:

(i) Assumption 2.1 holds for $\boldsymbol{d'}$ with some $\boldsymbol{c} \in \ell^3_\downarrow \setminus \ell^2_\downarrow$, and

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i \in [n]} d'_i = \mu_d, \quad \lim_{n \to \infty} \frac{1}{n} \sum_{i \in [n]} w_i = \mu_w, \quad \lim_{n \to \infty} \frac{1}{n} \sum_{i \in [n]} d'_i w_i = \mu_{d,w}. \quad (2.8)$$

(ii) $\max_{i \in [n]} w_i = O(n^\alpha)$, $\sum_{i \in [n]} w_i^2 = O(n)$ and
$\max\left\{ \sum_{i \in [n]} w_i^3, \sum_{i \in [n]} d'^2_i w_i, \sum_{i \in [n]} d'_i w_i^2 \right\} = O(n^{3\alpha})$.

(iii) (*Barely subcritical regime*) The configuration model is at the barely subcritical regime, i.e., there exists $0 < \delta < \eta$ and $\lambda_0 > 0$ such that

$$\nu'_n = \frac{\sum_{i \in [n]} d'_i(d'_i - 1)}{\sum_{i \in [n]} d'_i} = 1 - \lambda_0 n^{-\delta} + o(n^{-\delta}). \quad (2.9)$$

Let $\mathscr{C}'(j)$ denote the connected component of $\mathrm{CM}_n(\boldsymbol{d}')$ containing vertex j, and define

$$\mathscr{C}'_i = \begin{cases} \mathscr{C}'(i), & \text{if } i \leq j, \ \forall j \in \mathscr{C}'(i), \\ \varnothing, & \text{otherwise,} \end{cases} \tag{2.10}$$

and $\mathscr{W}_i = \sum_{k \in \mathscr{C}'_i} w_k$. Define the weight-based susceptibility functions as

$$s_r^\star = \frac{1}{n} \sum_{i \geq 1} \mathscr{W}_i^r \quad \text{for } r \geq 1, \qquad s_{pr}^\star = \frac{1}{n} \sum_{i \geq 1} \mathscr{W}_i \times |\mathscr{C}'_i|. \tag{2.11}$$

The definition in (2.10) takes care of the double counting in the definition of susceptibility functions. Also, define the weighted distance-based susceptibility as

$$\mathcal{D}_n^\star = \frac{1}{n} \sum_{i,j \in [n]} w_i w_j \mathrm{d}(i,j) \mathbb{1}\left\{i,j \text{ are in the same connected component}\right\}, \tag{2.12}$$

where d denotes the graph distance. The goal of the next result is to show that the component sizes and the susceptibility functions defined in (2.11) and (2.12) satisfy asymptotic conditions such as the entrance boundary conditions for the multiplicative coalescent [4]:

Theorem 2.10 (Susceptibility functions). *Under* Assumption 2.9, *as* $n \to \infty$,

$$n^{-\delta} s_2^\star \xrightarrow{\mathbb{P}} \frac{\mu_{d,w}^2}{\mu_d \lambda_0}, \quad n^{-\delta} s_{pr}^\star \xrightarrow{\mathbb{P}} \frac{\mu_{d,w}}{\lambda_0}, \quad n^{-(\alpha+\delta)} \mathscr{W}_j \xrightarrow{\mathbb{P}} \frac{\mu_{d,w}}{\mu_d \lambda_0} c_j,$$

$$n^{-3\alpha-3\delta+1} s_3^\star \xrightarrow{\mathbb{P}} \left(\frac{\mu_{d,w}}{\mu_d \lambda_0}\right)^3 \sum_{i=1}^\infty c_i^3, \quad n^{-2\delta} \mathcal{D}_n^\star \xrightarrow{\mathbb{P}} \frac{\mu_{d,w}^2}{\mu_d \lambda_0^2}. \tag{2.13}$$

For a connected graph G, $\Delta(G)$ denotes the diameter of the graph, and for any arbitrary graph G, $\Delta_{\max}(G) := \max \Delta(\mathscr{C})$, where the maximum is taken over all connected components $\mathscr{C} \subset G$. We simply write $\Delta_{\max}$ for $\Delta_{\max}(\mathrm{CM}_n(\boldsymbol{d}'))$. The asymptotics of $\Delta_{\max}$ is derived below:

Theorem 2.11 (Maximum diameter). *Under* Assumption 2.9, *as* $n \to \infty$, $\mathbb{P}(\Delta_{\max} > n^\delta(\log(n))^2) \to 0$.

Remark 2.12. By taking $w_i = 1$ for all $i \in [n]$, implies that $\mathscr{W}_i = |\mathscr{C}_i|$, and thus Theorem 2.10 hold also for the usual susceptibility functions defined in terms of the component sizes (cf. [48]). In the proof of Theorem 2.3, we will require a more general weight function, where w_i is taken to be the number of half-edges deleted from vertex i due to percolation.

Remark 2.13. Unlike Theorem 2.3, Theorems 2.10 and 2.11 yield statements about convergence in probability to constants. So, under Assumption 2.9, one can use the fact from [47] that $\liminf_{n \to \infty} \mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}') \text{ is simple}) > 0$, and thus it immediately follows that the results in Theorems 2.10 and 2.11 hold for $\mathrm{UM}_n(\boldsymbol{d}')$.

2.3 Overview of the proof

We now summarize the key ideas of the proofs at a heuristic level.

Universality theorem

As discussed earlier, we first prove a universality theorem (Theorem 5.4) which roughly states that if one replaces the vertices in a rank-one inhomogeneous random graph by

small metric spaces (called blobs), then the limiting metric space structure remains identical. The characterization of *blobs* leads to some asymptotic negligibility conditions, formally stated in Assumption 5.3, which simply says that the diameter of the individual blobs must be negligible compared to the typical distances in the whole graph. However, the typical distance can be cumulatively affected by the blobs, hence we get a different scaling factor for distances in Theorem 5.4 than in Theorem 5.2.

Mesoscopic or Blob-level analysis

Percolation on $\mathrm{CM}_n(\boldsymbol{d})$ can be viewed as a dynamic process by associating i.i.d. uniform $[0,1]$ weights U_e to each edge e, and keeping e if $U_e \leq p$. The parameter $p \in [0,1]$ can be interpreted as time. Now for $p_n = p_n(\lambda_n)$, for some $\lambda_n \to -\infty$, $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda_n))$ lies in the barely subcritical regime and the estimates for different functionals can be obtained using Theorem 2.10. We regard the components of $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda_n))$ as the blobs. Under the current scaling, the blobs shrink to zero, and the edges appearing in the dynamic process between the interval $[p_n(\lambda_n), p_n(\lambda)]$ connecting the blobs give rise to the macroscopic structure of the largest components of $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$. However, the effects of the blobs on the limiting structure are reflected via different functionals, which is the reason for referring to the properties of the blobs as *mesoscopic* properties.

Coupling to the multiplicative coalescent

Finally, the goal is to understand the macroscopic structure formed between blobs within the time interval $[p_n(\lambda_n), p_n(\lambda)]$. The merging dynamics of the components between $[p_n(\lambda_n), p_n(\lambda)]$ can be heuristically described as follows: Let p_0 be a time when an edge appears. Then the two half-edges corresponding to the new edge are chosen uniformly at random from the open half-edges (half-edges deleted due to percolation) of $\mathrm{CM}_n(\boldsymbol{d}, p_0-)$. Therefore, if $(\mathcal{O}_i(p))_{i \geq 1}$ denotes the vector of open half-edges in distinct components at time p, then the clusters corresponding to $\mathcal{O}_i(p)$ and $\mathcal{O}_j(p)$ merge at rate proportional to $\mathcal{O}_i(p) \times \mathcal{O}_j(p)$ and creates a new cluster with $\mathcal{O}_i(p) + \mathcal{O}_j(p) - 2$ open half-edges. Thus, the elements of the vector $(\mathcal{O}_i(p))_{i \geq 1}$, seen as masses, merge *approximately* as the multiplicative coalescent (see Definition 3), in the sense that the dynamics experience a depletion of half-edges in the components. Now, we can run a parallel process where the paired half-edges are replaced with new *dummy* open half-edges to the corresponding vertices [10, 31]. The dynamics in the latter process gives rise to an exact multiplicative coalescent and due to this fact, the modified graph $\bar{\mathcal{G}}_n$ can be shown to be distributed as a rank-one inhomogeneous random graph with the blobs being the mesoscopic components at time $p_n(\lambda_n)$. Now, the graph $\bar{\mathcal{G}}_n$ becomes the candidate for applying our universality theorem (see Theorem 7.15).

Structural comparison

Finally, we perform a structural comparison between $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$ and $\bar{\mathcal{G}}_n$ to conclude Theorem 2.3. Let us consider the largest component $\mathscr{C}^p_{(1)}$ (respectively $\bar{\mathscr{C}}^p_{(1)}$) of $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$ (respectively $\bar{\mathcal{G}}_n$). By the above coupling (with dummy half-edges being added), $\mathscr{C}^p_{(1)} \subset \bar{\mathscr{C}}^p_{(1)}$ and we know the asymptotic metric structure of $\bar{\mathscr{C}}^p_{(1)}$. Now, the idea is to show that (a) $|\bar{\mathscr{C}}^p_{(1)} \setminus \mathscr{C}^p_{(1)}| = o(|\mathscr{C}^p_{(1)}|)$ with high probability implying that the part of $\bar{\mathscr{C}}^p_{(1)}$ outside $\mathscr{C}^p_{(1)}$ is insignificant, (b) for any pair of vertices $u, v \in \mathscr{C}^p_{(1)}$, the shortest path between them in $\mathscr{C}^p_{(1)}$ and $\bar{\mathscr{C}}^p_{(1)}$ are identical. These two properties conclude the proof of Theorem 2.3 under the Gromov-weak topology.

3 Discussion

Optimality of assumptions and Gromov-weak topology

As mentioned in the introduction, our goal is not only to consider critical percolation, but to explore the universality class for the scaling limits in [15] in the same spirit as it was done in [10] for the Erdős-Rényi universality class. Our universality theorem (Theorem 5.4) holds under optimal assumptions, and does not require additional restrictions such as [10, Assumption 3.3]. However, it is worthwhile noting that the universality theorem (and consequently Theorem 2.3) holds with respect to the Gromov-weak topology instead of the stronger Gromov-Hausdorff-Prokhorov (GHP) topology. This is not a restriction that we impose, but in fact there is a conceptual barrier. If the convergence in Theorem 2.3 would hold in the GHP-topology only under Assumption 2.1, then the limiting metric space would be compact for any $\boldsymbol{\theta} \in \ell_\downarrow^3 \setminus \ell_\downarrow^2$, but additional restrictions are needed for the compactness of the limiting metric space and simply assuming $\boldsymbol{\theta} \in \ell_\downarrow^3 \setminus \ell_\downarrow^2$ does not suffice. See [5, Section 7], for an explicit conjecture about the compactness of such metric spaces by Aldous, Miermont and Pitman. In a follow-up work [12], we extend the scaling limit results in the GHP-topology by establishing the so-called global lower mass bound property [9, Theorem 6.1], which ensures that the components have sufficient mass everywhere and thus forbids the existence of long, thin paths, when the total mass of the component converges. However, one needs additional technical conditions in [12] on top of Assumption 2.1 to prove the global lower mass bound property.

Extensions, recent developments and open problems

(i) The universality theorem is applicable to dynamically evolving random networks with heavy-tailed degrees, which evolve (approximately) as the multiplicative coalescent over the critical window, and satisfy some nice properties such as Theorem 2.10 in the barely subcritical regime. For this reason, we believe that the universality theorem and the methods of this paper are applicable to many known inhomogeneous random graph models with suitable kernels [20], as well as Bohman-Frieze processes which satisfy different initial conditions so that one gets a heavy tailed-degree distribution at criticality. We leave these as interesting open problems.

(ii) In a recent work, Broutin, Duquesne and Wang [24] obtained structural limit laws for rank-one inhomogeneous random graphs which evolve as general multiplicative coalescent processes over the critical window. This framework unifies the scaling limits for the heavy-tailed and non heavy-tailed cases in terms of a single limit law. It will be interesting to prove a universality theorem for the limit laws in [24].

(iii) Recently, Conchon-Kerjan and Goldschmidt [29] derived the scaling limit of the maximal components at criticality for $\mathrm{CM}_n(\boldsymbol{d})$ when the degrees form an i.i.d. sample from a power-law distribution with $\tau \in (3, 4)$. The properties of the corresponding limiting object was investigated in a recent preprint by Goldschmidt, Haas, and Sénizergues [37]. The scaling limits in the i.i.d. setting has a completely different description of the limiting object compared to the one in this paper. It will be interesting to explore the connections between the results in the above paper and the current work.

(iv) It turns out that the study of the component structures corresponding to critical percolation plays a crucial role in the study of the metric structure of the minimal spanning tree (MST) [2]. In fact, a detailed understanding of the metric structures in the critical window obtained in [1] played a pivotal role in the proofs of [2]. The

connections to the MST-problem such as those outlined in [2] suggest that the scaling limit results in this paper will be useful in the study of metric structures for the MST for graphs with given degrees in the heavy-tailed regime. However, the MST problem in this regime is an open question.

4 Convergence of metric spaces, discrete structures and limit objects

The aim of this section is to define the proper notion of convergence relevant to this paper (Section 4.1), set up discrete structures required in the statement and in the proof of the universality result in Theorem 5.4 (Sections 4.2, 4.3, 4.4), and describe limit objects that arise in Theorem 2.3 (Sections 4.5 and 4.6).

4.1 Gromov-weak topology

A complete separable measured metric space (denoted by (X, d, μ)) is a complete, separable metric space (X, d) with an associated probability measure μ on the Borel sigma algebra $\mathcal{B}(X)$. The Gromov-weak topology is defined on $\mathscr{S}_0$, the space of all complete and separable measured metric spaces (see [38, 39], [15, Section 2.1.2]). The notion is formulated based on the philosophy of finite-dimensional convergence. Two measured metric spaces $(X_1, \mathrm{d}_1, \mu_1)$, $(X_2, \mathrm{d}_2, \mu_2)$ are considered to be equivalent if there exists an isometry $\psi : \mathrm{support}(\mu_1) \mapsto \mathrm{support}(\mu_2)$ such that $\mu_2 = \mu_1 \circ \psi^{-1}$. Let $\mathscr{S}_*$ be the space of all equivalence classes of $\mathscr{S}_0$. We (slightly) abuse the notation by not distinguishing between a metric space and its corresponding equivalence class. Fix $l \geq 2$ and $(X, \mathrm{d}, \mu) \in \mathscr{S}_*$. Given any collection of points $\mathbf{x} = (x_1, \ldots, x_l) \in X^l$, define $\mathrm{D}(\mathbf{x}) := (\mathrm{d}(x_i, x_j))_{i,j \in [l]}$ to be the matrix of pairwise distances of the points in $\mathbf{x}$. A function $\Phi : \mathscr{S}_* \mapsto \mathbb{R}$ is called a polynomial if there exists a bounded continuous function $\phi : \mathbb{R}^{l^2} \mapsto \mathbb{R}$ such that

$$\Phi((X, \mathrm{d}, \mu)) = \int \phi(\mathrm{D}(\mathbf{x})) d\mu^{\otimes l}, \tag{4.1}$$

where $\mu^{\otimes l}$ denotes the l-fold product measure. A sequence $\{(X_n, \mathrm{d}_n, \mu_n)\}_{n \geq 1} \subset \mathscr{S}_*$ is said to converge to $(X, \mathrm{d}, \mu) \in \mathscr{S}_*$ if and only if $\Phi((X_n, \mathrm{d}_n, \mu_n)) \to \Phi((X, \mathrm{d}, \mu))$ for all polynomials Φ on $\mathscr{S}_*$. By [38, Theorem 1], $\mathscr{S}_*$ is a Polish space under the Gromov-weak topology.

4.2 Super graphs

Our *super graphs* consist of three main ingredients: 1) A collection of metric spaces called *blobs*; 2) A graphical superstructure determining the connections between the blobs; 3) Connection points or junction points at each blob. In more detail, super graphs contain the following structures (see Figure 1):

(a) **Blobs**: A collection $\{(M_i, \mathrm{d}_i, \mu_i)\}_{i \in [m]}$ of connected, compact measured metric spaces.

(b) **Superstructure**: A (random) graph $\mathcal{G}$ with vertex set $[m]$. The graph has a weight sequence $\mathbf{p} = (p_i)_{i \in [m]}$ associated to the vertex set $[m]$. We regard M_i as the i-th vertex of $\mathcal{G}$.

(c) **Junction points**: An independent collection of random points $\mathbf{X} := (X_{i,j} : i, j \in [m])$ such that $X_{i,j} \sim \mu_i$ for all i, j. Further, $\mathbf{X}$ is independent of $\mathcal{G}$.

Using these three ingredients, define a metric space $(\bar{M}, \bar{\mathrm{d}}, \bar{\mu}) = \Gamma(\mathcal{G}, \mathbf{p}, \mathbf{M}, \mathbf{X})$, with $\bar{M} = \sqcup_{i \in [m]} M_i$ (the disjoint union of the M_i's) by putting an edge of *length* one between
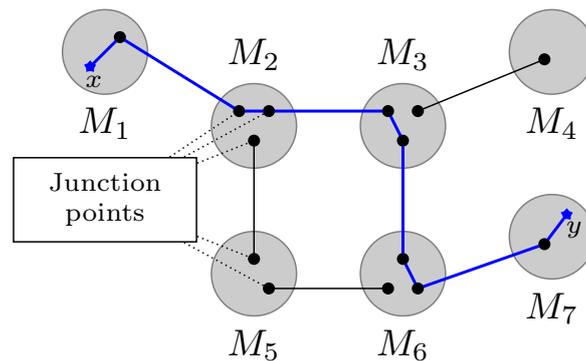
Figure 1: Construction of supergraphs described Section 4.2. The blue line represents the shortest path between x and y.

the pair of points $\{(X_{i,j}, X_{j,i}) : (i,j)$ is an edge of $\mathcal{G}\}$. The distance metric $\bar{\mathrm{d}}$ is the natural metric obtained from the graph distance and the inter-blob distance on a path. More precisely, for any $x, y \in \bar{M}$ with $x \in M_{j_1}$ and $y \in M_{j_2}$,

$$\bar{\mathrm{d}}(x,y) = \inf \left\{ k + \mathrm{d}_{j_1}(x, X_{j_1,i_1}) + \sum_{l=1}^{k-1} \mathrm{d}_{i_l}(X_{i_l,i_{l-1}}, X_{i_l,i_{l+1}}) + \mathrm{d}_{j_2}(X_{j_2,i_{k-1}}, y) \right\}, \qquad (4.2)$$

where the infimum is taken over all paths $(i_1, \ldots, i_{k-1})$ in $\mathcal{G}$ and all $k \geq 1$, and $i_0 = j_1$ and $i_k = j_2$. The measure $\bar{\mu}$ is given by $\bar{\mu}(A) := \sum_{i \in [m]} p_i \mu_i(A \cap M_i)$, for any measurable subset A of $\bar{M}$. Note that there is a one-to-one correspondence between the components of $\mathcal{G}$ and $\Gamma(\mathcal{G}, \mathbf{p}, \mathbf{M}, \mathbf{X})$ as the blobs are connected.

4.3 Space of trees with edge lengths, leaf weights, root-to-leaf measures, and blobs

In the proof of the main results we need the following spaces built on top of the space of discrete trees. The first space T_{IJ} was formulated in [6, 7] where it was used to study trees spanning a finite number of random points sampled from an inhomogeneous continuum random tree (as described in the next section).

4.3.1 The space T_{IJ}

Fix $I \geq 0$ and $J \geq 1$. Let T_{IJ} be the space of trees with each element $\mathbf{t} \in T_{IJ}$ having the following properties:

(a) There are exactly J leaves labeled $1+, \ldots, J+$, and the tree is rooted at the labeled vertex $0+$.

(b) There may be extra labeled vertices (called hubs) with labels in $\{1, \ldots, I\}$. (It is possible that only some, and not all, labels in $\{1, \ldots, I\}$ are used.)

(c) Every edge e has a strictly positive edge length l_e.

A tree $\mathbf{t} \in T_{IJ}$ can be viewed as being composed of two parts: (1) $\mathrm{shape}(\mathbf{t})$ describing the shape of the tree (including the labels of leaves and hubs) but ignoring edge lengths. The set of all possible shapes T_{IJ}^{shape} is obviously finite for fixed I, J. (2) The edge lengths $\mathbf{l}(\mathbf{t}) := (l_e : e \in \mathbf{t})$. We will consider the product topology on T_{IJ} consisting of the discrete topology on T_{IJ}^{shape} and the product topology on $\mathbb{R}^{\mathrm{E}(\mathbf{t})}$, where $\mathrm{E}(\mathbf{t})$ is the number of edges of $\mathbf{t}$.

4.3.2 The space T_{IJ}^*

Along with the three attributes above in T_{IJ}, the trees in T_{IJ}^* have two additional properties. Let $\mathcal{L}(\mathbf{t}) := \{1+, \ldots, J+\}$ denote the collection of leaves in $\mathbf{t}$. Then every leaf $v \in \mathcal{L}(\mathbf{t})$ has the following attributes:

(d) **Leaf weights:** A strictly positive number $A(v)$.

(e) **Root-to-leaf measures:** A probability measure $\nu_{\mathbf{t},v}$ on the path $[0+, v]$ connecting the root and the leaf v.

For each $v \in \mathcal{L}(\mathbf{t})$, the path $[0+, v]$ can be viewed as a compact measured metric space with the measure being $\nu_{\mathbf{t},v}$. Let $\mathcal{X}$ denote the space of compact measured metric spaces endowed with the Gromov-Hausdorff-Prokhorov topology (see [15, Section 2.1.1]). In addition to the topology on T_{IJ}, the space T_{IJ}^* with the additional two attributes inherits the product topology on $\mathbb{R}^J$ due to leaf weights and $\mathcal{X}^J$ due to the paths $[0+, v]$ endowed with $\nu_{\mathbf{t},v}$ for each $v \in \mathcal{L}(\mathbf{t})$. For consistency, we add a conventional state ∂ to the spaces T_{IJ} and T_{IJ}^*. Its use will be made clear in Section 5.

For all instances in this paper, the shape of a tree $\mathrm{shape}(\mathbf{t})$ will be viewed as a subgraph of a graph with m vertices. In that case, the tree will be assumed to inherit the vertex labels from the original graph. We will often write $\mathbf{t} \in T_{IJ}^{*m}$ to emphasize the fact that the vertices of $\mathbf{t}$ are labeled from a subset of $[m]$.

4.3.3 The space $\overline{T}_{IJ}^{*m}$

We enrich the space T_{IJ}^{*m} with some additional elements to accommodate the blobs. Consider $\mathbf{t} \in T_{IJ}^{*m}$ and construct $\bar{\mathbf{t}}$ as follows: Let $(M_i, \mathrm{d}_i, \mu_i)_{i \in [m]}$ be a collection of blobs and $\mathbf{X} = (X_{ij} : i, j \in [m])$ be the collection of junction points as defined in Section 4.2. Construct the metric space $\bar{\mathbf{t}}$ with elements in $\bar{M}(\mathbf{t}) = \sqcup_{i \in \mathbf{t}} M_i$, by putting an edge of *length* one between the pair of vertices $\{(X_{i,j}, X_{j,i}) : (i,j)$ is an edge of $\mathbf{t}\}$. The distance metric is given by (4.2). The path from the leaf v to the root $0+$ now contains blobs. Replace the root-to-leaf measure by $\bar{\nu}_{\mathbf{t},v}(A) := \sum_{i \in [0+,v]} \nu_{\mathbf{t},v}(i) \mu_i(M_i \cap A)$ for $A \subset \sqcup_{i \in [0+,v]} M_i$, where $\nu_{\mathbf{t},v}$ is the root-to-leaf measure on $[0+, v]$ for $\mathbf{t}$. Notice that T_{IJ}^{*m} can be viewed as a subset of $\overline{T}_{IJ}^{*m}$. In the proof of the universality theorem in Section 5, the blobs will be a fixed collection and, therefore, any $\mathbf{t} \in T_{IJ}^{*m}$ corresponds to a unique $\bar{\mathbf{t}} \in \overline{T}_{IJ}^{*m}$.

4.4 p-trees

For fixed $m \geq 1$, write $\mathbb{T}_m$ and $\mathbb{T}_m^{\mathrm{ord}}$ for the collection of all rooted trees with vertex set $[m]$ and rooted ordered trees with vertex set $[m]$ respectively. An ordered rooted tree is a rooted tree where children of each individual are assigned an order. We define a random tree model called **p-trees** [25, 59], and their corresponding limits, the so-called inhomogeneous continuum random trees, which play a key role in describing the limiting metric spaces. Fix $m \geq 1$, and a probability mass function $\mathbf{p} = (p_i)_{i \in [m]}$ with $p_i > 0$ for all $i \in [m]$. A **p-tree** is a random tree in $\mathbb{T}_m$, with law as follows: For any fixed $\mathbf{t} \in \mathbb{T}_m$ and $v \in \mathbf{t}$, write $d_v(\mathbf{t})$ for the number of children of v in the tree $\mathbf{t}$. Then the law of the **p**-tree, denoted by $\mathbb{P}_{\mathrm{tree}}$, is defined as

$$\mathbb{P}_{\mathrm{tree}}(\mathbf{t}) = \mathbb{P}_{\mathrm{tree}}(\mathbf{t}; \mathbf{p}) = \prod_{v \in [m]} p_v^{d_v(\mathbf{t})}, \quad \mathbf{t} \in \mathbb{T}_m. \tag{4.3}$$

Note that a normalizing constant is not required in (4.3) to make it a probability distribution (see [25, Lemma 1]). Generating a random **p**-tree $\mathscr{T} \sim \mathbb{P}_{\mathrm{tree}}$ and then assigning a

uniform random order on the children of every vertex $v \in \mathscr{T}$ gives a random element with law $\mathbb{P}_{\mathrm{ord}}(\cdot; \mathbf{p})$ given by

$$\mathbb{P}_{\mathrm{ord}}(\mathbf{t}) = \mathbb{P}_{\mathrm{ord}}(\mathbf{t}; \mathbf{p}) = \prod_{v \in [m]} \frac{p_v^{d_v(\mathbf{t})}}{(d_v(\mathbf{t}))!}, \quad \mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}. \tag{4.4}$$

4.4.1 The birthday construction of p-trees

We now describe a construction of **p**-trees, formulated in [25], that is relevant to this work. Let $\mathbf{Y} := (Y_0, Y_1, \ldots)$ be a sequence of i.i.d. random variables with distribution $\mathbf{p}$. Let $R_0 = 0$ and for $l \geq 1$, let R_l denote the l-th repeat time, i.e., $R_l = \min\{k > R_{l-1} : Y_k \in \{Y_0, \ldots, Y_{k-1}\}\}$. Now consider the directed graph formed via the edges $\mathcal{T}(\mathbf{Y}) := \{(Y_{j-1}, Y_j) : Y_j \notin \{Y_0, \ldots, Y_{j-1}\}, j \geq 1\}$. This gives a tree which we view as rooted at Y_0. The following striking result was shown in [25]:

Theorem 4.1 ([25, Lemma 1 and Theorem 2]). *The random tree $\mathcal{T}(\mathbf{Y})$, viewed as an element in $\mathbb{T}_m$, is distributed as a $\mathbf{p}$-tree with distribution (4.3) independently of $Y_{R_1-1}, Y_{R_2-1}, \ldots$ which are i.i.d. with distribution $\mathbf{p}$.*

Remark 4.1. The independence between the sequence $Y_{R_1-1}, Y_{R_2-1}, \ldots$ and the constructed $\mathbf{p}$-tree $\mathcal{T}(\mathbf{Y})$ is truly remarkable. In particular, let $\mathcal{T}_r \subset \mathcal{T}(\mathbf{Y})$ denote the subtree with vertex set $\{Y_0, Y_1, \ldots, Y_{R_r-1}\}$, i.e., the tree constructed in the first R_r steps. Further take $\tilde{\mathbf{Y}} = (\tilde{Y}_1, \ldots, \tilde{Y}_r)$ to be an i.i.d. sample from $\mathbf{p}$ and then construct the subtree $\mathcal{S}_r$ spanned by $\tilde{\mathbf{Y}}$. Then the above result (formalized as [25, Corollary 3]) implies that

$$(\tilde{Y}_1, \tilde{Y}_2, \ldots, \tilde{Y}_r; \mathcal{S}_r) \overset{d}{=} (Y_{R_1-1}, Y_{R_2-1}, \ldots, Y_{R_r-1}; \mathcal{T}_r). \tag{4.5}$$

We will use this fact in Section 5 to complete the proof of the universality theorem.

4.4.2 Tilted p-trees and connected components of $\mathrm{NR}_n(\boldsymbol{x}, t)$

Consider the vertex set $[n]$ and assign weight x_i to vertex i. Now, connect each pair of vertices i, j ($i \neq j$) independently with probability $q_{ij} := 1 - \exp(-t x_i x_j)$. The resulting random graph, denoted by $\mathrm{NR}_n(\boldsymbol{x}, t)$, is known as the Norros-Reittu model or the Poisson graph process [43]. For a connected component $\mathcal{C} \subseteq \mathrm{NR}_n(\boldsymbol{x}, t)$, let $\mathrm{mass}(\mathcal{C}) := \sum_{i \in \mathcal{C}} x_i$ and, for any $t \geq 0$, let $(\mathcal{C}_i(t))_{i \geq 1}$ denote the components in decreasing order of their mass sizes. In this section, we describe results from [14] that give a method of constructing connected components of $\mathrm{NR}_n(\boldsymbol{x}, t)$, conditionally on the vertices of the components. This construction involves tilted versions of $\mathbf{p}$-trees introduced in Section 4.4. Since these trees are parametrized via a driving probability mass function (pmf) $\mathbf{p}$, it will be easy to parametrize various random graph constructions in terms of pmfs as opposed to the vertex weights $\boldsymbol{x}$. Proposition 4.2 will relate vertex weights to pmfs.

Fix $n \geq 1$ and $\mathcal{V} \subset [n]$, and write $\mathbb{G}_{\mathcal{V}}^{\mathrm{con}}$ for the space of all simple connected graphs with vertex set $\mathcal{V}$. For fixed $a > 0$, and probability mass function $\mathbf{p} = (p_v)_{v \in \mathcal{V}}$, define probability distributions $\mathbb{P}_{\mathrm{con}}(\cdot; \mathbf{p}, a, \mathcal{V})$ on $\mathbb{G}_{\mathcal{V}}^{\mathrm{con}}$ as follows: For $i, j \in \mathcal{V}$, denote

$$q_{ij} := 1 - \exp(-a p_i p_j). \tag{4.6}$$

Then, for $G \in \mathbb{G}_{\mathcal{V}}^{\mathrm{con}}$,

$$\mathbb{P}_{\mathrm{con}}(G; \mathbf{p}, a, \mathcal{V}) := \frac{1}{Z(\mathbf{p}, a)} \prod_{(i,j) \in E(G)} q_{ij} \prod_{(i,j) \notin E(G)} (1 - q_{ij}), \tag{4.7}$$

where $Z(\mathbf{p}, a)$ is the normalizing constant. Now let $\mathcal{V}^{(i)}$ be the vertex set of $\mathcal{C}_i(t)$ for $i \geq 1$, and note that $(\mathcal{V}^{(i)})_{i \geq 1}$ denotes a random finite partition of the vertex set $[n]$. The next proposition yields a construction of the random (connected) graphs $(\mathcal{C}_i(t))_{i \geq 1}$:

Proposition 4.2 ([14, Proposition 6.1]). Given the partition $(\mathcal{V}^{(i)})_{i\geq 1}$, define, for $i \geq 1$,

$$\mathbf{p}_n^{(i)} := \left(\frac{x_v}{\sum_{v\in\mathcal{V}^{(i)}} x_v} : v \in \mathcal{V}^{(i)} \right), \quad a_n^{(i)} := t\left(\sum_{v\in\mathcal{V}_{(i)}} x_v \right)^2. \tag{4.8}$$

For each fixed $i \geq 1$, let $G_i \in \mathbb{G}_{\mathcal{V}^{(i)}}^{\mathrm{con}}$ be a connected simple graph with vertex set $\mathcal{V}^{(i)}$. Then

$$\mathbb{P}\left(\mathcal{C}_i(t) = G_i, \ \forall i \geq 1 \mid (\mathcal{V}^{(i)})_{i\geq 1}\right) = \prod_{i\geq 1} \mathbb{P}_{\mathrm{con}}(G_i; \mathbf{p}_n^{(i)}, a_n^{(i)}, \mathcal{V}^{(i)}). \tag{4.9}$$

Proposition 4.2 yields the following construction of $\mathrm{NR}_n(\boldsymbol{x}, t)$:

Algorithm 1. The random graph $\mathrm{NR}_n(\boldsymbol{x}, t)$ can be generated in two stages:

(S0) Generate the random partition $(\mathcal{V}^{(i)})_{i\geq 1}$ of the vertices into different components.

(S1) Conditionally on the partition, generate the internal structure of each component following the law of $\mathbb{P}_{\mathrm{con}}(\cdot; \mathbf{p}^{(i)}, a^{(i)}, \mathcal{V}^{(i)})$, independently across different components.

Let us now describe an algorithm to generate such connected components using distribution (4.7). To ease notation, let $\mathcal{V} = [m]$ for some $m \geq 1$ and fix a probability mass function $\mathbf{p}$ on $[m]$ and a constant $a > 0$ and write $\mathbb{P}_{\mathrm{con}}(\cdot) := \mathbb{P}_{\mathrm{con}}(\cdot; \mathbf{p}, a, [m])$ on $\mathbb{G}_m^{\mathrm{con}} := \mathbb{G}_{[m]}^{\mathrm{con}}$. To generate a sample G from $\mathbb{P}_{\mathrm{con}}$, one needs to first generate a $\mathbf{p}$-tree (with suitable tilt). The rest of the edges of G are *surplus edges*, which are generated by connecting the leaves to one of the vertices in their path to the root. Let us now describe this process formally. As a matter of convention, we view ordered rooted trees via their planar embedding using the associated ordering to determine the relative locations of siblings of an individual. We think of the left most sibling as the "oldest". Further, in a depth-first exploration, we explore the tree from left to right. Now given a planar rooted tree $\mathbf{t} \in \mathbb{T}_m$, let ρ denote the root, and for every vertex $v \in [m]$, let $[\rho, v]$ denote the path connecting ρ to v in the tree. Given this path and a vertex $i \in [\rho, v]$, write $\mathrm{RC}(i, [\rho, v])$ for the set of all children of i that fall to the right of $[\rho, v]$. Define $\mathfrak{P}(v, \mathbf{t}) := \cup_{i\in[m]}\mathrm{RC}(i, [\rho, v])$. In the terminology of [1, 15], $\mathfrak{P}(v, \mathbf{t})$ denotes the set of endpoints of all *permitted edges* emanating from v. Define

$$\mathfrak{G}_{(m)}(v) := \sum_{i\in[\rho,v]} \sum_{j\in[m]} p_j \mathbb{1}\left\{ j \in \mathrm{RC}(i, [\rho, v]) \right\}. \tag{4.10}$$

Let $(v(1), v(2), \ldots, v(m))$ denote the order of the vertices in the depth-first exploration of the tree $\mathbf{t}$. Let $y^*(0) = 0$ and $y^*(i) = y^*(i-1) + p_{v(i)}$ and define

$$A_{(m)}(u) = \mathfrak{G}_{(m)}(v(i)), \quad \text{for } u \in (y^*(i-1), y^*(i)], \quad \text{and} \quad \bar{A}_{(m)}(\cdot) := aA_{(m)}(\cdot), \tag{4.11}$$

where a is defined in (4.6). Define the function

$$\Lambda_{(m)}(\mathbf{t}) := a \sum_{v\in[m]} p_v \mathfrak{G}_{(m)}(v). \tag{4.12}$$

Finally, let $\mathrm{E}(\mathbf{t})$ denote the set of edges of $\mathbf{t}$, $\mathscr{T}_m^{\mathbf{P}}$ the $\mathbf{p}$-tree defined in (4.4), $\mathfrak{P}(\mathbf{t}) = \cup_{v\in[m]}\mathfrak{P}(v, \mathbf{t})$, and define the function $L : \mathbb{T}_m^{\mathrm{ord}} \to \mathbb{R}_+$ by

$$L(\mathbf{t}) = L_{(m)}(\mathbf{t}) := \prod_{(k,\ell)\in\mathrm{E}(\mathbf{t})} \left[\frac{\exp(ap_k p_\ell) - 1}{ap_k p_\ell} \right] \exp\left(\sum_{(k,\ell)\in\mathfrak{P}(\mathbf{t})} ap_k p_\ell \right), \tag{4.13}$$

for $\mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}$. Recall the (ordered) $\mathbf{p}$-tree distribution from (4.4). Let $\mathscr{T}_m^{\mathbf{P}}$ be a sample from $\mathbb{P}_{\mathrm{ord}}$. Using $L(\cdot)$ to tilt this distribution results in the distribution

$$\mathbb{P}_{\mathrm{ord}}^{\star}(\mathbf{t}) := \mathbb{P}_{\mathrm{ord}}(\mathbf{t}) \cdot \frac{L(\mathbf{t})}{\mathbb{E}_{\mathrm{ord}}[L(\mathscr{T}_m^{\mathbf{P}})]}, \qquad \mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}. \tag{4.14}$$

In the algorithm below, all the objects depend on the tree $\mathbf{t}$, but we often suppress this dependence to ease notation.

Algorithm 2. Let $\tilde{\mathcal{G}}_m(\mathbf{p}, a)$ denote a random graph sampled from $\mathbb{P}_{\mathrm{con}}(\cdot)$. This algorithm gives a construction of $\tilde{\mathcal{G}}_m(\mathbf{p}, a)$, proved in [15]:

(S1) **Tilted p-tree:** Generate a tilted ordered $\mathbf{p}$-tree $\mathscr{T}_m^{\mathbf{P},\star}$ with distribution (4.14). Now consider the (random) objects $\mathfrak{P}(v, \mathscr{T}_m^{\mathbf{P},\star})$ for $v \in [m]$ and the corresponding (random) functions $\mathfrak{G}_{(m)}(\cdot)$ on $[m]$ and $A_{(m)}(\cdot)$ on $[0, 1]$.

(S2) **Poisson number of possible surplus edges:** Let $\mathcal{P}$ denote a rate-one Poisson process on $\mathbb{R}_+^2$ and define

$$\bar{A}_{(m)} \cap \mathcal{P} := \left\{ (s, t) \in \mathcal{P} : s \in [0, 1], t \leq \bar{A}_{(m)}(s) \right\}. \tag{4.15}$$

Write $\bar{A}_{(m)} \cap \mathcal{P} := \{(s_j, t_j) : 1 \leq j \leq N_{(m)}^{\star}\}$ where $N_{(m)}^{\star} = |\bar{A}_{(m)} \cap \mathcal{P}|$. We next use the set $\{(s_j, t_j) : 1 \leq j \leq N_{(m)}^{\star}\}$ to generate pairs of points $\{(\mathcal{L}_j, \mathcal{R}_j) : 1 \leq j \leq N_{(m)}^{\star}\}$ in the tree that will be joined to form the surplus edges.

(S3) **First endpoints:** Fix j and suppose $s_j \in (y^*(i-1), y^*(i)]$ for some $i \geq 1$, where $y^*(i)$ is as given right above (4.11). Then the *first endpoint* of the surplus edge corresponding to (s_j, t_j) is $\mathcal{L}_j := v(i)$.

(S4) **Second endpoints:** Note that in the interval $(y^*(i-1), y^*(i)]$, the function $\bar{A}_{(m)}$ is of constant value $a\mathfrak{G}_{(m)}(v(i))$. We will view this value or height as being partitioned into sub-intervals of length ap_u for each $u \in \mathfrak{P}(v(i), \mathscr{T}_m^{\mathbf{P},\star})$, the collection of endpoints of permitted edges emanating from $\mathcal{L}_k$. (Assume that this partitioning is done according to some preassigned rule, e.g., using the order of the vertices in $\mathfrak{P}(v(i), \mathscr{T}_m^{\mathbf{P},\star})$). Suppose that t_j belongs to the interval corresponding to u. Then the *second endpoint* is $\mathcal{R}_j = u$. Form an edge between $(\mathcal{L}_j, \mathcal{R}_j)$.

(S5) In this construction, it is possible that one creates more than one surplus edge between two vertices. Remove any multiple surplus edges. This has vanishing probability in our applications.

Definition 1. Consider the connected random graph $\tilde{\mathcal{G}}_m(\mathbf{p}, a)$, given by Algorithm 2, viewed as a measured metric space via the graph distance and each vertex v is assigned measure p_v.

Lemma 4.2 ([15, Lemma 4.10]). *The random graph $\tilde{\mathcal{G}}_m(\mathbf{p}, a)$ generated by* Algorithm 2 *has the same law as $\mathbb{P}_{\mathrm{con}}(\cdot)$. Further, conditionally on $\mathscr{T}_m^{\mathbf{P},\star}$, the following hold:*

(a) $N_{(m)}^{\star}$ *has Poisson distribution with mean $\Lambda_{(m)}(\mathscr{T}_m^{\mathbf{P},\star})$, where $\Lambda_{(m)}$ is as in (4.12).*

(b) *Conditionally on $\mathscr{T}_m^{\mathbf{P},\star}$ and $N_{(m)}^{\star} = k$, the first endpoints $(\mathcal{L}_j)_{j \in [k]}$ can be generated in an i.i.d. fashion by sampling from the vertex set $[m]$ with probability distribution $\mathcal{J}^{(m)}(v) \propto p_v \mathfrak{G}_{(m)}(v), v \in [m]$.*

(c) *Conditionally on $\mathscr{T}_m^{\mathbf{P},\star}$, $N_{(m)}^{\star} = k$ and the first endpoints $(\mathcal{L}_j)_{j \in [k]}$, generate the second endpoints in an i.i.d. fashion where conditionally on $\mathcal{L}_j = v$, the probability distribution of $\mathcal{R}_j$ is given by*

$$Q_v^{(m)}(y) := \begin{cases} \sum_u p_u \mathbb{1}\left\{u \in \mathrm{RC}(y, [\rho, v])\right\} / \mathfrak{G}_{(m)}(v) & \text{if } y \in [\rho, v], \\ 0 & \text{otherwise.} \end{cases} \tag{4.16}$$

Create an edge between $\mathcal{L}_j$ and $\mathcal{R}_j$ for $1 \le j \le k$.

4.5 Inhomogeneous continuum random trees

In a series of papers [5, 6, 7] it was shown that **p**-trees, under various assumptions, converge to inhomogeneous continuum random trees (ICRTs) that we now describe. Recall from [34, 54] that a real tree is a metric space $(\mathscr{T}, d)$ that satisfies the following for every pair $a, b \in \mathscr{T}$:

(a) There is a **unique** isometric map $f_{a,b}: [0, d(a,b)] \to \mathscr{T}$ such that $f_{a,b}(0) = a$ and $f_{a,b}(d(a,b)) = b$.

(b) For any continuous one-to-one map $g : [0, 1] \to \mathscr{T}$ with $g(0) = a$ and $g(1) = b$, we have $g([0, 1]) = f_{a,b}([0, d(a,b)])$.

Construction of the ICRT: Given $\boldsymbol{\beta} \in \ell_{\downarrow}^2 \setminus \ell_{\downarrow}^1$ with $\sum_i \beta_i^2 = 1$, we will now define the inhomogeneous continuum random tree $\mathscr{T}^{\boldsymbol{\beta}}$. We mainly follow the notation in [7]. Assume that we are working on a probability space $(\Omega, \mathcal{F}, \mathbb{P}_{\boldsymbol{\beta}})$ rich enough to support the following:

(a) For each $i \ge 1$, let $\mathcal{P}_i := (\xi_{i,1}, \xi_{i,2}, \ldots)$ be rate β_i Poisson processes that are independent for different i. The first point of each process $\xi_{i,1}$ is special and is called a *joinpoint*, while the remaining points $\xi_{i,j}$ with $j \ge 2$ will be called *i-cutpoints* [7].

(b) Independently of the above, let $\boldsymbol{U} = (U_j^{(i)})_{i,j \ge 1}$ be a collection of i.i.d uniform $(0, 1)$ random variables. These are not required to construct the tree but will be used to define a certain function on the tree.

The **random** real tree (with marked vertices) $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$ is then constructed as follows:

(i) Arrange the cutpoints $\{\xi_{i,j} : i \ge 1, j \ge 2\}$ in increasing order as $0 < \eta_1 < \eta_2 < \cdots$. The assumption that $\sum_i \beta_i^2 < \infty$ implies that this is possible. For every cutpoint $\eta_k = \xi_{i,j}$, let $\eta_k^* := \xi_{i,1}$ be the corresponding joinpoint.

(ii) Next, build the tree inductively. Start with the branch $[0, \eta_1]$. Inductively assuming that we have completed step k, attach the branch $(\eta_k, \eta_{k+1}]$ to the joinpoint η_k^* corresponding to η_k.

Write $\mathscr{T}_0^{\boldsymbol{\beta}}$ for the corresponding tree after one has used up all the branches $[0, \eta_1]$, $\{(\eta_k, \eta_{k+1}] : k \ge 1\}$. Note that for every $i \ge 1$, the joinpoint $\xi_{i,1}$ corresponds to a vertex with infinite degree. Label this vertex i. The ICRT $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$ is the completion of the marked metric tree $\mathscr{T}_0^{\boldsymbol{\beta}}$. As argued in [7, Section 2], this is a real-tree as defined above which can be viewed as rooted at the vertex corresponding to zero. We call the vertex corresponding to joinpoint $\xi_{i,1}$ **hub** i.

The uniform random variables $(U_j^{(i)})_{i,j \ge 1}$ give rise to a natural ordering on $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$ (or a planar embedding of $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$) as follows: For $i \ge 1$, let $(\mathscr{T}_j^{(i)})_{j \ge 1}$ be the collection of subtrees hanging off the i-th hub. Associate $U_j^{(i)}$ with the subtree $\mathscr{T}_j^{(i)}$, and think of $\mathscr{T}_{j_1}^{(i)}$ appearing "to the right of" $\mathscr{T}_{j_2}^{(i)}$ if $U_{j_1}^{(i)} < U_{j_2}^{(i)}$. This is the natural ordering on $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$ when it is being viewed as a limit of ordered **p**-trees. We can think of the pair $(\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}, \boldsymbol{U})$ as the **ordered ICRT**.

4.6 Continuum limits of components

The aim of this section is to give an explicit description of the limiting (random) metric spaces in Theorem 2.3. We start by constructing a specific metric space using the tilted version of the ICRT in Section 4.6.1. Then we describe the limits of maximal components in Section 4.6.3.
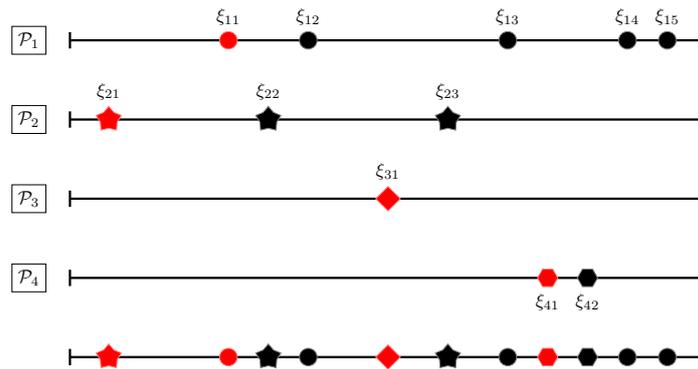
Figure 2: An illustration of the ICRT construction with four point process $\{\mathcal{P}_i : 1 \leq i \leq 4\}$. The red points represent the *joinpoint* of the corresponding point process and the blue points the corresponding cutpoints. The last line contains the union of the four point processes. See Figure 3 for the corresponding tree.
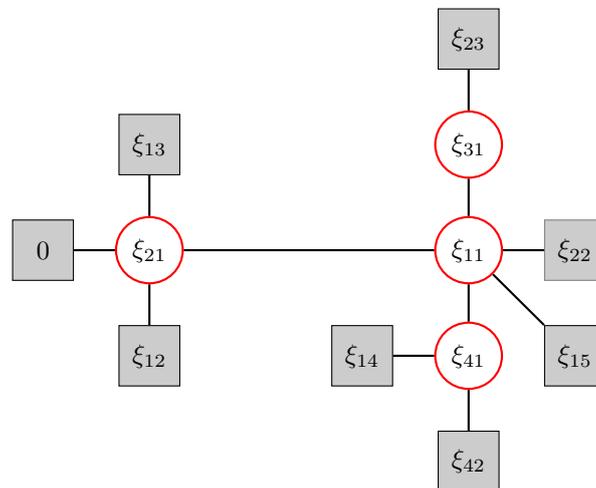
Figure 3: The tree constructed via the stick-breaking construction from Figure 2.

4.6.1   Tilted ICRTs and vertex identification

Let $(\Omega, \mathcal{F}, \mathbb{P}_\beta)$ and $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$ be as in Section 4.5. In [7], it was shown that one can associate a natural probability measure μ, called the **mass measure**, to $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$, satisfying $\mu(\mathcal{L}(\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}})) = 1$. Here we recall that $\mathcal{L}(\cdot)$ denotes the set of leaves. Before moving to the desired construction of the random metric space, we will need to define some more quantities that describe the asymptotic analogues of the quantities appearing in Algorithm 2. Similarly to (4.10), define $\mathfrak{G}_{(\infty)}(y) = \sum_{i \geq 1} \beta_i \big(\sum_{j \geq 1} U_j^{(i)} \times \mathbb{1}\{y \in \mathscr{T}_j^{(i)}\} \big)$. It was shown in [15] that $\mathfrak{G}_{(\infty)}(y)$ is finite for almost every realization of $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$ and for μ-almost every $y \in \mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$. For $y \in \mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$, let $[\rho, y]$ denote the path from the root ρ to y. For every y, define a probability measure on $[\rho, y]$ as

$$Q_y^{(\infty)}(v) := \frac{\beta_i U_j^{(i)}}{\mathfrak{G}_{(\infty)}(y)}, \quad \text{if } v \text{ is the } i\text{-th hub and } y \in \mathscr{T}_j^{(i)} \text{ for some } j. \tag{4.17}$$

Thus, this probability measure is concentrated on the hubs on the path from y to the root. Let $\gamma > 0$ be a constant. Informally, the construction goes as follows: We will first

tilt the distribution of the original ICRT $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}$ using the exponential functional

$$L_{(\infty)}(\mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}, \boldsymbol{U}) := \exp\left(\gamma \int_{y \in \mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}} \mathfrak{G}_{(\infty)}(y)\mu(dy)\right) \qquad (4.18)$$

to get a tilted tree $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta},\star}$. We then generate a random but finite number $N_{(\infty)}^{\star}$ of pairs of points $\{(x_k, y_k) : 1 \leq k \leq N_{(\infty)}^{\star}\}$ that will provide the surplus edges. The final metric space is obtained by creating *shortcuts* by identifying the points x_k and y_k. The construction mimics that of Algorithm 2. Formally the construction proceeds in four steps:

(a) **Tilted ICRT:** Define $\mathbb{P}_{\beta}^{\star}$ on Ω by

$$\frac{d\mathbb{P}_{\beta}^{\star}}{d\mathbb{P}_{\beta}} = \frac{\exp\left(\gamma \int_{y \in \mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}} \mathfrak{G}_{(\infty)}(y)\mu(dy)\right)}{\mathbb{E}\left[\exp\left(\gamma \int_{x \in \mathscr{T}_{(\infty)}^{\boldsymbol{\beta}}} \mathfrak{G}_{(\infty)}(x)\mu(dx)\right)\right]}. \qquad (4.19)$$

The expectation in the denominator is with respect to the original measure $\mathbb{P}_{\beta}$. Write $(\mathscr{T}_{(\infty)}^{\boldsymbol{\beta},\star}, \mu^{\star})$ and $\boldsymbol{U}^{\star} = (U_j^{(i),\star})_{i,j \geq 1}$ for the tree and the mass measure on it, and the associated random variables under this change of measure.

(b) **Poisson number of identification points:** Conditionally on $((\mathscr{T}_{(\infty)}^{\boldsymbol{\beta},\star}, \mu^{\star}), \boldsymbol{U}^{\star})$, generate $N_{(\infty)}^{\star}$ having a $\mathrm{Poisson}(\Lambda_{(\infty)}^{\star})$ distribution, where

$$\Lambda_{(\infty)}^{\star} := \gamma \int_{y \in \mathscr{T}_{(\infty)}^{\boldsymbol{\beta},\star}} \mathfrak{G}_{(\infty)}(y)\mu^{\star}(dy) = \gamma \sum_{i \geq 1} \beta_i \left[\sum_{j \geq 1} U_j^{(i),\star}\mu^{\star}(\mathscr{T}_j^{(i),\star})\right]. \qquad (4.20)$$

Here, $(\mathscr{T}_j^{(i),\star})_{j \geq 1}$ denotes the collection of subtrees of hub i in $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta},\star}$.

(c) **First endpoints (of shortcuts):** Conditionally on (a) and (b), sample x_k from $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta},\star}$ with density proportional to $\mathfrak{G}_{(\infty)}(x)\mu^{\star}(dx)$ for $1 \leq k \leq N_{(\infty)}^{\star}$.

(d) **Second endpoints (of shortcuts) and identification:** Having chosen x_k, choose y_k from the path $[\rho, x_k]$ joining the root ρ and x_k according to the probability measure $Q_{x_k}^{(\infty)}$ as in (4.17) but with $U_j^{(i),\star}$ replacing $U_j^{(i)}$. (Note that y_k is always a **hub** on $[\rho, x_k]$.) Identify x_k and y_k, i.e., form the quotient space by introducing the equivalence relation $x_k \sim y_k$ for $1 \leq k \leq N_{(\infty)}^{\star}$.

Definition 2. Fix $\gamma \geq 0$ and $\boldsymbol{\beta} \in \ell_{\downarrow}^2 \setminus \ell_{\downarrow}^1$ with $\sum_i \beta_i^2 = 1$. Let $\mathcal{G}_{(\infty)}(\boldsymbol{\beta}, \gamma)$ be the metric measure space constructed via the four steps above equipped with the measure inherited from the mass measure on $\mathscr{T}_{(\infty)}^{\boldsymbol{\beta},\star}$.

4.6.2 Scaling limit for the component sizes and surplus edges

Let us describe the scaling limit results for the component sizes and the surplus edges (#edges−#vertices+1) for the largest components of $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$ from [31]. Although we need to define the limiting object only for describing the limiting metric space, the convergence result will turn out to be crucial in Section 7 in the proof of Theorem 2.3, and therefore we state it here as well. Consider a decreasing sequence $\boldsymbol{\theta} \in \ell_{\downarrow}^3 \setminus \ell_{\downarrow}^2$. Denote by $\mathcal{I}_i(s) := \mathbb{1}\{\zeta_i \leq s\}$ where $\zeta_i \sim \mathrm{Exp}(\theta_i)$ independently, and $\mathrm{Exp}(r)$ denotes the exponential distribution with rate r. Consider the process

$$\bar{S}_{\infty}^{\lambda}(t) = \sum_{i=1}^{\infty} \theta_i \left(\mathcal{I}_i(t) - \theta_i t\right) + \lambda t, \qquad (4.21)$$

for some $\lambda \in \mathbb{R}$. Define the reflected version of $\bar{S}_\infty^\lambda(t)$ by $\text{refl}\big(\bar{S}_\infty^\lambda(t)\big) = \bar{S}_\infty^\lambda(t) - \inf_{0 \leq u \leq t} \bar{S}_\infty^\lambda(u)$. The processes of the form (4.21) were termed *thinned* Lévy processes in [17] since the summands are thinned versions of Poisson processes. Let $(\Xi_i(\boldsymbol{\theta}, \lambda))_{i \geq 1}$, $(\xi_i(\boldsymbol{\theta}, \lambda))_{i \geq 1}$, respectively, denote the vector of excursions and excursion-lengths of $\big(\text{refl}\big(\bar{S}_\infty^\lambda(t)\big)\big)_{t \geq 0}$, ordered according to the excursion lengths in a decreasing manner. Using [31, Fact 1], there are no ties among the excursion lengths almost surely. Denote the vector $(\xi_i(\boldsymbol{\theta}, \lambda))_{i \geq 1}$ by $\boldsymbol{\xi}(\boldsymbol{\theta}, \lambda)$. The fact that $\boldsymbol{\xi}(\boldsymbol{\theta}, \lambda)$ is always well defined follows from [4, Lemma 1]. Also, define the counting process of marks $\mathbf{N}$ to be a Poisson process that has intensity $\text{refl}\big(\bar{S}_\infty^\lambda(t)\big)$ at time t conditional on $(\text{refl}\big(\bar{S}_\infty^\lambda(u)\big))_{u \leq t}$. We use the notation $\mathscr{N}_i(\boldsymbol{\theta}, \lambda)$ to denote the number of marks within $\Xi_i(\boldsymbol{\theta}, \lambda)$.

For a connected graph G, let $\text{SP}(G) = \#\text{edges} - \#\text{vertices} + 1$ denote its surplus edges. In the context of this paper, we simply write ξ_i, $\boldsymbol{\xi}$ and $\mathscr{N}_i$ respectively for $\xi_i(\boldsymbol{\theta}/(\mu\nu), \lambda/\mu)$, $\boldsymbol{\xi}(\boldsymbol{\theta}/(\mu\nu), \lambda/\mu)$ and $\mathscr{N}_i(\boldsymbol{\theta}/(\mu\nu), \lambda/\mu)$.

Proposition 4.3 ([31, Theorem 4]). Under Assumption 2.1, as $n \to \infty$,

$$\big(n^{-\rho}|\mathscr{C}_{(i)}^p(\lambda)|, \text{SP}(\mathscr{C}_{(i)}^p(\lambda))\big)_{i \geq 1} \xrightarrow{d} \Big(\frac{1}{\nu}\xi_i, \mathscr{N}_i\Big)_{i \geq 1}, \tag{4.22}$$

with respect to the topology on the product space $\ell_\downarrow^2 \times \mathbb{N}^\mathbb{N}$.

The limiting object in [31, Theorem 4] is stated in a slightly different form compared to the right hand side of (4.22). However, the limiting objects are identical in distribution with suitable rescaling of time and space, and by observing that $r\text{Exp}(r) \stackrel{d}{=} \text{Exp}(1)$, where $\text{Exp}(r)$ denotes an exponential random variable with rate r (see Appendix A). In fact, the arguments in Appendix A establish the following lemma which will be used extensively in Section 7:

Lemma 4.3. For $\eta_1, \eta_2 > 0$, $\boldsymbol{\theta} \in \ell_\downarrow^3 \setminus \ell_\downarrow^2$ and $\lambda \in \mathbb{R}$, $\boldsymbol{\xi}(\eta_1\boldsymbol{\theta}, \eta_2\lambda) \stackrel{d}{=} \frac{1}{\eta_1}\boldsymbol{\xi}\big(\boldsymbol{\theta}, \frac{\eta_2}{\eta_1^2}\lambda\big)$.

4.6.3 Limiting component structures

We are now all set to describe the metric space M_i appearing in Theorem 2.3. Recall the graph $\mathcal{G}_\infty(\boldsymbol{\beta}, \gamma)$ from Definition 2. Using the notation of Section 4.6.2, write ξ_i^* for $\xi_i((\mu(\nu-1))^{-1}\boldsymbol{\theta}, (\mu(\nu-1)^2)^{-1}\nu^2\lambda)$ and Ξ_i^* for the excursion corresponding to ξ_i^*. Note that ξ_i^* has the same distribution as $(\nu-1)\xi_i/\nu$, where ξ_i is as in Proposition 4.3. Then the limiting space M_i is distributed as

$$M_i \stackrel{d}{=} \frac{\nu}{\nu-1} \frac{\xi_i^*}{\big(\sum_{v \in \Xi_i^*} \theta_v^2\big)^{1/2}} \mathcal{G}_\infty(\boldsymbol{\theta}^{(i)}, \gamma^{(i)}), \tag{4.23}$$

where $\boldsymbol{\theta}^{(i)} = \big(\frac{\theta_j}{\sum_{v \in \Xi_i^*} \theta_v^2} : j \in \Xi_i^*\big)$ and $\gamma^{(i)} = \frac{\xi_i^*}{\mu(\nu-1)}\big(\sum_{v \in \Xi_i^*} \theta_v^2\big)^{1/2}$.

5 Universality theorem

In this section, we develop universality principles that enable us to derive the scaling limits of the components for graphs that can be compared with the critical rank-one inhomogeneous random graph in a suitable sense. For the scaling limits in the basin of attraction of the Erdős-Rényi random graphs, such a universality theorem was proved in [10, Theorem 6.4], which was applied to deduce the scaling limits of the components for general inhomogeneous random graphs with a finite number of types and the configuration model with an exponential moment condition on the degrees. Here we focus on the universality class of the scaling limits in the heavy-tailed case. We first state the relevant result from [15] that was used in the context of rank-one inhomogeneous random graphs and then state our main result below. The convergence of metric spaces is with respect

to the Gromov-weak topology, unless stated otherwise. Recall the measured metric spaces $\tilde{\mathcal{G}}_m(\mathbf{p}, a)$ and $\mathcal{G}_\infty(\boldsymbol{\beta}, \gamma)$ defined in Definitions 1 and 2.

Assumption 5.1. (i) Let $\sigma(\mathbf{p}) := \left(\sum_i p_i^2\right)^{1/2}$. As $m \to \infty$, $\sigma(\mathbf{p}) \to 0$, and $p_i/\sigma(\mathbf{p}) \to \beta_i$ for each fixed $i \geq 1$, where $\boldsymbol{\beta} = (\beta_i)_{i \geq 1} \in \ell_\downarrow^2 \setminus \ell_\downarrow^1$, $\sum_i \beta_i^2 = 1$.

(ii) Recall a from (4.6). There exists a constant $\gamma > 0$ such that $a\sigma(\mathbf{p}) \to \gamma$.

Theorem 5.2 ([15, Theorem 4.5]). *Under Assumption 5.1, $\sigma(\mathbf{p})\tilde{\mathcal{G}}_m(\mathbf{p}, a) \xrightarrow{d} \mathcal{G}_{(\infty)}(\boldsymbol{\beta}, \gamma)$, as $m \to \infty$.*

For each $m \geq 1$, fix a collection of blobs $\mathbf{M}_m := \{(M_i, \mathrm{d}_i, \mu_i) : i \in [m]\}$. Recall the definition of super graphs from Section 4.2 and denote

$$\tilde{\mathcal{G}}_m^{\mathrm{bl}}(\mathbf{p}, a) = \Gamma(\tilde{\mathcal{G}}_m(\mathbf{p}, a), \mathbf{p}, \mathbf{M}_m, \mathbf{X}), \tag{5.1}$$

where $\mathbf{X} = (X_{ij})_{i,j \in [m]}$, $X_{ij} \sim \mu_i$ independently for each i. Moreover, $\mathbf{X}$ is independent of the graph $\tilde{\mathcal{G}}_m(\mathbf{p}, a)$. Let $u_i := \mathbb{E}[\mathrm{d}_i(X_i, X_i')]$ where $X_i, X_i' \sim \mu_i$ independently and $B_m := \sum_{i \in [m]} p_i u_i$. Let $\Delta_i := \mathrm{diam}(M_i)$ and $\Delta_{\max} := \max_{i \in [m]} \Delta_i$.

Assumption 5.3 (Maximum inter-blob-distance). $\lim_{m \to \infty} \frac{\sigma(\mathbf{p})\Delta_{\max}}{B_m + 1} = 0$.

Theorem 5.4 (Universality theorem). *Under Assumptions 5.1 and 5.3, as $m \to \infty$,*

$$\frac{\sigma(\mathbf{p})}{B_m + 1} \tilde{\mathcal{G}}_m^{\mathrm{bl}}(\mathbf{p}, a) \xrightarrow{d} \mathcal{G}_{(\infty)}(\boldsymbol{\beta}, \gamma). \tag{5.2}$$

Remark 5.5. Assumption 5.3 only assumes that the diameter of the blobs are negligible compared to the graph distances in $\tilde{\mathcal{G}}_m^{\mathrm{bl}}(\mathbf{p}, a)$. This, in a way, is a necessary condition to ensure that the inherent structure of the blobs does not affect the limit. Theorem 5.4 shows that only Assumption 5.3 is also sufficient and additional assumptions as in [10, Assumption 3.3] are not required to prove universality in the Gromov-weak topology.

The rest of this section is devoted to the proof of Theorem 5.4.

5.1 Completing the proof of the universality theorem in Theorem 5.4

To simplify notation, we write $\tilde{\mathcal{G}}_m$, $\tilde{\mathcal{G}}_m^{\mathrm{bl}}$, respectively, instead of $\tilde{\mathcal{G}}_m(\mathbf{p}, a)$ and $\tilde{\mathcal{G}}_m^{\mathrm{bl}}(\mathbf{p}, a)$.

Lemma 5.6 ([15, Lemma 4.11]). *Recall the definition of $N_{(m)}^\star$ from Algorithm 2. The sequence of random variables $(N_{(m)}^\star)_{m \geq 1}$ is tight.*

Recall the definition of Gromov-weak topology from Section 4.1. Fix some $l \geq 1$ and take any bounded continuous function $\phi : \mathbb{R}^{l^2} \mapsto \mathbb{R}$. We simply write $\Phi(X)$ for $\Phi((X, \mathrm{d}, \mu))$.

Key step 1 Let us write the scaled metric spaces as $\tilde{\mathcal{G}}_m^{\mathrm{s}} = \sigma(\mathbf{p})\tilde{\mathcal{G}}_m$ and $\tilde{\mathcal{G}}_m^{\mathrm{bl,s}} = \frac{\sigma(\mathbf{p})}{B_m + 1}\tilde{\mathcal{G}}_m^{\mathrm{bl}}$. Using Theorem 5.2 it is enough to show that

$$\lim_{m \to \infty} \left| \mathbb{E}\left[\Phi\left(\tilde{\mathcal{G}}_m^{\mathrm{bl,s}}\right)\right] - \mathbb{E}\left[\Phi\left(\tilde{\mathcal{G}}_m^{\mathrm{s}}\right)\right] \right| = 0. \tag{5.3}$$

The above step, together with Theorem 5.2, completes the proof of Theorem 5.4.

Key step 2 For any $K \geq 1$, the difference $\left| \mathbb{E}\left[\Phi\left(\tilde{\mathcal{G}}_m^{\mathrm{s}}\right)\right] - \sum_{k=0}^{K} \mathbb{E}\left[\Phi\left(\tilde{\mathcal{G}}_m^{\mathrm{s}}\right)\mathbb{1}\{N_{(m)}^\star = k\}\right] \right|$ is at most $\|\phi\|_\infty \mathbb{P}\left(N_{(m)}^\star \geq K + 1\right)$, and the same inequality also holds for $\tilde{\mathcal{G}}_m^{\mathrm{bl,s}}$. Thus, using Lemma 5.6, the proof of (5.3) reduces to showing that, for each fixed $k \geq 1$,

$$\lim_{m \to \infty} \left| \mathbb{E}\left[\Phi\left(\tilde{\mathcal{G}}_m^{\mathrm{bl,s}}\right)\mathbb{1}\{N_{(m)}^\star = k\}\right] - \mathbb{E}\left[\Phi\left(\tilde{\mathcal{G}}_m^{\mathrm{s}}\right)\mathbb{1}\{N_{(m)}^\star = k\}\right] \right| = 0. \tag{5.4}$$

Main aim of this section. Below, we define a function $g_\phi^k(\cdot)$ on the space $\overline{T}_{IJ}^*$ which captures the behavior of pairwise distances after creating k surplus edges. Under Assumption 5.3, we show that the introduction of blobs changes the distances within the tilted $\mathbf{p}$-trees and the g_ϕ^k values negligibly. This completes the proof of (5.4).

For any fixed $k \geq 0$, consider $\mathbf{t} \in T_{I,(k+l)}^*$ with root $0+$, leaves $\boldsymbol{i} = (1+, \dots, (k+l)+)$ and root-to-leaf measures $\nu_{\mathbf{t},i}$ on the path $[0+, i+]$ for all $1 \leq i \leq k+l$. We create a graph $G(\mathbf{t})$ by sampling, for each $1 \leq i \leq k$, points i_s on $[0+, i+]$ according $\nu_{\mathbf{t},i}$ and connecting $i+$ with i_s. Let $\mathrm{d}_{G(\mathbf{t})}$ denote the distance on $G(\mathbf{t})$ given by the sum of edge lengths in the shortest path. Then, the function $g_\phi^k : T_{I,(k+l)}^* \mapsto \mathbb{R}$ is defined as

$$g_\phi^k(\mathbf{t}) = \mathbb{E}\big[\phi\big(\mathrm{d}_{G(\mathbf{t})}(i+, j+) : k+1 \leq i, j \leq k+l\big)\big]\mathbb{1}\{\mathbf{t} \neq \partial\}, \qquad (5.5a)$$

where ∂ is a forbidden state defined as follows: Given any $\mathbf{t} \in T_{IJ}^*$, and a set of vertices $\boldsymbol{v} = (v_1, \dots, v_r)$, we denote by $\mathbf{t}(\boldsymbol{v})$, the subtree of $\mathbf{t}$ spanned by $\boldsymbol{v}$, i.e., $\mathbf{t}(\boldsymbol{v})$ is the subtree of $\mathbf{t}$ containing all vertices in $\boldsymbol{v}$ with minimal number of edges. We declare $\mathbf{t}(\boldsymbol{v}) = \partial$ if $[\rho, v_i] \subset [\rho, v_j]$ for some $i \neq j$, where ρ is the root of $\mathbf{t}$. Thus, if $\mathbf{t}(\boldsymbol{v}) \neq \partial$, the tree $\mathbf{t}(\boldsymbol{v})$ necessarily has r leaves. Notice that the expectation in (5.5a) is over the choices of i_s-values only. In our context, $\mathbf{t}$ is always considered as a subgraph of the graph on the vertex set $[m]$ and thus we assume that $\mathbf{t}$ has inherited the labels from the corresponding graph. Thus $\mathbf{t} \in T_{I,(k+l)}^{*m}$. There is a natural way to extend $g_\phi^k(\cdot)$ to $\overline{T}_{I,(k+l)}^{*m}$ as follows: Consider $\bar{\mathbf{t}} \in \overline{T}_{I,(k+l)}^{*m}$ and the corresponding $\mathbf{t} \in T_{I,(k+l)}^{*m}$ (see Section 4.3.3). Let $0+$, $\boldsymbol{i}$, $(\nu_{\mathbf{t},i})_{i \in [k+l]}$ and $(i_s)_{i \in [k+l]}$ be as defined above. Let $\bar{G}(\bar{\mathbf{t}})$ denote the metric space obtained by introducing an edge of length one between X_{i+i_s} and $X_{i_s i+}$, where X_{ij} has distribution μ_i for all $j \geq 1$, independently of each other and other shortcuts. For $k+1 \leq i \leq k+l$, $X_i \in M_{x_i}$ have distribution μ_{x_i} independently for all $i \geq 1$. Let $\bar{\mathrm{d}}_{\bar{G}(\bar{\mathbf{t}})}$ denote the distance on $\bar{G}(\bar{\mathbf{t}})$. Then, let

$$g_\phi^k(\bar{\mathbf{t}}) = \mathbb{E}\big[\phi\big(\bar{\mathrm{d}}_{\bar{G}(\bar{\mathbf{t}})}(X_i, X_j) : k+1 \leq i, j \leq k+l\big)\big]\mathbb{1}\{\mathbf{t} \neq \partial\}, \qquad (5.5b)$$

where the expectation is taken over the collection of random variables X_{i+i_s} and $X_{i_s i+}$. At this moment, we urge the reader to recall the construction in Algorithm 2, Lemma 4.2 and all the associated notations. Now, conditionally on $\mathscr{T}_m^{\mathbf{P},\star}$, we can construct the tree $\mathscr{T}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l})$, where

(a) $\tilde{\mathbf{V}}_m^{k,k+l} = (\tilde{V}_1, \dots, \tilde{V}_k, V_{k+1}, \dots, V_{k+l})$ is an independent collection of vertices from the vertex set of $\mathscr{T}_m^{\mathbf{P},\star}$;

(b) $\tilde{V}_i$ is distributed as $\mathcal{J}^{(m)}(\cdot)$, for $1 \leq i \leq k$ and V_i is distributed as $\mathbf{p}$, for $k+1 \leq i \leq k+l$.

Note that, by [15, (4.25)], $\lim_{m \to \infty} \mathbb{P}(\mathscr{T}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l}) = \partial) = 0$. Whenever $\mathscr{T}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l}) \neq \partial$, $\mathscr{T}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l})$ can be considered as an element of $T_{I,k+l}^{*m}$ using the leaf-weights $(\mathfrak{G}_{(m)}(\tilde{V}_i))_{i=1}^k$, $(\mathfrak{G}_{(m)}(V_i))_{i=k+1}^{k+l}$ and root-to-leaf measures given by $(Q_{\tilde{V}_i}^m(\cdot))_{i=1}^k$, $(Q_{V_i}^m(\cdot))_{i=k+1}^{k+l}$. Let $\bar{\mathscr{T}}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l})$ denote the element corresponding to $\mathscr{T}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l})$ with blobs. Thus, $\bar{\mathscr{T}}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l})$ is viewed as an element of $\overline{T}_{I,(k+l)}^{*m}$. Let $\mathbf{V}_m = (V_1, \dots, V_{k+l})$ be an i.i.d. collection of random variables with distribution $\mathbf{p}$. Let $\mathbb{E}_{\mathbf{p},\star}$ denote the expectation conditionally on $\mathscr{T}_m^{\mathbf{P},\star}$ and $N_{(m)}^\star$. The proof of (5.4) now reduces to

$$\left| \mathbb{E}\big[\Phi(\tilde{\mathcal{G}}_m^{\mathrm{bl,s}})\mathbb{1}\{N_{(m)}^\star = k\}\big] - \mathbb{E}\big[\Phi(\tilde{\mathcal{G}}_m^{\mathrm{s}})\mathbb{1}\{N_{(m)}^\star = k\}\big] \right|$$

$$= \left| \mathbb{E}\left[\mathbb{E}_{\mathbf{p},\star}\left[g_\phi^k\Big(\frac{\sigma(\mathbf{p})}{B_m+1} \bar{\mathscr{T}}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l})\Big)\right]\mathbb{1}\{N_{(m)}^\star = k\}\right] \right. \qquad (5.6)$$

$$\left. - \mathbb{E}\left[\mathbb{E}_{\mathbf{p},\star}\big[g_\phi^k\big(\sigma(\mathbf{p})\mathscr{T}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l})\big)\big]\mathbb{1}\{N_{(m)}^\star = k\}\right] \right| + o(1).$$

Notice that the tilting does not affect the blobs themselves but only the superstructure. Recall also the definition of the tilting function $L(\cdot)$ from (4.13). Using the fact that $\mathscr{J}^{(m)}(v) \propto p_v \mathfrak{G}_{(m)}(v)$,

$$\mathbb{E}_{\mathbf{p},\star}\big[\mathsf{g}_\phi^k\big(\sigma(\mathbf{p})\mathscr{T}_m^{\mathbf{P},\star}(\tilde{\mathbf{V}}_m^{k,k+l})\big)\big] = \frac{\mathbb{E}_{\mathbf{p},\star}\big[\prod_{i=1}^k \mathfrak{G}_{(m)}(V_i)\mathsf{g}_\phi^k\big(\sigma(\mathbf{p})\mathscr{T}_m^{\mathbf{P},\star}(\mathbf{V}_m)\big)\big]}{\big(\mathbb{E}_{\mathbf{p},\star}[\mathfrak{G}_{(m)}(V_1)]\big)^k}. \tag{5.7}$$

and an identical expression holds by replacing $\sigma(\mathbf{p})\mathscr{T}_m^{\mathbf{P},\star}$ by $\frac{\sigma(\mathbf{p})}{B_m+1}\bar{\mathscr{T}}_m^{\mathbf{P}}$. Denote the expectation conditionally on $\mathscr{T}_m^{\mathbf{P}}$ and $N_{(m)}$ by $\mathbb{E}_{\mathbf{p}}$ and simply write $\bar{\mathscr{T}}_m^{\mathbf{P},\mathrm{s}}$, $\mathscr{T}_m^{\mathbf{P},\mathrm{s}}$ for $\frac{\sigma(\mathbf{p})}{B_m+1}\bar{\mathscr{T}}_m^{\mathbf{P}}(\mathbf{V}_m)$, $\sigma(\mathbf{p})\mathscr{T}_m^{\mathbf{P}}(\mathbf{V}_m)$ respectively. Now, (5.6) simplifies to

$$\Big|\mathbb{E}\big[\Phi(\tilde{\mathcal{G}}_m^{\mathrm{bl,s}})\mathbb{1}\{N_{(m)}^\star = k\}\big] - \mathbb{E}\big[\Phi(\tilde{\mathcal{G}}_m^{\mathrm{s}})\mathbb{1}\{N_{(m)}^\star = k\}\big]\Big|$$
$$\leq \frac{1}{\mathbb{E}\left[L(\mathscr{T}_m^{\mathbf{P}})\right]}\left|\mathbb{E}\left[\frac{\mathbb{E}_{\mathbf{p}}\big[\prod_{i=1}^k \mathfrak{G}_{(m)}(V_i)\mathsf{g}_\phi^k\big(\bar{\mathscr{T}}_m^{\mathbf{P},\mathrm{s}}\big)\big]}{\big(\mathbb{E}_{\mathbf{p}}[\mathfrak{G}_{(m)}(V_1)]\big)^k}L(\mathscr{T}_m^{\mathbf{P}})\mathbb{1}\left\{N_{(m)} = k\right\}\right] \right. \tag{5.8}$$
$$\left. - \mathbb{E}\left[\frac{\mathbb{E}_{\mathbf{p}}\big[\prod_{i=1}^k \mathfrak{G}_{(m)}(V_i)\mathsf{g}_\phi^k\big(\mathscr{T}_m^{\mathbf{P},\mathrm{s}}\big)\big]}{\big(\mathbb{E}_{\mathbf{p}}[\mathfrak{G}_{(m)}(V_1)]\big)^k}L(\mathscr{T}_m^{\mathbf{P}})\mathbb{1}\left\{N_{(m)} = k\right\}\right]\right|.$$

Proposition 5.7. *As $m \to \infty$, $\big|\mathsf{g}_\phi^k\big(\bar{\mathscr{T}}_m^{\mathbf{P},\mathrm{s}}\big) - \mathsf{g}_\phi^k\big(\mathscr{T}_m^{\mathbf{P},\mathrm{s}}\big)\big| \xrightarrow{\mathbb{P}} 0$.*

We first show that it is enough to prove Proposition 5.7 to complete the proof of (5.8), but before that we first need to state some results. The proofs of Facts 5.9 and 5.10 below are elementary and we omit the proof here. The proof of Proposition 5.7 is deferred to Section 5.2.

Lemma 5.8 ([15, Proposition 4.8, Theorem 4.15])**.** *$(L(\mathscr{T}_m^{\mathbf{P}}))_{m\geq 1}$ is uniformly integrable. Also, for each $k \geq 0$, the quantity*

$$\left(\mathbb{E}_{\mathbf{p}}\left[\frac{\mathfrak{G}_{(m)}(V_1)}{\sigma(\mathbf{p})}\right], \ \mathbb{E}_{\mathbf{p}}\left[\left(\prod_{i=1}^k \frac{\mathfrak{G}_{(m)}(V_i)}{\sigma(\mathbf{p})}\right)\mathsf{g}_\phi^k\big(\mathscr{T}_m^{\mathbf{P},\mathrm{s}}\big)\right]\right) \tag{5.9}$$

converges in distribution to some random variable.

Fact 5.9. *Consider three sequences of random variables $(X_m)_{m\geq 1}$, $(Y_m)_{m\geq 1}$ and $(Y'_m)_{m\geq 1}$ such that (i) $(X_m)_{m\geq 1}$ is uniformly integrable, (ii) $(Y_m)_{m\geq 1}$ and $(Y'_m)_{m\geq 1}$ are almost surely bounded and (iii) $Y_m - Y'_m \xrightarrow{\mathbb{P}} 0$. Then, as $m \to \infty$, $\mathbb{E}\left[|X_m Y_m - X_m Y'_m|\right] \to 0$.*

Fact 5.10. *Suppose that $(X_m)_{m\geq 1}$ is a sequence of random variables such that for every $m \geq 1$, there exists a further sequence $(X_{m,r})_{r\geq 1}$ satisfying (i) for each fixed $r \geq 1$, $X_{m,r} \xrightarrow{\mathbb{P}} 0$ as $m \to \infty$, and (ii) $\lim_{r\to\infty}\limsup_{m\to\infty}\mathbb{P}(|X_m - X_{m,r}| > \varepsilon) = 0$ for any $\varepsilon > 0$. Then $X_m \xrightarrow{\mathbb{P}} 0$ as $m \to \infty$.*

Proof of (5.8) from Proposition 5.7. We apply Fact 5.9 with $X_m = L(\mathscr{T}_m^{\mathbf{P}})\mathbb{1}\{N_{(m)} = k\}$, which is uniformly integrable by Lemma 5.8. Thus it is enough to show that

$$\left|\frac{\mathbb{E}_{\mathbf{p}}\big[\prod_{i=1}^k \mathfrak{G}_{(m)}(V_i)\mathsf{g}_\phi^k\big(\bar{\mathscr{T}}_m^{\mathbf{P},\mathrm{s}}\big)\big]}{\big(\mathbb{E}_{\mathbf{p}}[\mathfrak{G}_{(m)}(V_1)]\big)^k} - \frac{\mathbb{E}_{\mathbf{p}}\big[\prod_{i=1}^k \mathfrak{G}_{(m)}(V_i)\mathsf{g}_\phi^k\big(\mathscr{T}_m^{\mathbf{P},\mathrm{s}}\big)\big]}{\big(\mathbb{E}_{\mathbf{p}}[\mathfrak{G}_{(m)}(V_1)]\big)^k}\right| \xrightarrow{\mathbb{P}} 0. \tag{5.10}$$

Applying Lemma 5.8 again, the above reduces to showing

$$\mathbb{E}_{\mathbf{p}}\left[\left(\prod_{i=1}^k \frac{\mathfrak{G}_{(m)}(V_i)}{\sigma(\mathbf{p})}\right)\big(\mathsf{g}_\phi^k\big(\bar{\mathscr{T}}_m^{\mathbf{P},\mathrm{s}}\big) - \mathsf{g}_\phi^k\big(\mathscr{T}_m^{\mathbf{P},\mathrm{s}}\big)\big)\right] \xrightarrow{\mathbb{P}} 0. \tag{5.11}$$

We now apply Fact 5.10. Let Y_m denote the term inside the expectation in (5.11). Further, sample the set of leaves $\mathbf{V}_m$ independently r times on the same tree $\mathscr{T}_m^{\mathbf{P}}$ and

let Y_m^i denote the observed value of Y_m in the i-th sample. Now, let $X_m = \mathbb{E}_\mathbf{p}[Y_m]$ and $X_{m,r} = r^{-1} \sum_{i=1}^r Y_m^i$. First, to verify condition (ii), note that $\mathbb{E}_\mathbf{p}(X_{m,r}) = X_m$ and therefore Chebyshev's inequality yields

$$\mathbb{P}\left(|X_m - X_{m,r}| > \varepsilon\right) \le \frac{\mathbb{E}[X_m^2]}{\varepsilon^2 r} \le \frac{4\|\phi\|_\infty^2}{\varepsilon^2 r} \frac{\mathbb{E}\left[\mathfrak{G}_{(m)}(V_1)^{2k}\right]}{\sigma(\mathbf{p})^{2k}}. \tag{5.12}$$

By (4.11), $\|\mathfrak{G}_{(m)}\|_\infty \le \|A_{(m)}\|_\infty$, and thus an application of [15, Lemma 4.9, (4.12)] yields that for any $x \ge \mathrm{e}$ and $m \ge 1$,

$$\mathbb{P}(\|\mathfrak{G}_{(m)}\|_\infty \ge x\sigma(\mathbf{p})) \le \mathrm{e}^{-Cx \log(\log x)}, \tag{5.13}$$

where $C > 0$ is a constant. Combining (5.12) and (5.13), the condition (ii) is verified. Next, condition (i) in Fact 5.10 is satisfied by Proposition 5.7 and (5.13). An application of Fact 5.10 concludes the proof of (5.11), and hence the proof of (5.8) follows. □

5.2 Comparing distances with and without blobs: Proof of Proposition 5.7

In this section, we will use the notion of Gromov-Hausdorff-Prokhorov topology on the collection of measured metric spaces (X, d, μ), where (X, d) is a compact metric space and μ is a probability measure on corresponding Borel sigma algebra. Without re-defining all the required notions, we refer the reader to [15, Section 2.1.1]. Let $\mathrm{d}_{\mathrm{GHP}}$ denote the distances in this topology. We further recall the notation dis for distortion and $D(\mu; \mu_1, \mu_2)$ for discrepancy of measures as defined in [15, Section 2.1.1]. Denote the root of $\mathscr{T}_m^\mathbf{P}(\mathbf{V}_m)$ by $0+$ and the jth leaf by $j+$. Let $\mathscr{M}_j^m := \left([0+, j+], \mathrm{d}, \nu_j\right)$ be the (random) measured metric space, where ν_j is any probability measure on the Borel sigma-algebra of $\{[0+, j+], \mathrm{d}\}$. In particular, we can take ν_j's to be the corresponding root-to-leaf measures. Let $\bar{\mathscr{M}}_j^m := \{\bar{M}_j, \bar{\mathrm{d}}, \bar{\nu}_j\}$ be the measured metric space with $\bar{M}_j := \sqcup_{i \in [0+, j+]} M_i$ and the induced root-to-leaf measure $\bar{\nu}_j(A) = \sum_{i \in [0+, j+]} \nu_j(\{i\}) \mu_i(A \cap M_i)$. For convenience, we have suppressed the dependence on $\mathscr{T}_m^\mathbf{P}(\mathbf{V}_m)$ in the notation. Note that $\mathscr{M}_j^m$ is coupled to $\bar{\mathscr{M}}_j^m$ in the obvious way that the superstructure of $\bar{\mathscr{M}}_j^m$ is given by $\mathscr{M}_j^m$. We need the following lemma to prove Proposition 5.7:

Lemma 5.11. *For $j \ge 1$, as $m \to \infty$, $\mathrm{d}_{\mathrm{GHP}}\left(\sigma(\mathbf{p})\mathscr{M}_j^m, \frac{\sigma(\mathbf{p})}{B_m+1}\bar{\mathscr{M}}_j^m\right) \xrightarrow{\mathbb{P}} 0$.*

Proof. We prove this for $j = 1$ only. The proof for $j \ge 2$ is identical. For $x \in \bar{M}_1$, we denote its corresponding vertex label by $i(x)$, i.e., $i(x) = k$ if and only if $x \in M_k$. Consider the correspondence C_m and the measure $\mathfrak{m}$ on the product space $[0+, 1+] \times \bar{M}_1$ defined as

$$C_m := \{(i, x) : i \in [0+, 1+], x \in M_i\}, \quad \mathfrak{m}(\{i\} \times A) = \nu_1(\{i\})\mu_i(A \cap M_i). \tag{5.14}$$

Note that the discrepancy of $\mathfrak{m}$ satisfies $D(\mathfrak{m}; \nu_1, \bar{\nu}_1) = 0$, since the marginals are exactly equal to ν_1 and $\bar{\nu}_1$. Further, $\mathfrak{m}(C_m^c) = 0$. Therefore, Lemma 5.11 follows if we can prove that

$$\mathrm{dis}(C_m) := \sup_{x,y \in \bar{M}_1} \left\{ \sigma(\mathbf{p})\mathrm{d}(i(x), i(y)) - \frac{\sigma(\mathbf{p})}{B_m + 1}\bar{\mathrm{d}}(x, y) \right\} \xrightarrow{\mathbb{P}} 0. \tag{5.15}$$

To simplify the expression for $\mathrm{dis}(C_m)$, suppose that $i(x)$ is an ancestor of $i(y)$ on the path from $0+$ to $1+$. Then,

$$\mathrm{d}(i(x), i(y)) = \mathrm{d}(0+, i(y)) - \mathrm{d}(0+, i(x)),$$
$$\bar{\mathrm{d}}(x_0, y) - \bar{\mathrm{d}}(x_0, x) \le \bar{\mathrm{d}}(x, y) \le \bar{\mathrm{d}}(x_0, y) - \bar{\mathrm{d}}(x_0, x) + 2\Delta_{\max},$$

for any $x_0 \in M_{0+}$. This implies that

$$
\begin{aligned}
\sup_{x,y \in \bar{M}_1} & \left\{ \sigma(\mathbf{p}) \mathrm{d}(i(x), i(y)) - \frac{\sigma(\mathbf{p})}{B_m + 1} \bar{\mathrm{d}}(x,y) \right\} \\
& \leq 2 \sup_{y \in \bar{M}_1} \left\{ \sigma(\mathbf{p}) \mathrm{d}(0+, i(y)) - \frac{\sigma(\mathbf{p})}{B_m + 1} \bar{\mathrm{d}}(x_0, y) \right\} + \frac{2\sigma(\mathbf{p}) \Delta_{\max}}{B_m + 1}.
\end{aligned}
\tag{5.16}
$$

Further, replacing y by any other point y' in the right hand side in (5.16) incurs an error of at most $\sigma(\mathbf{p}) \Delta_{\max}/(B_m + 1)$. Now, write the path $[0+, 1+]$ as $0+ = i_0 \to i_1 \to \cdots \to i_{R^*-2} \to i_{R^*-1} = 1+$. Then

$$
\mathrm{dis}(C_m) \leq 2 \sup_{k \leq R^*-1} \left| \sigma(\mathbf{p}) \mathrm{d}(i_0, i_k) - \frac{\sigma(\mathbf{p})}{B_m + 1} \bar{\mathrm{d}}(X_{i_0, i_1}, X_{i_k, i_{k+1}}) \right| + \frac{6\sigma(\mathbf{p}) \Delta_{\max}}{B_m + 1},
\tag{5.17}
$$

where $(X_{i,j})_{i,j \in [m]}$ are the junction-points. Using Assumption 5.3 and (5.17), it is now enough to show that for any $\varepsilon > 0$,

$$
\lim_{m \to \infty} \mathbb{P} \left(\sup_{k \leq R^*-1} \left| \sigma(\mathbf{p}) \mathrm{d}(i_0, i_k) - \frac{\sigma(\mathbf{p})}{B_m + 1} \bar{\mathrm{d}}(X_{i_0, i_1}, X_{i_k, i_{k+1}}) \right| > \varepsilon \right) = 0.
\tag{5.18}
$$

Denote the term inside $\sup$ above by Q_k. Then,

$$
\begin{aligned}
Q_k & := \left[\sigma(\mathbf{p}) \mathrm{d}(i_0, i_k) - \frac{\sigma(\mathbf{p})}{B_m + 1} \bar{\mathrm{d}} \big(X_{i_0, i_1}, X_{i_k, i_{k+1}} \big) \right] \\
& = \left[\sigma(\mathbf{p}) k - \frac{\sigma(\mathbf{p})}{B_m + 1} \left(k + \sum_{j=1}^{k} \mathrm{d}_{i_j} \big(X_{i_j, i_{j-1}}, X_{i_j, i_{j+1}} \big) \right) \right] \\
& = \frac{\sigma(\mathbf{p})}{B_m + 1} \left[\sum_{j=1}^{k} \Big(B_m - \mathrm{d}_{i_j} \big(X_{i_j, i_{j-1}}, X_{i_j, i_{j+1}} \big) \Big) \right].
\end{aligned}
\tag{5.19}
$$

Recall the construction of the path $[0+, 1+]$ via the birthday problem from Section 4.4.1. Take $\mathbf{J} := (J_i)_{i \geq 1}$ such that J_i are an i.i.d. sample from $\mathbf{p}$. Further let $\boldsymbol{\xi} := (\xi_i)_{i \in [m]}$ be an independent sequence such that ξ_i is the distance between two points, chosen randomly from M_i according to μ_i. Further, let $\mathbf{J}$ and $\boldsymbol{\xi}$ be independent. Then R^* can be thought of as the first repeat time of the sequence $\mathbf{J}$. Thus, $(Q_k)_{k=1}^{R^*-1}$ in (5.19) has the same distribution as $(\hat{Q}_k)_{k=1}^{R^*-1}$, where

$$
\hat{Q}_k := \frac{\sigma(\mathbf{p})}{B_m + 1} \sum_{i=1}^{k} \big(B_m - \xi_{J_i} \big).
\tag{5.20}
$$

From the birthday construction $\mathbb{E}\left[\xi_{J_1}\right] = \sum_{i \in [m]} p_i u_i = B_m$ and $(\xi_{J_i})_{i \geq 1}$ is an independent sequence. Therefore, $(\hat{Q}_k)_{k \geq 0}$ is a martingale. Further,

$$
\mathrm{Var}(\hat{Q}_k) \leq \left(\frac{\sigma(\mathbf{p})}{B_m + 1} \right)^2 k \Delta_{\max} \sum_{i \in [m]} p_i u_i = \frac{\sigma(\mathbf{p})^2 k \Delta_{\max} B_m}{(B_m + 1)^2}.
\tag{5.21}
$$

Thus, by Doob's inequality (for example applying [61, Chapter II, Lemma 54.5] to $\hat{Q}_k^2$), it follows that, for any $\varepsilon > 0$ and $T > 0$,

$$
\mathbb{P} \left(\sup_{k \leq T} |\hat{Q}_k| > \varepsilon \right) \leq \frac{T \sigma(\mathbf{p})^2 \Delta_{\max} B_m}{(B_m + 1)^2 \varepsilon^2}.
\tag{5.22}
$$

Recall from [25, Theorem 4] that $(\sigma(\mathbf{p}) R^*)_{m \geq 1}$ is a tight sequence of random variables. The proof now follows using Assumption 5.3. $\qquad \square$

Proof of Proposition 5.7 using Lemma 5.11. We use the objects defined in (5.14), (5.15) in the proof of Lemma 5.11 for all the path metric spaces with $j \leq k$. We assume that we are working on a probability space such that the convergence (5.15) holds almost surely for all $j \leq k$. To summarize, for fixed $\varepsilon > 0$ and for each $j \leq k$, we can choose a correspondence C_m^j and a measure $\mathfrak{m}_j$ of $[0+, j+] \times \bar{M}_j$ satisfying (i) $(i, X_{ik}) \in C_m^j$, for all $i, k \in [0+, j+]$, (ii) $\mathrm{dis}(C_m^j) < \varepsilon/2k$ almost surely, and (iii) $D(\mathfrak{m}_j; \nu_j, \bar{\nu}_j) = 0$ and $\mathfrak{m}_j((C_m^j)^c) = 0$. Recall the definitions of the function g_ϕ^k from (5.5a), (5.5b) and the associated graphs $G(\cdot)$, $\bar{G}(\cdot)$. We simply write G and $\bar{G}$ for $G(\sigma(\mathbf{p})\mathscr{T}_m^\mathbf{P}(\mathbf{V}_m))$ and $\bar{G}(\frac{\sigma(\mathbf{p})}{B_m+1}\bar{\mathscr{T}}_m^\mathbf{P}(\mathbf{V}_m))$, respectively. Let $\mathfrak{m}^{\otimes k}$ denote the k-fold product measure of $\mathfrak{m}_j$ for $j \leq k$. We denote the graph distance on a graph H by d_H. Note that

$$\left| \mathrm{g}_\phi^k\big(\sigma(\mathbf{p})\mathscr{T}_m^\mathbf{P}(\mathbf{V}_m)\big) - \mathrm{g}_\phi^k\Big(\frac{\sigma(\mathbf{p})}{B_m+1}\bar{\mathscr{T}}_m^\mathbf{P}(\mathbf{V}_m)\Big) \right|$$
$$\leq \mathbb{E}\big[\big|\phi\big((\mathrm{d}_G(i+, j+))_{i,j=k+1}^{k+l}\big) - \phi\big((\mathrm{d}_{\bar{G}}(X_i, X_j))_{i,j=k+1}^{k+l}\big)\big|\big], \tag{5.23}$$

where $X_i \sim \mu_i$ independently for $i \in [m]$, and the above expectation is with respect to the measure $\mathfrak{m}^{\otimes k}$. Recall the notation while defining $\mathrm{g}_\phi^k(\cdot)$ in (5.5a), (5.5b). Notice that for any point $k \in [0+, i+]$ and $x_k \in M_k$ and $x_{i_s} \in M_{i_s}$,

$$|\mathrm{d}_\mathfrak{t}(k, i_s) - \mathrm{d}_{\bar{\mathfrak{t}}}(x_k, x_{i_s})| \leq \frac{\varepsilon}{2k}. \tag{5.24}$$

Now, for any path $i+$ to $j+$ in G, we can essentially take the same path from X_i to X_j in $\bar{G}$ and take the corresponding inter-blob paths on the way. The distance traversed in $\bar{G}$ in this way gives an upper bound on $\mathrm{d}_{\bar{G}}(X_i, X_j)$. Notice that, by (5.24), taking a shortcut contributes at most $\varepsilon/2k$ to the difference of the distance traveled in G and $\bar{G}$. Also, traversing a shortcut edge contributes $\sigma(\mathbf{p})B_m/(B_m+1)$ and there are at most k shortcuts on the path. Furthermore, it may be required to reach the relevant junction points from X_i and X_j and that contributes at most $2\sigma(\mathbf{p})\Delta_{\max}/(B_m+1)$. Thus, for $k+1 \leq i, j \leq k+l$, and sufficiently large m,

$$\mathrm{d}_{\bar{G}}(X_i, X_j) \leq \mathrm{d}_G(i+, j+) + \frac{\varepsilon}{2} + \frac{k\sigma(\mathbf{p})B_m}{B_m+1} + \frac{2\sigma(\mathbf{p})\Delta_{\max}}{B_m+1} \leq \mathrm{d}_G(i+, j+) + \varepsilon. \tag{5.25}$$

By symmetry we can conclude the lower bound also, and the continuity of $\phi(\cdot)$ (see [15, Theorem 4.18]) along with (5.23) completes the proof of Proposition 5.7. $\qquad\square$

6 Mesoscopic properties: Proofs of Theorems 2.10 and 2.11

At this moment, we urge the reader to recall the definitions from (2.9), (2.10), (2.11), and (2.12). The configuration model graphs considered in this section will be assumed to have degree sequence $\boldsymbol{d}'$ and the vertices have an associated weight sequence $\boldsymbol{w}$ such that Assumption 2.9 is satisfied. We use the notation C, C' to denote generic positive constants, whose values can be different in different lines. The rest of the section is organized as follows. In Section 6.1, we start by proving the required bound on the diameter in Theorem 2.11. In order to deal with the different terms in Theorem 2.10, we first obtain some moment estimates in Section 6.2, and these estimates are then used to prove asymptotics of $s_2^\star$ in Section 6.3. The individual component weights are estimated in Section 6.4. In Section 6.5, we prove asymptotics of $s_3^\star$, and finally the mesoscopic typical distance is computed in Section 6.6.

6.1 Maximum diameter: Proof of Theorem 2.11

Let $\ell_n' = \sum_{i \in [n]} d_i'$. We will use path-counting estimates for the configuration model from [48, Lemma 5.1]. Let P_l denote the number of paths of length l in $\mathrm{CM}_n(\boldsymbol{d}')$. Then

[48, Lemma 5.1] shows that for any $l \geq 1$

$$\mathbb{E}[P_l] \leq \ell'_n (\nu'_n)^{l-1}. \tag{6.1}$$

If the maximum diameter is at least $n^\delta (\log n)^2$, then there exists a path of length $n^\delta (\log n)^2$, and therefore

$$\mathbb{P}(\Delta_{\max} > n^\delta (\log n)^2) \leq \sum_{l \geq n^\delta (\log n)^2} \mathbb{E}[P_l] \leq \frac{\ell'_n (\nu'_n)^{n^\delta (\log(n))^2}}{1 - \nu'_n} \leq C n^{1+\delta} e^{-C'(\log n)^2}, \tag{6.2}$$

where the second step follows using (6.1). Thus the proof of Theorem 2.11 follows. $\square$

6.2 Moment bounds for total weights

Consider the size-biased distribution on the vertex set $[n]$ with sizes $(w_i)_{i \in [n]}$. Let V_n and V_n^*, respectively, denote a vertex chosen uniformly at random and according to the size-biased distribution with respect to the sizes $(w_i)_{i \in [n]}$, independently of the underlying graph $\mathrm{CM}_n(\boldsymbol{d}')$. Let D'_n, W_n (respectively D_n^*, W_n^*) denote the degree and weight of V_n (respectively V_n^*). For a vertex $v \in [n]$, let $\mathscr{W}(v) := \sum_{k \in \mathscr{C}'(v)} w_k$, where $\mathscr{C}'(v)$ denotes the component of $\mathrm{CM}_n(\boldsymbol{d}')$ containing v. The reader should note the difference in notation that terms such as $\mathscr{W}_i$, $\mathscr{C}'_i$ with i in the subscript refer to the quantities defined in (2.10).

In this section, we prove the following moment bounds for $\mathscr{W}(V_n^*)$, which will help us compute the expectation and variance of $s_2^\star$:

Lemma 6.1. *Under* Assumption 2.9, *the following holds:*

(i) $\mathbb{E}\left[\mathscr{W}(V_n^*)\right] = \frac{\mathbb{E}[D_n^*]\mathbb{E}[D'_n W_n]}{\mathbb{E}[D'_n](1-\nu'_n)}(1 + o(1))$,

(ii) $\mathbb{E}\left[(\mathscr{W}(V_n^*))^2\right] \leq \frac{\mathbb{E}[D_n^*](\mathbb{E}[D'_n W_n])^2 \sigma_3(n)}{(\mathbb{E}[D'_n])^2(1-\nu'_n)^3}(1 + o(1))$, *where* $\sigma_3(n) = \frac{1}{\ell'_n} \sum_{k \in [n]} d'_k (d'_k - 1)(d'_k - 2)$,

(iii) $\mathbb{E}\left[(\mathscr{W}(V_n^*))^3\right] = o(n^{1+2\delta})$.

Proof of Lemma 6.1 (i). We use path-counting techniques for configuration models from [48, Lemma 5.1]. Let $\mathfrak{I}_l(v, k)$ denote the collection of $\boldsymbol{x} = (x_i)_{0 \leq i \leq l}$ such that $x_0 = v$, $x_l = k$ and x_i's are distinct. Then, an identical argument to the proof of [48, Lemma 5.1] shows that, for any $l \geq 1$, the expected number of paths of length exactly l starting from vertex v and ending at k is given by

$$\sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} \frac{d'_{x_0} d'_{x_l} \prod_{i=1}^{l-1} d'_{x_i}(d'_{x_i} - 1)}{(\ell'_n - 1) \cdots (\ell'_n - 2l + 1)} \leq \frac{d'_v \ell'_n}{\ell'_n - 2l + 3} \nu_n'^{l-1} = \left(1 + O\left(\frac{l}{n}\right)\right) d'_v \nu_n'^{l-1}, \tag{6.3}$$

where the last step holds for $l = o(n)$. Let $\mathcal{A}_l(v, k)$ denote the event that there exists a path of length l from v to k and let $\mathcal{A}'_l(v, k)$ denote the event that there exist two different paths from v to k, one of length l and another one of length at most $l - 1$. Notice that

$$\mathbb{E}\left[\mathscr{W}(V_n^*)|V_n^* = v\right] = \mathbb{E}\left[\sum_{k \in [n]} w_k \mathbb{1}\{v \rightsquigarrow k\}\right] \leq w_v + \sum_{l \geq 1} \sum_{k \in [n]} w_k \mathbb{P}\left(\mathcal{A}_l(v, k)\right), \tag{6.4a}$$

$$\mathbb{E}\left[\mathscr{W}(V_n^*)|V_n^* = v\right] \geq \sum_{l \geq 1} \sum_{k \in [n]} w_k \mathbb{P}\left(\mathcal{A}_l(v, k)\right) - \sum_{l \geq 1} \sum_{k \in [n]} w_k \mathbb{P}\left(\mathcal{A}'_l(v, k)\right). \tag{6.4b}$$

Now, using (6.3) and Assumption 2.9, (6.4a) yields

$$\mathbb{E}\left[\mathscr{W}(V_n^*)|V_n^* = v\right]$$

$$\leq w_v + \sum_{l=1}^{n^\delta(\log n)^2} \sum_{k \in [n]} w_k \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} \frac{d'_{x_0} d'_{x_k} \prod_{i=1}^{l-1} d'_{x_i}(d'_{x_i}-1)}{(\ell'_n - 1) \cdots (\ell'_n - 2l + 1)} + C\ell_n^w n^{1+\delta} \mathrm{e}^{-C'(\log n)^2} \tag{6.5}$$

$$\leq w_v + (1 + o(1)) \frac{d'_v \mathbb{E}\left[D'_n W_n\right]}{\mathbb{E}\left[D'_n\right]} \sum_{l=1}^\infty \nu_n'^{l-1} + o(1),$$

where in the second step we have used (6.2) and (6.3), and in the last step, we have used the facts that $\ell'_n = n\mathbb{E}[D'_n]$ and $\sum_{k \in [n]} d'_k w_k = n\mathbb{E}[D'_n W_n]$. Thus,

$$\mathbb{E}\left[\mathscr{W}(V_n^*)\right] \leq \frac{\mathbb{E}\left[D_n^*\right] \mathbb{E}\left[D'_n W_n\right]}{\mathbb{E}\left[D'_n\right](1 - \nu'_n)}(1 + o(1)), \tag{6.6}$$

where the multiplicative $(1 + o(1))$ term in the final expression comes observing that $(1 - \nu'_n)^{-1} = \Theta(n^\delta)$, and $\lim_{n \to \infty} \mathbb{E}[W_n^*] < \infty$. For the computation of the lower bound in (6.4b), we note that

$$\mathbb{P}(\mathcal{A}_l(v,k)) \geq \mathbb{P}(\exists \text{ a unique path of length } l \text{ from } v \text{ to } k)$$

$$= \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} \mathbb{P}(\boldsymbol{x} \text{ is the unique path between } v \text{ and } k)$$

$$\geq \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} \mathbb{P}(\boldsymbol{x} \text{ is a path from } v \text{ and } k)$$

$$- \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} \mathbb{P}(\exists \boldsymbol{y} \in \mathfrak{I}_l(v,k) \setminus \{\boldsymbol{x}\} : \boldsymbol{x}, \boldsymbol{y} \text{ both create paths from } v \text{ and } k),$$

$$\tag{6.7}$$

where the final step follows from using the inclusion-exclusion principle. Also, the first term inside the sum in (6.3) is the probability that $\boldsymbol{x} = (x_0, \ldots, x_l)$ creates a path for some $\boldsymbol{x}$. Thus,

$$\mathbb{P}(\mathcal{A}_l(v,k)) \geq \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} \frac{d'_{x_0} d'_{x_k} \prod_{i=1}^{l-1} d'_{x_i}(d'_{x_i}-1)}{(\ell'_n - 1) \cdots (\ell'_n - 2l + 1)}$$

$$- \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} \mathbb{P}(\exists \boldsymbol{y} \in \mathfrak{I}_l(v,k) \setminus \{\boldsymbol{x}\} : \boldsymbol{x}, \boldsymbol{y} \text{ both create paths from } v \text{ and } k).$$

$$\tag{6.8}$$

If we have two distinct paths, one being $\boldsymbol{x}$, and another of length at most l (say $\boldsymbol{y}$), then there must be two distinct vertices a, b in $\boldsymbol{x}$ such that the path between a, b on $\boldsymbol{x}$ is disjoint from that of $\boldsymbol{y}$. Depending on the choices of a, b, one of the structures in Figure 4 occurs. Let r be the length of the path between a, b that is disjoint of $\boldsymbol{x}$. Denote by $\mathcal{A}'_l(v, k, \boldsymbol{x}, i)$ the event that the structure of type i (i=I, II, III, IV) in Figure 4 appears, where $\boldsymbol{x} \in \mathfrak{I}_l(v,k)$. Using an argument identical to (6.5), and applying Assumption 2.9, it follows that

$$\sum_{l=1}^{n^\delta(\log n)^2} \sum_{k \in [n]} w_k \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} (l-1)(l-2)\mathbb{P}\left(\mathcal{A}'_l(v, k, \boldsymbol{x}, \mathrm{I})\right)$$

$$\leq Cd'_v \mathbb{E}\left[D'_n W_n\right] \sum_{l \geq 1} \sum_{r \geq 1} \frac{\sigma_3(n)^2}{\ell'_n}(l-1)(l-2)\nu_n'^{l+r-4} \tag{6.9a}$$

$$\leq C \frac{d'_v n^{6\alpha-3}}{1 - \nu'_n} \sum_{l \geq 3} (l-1)(l-2)\nu_n'^{l-3} \leq C \frac{d'_v n^{6\alpha-3}}{(1 - \nu'_n)^4} = o(d'_v n^\delta),$$
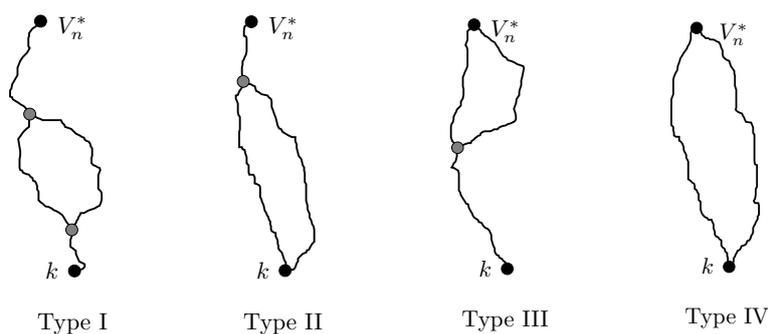
Figure 4: Possible structures for two distinct paths from V_n^* to k.

where the $(l-1)(l-2)$ factor is due to the possible choices of a, b, $\sigma_3(n)^2$ is due to the two branch points for which three half-edges needs to be paired, and in the last step we have used the fact $6\alpha - 3 + 3\delta < 6\alpha - 3 + 3\eta = 0$ since $\delta < \eta$. Similarly, with $b = k$, we get the Type-II structures in Figure 4, and thus

$$
\sum_{l=1}^{n^\delta (\log n)^2} \sum_{k \in [n]} w_k \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} (l-1) \mathbb{P}\left(\mathcal{A}_l'(v,k,\boldsymbol{x},\mathrm{II})\right)
$$
$$
\leq C d_v' \sigma_3(n) \left(\frac{1}{\ell_n'^2} \sum_{k \in [n]} w_k d_k'(d_k'-1)\right) \sum_{l \geq 1} \sum_{r \geq 1} (l-1)(\nu_n')^{l+r-3} \tag{6.9b}
$$
$$
\leq C \frac{d_v' n^{6\alpha-3}}{(1-\nu_n')^3} = o(d_v' n^\delta),
$$

Again,

$$
\sum_{l=1}^{n^\delta (\log n)^2} \sum_{k \in [n]} w_k \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} (l-1) \mathbb{P}\left(\mathcal{A}_l'(v,k,\boldsymbol{x},\mathrm{III})\right)
$$
$$
\leq C d_v'(d_v'-1)\sigma_3(n) \left(\frac{1}{\ell_n'^2} \sum_{k \in [n]} w_k d_k'\right) \sum_{l \geq 1} \sum_{r \geq 1} (l-1)(\nu_n')^{l+r-3} \tag{6.9c}
$$
$$
\leq C \frac{d_v'^2 n^{6\alpha-3}}{n^{3\alpha-1}(1-\nu_n')^3} = o(d_v'^2 n^{\delta+1-3\alpha}),
$$

and

$$
\sum_{l=1}^{n^\delta (\log n)^2} \sum_{k \in [n]} w_k \sum_{\boldsymbol{x} \in \mathfrak{I}_l(v,k)} \mathbb{P}\left(\mathcal{A}_l'(v,k,\boldsymbol{x},\mathrm{IV})\right)
$$
$$
\leq C d_v'(d_v'-1) \left(\frac{1}{\ell_n'^2} \sum_{k \in [n]} w_k d_k'(d_k'-1)\right) \sum_{l \geq 1} \sum_{r \geq 1} (\nu_n')^{l+r-2} \tag{6.9d}
$$
$$
\leq C \frac{d_v'^2 n^{6\alpha-3}}{n^{3\alpha-1}(1-\nu_n')^2} = o(d_v'^2 n^{\delta+1-3\alpha}).
$$

Taking expectations with respect to V_n^*, all the terms in (6.9a), (6.9b), (6.9c) and (6.9d) are $o(n^\delta)$, where we use that $\mathbb{E}[(D_n^*)^2] = O(n^{3\alpha-1})$. Note that the estimates in (6.9a), (6.9b), (6.9c) and (6.9d) can be used to upper bound $\mathbb{P}(\mathcal{A}_l'(v,k))$ in (6.4b) as well. To

compute the leading contribution to (6.4b), using (6.2) and (6.8), we lower bound

$$
d'_v \sum_{l=1}^{n^\delta (\log n)^2} \sum_{k\in[n]} w_k \sum_{\boldsymbol{x}\in\mathfrak{I}_l(v,k)} \frac{1}{\ell_n'^{l-1}} \prod_{i=1}^{l-1} d'_{x_i}(d'_{x_i}-1)d'_k + o(1)
$$

$$
\geq \frac{d'_v \mathbb{E}[D'_n W_n]}{\mathbb{E}[D'_n]} \sum_{l=1}^{n^\delta (\log n)^2} \left(\nu_n'^{l-1} - \frac{d'_1 n^\delta (\log n)^2}{\ell_n'^{l-1}} \left(\sum_{i\in[n]} d'_i(d'_i-1) \right)^{l-2} \right) + o(1) \tag{6.10}
$$

$$
= \frac{d'_v \mathbb{E}[D'_n W_n] \left(1-(\nu'_n)^{n^\delta (\log n)^2}\right)}{\mathbb{E}[D'_n](1-\nu'_n)} (1+o(1)) = \frac{d'_v \mathbb{E}[D'_n W_n]}{\mathbb{E}[D'_n](1-\nu'_n)}(1+o(1)),
$$

where we have used the fact that $d'_1 l \leq d'_1 n^\delta (\log n)^2$ and inclusion-exclusion to obtain the third step, and (2.9), $d'_1 n^\eta/\ell'_n = c_1/\mu_d(1+o(1))$ and the fact $(\nu'_n)^{n^\delta (\log n)^2} \leq \mathrm{e}^{-C(\log n)^2} = o(1)$ in the last step. Thus, it follows that

$$
\mathbb{E}[\mathscr{W}(V_n^*)] \geq \frac{\mathbb{E}[D_n^*]\mathbb{E}[D'_n W_n]}{\mathbb{E}[D'_n](1-\nu'_n)}(1+o(1)), \tag{6.11}
$$

and the proof of Lemma 6.1 (i) is now complete using (6.6). $\qquad\square$

Remark 6.2. It may be worthwhile to point out that the upper bound (6.6) holds for any configuration model satisfying $\nu'_n < 1 - n^{-\varepsilon_0}$ for some $\varepsilon_0 > 0$ and $\sum_{i\in[n]} w_i^2 = O(n)$. The rest of Assumption 2.9 is not required in the proof of this upper bound.

Proof of Lemma 6.1 (ii). Note that

$$
\left(\mathscr{W}(V_n^*)\right)^r = \sum_{k_1,\ldots,k_r\in[n]} w_{k_1}\cdots w_{k_r} \mathbb{1}\{V_n^* \rightsquigarrow k_1,\ldots,V_n^* \rightsquigarrow k_r\} = D_n(V_n^*,r) + E_n(V_n^*,r),
$$

$$\tag{6.12}$$

where

$$
D_n(V_n^*,r) := \sum_{\substack{k_1,\ldots,k_r\in[n] \\ V_n^*,k_1,\ldots,k_r \text{ distinct}}} w_{k_1}\cdots w_{k_r} \mathbb{1}\{V_n^* \rightsquigarrow k_1,\ldots,V_n^* \rightsquigarrow k_r\}. \tag{6.13}
$$

Let us formulate a general upper bound on $\mathbb{E}[D_n(V_n^*,r)]$ using similar computations as in (6.6). Note that if $V_n^* \rightsquigarrow k_i$ for all $i\in[r]$, then we must have a tree T with V_n^* as root and $(k_i)_{i\in[r]}$ as leaves. Let us "collapse" all the degree-two vertices in T except V_n^*. More precisely, we sequentially take a degree-two vertex (except V_n^*), delete it, and create an edge between its neighbors. Denote the obtained tree by $T' = (\mathrm{V}(T'), \mathrm{E}(T'))$. Thus, T' can be thought of as a rooted tree with V_n^* being its root, and $(k_i)_{i\in[r]}$ being its leaves. Also, T' does not have any degree two vertices except possibly V_n^*. Further, note that $r+1 \leq |\mathrm{V}(T')| \leq 2r$, and thus $r \leq |\mathrm{E}(T')| \leq 2r-1$. Let $m_i(T')$ denote the number of degree-i vertices in T' and let $d_0(T')$ be the degree of V_n^* in T'.

Let l_e be the number of edges that are collapsed to create $e \in \mathrm{E}(T')$. In that case, exactly $l_e - 1$ degree-two vertices get collapsed in T'. Using (6.2), we can restrict ourselves to the case $l_e \leq n^\delta (\log n)^2$, and the error due to such a restriction is given by $(\ell_n^w)^r n^{1+\delta} \mathrm{e}^{-C'(\log n)^2}$. For each T' described above, i half-edges of $V_n^* = v$ are being paired, for which there are $\prod_{i=1}^{d_0(T')}(d'_v - i + 1)$ possible ways. Moreover, the number of choices of the r distinct leaves $(k_i)_{i\in[r]}$ gives rise to the factor $\left(\sum_{k\in[n]} d'_k w_k\right)^r$ in the path counting. Define

$$
Q_n(T') := \prod_{j\geq 3} (\sigma_j(n))^{m_j(T')-\mathbb{1}\{j=d_0(T')\}}, \quad R_n(T') := \sum_{\substack{(l_1,\ldots,l_{|\mathrm{E}(T')|}): \\ 1\leq l_e \leq n^\delta (\log n)^2}} (\nu'_n)^{\sum_{e=1}^{|\mathrm{E}(T')|}(l_e-1)}. \tag{6.14}
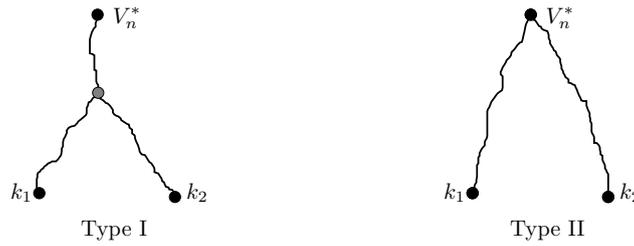$$

Figure 5: Possible paths when $V_n^* \rightsquigarrow k_1$, and $V_n^* \rightsquigarrow k_2$.

Note that $Q_n(T')$ gives the contribution due to the pairing of the half-edges of the vertices in $V(T') \setminus \{V_n^*, k_1, \ldots, k_r\}$, and $R_n(T')$ is the total contribution due to degree-two vertices of possible trees T that could give rise to T' after collapsing. Thus,

$$
\mathbb{E}\big[D_n(V_n^*, r) | V_n^* = v\big] \leq (1 + o(1)) \sum_{T'} \prod_{i=1}^{d_0(T')} (d_v' - i + 1)
$$
$$
\times \left(\frac{1}{\ell_n'} \sum_{k \in [n]} d_k' w_k \right)^r Q_n(T') R_n(T') + C(\ell_n^w)^r n^{1+\delta} \mathrm{e}^{-C'(\log n)^2}, \tag{6.15}
$$

for constants $C, C' > 0$.

Let us now apply (6.12) and (6.15) for the special case $r = 2$. Figure 5 describes the possible structures T'. Application of (6.15) yields

$$
\mathbb{E}\big[D_n(V_n^*, 2)\big] \leq \frac{\mathbb{E}\left[D_n^*\right] \left(\sum_{k \in [n]} d_k' w_k\right)^2 \sigma_3(n)}{\ell_n'^2 (1 - \nu_n')^3} + \frac{\mathbb{E}\left[D_n^*(D_n^* - 1)\right] \left(\sum_{k \in [n]} d_k' w_k\right)^2}{\ell_n'^2 (1 - \nu_n')^2}. \tag{6.16}
$$

The two terms are $O(n^{3\alpha + 3\delta - 1})$ and $O(n^{3\alpha + 2\delta - 1})$ respectively. Also,

$$
\mathbb{E}\big[E_n(V_n^*, 2)\big] \leq \mathbb{E}\left[(W_n^*)^2\right] + 2\mathbb{E}[W_n^* \mathscr{W}(V_n^*)] + \mathbb{E}\left[\sum_{k \in [n]} w_k^2 \mathbb{1}\left\{V_n^* \rightsquigarrow k\right\}\right], \tag{6.17}
$$

where the first term is due to $V_n^* = k_1 = k_2$, the second term is due to $V_n^* = k_1$ but $V_n^* \neq k_2$ or $V_n^* = k_2$ but $V_n^* \neq k_1$, while the third term due to $k_1 = k_2$ but $V_n^* \neq k_1$. The three terms are respectively $O(n^{3\alpha - 1})$, $O(n^{3\alpha + \delta - 1})$ and $O(n^{3\alpha + \delta - 1})$, where we have used (6.5) to compute the second term, and an analogous computation as in (6.5) to compute the final term (replacing w_k by w_k^2 in (6.5)). This proves our required upper bound that

$$
\mathbb{E}\big[\big(\mathscr{W}(V_n^*)\big)^2\big] \leq (1 + o(1)) \frac{\mathbb{E}\left[D_n^*\right] \left(\mathbb{E}\left[D_n' W_n\right]\right)^2 \sigma_3(n)}{\left(\mathbb{E}\left[D_n'\right]\right)^2 (1 - \nu_n')^3}. \tag{6.18}
$$

$\square$

Proof of Lemma 6.1 (iii). We again use (6.12) and (6.15). For the third moment, the leading contributions to $\mathbb{E}[D_n(V_n^*, 3)]$ arise from one of the structures given in Figure 6. We will use the fact that $\mathbb{E}[(D_n^*)^{r-1}], \sigma_r(n) = O(n^{r\alpha - 1})$ for $r \geq 3$. The contributions on $\mathbb{E}[D_n(V_n^*, 3)]$ due to the first type of tree in Figure 6 are upper bounded by

$$
\frac{C\sigma_3(n)^2}{(1 - \nu_n')^5} = O(n^{6\alpha - 2 + 5\delta}) = O(n^{1 + 2\delta + (6\alpha - 3 + 3\delta)}) = o(n^{1 + 2\delta}), \tag{6.19}
$$

where in the last step we have used the fact that $6\alpha - 3 + \delta < 6\alpha - 3 + 1 - 2\alpha = 2(2\alpha - 1) < 0$. The contributions due to the other three types of trees are respectively upper bounded
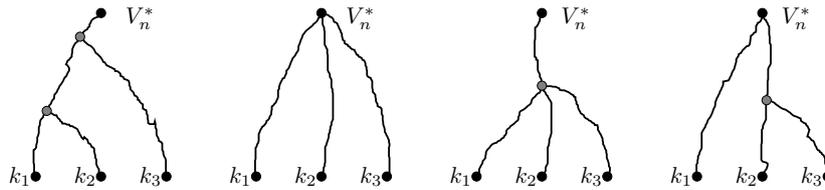
Figure 6: Possible paths when $V_n^* \rightsquigarrow k_1$, $V_n^* \rightsquigarrow k_2$, and $V_n^* \rightsquigarrow k_3$.

by

$$\frac{C\mathbb{E}[(D_n^*)^3]}{(1-\nu_n')^3} = O(n^{4\alpha-1+3\delta}), \quad \frac{C\sigma_4(n)}{(1-\nu_n')^4} = O(n^{4\alpha-1+4\delta}), \quad \frac{C\mathbb{E}[(D_n^*)^2]\sigma_3(n)}{(1-\nu_n')^4} = O(n^{6\alpha-2+4\delta}),$$

all of which are $o(n^{1+2\delta})$. Also,

$$\mathbb{E}[E_n(V_n^*,3)] \leq \mathbb{E}[(W_n^*)^3]$$
$$+ \mathbb{E}\left[\sum_{k\in[n]} w_k^3 \mathbb{1}\{V_n^* \rightsquigarrow k\}\right] + 3\mathbb{E}\left[W_n^* \sum_{k\in[n]} w_k^2 \mathbb{1}\{V_n^* \rightsquigarrow k\}\right] + 3\mathbb{E}\left[(W_n^*)^2 \mathscr{W}(V_n^*)\right]$$
$$+ 3\mathbb{E}[W_n^*(\mathscr{W}(V_n^*))^2] + 3\mathbb{E}\left[\sum_{k_1,k_2\in[n]} w_{k_1}^2 w_{k_2} \mathbb{1}\{V_n^* \rightsquigarrow k_1, V_n^* \rightsquigarrow k_2\}\right],$$

$$(6.20)$$

where the first term is due to $|\{V_n^*, k_1, k_2, k_3\}| = 1$, the three cases in the second line are due to $|\{V_n^*, k_1, k_2, k_3\}| = 2$ and the final two cases are due to $|\{V_n^*, k_1, k_2, k_3\}| = 3$. Using the fact that $\max_{i\in[n]} w_i = O(n^\alpha)$, we can use the estimates in (6.16) and (6.17) to show that the first term is $O(n^{4\alpha-1})$, the next three terms are $O(n^{4\alpha+\delta-1})$, and the last two terms are $O(n^{4\alpha+3\delta-1})$. All these contributions are $o(n^{1+2\delta})$ and hence we conclude that $\mathbb{E}[(\mathscr{W}(V_n^*))^3] = o(n^{1+2\delta})$. $\qquad\square$

6.3 Analysis of the susceptibility function $s_2^\star$

Asymptotics of $s_2^\star$. The asymptotics of $s_2^\star$ is a consequence of the Chebyshev inequality. Denote $\ell_n^w = \sum_{i\in[n]} w_i$. First, if $\mathbb{E}_{\boldsymbol{d}'}$ denotes the conditional expectation given $\mathrm{CM}_n(\boldsymbol{d}')$, then for any $r \geq 1$,

$$\mathbb{E}_{\boldsymbol{d}'}\left[(\mathscr{W}(V_n^*))^{r-1}\right] = \sum_{i\geq 1} \sum_{k\in\mathscr{C}_i'} \frac{w_k}{\ell_n^w}\left(\sum_{l\in\mathscr{C}_i'} w_l\right)^{r-1} = \frac{1}{\ell_n^w}\sum_{i\geq 1}(\mathscr{W}_i)^r. \quad (6.21)$$

Therefore, using Lemma 6.1 and (2.9), it follows from Assumption 2.9 that

$$n^{-\delta}\mathbb{E}[s_2^\star] = \frac{\ell_n^w}{n} n^{-\delta}\mathbb{E}[\mathscr{W}(V_n^*)] \rightarrow \frac{\mu_{d,w}^2}{\mu_d\lambda_0}, \quad (6.22)$$

where we have used the fact that $\mathbb{E}[D_n^*] \rightarrow \mu_{d,w}/\mu_w$. It remains to compute the variance. Let U_n^* denote another vertex chosen in a size-biased way with the sizes being $(w_i)_{i\in[n]}$, independently of the graph $\mathrm{CM}_n(\boldsymbol{d}')$ and V_n^*. Then (6.21) yields

$$\mathbb{E}\left[(s_2^\star)^2\right] = \frac{1}{n^2}\mathbb{E}\left[\sum_{i,j\geq 1} \mathscr{W}_i^2\mathscr{W}_j^2\right] = \frac{1}{n^2}\mathbb{E}\left[\sum_{i\geq 1} \mathscr{W}_i^4\right] + \frac{1}{n^2}\mathbb{E}\left[\sum_{i\neq j} \mathscr{W}_i^2\mathscr{W}_j^2\right]$$
$$= \frac{\ell_n^w}{n}\frac{1}{n}\mathbb{E}\left[(\mathscr{W}(V_n^*))^3\right] + \left(\frac{\ell_n^w}{n}\right)^2 \mathbb{E}\left[\mathscr{W}(U_n^*)\mathscr{W}(V_n^*)\mathbb{1}\{U_n^* \notin \mathscr{C}'(V_n^*)\}\right],$$

$$(6.23)$$

where the equality of the second term in the third equality follows using similar arguments as in (6.21). Denote the last two terms of (6.23) by (**I**) and (**II**) respectively. To estimate (**II**), observe that, conditionally on the graph $\mathscr{C}'(V_n^*)$, the graph obtained by removing $\mathscr{C}'(V_n^*)$ from $\mathrm{CM}_n(\boldsymbol{d}')$ is again a configuration model with the induced degree sequence $\tilde{\boldsymbol{d}}$ and number of vertices $\tilde{n}$. Let $\tilde{\nu}_n$ denote the corresponding criticality parameter. In the proof of Lemma 6.1 (i), we observed that the upper bound holds whenever $\tilde{\nu}_n < 1 - n^{-\varepsilon}$ with $\varepsilon \in (0,1)$ (see Remark 6.2). To this end, let us show that there exists $\varepsilon_0 \in (0,1)$ and $c_1 > 0$ such that for all sufficiently large n,

$$\mathbb{P}\left(\tilde{\nu}_n < 1 - n^{-\varepsilon_0} \mid \mathscr{C}'(V_n^*)\right) = 1, \text{ with probability at least } 1 - \mathrm{e}^{-n^{c_1}}. \tag{6.24}$$

Denote $\ell_n' = \sum_{i \in [n]} d_i'$. To see (6.24), first notice that

$$\tilde{\nu}_n - 1 = \nu_n' - 1 - \frac{\sum_{j \in \mathscr{C}'(V_n^*)} d_j'(d_j' - 2)}{\sum_{j \in [n]} d_j'} + (\tilde{\nu}_n - 1)\frac{\sum_{j \in \mathscr{C}'(V_n^*)} d_j'}{\ell_n'}. \tag{6.25}$$

Moreover, for any connected graph $\mathcal{G}$, $\sum_{i \in \mathcal{G}} d_i'(d_i' - 2) \geq -2$ (this can be proved by induction) so that

$$(\tilde{\nu}_n - 1)\left(1 - \sum_{j \in \mathscr{C}'(V_n^*)} \frac{d_j'}{\ell_n'}\right) \leq \nu_n' - 1 + \frac{2}{\ell_n'}. \tag{6.26}$$

Next we use the following:

Fact 6.3. *There exists $c_0, c_1 > 0$ (sufficiently small), and $n_0 \geq 1$ such that for all $n \geq n_0$, $\mathbb{P}(\sum_{j \in \mathscr{C}'(V_n^*)} d_j' \geq n^{\alpha + \delta + c_0}) \leq \mathrm{e}^{-n^{c_1}}$.*

The proof of Fact 6.3 follows using the exploration process in Section 6.4, and martingale concentration inequalities such as [36] (see Appendix C for a detailed proof). The proof of (6.24) now follows using $\ell_n' = \Theta(n)$, (2.9), and Fact 6.3.

As mentioned in Remark 6.2, now we can apply the upper bound from (6.5). Therefore,

$$\mathbb{E}\left[\mathscr{W}(U_n^*)\mathbb{1}\left\{U_n^* \notin \mathscr{C}'(V_n^*)\right\}\big|\mathscr{C}'(V_n^*)\right]$$
$$= \frac{\sum_{i \notin \mathscr{C}'(V_n^*)} w_i}{\ell_n^w}\mathbb{E}\left[\mathscr{W}(U_n^*)\big|\mathscr{C}'(V_n^*), U_n^* \notin \mathscr{C}'(V_n^*)\right]$$
$$\leq \frac{\sum_{i \in [n]} w_i^2}{\ell_n^w} + \frac{\sum_{i \notin \mathscr{C}'(V_n^*)} w_i}{\ell_n^w} \times \frac{\left(\sum_{i \notin \mathscr{C}'(V_n^*)} d_i' w_i\right)^2 \sum_{i \notin \mathscr{C}'(V_n^*)} d_i'}{\sum_{i \notin \mathscr{C}'(V_n^*)} w_i \sum_{i \notin \mathscr{C}'(V_n^*)} d_i'(-\sum_{i \notin \mathscr{C}'(V_n^*)} d_i'(d_i' - 2))}$$
$$\leq \frac{\left(\sum_{i \in [n]} d_i' w_i\right)^2}{\ell_n^w \ell_n'(1 - \nu_n' + o_{\mathbb{P}}(1))} + O(1) = \mathbb{E}\left[\mathscr{W}(V_n^*)\right]\left(1 + o_{\mathbb{P}}(1)\right), \tag{6.27}$$

where the penultimate step again follows from (6.26). Thus,

$$\mathbb{E}\left[\mathscr{W}(U_n^*)\mathscr{W}(V_n^*)\mathbb{1}\left\{U_n^* \notin \mathscr{C}'(V_n^*)\right\}\right] \leq \left(\mathbb{E}\left[\mathscr{W}(V_n^*)\right]\right)^2\left(1 + o(1)\right). \tag{6.28}$$

Now, (6.23), (6.28) together with Lemma 6.1 implies that $\mathrm{Var}\left(s_2^\star\right) = o(n^{2\delta})$. We can use the Chebyshev inequality and (6.22) to conclude that $n^{-\delta}s_2^\star \xrightarrow{\mathbb{P}} \mu_{d,w}^2/(\mu_d\lambda_0)$. $\qquad\square$

Remark 6.4. The method used to obtain the asymptotics of $s_2^\star$ can also be followed verbatim to obtain the asymptotics of $s_{pr}^\star$. Indeed, notice that

$$\mathbb{E}\left[s_{pr}^\star\right] = \frac{1}{n}\mathbb{E}\left[\sum_{i \geq 1} \mathscr{W}_i|\mathscr{C}_i'|\right] = \mathbb{E}\left[\mathscr{W}(V_n)\right]. \tag{6.29}$$

A similar identity for the second moment of $s_{pr}^\star$ also holds.

6.4 Barely subcritical masses

We now prove the asymptotics of $\mathscr{W}_j$ in Theorem 2.10. The idea is to obtain the asymptotics for $\mathscr{W}(j)$ for each fixed j, and then show that $\mathscr{W}_j = \mathscr{W}(j)$ with high probability. Consider the breadth-first exploration of the graph starting from vertex j as follows:

Algorithm 3. The algorithm carries along three disjoint sets of half-edges: *active, neutral, dead.*

(S0) At stage $i = 0$, the half-edges incident to j are active and all the other half-edges are neutral. Order the initially active half-edges arbitrarily.

(S1) At each stage, take the *smallest* half-edge e and pair it with another half-edge f, chosen uniformly at random from the set of half-edges that are either active or neutral. If f is neutral, then the vertex v to which f is incident, is not discovered yet. Declare the half-edges incident to v to be active and *larger* than all other active vertices (choose any order between the half-edges incident to v). Declare e, f to be dead.

(S2) Repeat from (S1) until the set of active half-edges is empty.

Define the process $\mathbf{S}_n^j$ by $S_n^j(l) = S_n^j(l-1) + d'_{(l)}J_l - 2$, and $S_n^j(0) = d'_j$, where J_l is the indicator that a new vertex is discovered at time l and $d'_{(l)}$ is the degree of the discovered vertex, if any. Thus, when the exploration starts from vertex j, then S_n^j tracks the number of active half-edges. Let $L := \inf\{l \geq 1 : S_n^j(l) = 0\}$. By convention, we assume that $S_n^j(l) = 0$ for $l > L$. Let $\mathscr{V}_l$ denote the vertex set discovered up to time l excluding j and $\mathcal{I}_i^n(l) := \mathbb{1}\{i \in \mathscr{V}_l\}$. Define $\mathcal{I}_j^n(l) \equiv 0$. Also, let $\mathscr{F}_l$ denote the sigma-field containing all the information upto time l in Algorithm 3. Note that

$$S_n^j(l) = d'_j + \sum_{i \in [n]} d'_i \mathcal{I}_i^n(l) - 2l = d'_j + \sum_{i \in [n]} d'_i \left(\mathcal{I}_i^n(l) - \frac{d'_i}{\ell'_n} l \right) + (\nu'_n - 1)\, l. \tag{6.30}$$

Consider the re-scaled process $\bar{\mathbf{S}}_n^j$ defined as $\bar{S}_n^j(t) = n^{-\alpha} S_n^j(\lfloor tn^{\alpha+\delta} \rfloor)$. Then, using Assumption 2.9,

$$\bar{S}_n^j(t) = c_j + n^{-\alpha} \sum_{i \in [n]} d'_i \left(\mathcal{I}_i^n(tn^{\alpha+\delta}) - \frac{d'_i}{\ell'_n} tn^{\alpha+\delta} \right) - \lambda_0 t + o(1). \tag{6.31}$$

The following three lemmas determine the asymptotics of $\mathscr{W}_i$ and $s_3^\star$:

Lemma 6.5. *Let L_j be the function with $L_j(t) = c_j - \lambda_0 t$ for $t \in [0, c_j\lambda_0^{-1}]$ and $L_j(t) = 0$ for $t > c_j\lambda_0^{-1}$. Then, under Assumption 2.9, as $n \to \infty$,*

$$\sup_{t \leq c_j\lambda_0^{-1}} \left| \bar{S}_n^j(t) - L_j(t) \right| \xrightarrow{\mathbb{P}} 0. \tag{6.32}$$

Lemma 6.6. *For any $T > 0$, $\sup_{l \leq Tn^{\alpha+\delta}} \left| \sum_{i \in [n]} w_i \mathcal{I}_i^n(l) - \frac{\sum_{i \in [n]} d'_i w_i}{\sum_{i \in [n]} d'_i} l \right| = o_{\mathbb{P}}(n^{\alpha+\delta})$.*

Lemma 6.7. *Fix any $j \geq 1$. Then with high probability $\mathscr{W}(j) = \mathscr{W}_j$.*

Asymptotics of $\mathscr{W}_j$. Note that, since the exploration process explores one edge at each time, Lemma 6.5 implies that (see e.g. [62, Theorem 13.6.4])

$$\frac{1}{2n^{\alpha+\delta}} \sum_{k \in \mathscr{C}'(j)} d'_k \xrightarrow{\mathbb{P}} \frac{c_j}{\lambda_0}. \tag{6.33}$$

Moreover, Lemma 6.6 yields that

$$\frac{1}{n^{\alpha+\delta}} \mathscr{W}(j) = \frac{1}{n^{\alpha+\delta}} \sum_{k \in \mathscr{C}'(j)} w_k = \frac{\sum_{i \in [n]} d'_i w_i}{\ell'_n n^{\alpha+\delta}} \frac{1}{2} \sum_{k \in \mathscr{C}'(j)} d'_k + o_{\mathbb{P}}(1) \xrightarrow{\mathbb{P}} \frac{\mu_{d,w}}{\mu_d \lambda_0} c_j. \tag{6.34}$$

Now the asymptotics of $\mathscr{W}_j$ in Theorem 2.10 follows by an application of Lemma 6.7. $\quad\square$

Next we provide a proof for Lemma 6.7. The proofs of Lemmas 6.5 and 6.6 follow using similar techniques as in [31], and thus are provided in Appendix B.

Proof of Lemma 6.7. Recall the definition of $\mathscr{C}_j'$, $\mathscr{W}_j$ from (2.10). For fixed $K \geq 1$, if all the components $(\mathscr{C}'(j))_{j \in [K]}$ are disjoint, then $j = \min\{k : k \in \mathscr{C}'(j)\}$, i.e., j is the minimum index among the vertices in $\mathscr{C}'(j)$. In that case, $\mathscr{C}'(j) = \mathscr{C}_j'$ and thus $\mathscr{W}(j) = \mathscr{W}_j$. Thus, it is enough to show that, for each fixed $i, j \geq 1$,

$$\mathbb{P}(i \text{ and } j \text{ are in same connected component}) \to 0. \tag{6.35}$$

If i, j are in the same connected component, then the $\bar{\mathbf{S}}_n^j$ will have a jump of size $d_i' = (1 + o(1))c_i n^\alpha$. By Lemma 6.5, and the fact that L_j is continuous, it follows that the probability of $\bar{\mathbf{S}}_n^j$ having a jump of size at least εn^α tends to zero for any fixed $\varepsilon > 0$. Thus we conclude (6.35) and the proof follows. □

6.5 Analysis of the susceptibility function $s_3^\star$

The aim of this section is to prove the following proposition which estimates the contribution on $s_3^\star$ due to components $(\mathscr{C}_i')_{i > K}$:

Proposition 6.8. *Suppose that* Assumption 2.9 *holds. For any $\varepsilon > 0$,*

$$\lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\sum_{i > K} \left(\mathscr{W}_i \right)^3 > \varepsilon n^{3(\alpha+\delta)} \right) = 0. \tag{6.36}$$

Proof. Let $\mathcal{G}^K$ denote the graph obtained by deleting all the edges incident to the vertices in $[K]$. In this proof, a superscript K to any previously defined object will correspond to the object in $\mathcal{G}^K$. Note that $\mathcal{G}^K$ is again distributed as a configuration model conditioned on the new degree sequence $\boldsymbol{d}^K = (d_i^K)_{i \in [n]}$. We also augment a previously defined notation with K in the superscript to denote the corresponding quantity for $\mathcal{G}^K$. Note that $\sum_{i \in [n]} d_i^K \geq \ell_n' - 2 \sum_{i \in [K]} d_i' = \ell_n'(1 + O(n^{\alpha-1}))$. Also, $d_i^K = 0$ for all $i \in [K]$, and $d_i^K \leq d_i'$ for all $i \in [n] \setminus [K]$. Recall the definition of c_i's from Assumption 2.9. First, for each fixed $K \geq 1$,

$$
\begin{aligned}
\nu_n^K &:= \frac{\sum_{i \in [n]} d_i^K (d_i^K - 1)}{\sum_{i \in [n]} d_i^K} \leq \frac{\sum_{i > K} d_i'(d_i' - 1)}{\ell_n'(1 + O(n^{\alpha-1}))} \\
&= \frac{\sum_{i \in [n]} d_i'(d_i' - 1)}{\ell_n'} - \frac{\sum_{i \in [K]} d_i'(d_i' - 1)}{\ell_n'} + O(n^{\alpha-1}) \\
&= \nu_n' + O(n^{2\alpha-1}) + O(n^{\alpha-1}) = 1 - \lambda_0 n^{-\delta} + o(n^{-\delta}),
\end{aligned} \tag{6.37}
$$

where we have used the fact that $\delta < \eta = 1 - 2\alpha < 1 - \alpha$ in the last step. We aim to apply the upper bound (6.18). Since we have only deleted $K = O(1)$ many vertices and $\Theta(n^\alpha)$ many half-edges to obtain $\mathcal{G}^K$, it follows that $(d_i^K)_{i \in [n]}$ also satisfies Assumption 2.9. We can apply the upper bound in (6.18), and thus

$$
\begin{aligned}
\frac{1}{n} \mathbb{E}\left[\sum_i \left(\mathscr{W}_i^K \right)^3 \right] = \frac{\ell_n^w}{n} \mathbb{E}\left[\left(\mathscr{W}^K(V_n^\star) \right)^2 \right] &\leq C \frac{\sum_{i > K} d_i'(d_i' - 1)(d_i' - 2)}{\sum_{i > K} d_i'(1 - \nu_n^{>K})^3} \\
&\leq C n^{3\alpha + 3\delta - 1}\left(n^{-3\alpha} \sum_{i > K} d_i'^3 \right),
\end{aligned} \tag{6.38}
$$

which tends to zero in the iterated limit $\lim_{K \to \infty} \limsup_{n \to \infty}$. Therefore, using the Markov inequality and the fact that $\boldsymbol{c} \in \ell_\downarrow^3 \setminus \ell_\downarrow^2$, it follows that, for any $\varepsilon > 0$,

$$\lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\sum_{i \geq 1} \left(\mathscr{W}_i^K \right)^3 > \varepsilon n^{3(\alpha+\delta)} \right) = 0. \tag{6.39}$$

Now, the proof is complete by observing that $\sum_{i>K} \mathscr{W}_i^3 \leq \sum_{i \geq 1} (\mathscr{W}_i^K)^3$. $\qquad \square$

Remark 6.9. Notice that the proof of Proposition 6.8 can be modified to conclude the similar results for $\sum_{i>K} (\mathscr{W}_i)^2 |\mathscr{C}_i'|$ and $\sum_{i>K} \mathscr{W}_i |\mathscr{C}_i'|^2$. Indeed, an analogue of (6.38) can be computed by observing that $\mathbb{E}[\sum_{i \geq 1} (\mathscr{W}_i^K)^2 |\mathscr{C}_i'^K|] = n\mathbb{E}[\mathscr{W}^K(V_n)]$, and $\mathbb{E}[\sum_{i \geq 1} \mathscr{W}_i^K (|\mathscr{C}_i'^K|)^2] = \ell_n^w \mathbb{E}[|\mathscr{C}'^K(V_n^*)|^2]$.

Finally we prove the asymptotics of $s_3^\star$ stated in Theorem 2.10:

Asymptotics of $s_3^\star$. The proof follows by combining the asymptotics of $\mathscr{W}_j$ and Proposition 6.8. $\qquad \square$

Remark 6.10. The argument for $s_3^\star$ can be followed verbatim to also conclude that

$$n^{-3\alpha-3\delta} \sum_{i=1}^{\infty} (\mathscr{W}_i)^2 |\mathscr{C}_i'| \xrightarrow{\mathbb{P}} \frac{\mu_{d,w}^2}{\mu_d^2 \lambda_0^3} \sum_{i=1}^{\infty} c_i^3, \quad n^{-3\alpha-3\delta} \sum_{i=1}^{\infty} \mathscr{W}_i |\mathscr{C}_i'|^2 \xrightarrow{\mathbb{P}} \frac{\mu_{d,w}}{\mu_d \lambda_0^3} \sum_{i=1}^{\infty} c_i^3. \quad (6.40)$$

6.6 Mesoscopic typical distances

In this section, we obtain the asymptotics of $\mathcal{D}_n^\star$ in Theorem 2.10 using a similar analysis as in Section 6.3. Again the proof involves the Chebyshev inequality where the moments are estimated using path counting. We sketch the computation of $\mathbb{E}[\mathcal{D}_n^\star]$. Recall the notations U_n^*, V_n^*, $\mathcal{A}_l(v,k)$ and $\mathcal{A}_l'(v,k)$ from Section 6.3. Note that

$$\mathbb{E}[\mathcal{D}_n^\star] = \frac{1}{n}\mathbb{E}\left[\sum_{i,k\in[n]} w_i w_k \mathrm{d}(i,k)\mathbb{1}\{k \in \mathscr{C}'(i)\} \right] = \frac{\ell_n^w}{n}\mathbb{E}\left[\sum_{k\in[n]} w_k \mathrm{d}(V_n^*,k)\mathbb{1}\{k \in \mathscr{C}'(V_n^*)\} \right]$$

$$\leq \frac{\ell_n^w}{n} \sum_{k\in[n]} w_k \sum_{l \geq 1} l\mathbb{P}(\mathcal{A}_l(V_n^*,k)) = \frac{\ell_n^w}{n} \sum_{l \geq 1} l \sum_{k\in[n]} w_k \mathbb{P}(\mathcal{A}_l(V_n^*,k)),$$

and also

$$\mathbb{E}[\mathcal{D}_n^\star] \geq \frac{\ell_n^w}{n} \sum_{l \geq 1} l \left(\sum_{k\in[n]} w_k \big(\mathbb{P}(\mathcal{A}_l(V_n^*,k)) - \mathbb{P}(\mathcal{A}_l'(V_n^*,k)) \big) \right). \quad (6.41)$$

Now compare the terms above to (6.4a), (6.4b). The only difference is that there is an extra multiplicative l here. Thus we can follows identical arguments as in the proof of (6.5), (6.11), and at the final step, we can use that $\sum_{l \geq 1} l(\nu_n')^{l-1} = (1 - \nu_n')^{-2}$. Thus,

$$\mathbb{E}[\mathcal{D}_n^\star] = \frac{\mathbb{E}[W_n]\mathbb{E}[D_n^*]\mathbb{E}[D_n' W_n]}{\mathbb{E}[D_n'](1 - \nu_n')^2}(1 + o(1)) = \frac{(\mathbb{E}[D_n' W_n])^2}{\mathbb{E}[D_n'](1 - \nu_n')^2}(1 + o(1)). \quad (6.42)$$

The variance terms can also be computed similarly. Due to the presence of l^2 in the second moment, we can use $\sum_{l \geq 1} l(l-1)(\nu_n')^{l-2} = (1 - \nu_n')^3$. This gives rise to an additional fact $1/(1 - \nu_n')^2 = O(n^{2\delta})$. Again, the identical arguments as (6.27) can be applied to show that $\mathrm{Var}(\mathcal{D}_n^\star) = o(n^{4\delta})$. Thus the proof of the asymptotics of $\mathcal{D}_n^\star$ follows. $\qquad \square$

7 Metric space limit for critical percolation clusters

The aim of this section is to complete the proof of Theorem 2.3. We start by defining the multiplicative coalescent process [3, 4] that will play a pivotal role in this section.

Definition 3 (Multiplicative coalescent)**.** Consider a (possibly infinite) collection of particles and let $\mathbf{X}(s) = (X_i(s))_{i \geq 1}$ denote the collection of masses of those particles at time s. Thus the i-th particle has mass $X_i(s)$ at time s. The evolution of the system takes place according to the following rule at time s: At rate $X_i(s)X_j(s)$, particles i and j merge into a new particle of mass $X_i(s) + X_j(s)$.

Before going into the details, let us describe the general idea and the organization of this section. The proof combines many ingredients and ideas from [10] and [31]. In Section 7.1 we consider a dynamically growing process of graphs that approximates the percolation clusters in the critical window. Now, the graphs generated by this dynamic evolution satisfy the following properties: (i) In the critical window, the components merge *approximately* as the multiplicative coalescent where the *mass* of each component is approximately proportional to the component size; (ii) The masses of the barely subcritical clusters satisfy *nice* properties due to Theorem 2.10. In Section 7.2, we derive the required properties in the barely subcritical regime for the dynamically growing graph process using Theorems 2.10 and 2.11. Section 7.3 is devoted to deriving scaling limits of functionals of $\mathcal{G}_n(t_c(\lambda))$. In Section 7.4, we modify the dynamic process in such a way that the components merge exactly as multiplicative coalescent. Since the exact multiplicative coalescent corresponds to the rank-one inhomogeneous random graphs, thinking of these barely subcritical clusters as blobs, we use the universality theorem (Theorem 5.4) in Section 7.5 to determine the metric space limits of the largest components of the modified graph (Theorem 7.15). We finally complete the proof of Theorem 2.3 in Section 7.6. The proof of Theorem 2.4 is given in Section 7.7.

7.1 The dynamic construction and its properties

Algorithm 4 (The dynamic construction). Let $\mathcal{G}_n(t)$ be the graph obtained up to time t by the following dynamic construction:

(S0) Initially, any vertex i has d_i incident half-edges and all the half-edges are alive. During the construction, a half-edge can be in one of the following two sets: alive or dead. All the half-edges have an independent unit rate exponential clock attached to them.

(S1) Whenever a clock rings, we take the corresponding half-edge, kill it and pair it with a half-edge chosen uniformly at random among the alive half-edges. The paired half-edge is also killed and the exponential clocks associated with killed half-edges are discarded.

Since a half-edge is paired with another unpaired half-edge, chosen uniformly at random from the set of all unpaired half-edges, the final graph $\mathcal{G}_n(\infty)$ is distributed as $\mathrm{CM}_n(\boldsymbol{d})$. Define

$$t_c(\lambda) = \frac{1}{2} \log\left(\frac{\nu_n}{\nu_n - 1}\right) + \frac{\nu_n}{2(\nu_n - 1)} \frac{\lambda}{n^\eta}. \tag{7.1}$$

We denote the i-th largest component of $\mathcal{G}_n(t)$ by $\mathscr{C}_{(i)}(t)$. In the subsequent part of this paper, we will derive the metric space limit of $(\mathscr{C}_{(i)}(t_c(\lambda)))_{i \geq 1}$. The following lemma enables us to switch to the conclusions for the largest clusters of $\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda))$:

Lemma 7.1 ([31, Proposition 24]). *There exists $\varepsilon_n = o(n^{-\eta})$ and a coupling such that, with high probability,*

$$\mathcal{G}_n(t_c(\lambda) - \varepsilon_n) \subset \mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda)) \subset \mathcal{G}_n(t_c(\lambda) + \varepsilon_n), \tag{7.2}$$

$$\mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda) - \varepsilon_n) \subset \mathcal{G}_n(t_c(\lambda) \subset \mathrm{CM}_n(\boldsymbol{d}, p_n(\lambda) + \varepsilon_n). \tag{7.3}$$

Let $\omega_i(t)$ denote the number of unpaired/open half-edges incident to vertex i at time t in Algorithm 4. We end this section by understanding the evolution of some functionals of the degrees and the open half-edges in the graph $\mathcal{G}_n(t)$. Let $s_1(t)$ denote the total number of unpaired half-edges at time t. Denote also $s_2(t) = \sum_{i \in [n]} \omega_i(t)^2$, $s_{d,\omega}(t) = \sum_{i \in [n]} d_i \omega_i(t)$. Further, we write $\mu_n = \ell_n/n$.

Lemma 7.2. *Under Assumption 2.1, the quantities* $\sup_{t \leq T} |\frac{1}{n} s_1(t) - \mu_n \mathrm{e}^{-2t}|$, $\sup_{t \leq T} |\frac{1}{n} s_2(t) - \mu_n \mathrm{e}^{-4t}(\nu_n + \mathrm{e}^{2t})|$, $\sup_{t \leq T} |\frac{1}{n} s_{d,\omega}(t) - \mu_n(1 + \nu_n)\mathrm{e}^{-2t}|$ *are all* $O_{\mathbb{P}}(n^{-1/2})$, *for any $T > 0$.*

Proof. The proof uses the differential equation method [63]. Notice that, after each ring of an exponential clock in Algorithm 4, $s_1(t)$ decreases by two. Let Y denote a unit-rate Poisson process. Using the random time change representation [33],

$$s_1(t) = \ell_n - 2Y\left(\int_0^t s_1(u)\mathrm{d}u\right) = \ell_n + M_n(t) - 2\int_0^t s_1(u)\mathrm{d}u, \qquad (7.4)$$

where M_n is a martingale. Now, the quadratic variation of M_n satisfies $\langle M_n \rangle(t) \leq 4t\ell_n = O(n)$, which implies that $\sup_{t\leq T}|M_n(t)| = O_{\mathbb{P}}(\sqrt{n})$. Moreover, notice that the function $f(t) = \mu_n \mathrm{e}^{-2t}$ satisfies $f(t) = \mu_n - 2\int_0^t f(u)\mathrm{d}u$. Therefore,

$$\sup_{t\leq T}\left|\frac{1}{n}s_1(t) - \mu_n\mathrm{e}^{-2t}\right| \leq \sup_{t\leq T}\frac{|M_n(t)|}{n} + 2\int_0^T \sup_{t\leq u}\left|\frac{1}{n}s_1(t) - \mu_n\mathrm{e}^{-2t}\right|\mathrm{d}u. \qquad (7.5)$$

Using Grőnwall's inequality [55, Proposition 1.4, page 204], it follows that

$$\sup_{t\leq T}\left|\frac{1}{n}s_1(t) - \mu_n\mathrm{e}^{-2t}\right| \leq \mathrm{e}^{2T}\sup_{t\leq T}\frac{|M_n(t)|}{n} = O_{\mathbb{P}}(n^{-1/2}), \qquad (7.6)$$

as required. For $s_2(t)$, note that if half-edges corresponding to vertices i and j are paired, s_2 changes by $-2\omega_i - 2\omega_j + 2$ and if two half-edges corresponding to i are paired, s_2 changes by $-4\omega_i + 4$. Thus,

$$\sum_{i\in[n]}\omega_i(t)^2 = \sum_{i\in[n]}d_i^2 + M_n'(t) + \int_0^t\sum_{i\neq j}\frac{\omega_i(u)\omega_j(u)(-2\omega_i(u) - 2\omega_j(u) + 2)}{s_1(u) - 1}\mathrm{d}u$$
$$+ \int_0^t\sum_{i\in[n]}\frac{\omega_i(u)(\omega_i(u) - 1)(-4\omega_i(u) + 4)}{s_1(u) - 1}\mathrm{d}u \qquad (7.7)$$
$$= n\mu_n(1 + \nu_n) + M_n'(t) + \int_0^t(-4s_2(u) + 2s_1(u))\mathrm{d}u + O(1),$$

where M_n' is a martingale with quadratic variation given by $\langle M_n'\rangle(t) = O(n)$. Again, an estimate equivalent to (7.6) follows using Grőnwall's inequality. Notice also that when a clock corresponding to vertex i rings and it is paired to vertex j, then $s_{d,\omega}$ decreases by $d_i + d_j$. Thus,

$$s_{d,\omega}(t)$$
$$= \sum_{i\in[n]}d_i^2 + M_n''(t) - \int_0^t\sum_{i\neq j}\frac{\omega_i(u)\omega_j(u)(d_i + d_j)}{s_1(u) - 1}\mathrm{d}u - \int_0^t\sum_{i\in[n]}\frac{\omega_i(u)(\omega_i(u) - 1)2d_i}{s_1(u) - 1}\mathrm{d}u$$
$$= n\mu_n(1 + \nu_n) + M_n''(t) - 2\int_0^t s_{d,\omega}(u)\mathrm{d}u,$$
$$\qquad (7.8)$$

where M_n'' is a martingale with quadratic variation given by $\langle M_n''\rangle(t) \leq 2t\sum_{i\in[n]}d_i^2 = O(n)$. We can now apply Grőnwall's inequality as before. The proof of Lemma 7.2 is complete. $\square$

7.2 Entrance boundary for open half-edges

Define

$$t_n = \frac{1}{2}\log\left(\frac{\nu_n}{\nu_n - 1}\right) - \frac{\nu_n}{2(\nu_n - 1)}\frac{1}{n^\delta}, \quad 0 < \delta < \eta. \qquad (7.9)$$

The goal is to show that the open half-edges satisfy the entrance boundary conditions. Let $\boldsymbol{d}(t) = (d_i(t))_{i\in[n]}$ denote the degree sequence of $\mathcal{G}_n(t)$ constructed by Algorithm 4.

Recall that $\mathcal{G}_n(t)$ is a configuration model conditionally on $\boldsymbol{d}(t)$. Let us first derive the asymptotics of $\nu_n(t_n)$. Recall that $\omega_i(t)$ denotes the number of open half-edges adjacent to vertex i in $\mathcal{G}_n(t)$. Notice that

$$\nu_n(t_n) = \frac{\sum_{i \in [n]}(d_i - \omega_i(t_n))^2}{\ell_n - s_1(t_n)} - 1 = \frac{\sum_{i \in [n]} d_i^2 - 2s_{d,\omega}(t_n) + s_2(t_n)}{\ell_n - s_1(t_n)} - 1. \qquad (7.10)$$

Using Lemma 7.2 and Assumption 2.1,

$$\frac{1}{n}(\ell_n - s_1(t_n)) = \mu_n(1 - \mathrm{e}^{-2t_n}) + o_{\mathbb{P}}(n^{-\delta}) = \frac{\mu_n}{\nu_n}\left(1 - \frac{\nu_n}{n^\delta}\right) + o_{\mathbb{P}}(n^{-\delta}), \qquad (7.11)$$

$$\frac{1}{n}\left(\sum_{i \in [n]} d_i^2 - 2s_{d,\omega}(t_n) + s_2(t_n)\right) = \frac{\mu_n}{\nu_n}\left(2 - \frac{3\nu_n}{n^\delta}\right) + o_{\mathbb{P}}(n^{-\delta}). \qquad (7.12)$$

Thus, (7.11) and (7.12) yields that $\nu_n(t_n) = 1 - \nu_n n^{-\delta} + o_{\mathbb{P}}(n^{-\delta})$. Further, using the differential equation method again, the evolution of $(\omega_i(t))_{t \geq 0}$ is given by

$$\begin{aligned}
\omega_i(t) &= d_i + M_n(t) + \int_0^t \left[\frac{2\omega_i(u)(s_1(u) - \omega_i(u))}{s_1(u) - 1} + \frac{2\omega_i(u)(\omega_i(u) - 1)}{s_1(u) - 1}\right]\mathrm{d}u \\
&= d_i + M_n(t) + 2\int_0^t \omega_i(u)\mathrm{d}u,
\end{aligned} \qquad (7.13)$$

and Assumption 2.1 yields that, for all $T > 0$,

$$\sup_{t \leq T}|n^{-\alpha}\omega_i(t) - \theta_i \mathrm{e}^{-2t}| \xrightarrow{\mathbb{P}} 0. \qquad (7.14)$$

We aim to apply the results for the barely subcritical regime in Theorem 2.10 to the number of open half-edges $\boldsymbol{\omega}(t_n) = (\omega_i(t_n))_{i \in [n]}$. Notice that, by Lemma 7.2, (7.14) and Assumption 2.1, $\boldsymbol{\omega}(t_n)$ and $\boldsymbol{d}(t_n)$ satisfy Assumption 2.9 with

$$\mu_\omega = \frac{\mu(\nu - 1)}{\nu}, \quad \mu_d = \frac{\mu}{\nu}, \quad \mu_{d,\omega} = \frac{\mu(\nu - 1)}{\nu}, \quad c_i = \frac{\theta_i}{\nu}. \qquad (7.15)$$

Let $\mathscr{C}_i(t)$ be defined analogously as (2.10) for the graph $\mathcal{G}_n(t)$. Denote $f_i(t) = \sum_{k \in \mathscr{C}_i(t)} \omega_k(t)$ and $\boldsymbol{f}(t) = (f_i(t))_{i \geq 1}$. The following theorem summarizes the entrance boundary conditions for $\boldsymbol{f}(t)$. Let s_2^ω, s_3^ω, $\mathcal{D}_n^\omega$ respectively denote the quantities $s_2^\star$, $s_3^\star$, $\mathcal{D}_n^\star$ respectively with the weights being the number of open half-edges, and the underlying graph being $\mathcal{G}_n(t_n)$.

Theorem 7.3. *Under* Assumption 2.1, *as* $n \to \infty$,

$$n^{-\delta}s_2^\omega \xrightarrow{\mathbb{P}} \frac{\mu(\nu - 1)^2}{\nu^2}, \quad n^{-\delta}s_{pr}^\omega \xrightarrow{\mathbb{P}} \frac{\mu(\nu - 1)}{\nu^2}, \quad n^{-(\alpha + \delta)}f_i(t_n) \xrightarrow{\mathbb{P}} \left(\frac{\nu - 1}{\nu^2}\right)\theta_i \qquad (7.16)$$

$$n^{-3\alpha - 3\delta + 1}s_3^\omega \xrightarrow{\mathbb{P}} \left(\frac{\nu - 1}{\nu^2}\right)^3 \sum_{i=1}^\infty \theta_i^3, \quad n^{-2\delta}\mathcal{D}_n^\omega \xrightarrow{\mathbb{P}} \frac{\mu(\nu - 1)^2}{\nu^3}. \qquad (7.17)$$

Remark 7.4. Setting $w_i = 1$ for all i, we get the entrance boundary conditions for the component sizes also. In this case $\mu_d = \mu_{d,w} = \mu/\nu$. Replacing ω by c in the above notation to denote the component susceptibilities, it follows that

$$n^{-\delta}s_2^c \xrightarrow{\mathbb{P}} \frac{\mu}{\nu^2}, \quad n^{-(\alpha + \delta)}|\mathscr{C}_i(t_n)| \xrightarrow{\mathbb{P}} \frac{\theta_i}{\nu^2}, \quad n^{-3\alpha - 3\delta + 1}s_3^c \xrightarrow{\mathbb{P}} \frac{1}{\nu^6}\sum_{i=1}^\infty \theta_i^3.$$

7.3 Components of the dynamically constructed graph

The idea is to regard $(\mathscr{C}_i(t_n))_{i\geq 1}$, the connected components at time t_n, as blobs. For $t \geq t_n$, the graph $\mathcal{G}_n(t)$ should be viewed as a super-graph with the superstructure being determined by the edges appearing after time t_n. Thus, the components of $\mathcal{G}_n(t)$ can be regarded as a union of the blobs. For a component $\mathscr{C}$, we use the notation $\mathfrak{B}(\mathscr{C})$ to denote the collection of indices corresponding to the blobs within $\mathscr{C}$ given by $\{b : \mathscr{C}_b(t_n) \subset \mathscr{C}, \mathscr{C}_b(t_n) \neq \varnothing\}$. Denote

$$\mathcal{F}_i(t) = \sum_{b \in \mathfrak{B}(\mathscr{C}_{(i)}(t))} f_b(t_n).$$

Let us denote the ordered components and the $\mathcal{F}$-values of $\mathcal{G}_n(t_c(\lambda))$ simply by $(\mathscr{C}_{(i)}(\lambda))_{i\geq 1}$ and $(\mathcal{F}_i(\lambda))_{i\geq 1}$ respectively. The goal in this section is to obtain the scaling of these component functionals, and also understand structural properties related to the surplus edges. Recall that $\mathrm{SP}(\mathscr{C})$ denotes the number of surplus edges in the component $\mathscr{C}$, i.e. $\mathrm{SP}(\mathscr{C}) = \#$edges in $\mathscr{C} - |\mathscr{C}| + 1$. The following result gives the scaling limits of the rescaled component sizes and surplus edges of $\mathcal{G}_n(t_c(\lambda))$:

Proposition 7.5. *Let* $(\mathscr{C}_{(i)}(\lambda))_{i\geq 1}$ *denote the ordered vector of components sizes of the graph* $\mathcal{G}_n(t_c(\lambda))$. *Then,* $\left(n^{-\rho}|\mathscr{C}_{(i)}(\lambda)|, \mathrm{SP}(\mathscr{C}_{(i)}(\lambda))\right)_{i\geq 1} \xrightarrow{d} (\frac{1}{\nu}\xi_i, \mathcal{N}_i)_{i\geq 1}$ *as* $n \to \infty$, *with respect to the topology on* $\ell^2_\downarrow \times \mathbb{N}^{\mathbb{N}}$, *where the limiting objects are defined in Proposition 4.3.*

The proof is a direct consequence of Lemma 7.1 and Proposition 4.3. See for example [31, Proposition 25]. The components consist of surplus edges within the blobs and the surplus edges in the superstructure. Next, let $\mathrm{SP}'(\mathscr{C}_{(i)}(\lambda))$ denote the number of surplus edges in the superstructure of $\mathscr{C}_{(i)}(\lambda)$. Thus $\mathrm{SP}'(\mathscr{C}_{(i)}(\lambda))$ denotes the macroscopic surplus edges which are not inside some blob. The next result proves that all the surplus edges in the critical components are macroscopic. Further, it relates the component sizes and the $\mathcal{F}$-values of $\mathcal{G}_n(t_c(\lambda))$:

Proposition 7.6. *Assume that* $\eta/2 < \delta < \eta$. *Then, for each* $1 \leq i \leq K$, *the following hold:*

(a) *With high probability,* $\mathrm{SP}'(\mathscr{C}_{(i)}(\lambda)) = \mathrm{SP}(\mathscr{C}_{(i)}(\lambda))$. *Consequently, there are no surplus edges within blobs in* $\mathscr{C}_{(i)}(\lambda)$ *with high probability;*

(b) $\mathcal{F}_i(\lambda)/|\mathscr{C}_{(i)}(\lambda)| \xrightarrow{\mathbb{P}} \nu - 1$. *Consequently,* $\left(n^{-\rho}\mathcal{F}_i(\lambda)\right)_{i\geq 1} \xrightarrow{d} \frac{\nu-1}{\nu}\boldsymbol{\xi}$ *with respect to the product topology.*

Since $\mathrm{SP}'(\mathscr{C}_{(i)}(\lambda)) \leq \mathrm{SP}(\mathscr{C}_{(i)}(\lambda))$ almost surely, for Part (a) it suffices to show that

$$\mathrm{SP}'(\mathscr{C}_{(i)}(\lambda)) \text{ and } \mathrm{SP}(\mathscr{C}_{(i)}(\lambda)) \text{ have the same distributional limit.} \tag{7.18}$$

Let $\mathcal{G}'_n$ denote the graph obtained from $\mathcal{G}_n(t_c(\lambda))$ by shrinking each blob to a single node. Then, $\mathrm{SP}'(\cdot)$ can be viewed as the surplus edges in the components of $\mathcal{G}'_n$. The graph $\mathcal{G}'_n$ can also be viewed to be constructed dynamically as in Algorithm 4 with the degree sequence being $(f_i(t_n))_{i\geq 1}$. In the following, we investigate the relations between $\mathcal{G}_n(t_n)$ and $\mathcal{G}'_n$. Lemma 7.2 implies that the number of unpaired half-edges in $\mathcal{G}_n(t_n)$ that are paired in $\mathcal{G}_n(t_c(\lambda))$ is given by

$$s_1(t_n) - s_1(t_c(\lambda)) = n\mu_n(n^{-\delta} + \lambda n^{-\eta}) + o_{\mathbb{P}}(n^{1-\gamma}), \quad \text{for some } \eta < \gamma. \tag{7.19}$$

Note that the we have used $\delta > \eta/2$ in (7.19).

Algorithm 5. Define $\pi_n = \frac{\nu_n}{\nu_n - 1}(n^{-\delta} + \lambda n^{-\eta})$ and associate $f_i(t_n)$ half-edges to the vertex i of $\mathcal{G}'_n$. Construct the graph $\mathcal{G}'_n(\pi_n)$ as follows:

(S1) Retain each half-edge independently with probability π_n.

(S2) Create a uniform perfect matching between the retained half-edges and obtain $\mathcal{G}'_n(\pi_n)$ by creating edges corresponding to any two pair of matched half-edges.

In (S1), if the total number of retained half-edges is odd, then add an extra half-edge to vertex 1. However, this possible addition of 1 extra half-edge will be ignored since it does not make any difference in the asymptotic computations. Notice that a_i, the number of half-edges attached to i that are retained by Algorithm 5 (S1), is distributed as $\mathrm{Bin}(f_i(t_n), \pi_n)$, independently for each i. Thus the number of half-edges in the graph $\mathcal{G}'_n(\pi_n)$ is distributed as a $\mathrm{Bin}(s_1(t_n), \pi_n)$ random variable. We claim that there exists $\varepsilon_n = o(n^{-\eta})$ and a coupling such that, with high probability

$$\mathcal{G}'_n(\pi_n - \varepsilon_n) \subset \mathcal{G}'_n \subset \mathcal{G}'_n(\pi_n + \varepsilon_n). \tag{7.20}$$

The proof follows using an identical argument as [31, Proposition 24] using the estimate (7.19) and standard concentration inequalities for binomial random variables. We skip the proof here. We now continue to analyze $\mathcal{G}'_n(\pi_n)$, keeping in mind that the relation (7.20) allows us make conclusions for $\mathcal{G}'_n$. To analyze the component sizes and the surplus edges of the components of $\mathcal{G}'(\pi_n)$ we first need some regularity conditions on $\boldsymbol{a}$, the degree sequence of $\mathcal{G}'_n(\pi_n)$, as summarized in the following lemma:

Lemma 7.7. *For any* $0 < \delta < \eta$, *as* $n \to \infty$, $n^{-\alpha} a_i \xrightarrow{\mathbb{P}} \frac{\theta_i}{\nu}$, $\frac{a_i}{\sum_i a_i} n^{\rho-\delta} \xrightarrow{\mathbb{P}} \frac{\theta_i}{\mu\nu}$, $\nu_n(\boldsymbol{a}) = \frac{\sum_i a_i(a_i-1)}{\sum_i a_i} = 1 + \lambda n^{-\eta+\delta} + o_{\mathbb{P}}(n^{-\eta+\delta})$, *and for any* $\varepsilon > 0$,

$$\lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\sum_{i>K} a_i^3 > \varepsilon n^{3\alpha} \right) = 0. \tag{7.21}$$

Proof. Using Theorem 7.3 and the fact that $a_i \sim \mathrm{Bin}(f_i(t_n), \pi_n)$, one gets $n^{-\alpha} a_i = (1 + o_{\mathbb{P}}(1))\frac{\theta_i}{\nu}$. Moreover, $\sum_i a_i \sim \mathrm{Bin}(\sum_i f_i(t_n), \pi_n)$ and $\sum_i a_i = (1 + o_{\mathbb{P}}(1))\pi_n \sum_i f_i(t_n)$ yield the required asymptotics for $a_i / \sum_i a_i$. Next note that if $X \sim \mathrm{Bin}(r, \pi)$, then $\mathrm{Var}(X(X-1)) = 2r(r-1)\pi^2(1-\pi)(1+(2r-3)\pi)$. Thus,

$$\mathrm{Var}\left(\sum_i a_i(a_i-1) \Big| (f_i(t_n))_{i \geq 1} \right) = \sum_i \mathrm{Var}\left(a_i(a_i-1) \big| (f_i(t_n))_{i \geq 1} \right)$$
$$= O_{\mathbb{P}}\left(\pi_n^2 \sum_i f_i^2(t_n) + \pi_n^3 \sum_i f_i^3(t_n) \right) = O_{\mathbb{P}}(n^{1-\delta} + n^{3\alpha}) = O_{\mathbb{P}}(n^{3\alpha}). \tag{7.22}$$

Therefore, for any $\varepsilon > 0$,

$$\mathbb{P}\left(\Big| \sum_i a_i(a_i-1) - \pi_n^2 \sum_i f_i(t_n)(f_i(t_n)-1) \Big| > \varepsilon n^{1-\delta} \Big| (f_i(t_n))_{i \geq 1} \right) = O_{\mathbb{P}}(n^{3\alpha-2+2\delta}), \tag{7.23}$$

which is $o_{\mathbb{P}}(1)$ since $\delta < \eta$ and $3\alpha - 2 + 2\eta = 3\alpha - 2 + 2 - 4\alpha < 0$. We conclude that

$$\sum_i a_i(a_i-1) = (1 + o_{\mathbb{P}}(1))\pi_n^2 \sum_i f_i(t_n)(f_i(t_n)-1), \tag{7.24}$$

and the required asymptotics for $\nu_n(\boldsymbol{a})$ follows. To see (7.21), note that $\mathbb{E}\left[\sum_{i>K} a_i(a_i-1)(a_i-2) \big| (f_i(t_n))_{i \geq 1} \right] = \pi_n^3 \sum_{i>K} f_i(t_n)^3$, and the proof follows again by using the condition on s_3^ω in Theorem 7.3. $\qquad \square$

Consider the exploration of the graph $\mathcal{G}'_n(\pi_n)$ via Algorithm 3, but now the first vertex is chosen proportional to its degree. Define the exploration process by $\boldsymbol{S}_n$ similarly as the process $\boldsymbol{S}^j_n(l)$ in Section 6.4. Call a vertex *discovered* if it is either active or killed. Let $\mathcal{V}_l$ denote the set of vertices discovered up to time l and $\mathcal{I}^n_i(l) := \mathbb{1}\{i \in \mathcal{V}_l\}$. Note that

$$S_n(l) = \sum_i a_i \mathcal{I}^n_i(l) - 2l = \sum_i a_i \left(\mathcal{I}^n_i(l) - \frac{a_i}{\ell^a_n} l \right) + (\nu_n(\boldsymbol{a}) - 1) \, l, \qquad (7.25)$$

where $\ell^a_n = \sum_i a_i$. Consider the re-scaled version $\bar{\mathbf{S}}_n$ defined as $\bar{S}_n(t) = n^{-\alpha} S_n(\lfloor tn^{\rho-\delta} \rfloor)$. Define the limiting process

$$S(t) = \sum_{i=1}^{\infty} \frac{\theta_i}{\nu} \left(\mathbb{1}\{\mathrm{Exp}(\theta_i/(\mu\nu)) \le t\} - \frac{\theta_i}{\mu\nu} t \right) + \lambda t. \qquad (7.26)$$

Proposition 7.8. *As $n \to \infty$, $\bar{\mathbf{S}}_n \xrightarrow{d} \mathbf{S}$ with respect to the Skorohod J_1 topology.*

The proof of Proposition 7.8 can be carried out using similar ideas as [31, Theorem 8]. A sketch of the proof is given in Appendix D. The excursion lengths of the exploration process give the number of edges in the explored components. Now, at each step l, the probability of discovering a surplus edge, conditioned on the past, is approximately the proportion of half-edges that are active. Note that the number of active half-edges is the reflected version of $\mathbf{S}_n$ given by $\mathrm{refl}(S_n(t)) = S_n(t) - \inf_{u \le t} S_n(u)$. Thus, conditional on $(S_n(l))_{l \le tn^{\rho-\delta}}$, the rate at which a surplus edge appears at time $tn^{\rho-\delta}$ is approximately $n^{\rho-\delta} \frac{\mathrm{refl}(S_n(tn^{\rho-\delta}))}{\sum_i a_i} = \frac{1}{\mu} \mathrm{refl}\left(\bar{S}_n(t)\right)(1 + o_{\mathbb{P}}(1))$. Therefore, Proposition 7.8 implies that for each $K \ge 1$, there exists components $C_1, \ldots, C_K \subset \mathcal{G}'_n(\pi_n)$ such that

$$\left(n^{-\rho+\delta} |C_i|, \mathrm{SP}(C_i) \right)_{i \in [K]} \xrightarrow{d} \left(\xi_i, \mathcal{N}_i \right)_{i \in [K]}, \qquad (7.27)$$

where ξ_i and $\mathcal{N}_i$ are defined in Proposition 4.3. We refer to [31, Section 5.4] for more details regarding the proof of (7.27). Here we have also used the fact that the ordered excursion lengths of the process $(S(t))_{t \ge 0}$, defined in (7.26), are identically distributed as the ordered excursion lengths of $(S(t)/\mu)_{t \ge 0}$. We can now combine (7.20) and (7.27) to obtain the asymptotics for the number of blobs in the largest connected components and $\mathrm{SP}'(\cdot)$. Denote $\mathscr{B}(\mathscr{C}) = |\mathfrak{B}(\mathscr{C})|$ for a component $\mathscr{C} \subset \mathcal{G}_n(t_c(\lambda))$.

Lemma 7.9. *For $K \ge 1$, there exist components $\mathscr{C}^1, \ldots, \mathscr{C}^K \subset \mathcal{G}_n(t_c(\lambda))$ such that the following convergence holds:* $(n^{-\rho+\delta} \mathscr{B}(\mathscr{C}^i), \mathrm{SP}'(\mathscr{C}^i))_{i \in [K]} \xrightarrow{d} (\xi_i, \mathcal{N}_i)_{i \in [K]}$.

Lemma 7.10. *For any $K \ge 1$, $\mathscr{C}^i = \mathscr{C}_{(i)}(\lambda), \forall i \in [K]$ with high probability.*

Proof. Notice that, $\sum_{j \le i} |\mathscr{C}^j| \le \sum_{j \le i} |\mathscr{C}_{(j)}(\lambda)|$ for all $i \in [K]$, almost surely. Thus, it is enough to prove that $|\mathscr{C}^i|$ and $|\mathscr{C}_{(i)}(\lambda)|$ involve the same re-scaling factor and have the same scaling limit. We again make use of the inclusions in graphs in (7.20). Algorithm 3 explores the components of $\mathcal{G}'_n(\pi_n)$ in a size-biased manner with the sizes being $(a_i)_{i \ge 1}$. An application of Lemma 7.14 with $y_i = \mathscr{C}_i(t_n)$ yields that, for any $t > 0$, uniformly for $l \le tn^{\rho-\delta}$,

$$\sum_i |\mathscr{C}_i(t_n)| \mathcal{I}^n_i(l) = \sum_i |\mathscr{C}_i(t_n)| \frac{a_i}{\sum_i a_i} l + o_{\mathbb{P}}(n^\rho). \qquad (7.28)$$

Since $a_i \sim \mathrm{Bin}(f_i(t_n), \pi_n)$, we can apply concentration inequalities like [51, Corollary 2.27] and use the asymptotics from Theorem 7.3 to conclude that

$$n^{-\delta} \frac{\sum_i a_i |\mathscr{C}_i(t_n)|}{\sum_i a_i} = \frac{\frac{\mu(\nu-1)}{\nu^2}}{\frac{\mu(\nu-1)}{\nu}} (1 + o_{\mathbb{P}}(1)) = \frac{1}{\nu}(1 + o_{\mathbb{P}}(1)). \qquad (7.29)$$

Thus, (7.28) and (7.29), together with (7.20), imply that $\frac{\nu|\mathscr{C}^i|}{n^\delta \mathscr{B}(\mathscr{C}^i)} \xrightarrow{\mathrm{P}} 1$, and it follows from Lemma 7.9 and Lemma 4.3 that $(n^{-\rho}|\mathscr{C}^i|)_{i\in[K]} \xrightarrow{d} (\frac{1}{\nu}\xi_i)_{i\in[K]}$. $\qquad\square$

Proof of Proposition 7.6. We are now finally in the position to prove Proposition 7.6. Using Lemmas 7.9, 7.10, and Proposition 7.5 together with (7.20), we directly conclude Part (a) from (7.18). For Part (b), we can follow the same arguments as (7.28) to conclude that, uniformly for $l \le tn^{\rho-\delta}$,

$$\sum_i f_i(t_n)\mathcal{I}_i^n(l) = \sum_i f_i(t_n)\frac{a_i}{\sum_i a_i}l + o_{\mathrm{P}}(n^\rho), \qquad (7.30)$$

where $n^{-\delta}\frac{\sum_i a_i f_i(t_n)}{\sum_i a_i} = \frac{\nu-1}{\nu}(1 + o_{\mathrm{P}}(1))$. Now, (7.28) and (7.30) together with (7.20) prove Part (b). $\qquad\square$

7.4 Coupling with the multiplicative coalescent

Recall the definitions of $t_c(\lambda)$ and t_n from (7.1) and (7.9). Now, let us investigate the dynamics of $\boldsymbol{f}(t)$ starting from time t_n. Notice that, in the time interval $[t_n, t_c(\lambda)]$, components with masses $f_i(t)$ and $f_j(t)$ merge at rate

$$f_i(t)\frac{f_j(t)}{s_1(t)-1} + f_j(t)\frac{f_i(t)}{s_1(t)-1} = \frac{2f_i(t)f_j(t)}{s_1(t)-1} \approx \frac{2\nu f_i(t)f_j(t)}{\mu(\nu-1)n}, \qquad (7.31)$$

and create a component with $f_i(t) + f_j(t) - 2$ open half-edges. Thus $\boldsymbol{f}(t)$ does not exactly evolve as a multiplicative coalescent, but it is close. We define an exact multiplicative coalescent that approximates the above process:

Algorithm 6 (Modified process). Conditionally on $\mathcal{G}_n(t_n)$, associate a rate $2/(s_1(t_n)-1)$ Poisson process $\mathcal{P}(e,f)$ to each pair of unpaired-half-edges (e,f). An edge (e,f) is created between the vertices incident to e and f at the instance when $\mathcal{P}(e,f)$ rings. However, the half-edges are not discarded after the pairing. At time $t > t_n$, the obtained *modified* graph $\bar{\mathcal{G}}_n(t)$ consists of the edges of $\mathcal{G}_n(t_n)$, and the edges created by this algorithm between times t_n and t.

Proposition 7.11. *There exists a coupling such that $\mathcal{G}_n(t) \subset \bar{\mathcal{G}}_n(t)$ for all $t > t_n$ with probability one.*

Proof. Recall the construction of $\mathcal{G}_n(t)$ from Algorithm 4. We modify (S1) as follows: whenever two half-edges are paired, we do not kill the corresponding half-edges and do not discard the associated exponential clocks. Instead we reset the corresponding exponential clocks. The graphs generated by this modification of Algorithm 4 have the same distribution as $\bar{\mathcal{G}}_n(t)$, conditionally on $\mathcal{G}_n(t_n)$. Moreover, the above also gives a natural coupling such that $\mathcal{G}_n(t) \subset \bar{\mathcal{G}}_n(t)$, by viewing the event times of Algorithm 4 as a thinning of the event times of the modified process. $\qquad\square$

Henceforth, we will always assume that we are working on a probability space such that Proposition 7.11 holds. Recall that the connected components at time t_n, $(\mathscr{C}_i(t_n))_{i\ge1}$, are regarded as blobs, and we can also view $\bar{\mathcal{G}}_n(t)$ as a super-graph with the superstructure being determined by the edges appearing after time t_n in Algorithm 6. Let us denote the ordered connected components of $\bar{\mathcal{G}}_n(t)$ by $(\bar{\mathscr{C}}_{(i)}(t))_{i\ge1}$. Define

$$\bar{\mathcal{F}}_i(t) = \sum_{b\in\mathfrak{B}(\bar{\mathscr{C}}_{(i)}(t))} f_b(t_n),$$

where $\mathfrak{B}(\mathscr{C}) := \{b : \mathscr{C}_b(t_n) \subset \mathscr{C}, \mathscr{C}_b(t_n) \ne \varnothing\}$. The $\bar{\mathcal{F}}$-value is regarded as the mass of component $\bar{\mathscr{C}}_{(i)}(t)$ at time t. Note that for the modified process in Algorithm 6,

conditionally on $\mathcal{G}_n(t_n)$, at time $t \in [t_n, t_c(\lambda)]$, $\bar{\mathscr{C}}_{(i)}(t)$ and $\bar{\mathscr{C}}_{(j)}(t)$ merge at exact rate $2\bar{\mathcal{F}}_i(t)\bar{\mathcal{F}}_j(t)/(s_1(t_n) - 1)$ and the new component has mass $\bar{\mathcal{F}}_i(t) + \bar{\mathcal{F}}_j(t)$. Thus, the vector of masses $(\bar{\mathcal{F}}_i(t))_{i \geq 1}$ merge as an exact multiplicative coalescent.

7.5 Properties of the modified process

Notice that, conditionally on $\mathcal{G}_n(t_n)$, blobs b_i and b_j are connected in $\bar{\mathcal{G}}_n(t_c(\lambda))$ with probability

$$p_{ij} = 1 - \exp\left(- f_{b_i}(t_n)f_{b_j}(t_n)\left[\frac{1}{n^{1+\delta}}\frac{\nu^2}{\mu(\nu-1)^2} + \frac{1}{n^{1+\eta}}\frac{\nu^2}{\mu(\nu-1)^2}\lambda\right](1 + o_{\mathbb{P}}(1))\right), \quad (7.32)$$

where the $o_{\mathbb{P}}(\cdot)$ term appearing above is uniform in i, j. Thus, using Theorem 7.3, (7.32) is of the form $1 - e^{-qx_ix_j(1+o_{\mathbb{P}}(1))}$ with

$$x_i^n = n^{-\rho}f_{b_i}(t_n), \quad q = \frac{1}{\sigma_2(\boldsymbol{x}^n)} + \frac{\nu^2}{\mu(\nu-1)^2}\lambda, \quad (7.33)$$

where $\sigma_r(\boldsymbol{x}^n) = \sum_{i \geq 1}(x_i^n)^r$. By Theorem 2.10, the sequence $\boldsymbol{x}^n$ satisfies the entrance boundary conditions of [4], i.e.,

$$\frac{\sigma_3(\boldsymbol{x}^n)}{(\sigma_2(\boldsymbol{x}^n))^3} \xrightarrow{\mathbb{P}} \frac{1}{\mu^3(\nu-1)^3}\sum_{i=1}^{\infty}\theta_i^3, \quad \frac{x_i^n}{\sigma_2(\boldsymbol{x}^n)} \xrightarrow{\mathbb{P}} \frac{1}{\mu(\nu-1)}\theta_i, \quad \sigma_2(\boldsymbol{x}^n) \xrightarrow{\mathbb{P}} 0. \quad (7.34)$$

To simplify the notation, we write $\bar{\mathcal{F}}_i(\lambda)$ for $\bar{\mathcal{F}}_i(t_c(\lambda))$ and $\bar{\mathscr{C}}_{(i)}(\lambda)$ for $\bar{\mathscr{C}}_{(i)}(t_c(\lambda))$. The following result is a consequence of [4, Proposition 7], [15, Lemma 5.3], and Lemma 4.3:

Proposition 7.12. As $n \to \infty$, $\left(n^{-\rho}\bar{\mathcal{F}}_i(\lambda)\right)_{i \geq 1} \xrightarrow{d} \frac{\nu-1}{\nu}\boldsymbol{\xi}$ with respect to the $\ell_{\downarrow}^2$ topology, where $\boldsymbol{\xi}$ is defined in Proposition 4.3.

We next relate $(\bar{\mathcal{F}}_i(\lambda))_{i \geq 1}$ to $(\bar{\mathscr{C}}_{(i)}(\lambda))_{i \geq 1}$, for each fixed i.

Proposition 7.13. As $n \to \infty$, $\bar{\mathcal{F}}_i(\lambda) = (\nu - 1)|\bar{\mathscr{C}}_{(i)}(\lambda)| + o_{\mathbb{P}}(n^\rho)$. Consequently, $\left(n^{-\rho}|\bar{\mathscr{C}}_{(i)}(\lambda)|\right)_{i \geq 1} \xrightarrow{d} \frac{1}{\nu}\boldsymbol{\xi}$ with respect to the product topology.

We will need the following lemma, the proof of which is same as [14, Lemma 8.2]:

Lemma 7.14 ([14, Lemma 8.2]). *Consider two ordered weight sequences $\boldsymbol{x} = (x_i)_{i \in [m]}$ and $\boldsymbol{y} = (y_i)_{i \in [m]}$. Consider the size-biased reordering $(v(1), v(2), \dots)$ of $[m]$ with respect to the weights $\boldsymbol{x}$ and let $V(i) := \{v(1), \dots, v(i)\}$. Denote $m_{rs} = \sum_i x_i^r y_i^s$, define $\mathrm{M}_n = m_{11}/m_{10}$ and assume that $\mathrm{M}_n > 0$ for each n. Suppose that the following conditions hold:*

$$\frac{lm_{21}}{m_{10}m_{11}} \to 0, \quad \frac{m_{12}m_{10}}{lm_{11}^2} \to 0, \quad \frac{lm_{20}}{m_{10}^2} \to 0, \quad \text{as } n \to \infty. \quad (7.35)$$

Then, as $n \to \infty$, $\sup_{k \leq l}\left|\frac{1}{l\mathrm{M}_n}\sum_i y_i\mathbb{1}\left\{i \in V(k)\right\} - \frac{k}{l}\right| \xrightarrow{\mathbb{P}} 0$.

Proof of Proposition 7.13. We only prove the asymptotic relation of $\bar{\mathcal{F}}_1(\lambda)$ and $|\bar{\mathscr{C}}_{(1)}(\lambda)|$. Consider the breadth-first exploration of the *superstructure* of graph $\bar{\mathcal{G}}_n(t_c(\lambda))$ (which is also a rank-one inhomogeneous random graph) using the Aldous-Limic construction from [4, Section 2.3]. Notice that the vertices are explored in a size-biased manner with the sizes being $\boldsymbol{x} = (x_i)_{i \geq 1}$, where $x_i = x_i^n = n^{-\rho}f_{b_i}(t_n)$ are as defined in (7.33). Let $v(i)$ be the i-th vertex explored. Further, let $\bar{\mathscr{C}}_{(i)}^{\mathrm{st}}(\lambda)$ denote the component $\bar{\mathscr{C}}_{(i)}(\lambda)$, where the blobs have been shrunk to single vertices. Then, from [4], one has the following:

(i) there exists random variables m_L, m_R such that $\bar{\mathscr{C}}_{(i)}^{\mathrm{st}}(\lambda)$ is explored between $m_L + 1$ and m_R;

(ii) $\sum_{i \leq m_R} x_{v(i)}$ is tight;

(iii) $\sum_{i=m_L+1}^{m_R} x_{v(i)} \xrightarrow{d} \gamma$, where γ is some non-degenerate, positive random variable.

Let $y_i = n^{-\rho}|\mathscr{C}_{b_i}(t_n)|$. Using Theorem 7.3, Remark 6.10 and Remark 7.4, it follows that $\sum_i x_i^r y_i^s = O_{\mathbb{P}}(n^{3\delta-3\eta})$; for $r + s = 3$, $\sum_i x_i = O_{\mathbb{P}}(n^{1-\rho})$, and $\sum_i x_i^r y_i^s = O_{\mathbb{P}}(n^{-2\rho+1+\delta})$; for $r + s = 2$. Below, we show that

$$\frac{\sum_{i=m_L+1}^{m_R} y_{v(i)}}{\sum_{i=m_L+1}^{m_R} x_{v(i)}} \times \frac{\sum_i x_i^2}{\sum_i x_i y_i} \xrightarrow{\mathbb{P}} 1. \tag{7.36}$$

The proof of Proposition 7.13 follows from (7.36) by using Theorem 7.3 and observing that $\frac{\sum_i x_i^2}{\sum_i x_i y_i} = \frac{s_2^\omega(t_n)}{s_{pr}^\omega(t_n)} \xrightarrow{\mathbb{P}} \nu - 1$. To prove (7.36), we will now apply Lemma 7.14. Denote $m_0 = \sum_i x_i / \sum_i x_i^2$ and consider $l = 2Tm_0$ for some fixed $T > 0$. Using Theorem 7.3, an application of Lemma 7.14 yields

$$\sup_{k \leq 2Tm_0} \left| \sum_{i=1}^k x_{v(i)} - \frac{k}{m_0} \right| \xrightarrow{\mathbb{P}} 0. \tag{7.37}$$

Now, for any $\varepsilon > 0$, $T > 0$ can be chosen large enough such that $\sum_{i=1}^{m_R} x_{v(i)} > T$ has probability at most ε and on the event $\left\{ \sup_{k \leq 2Tm_0} \left| \sum_{i=1}^k x_{v(i)} - \frac{k}{m_0} \right| \leq \varepsilon \right\} \cap \left\{ \sum_{i=1}^{m_R} x_{v(i)} \leq T \right\}$, one has $m_L < m_R < 2Tm_0$. Thus, it follows that

$$\left| \sum_{i=m_L+1}^{m_R} x_{v(i)} - \frac{m_R - m_L}{m_0} \right| \xrightarrow{\mathbb{P}} 0. \tag{7.38}$$

An identical argument as above shows that

$$\left| \sum_{i=m_L+1}^{m_R} y_{v(i)} - \frac{m_R - m_L}{m_0'} \right| \xrightarrow{\mathbb{P}} 0, \tag{7.39}$$

where $m_0' = \sum_i x_i / \sum_i x_i y_i$. The proof of (7.36) now follows from (7.38) and (7.39). The asymptotic distribution for $(n^{-\rho}|\bar{\mathscr{C}}_{(i)}(\lambda)|)_{i \geq 1}$ can be obtained using Proposition 7.12. $\qquad \square$

Recall that $\omega_i(t_n)$ denotes the number of open-half edges attached to vertex i in the graph $\mathcal{G}_n(t_n)$. We now equip $\bar{\mathscr{C}}_{(i)}(\lambda)$ with the probability measure μ_{fr}^i given by $\mu_{\mathrm{fr}}^i(A) = \sum_{k \in A} \omega_k(t_n)/\mathcal{F}_i(\lambda)$ for $A \subset \bar{\mathscr{C}}_{(i)}(\lambda)$, and denote the corresponding measured metric space by $\bar{\mathscr{C}}_{(i)}^{\mathrm{fr}}(\lambda)$.

Theorem 7.15. *Under* Assumption 2.1, *as* $n \to \infty$,

$$\left(n^{-\eta} \bar{\mathscr{C}}_{(i)}^{\mathrm{fr}}(\lambda) \right)_{i \geq 1} \xrightarrow{d} (M_i)_{i \geq 1}, \tag{7.40}$$

with respect to the $\mathscr{S}_*^{\mathbb{N}}$ *topology, where M_i is defined in Section 4.6.3.*

Proof. We just consider the metric space limit of $\bar{\mathscr{C}}_{(i)}^{\mathrm{fr}}(\lambda)$ for each fixed $i \geq 1$ and the joint convergence in (7.40) follows using the joint convergence of different functionals used throughout the proof. Recall the notation $\mathfrak{B}(\mathscr{C}) := \{b : \mathscr{C}_b(t_n) \subset \mathscr{C}, \mathscr{C}_b(t_n) \neq \varnothing\}$ for a component $\mathscr{C}$. Now, $\bar{\mathscr{C}}_{(i)}^{\mathrm{fr}}(\lambda)$ can be seen as a super-graph as defined in Section 4.2 with

(i) the collection of blobs $\{\mathscr{C}_b(t_n) : b \in \mathfrak{B}(\bar{\mathscr{C}}_{(i)}(\lambda))\}$ and within blob measure μ_b given by $\mu_b(A) = \sum_{k \in A} \omega_k(t_n)/f_b(t_n)$, $A \subset \mathscr{C}_b(t_n)$, $b \in \mathfrak{B}(\bar{\mathscr{C}}_{(i)}(\lambda))$;

(ii) the superstructure consisting of the edges appearing during $[t_n, t_c(\lambda)]$ in Algorithm 6 and weight sequence $(f_b(t_n)/\bar{\mathcal{F}}_i(\lambda) : b \in \mathfrak{B}(\bar{\mathscr{C}}_{(i)}(\lambda)))$.

Recall $x_i = x_i^n = n^{-\rho} f_{b_i}(t_n)$ defined in (7.33). Let $\mathrm{d}(\cdot, \cdot)$ denote the graph distance on $\bar{\mathscr{C}}_{(i)}(\lambda)$ and define

$$u_i = \sum_{v_1, v_2 \in \mathscr{C}_{b_i}(t_n)} \frac{\omega_{v_1} \omega_{v_2}}{f_{b_i}^2(t_n)} \mathrm{d}(v_1, v_2), \quad B_n^{(i)} = \frac{\sum_{b_j \in \mathfrak{B}(\bar{\mathscr{C}}_{(i)}(\lambda))} x_j u_j}{\sum_{b_j \in \mathfrak{B}(\bar{\mathscr{C}}_{(i)}(\lambda))} x_j}. \tag{7.41}$$

Here u_i gives the average distance within blob $\mathscr{C}_{b_i}(t_n)$. Using Lemma 7.14, we will show

$$B_n^{(i)} \times \frac{\sum_i x_i^2}{\sum_i x_i^2 u_i} \xrightarrow{\mathbb{P}} 1. \tag{7.42}$$

The argument is same as the proof of (7.36). We only have to ensure that (7.35) holds with $y_i = x_i u_i$. Thus, we need to show that

$$\frac{n^{\rho - \delta} \sum_i x_i^3 u_i}{\sum_i x_i \sum_i x_i^2 u_i} \xrightarrow{\mathbb{P}} 0, \quad \text{and} \quad \frac{\sum_i x_i^3 u_i^2 \sum_i x_i}{n^{\rho - \delta} \left(\sum_i x_i^2 u_i \right)^2} \xrightarrow{\mathbb{P}} 0. \tag{7.43}$$

First, notice that, by Lemma 7.2 and Theorem 7.3,

$$\begin{aligned} \mathrm{M}_n = \frac{\sum_i x_i^2 u_i}{\sum_i x_i} &= (1 + o_{\mathbb{P}}(1)) \frac{\nu n^{-1+\rho}}{\mu(\nu - 1)} n^{-2\rho} \sum_b f_b^2(t_n) \sum_{v_1, v_2 \in \mathscr{C}_b(t_n)} \frac{\omega_{v_1} \omega_{v_2}}{f_b^2(t_n)} \mathrm{d}(v_1, j) \\ &= \frac{\nu n^{-1+\rho}}{\mu(\nu - 1)} n^{1-2\rho} \mathcal{D}_n^\omega = n^{2\delta - \rho} \frac{\nu - 1}{\nu^2} (1 + o_{\mathbb{P}}(1)). \end{aligned} \tag{7.44}$$

Also, recall from Theorem 2.11 that $u_{\max} = \max_b u_b = O_{\mathbb{P}}(n^\delta \log(n))$. Now,

$$\begin{aligned} \frac{n^{\rho - \delta} \sum_i x_i^3 u_i}{\sum_i x_i \sum_i x_i^2 u_i} &\leq \frac{n^{\rho - \delta} u_{\max} \sum_i x_i^3}{\sum_i x_i \sum_i x_i^2 u_i} \\ &= O_{\mathbb{P}} \left(\frac{n^{\rho - \delta} n^\delta \log^2(n) n^{-3\rho} n^{3\alpha + 3\delta}}{n^{1-\rho} n^{2\delta - \rho} n^{1-\rho}} \right) = O_{\mathbb{P}}(n^{\delta - \eta} \log^2(n)) = o_{\mathbb{P}}(1), \\ \frac{\sum_i x_i^3 u_i^2 \sum_i x_i}{n^{\rho - \delta} \left(\sum_i x_i^2 u_i \right)^2} &\leq \frac{x_{\max} u_{\max} \sum_i x_i}{n^{\rho - \delta} \sum_i x_i^2 u_i} \\ &= O_{\mathbb{P}} \left(\frac{n^{-\rho} n^{\alpha + \delta} n^\delta \log^2(n)}{n^{\rho - \delta} n^{2\delta - \rho}} \right) = O_{\mathbb{P}}(n^{\delta - \eta} \log^2(n)) = o_{\mathbb{P}}(1), \end{aligned}$$

and (7.43) follows, and hence the proof of (7.42) also follows. Recall that the superstructure of $\bar{\mathcal{G}}_n(t_c(\lambda))$ has the same distribution as a Norros-Reittu random graph $\mathrm{NR}_n(\boldsymbol{x}, q)$ with the parameters given by (7.33). Thus, using Proposition 4.2, we now aim to use Theorem 5.4 on $\mathscr{C}_{(i)}^{\mathrm{fr}}(\lambda)$ with the blobs being $(\mathscr{C}_i(t_n))_{i \geq 1}$, and $\mathbf{p}_n^{(i)}$, $a_n^{(i)}$ given by (4.8). Define $\Upsilon_n^{(i)} = \left(p_b / \sigma(\mathbf{p}_n^{(i)}) : b \in \mathfrak{B}(\bar{\mathscr{C}}_{(i)}(\lambda)) \right)$. Let $\mathcal{N}(\mathbb{R}_+)$ denote the space of all counting measures equipped with the vague topology and let $\mathbb{S} := \mathbb{R}_+^3 \times \mathcal{N}(\mathbb{R}_+)$ denote the product space. Define

$$\mathscr{P}_n = \left(a_n^{(i)} \sigma(\mathbf{p}_n^{(i)}), \sum_{b \in \mathfrak{B}(\bar{\mathscr{C}}_{(i)}(\lambda))} x_b, \frac{1}{\sigma_2^2(\boldsymbol{x}^n)} \sum_{b \in \mathfrak{B}(\bar{\mathscr{C}}_{(i)}(\lambda))} x_b^2, \Upsilon_n^{(i)} \right)_{i \geq 1}, \tag{7.45}$$

viewed as an element of $\mathbb{S}^{\mathbb{N}}$. Recall the definition of ξ_i^* and Ξ_i^* from Section 4.6.3. Define

$$\mathscr{P}^\infty = \left(\frac{\xi_i^*}{\mu(\nu - 1)} \left(\sum_{v \in \Xi_i^*} \theta_v^2 \right)^{1/2}, \xi_i^*, \frac{1}{\mu^2(\nu - 1)^2} \sum_{v \in \Xi_i^*} \theta_v^2, \left(\frac{\theta_j}{\sum_{v \in \Xi_i^*} \theta_v^2} : j \in \Xi_i^* \right) \right)_{i \geq 1} \tag{7.46}$$

The following is a consequence of [15, Proposition 5.1, Lemma 5.4]:

$$\sigma(\mathbf{p}_n^{(i)}) \xrightarrow{\mathbb{P}} 0, \quad \text{and} \quad \mathscr{P}_n \xrightarrow{d} \mathscr{P}_\infty \text{ on } \mathbb{S}^{\mathbb{N}}. \tag{7.47}$$

Without loss of generality, we assume that the convergence in (7.47) holds almost surely. Now, using (7.42), it follows that

$$\frac{\sigma(\mathbf{p}_n^{(i)})}{1 + B_n^{(i)}} = \frac{\sigma_2(\boldsymbol{x}^n)\big(\sum_{v \in \Xi_i^*} \theta_v^2\big)^{1/2}}{\mu(\nu-1)\xi_i^*} \times \frac{\sum_i x_i^2}{\sum_i x_i^2 u_i}(1 + o(1))$$

$$= \frac{\sigma_2^2(\boldsymbol{x}^n)\big(\sum_{v \in \Xi_i^*} \theta_v^2\big)^{1/2}}{\mu(\nu-1)\xi_i^* \sum_i x_i^2 u_i}(1 + o(1)) = n^{-\eta}\frac{\nu-1}{\nu}\frac{1}{\xi_i^*}\bigg(\sum_{v \in \Xi_i^*} \theta_v^2\bigg)^{1/2}(1 + o(1)),$$

where the last step follows from Theorem 7.3, (7.44) and (7.47). The proof of Theorem 7.15 is now complete using Theorem 5.4. □

In the final part of the proof, we will also need an estimate of the surplus edges in the components $\bar{\mathscr{C}}_{(i)}(\lambda)$, that can be obtained by following the exact same argument as the proof outline of Lemma 7.9. Recall that the superstructure on the graph $\bar{\mathcal{G}}_n(t_c(\lambda))$ is a rank-one inhomogeneous random graph $\mathrm{NR}_n(\boldsymbol{x}, q)$. The connection probabilities given by (7.32) can be written as $1 - \exp(-z_i(\lambda)z_j(\lambda)/\sum_k z_k(\lambda))$, where $z_i(\lambda) = \frac{f_i(t_n)\sum_j f_j(t_n)}{\sum_j f_j^2(t_n)}\big(1 + \lambda n^{-\eta+\delta} + o_{\mathbb{P}}(n^{-\eta+\delta})\big)$. Moreover, using Theorem 7.3, it follows that

$$n^{-\alpha}z_i(\lambda) \xrightarrow{\mathbb{P}} \frac{\theta_i}{\nu}, \quad \frac{z_i(\lambda)}{\sum_j z_j(\lambda)} \xrightarrow{\mathbb{P}} \frac{\theta_i}{\mu\nu}, \quad \nu_n(\boldsymbol{z}) = \frac{\sum_i z_i^2(\lambda)}{\sum_i z_i(\lambda)} = 1 + \lambda n^{-\eta+\delta} + o_{\mathbb{P}}(n^{-\eta+\delta}). \tag{7.48}$$

Now, we may consider the breadth-first exploration of the above graph and define the exploration process $S_n^{\mathrm{NR}}(l) = \sum_i z_i(\lambda)\mathcal{I}_i^n(l) - l$, as in (7.25). The only thing to note here is that the component sizes are not necessarily encoded by the excursion lengths above the past minima of $\mathbf{S}_n^{\mathrm{NR}}$. However, if $\tilde{S}_n^{\mathrm{NR}}(l) = \sum_i \mathcal{I}_i^n(l) - l$, then it can be shown that (see [17, Lemma 3.1]) $\tilde{\mathbf{S}}_n^{\mathrm{NR}}$ and $\mathbf{S}_n^{\mathrm{NR}}$ have the same distributional limit. Thus, a conclusion identical to Proposition 7.8 follows for $\bar{\mathcal{G}}_n(t_c(\lambda))$. Due to the size-biased exploration of the components one can also obtain analogues of Lemmas 7.9 and 7.10 for $\bar{\mathcal{G}}_n(t_c(\lambda))$.

Proposition 7.16. For fixed $K \geq 1$, $(n^{-\rho+\delta}\mathscr{B}(\bar{\mathscr{C}}_{(i)}(\lambda)), \mathrm{SP}'(\bar{\mathscr{C}}_{(i)}(\lambda)))_{i \in [K]} \xrightarrow{d} (\xi_i, \mathscr{N}_i)_{i \in [K]}$, as $n \to \infty$.

7.6 Comparing modified and the original process: Completing the proof of Theorem 2.3

In this section, we finally conclude the proof of Theorem 2.3. We start with the following:

Lemma 7.17. For each fixed $i \geq 1$, $\mathscr{C}_{(i)}(\lambda) \subset \bar{\mathscr{C}}_{(i)}(\lambda)$ with high probability.

The proof is identical to [31, Lemma 31]. Recall Theorem 7.15 and the terminologies therein. Let $n^{-\eta}\mathscr{C}_{(i)}^{\mathrm{fr}}(\lambda)$ denote the measured metric space with measure μ_{fr}^i and the distances multiplied by $n^{-\eta}$. At this moment, let us bring together the relevant properties $\mathscr{C}_{(i)}(\lambda)$ and $\bar{\mathscr{C}}_{(i)}(\lambda)$:

(A) By Lemma 7.17, it follows that with high probability $\mathscr{C}_{(i)}(\lambda) \subset \bar{\mathscr{C}}_{(i)}(\lambda)$ for any fixed $i \geq 1$.

(B) Part (A) implies $\bar{\mathcal{F}}_i(\lambda) \geq \mathcal{F}_i(\lambda)$ with high probability. The same scaling limits from Propositions 7.6 (b) and 7.12 implies $\bar{\mathcal{F}}_i(\lambda) - \mathcal{F}_i(\lambda) \xrightarrow{\mathbb{P}} 0$ and consequently $\mu_{\mathrm{fr}}^i(\bar{\mathscr{C}}_{(i)}(\lambda) \setminus \mathscr{C}_{(i)}(\lambda)) \xrightarrow{\mathbb{P}} 0$.

(C) By Propositions 7.6 (a) and 7.16, there are no surplus edges with one endpoint in $\bar{\mathscr{C}}_{(i)}(\lambda) \setminus \mathscr{C}_{(i)}(\lambda)$ with high probability. Moreover, with high probability there is no surplus edge within the blobs by (7.18). This implies that, for any pair of vertices $u, v \in \mathscr{C}_{(i)}(\lambda)$, with high probability, the shortest path between them is exactly the same in $\mathscr{C}_{(i)}(\lambda)$ and $\bar{\mathscr{C}}_{(i)}(\lambda)$.

Thus, from the definition of Gromov-weak convergence in Section 4.1, an application of Theorem 7.15 yields that $\left(n^{-\eta}\mathscr{C}_{(i)}^{\text{fr}}\right)_{i \geq 1} \xrightarrow{d} (M_i)_{i \geq 1}$, The only thing remaining to show is that we can replace the measure μ_{fr}^i by $\mu_{\text{ct},i}$. Now, using Propositions 7.5 and 7.6 (b), it is enough to show that

$$\sum_{b \in \mathfrak{B}(\mathscr{C}_{(i)}(\lambda))} \left| f_b(t_n) - (\nu - 1)|\mathscr{C}_b(t_n)| \right| = o_{\mathbb{P}}(n^\rho). \tag{7.49}$$

Indeed, during the breadth-first exploration of the superstructure of $\mathcal{G}_n(t_c(\lambda))$, the blobs are explored in a size-biased manner with the sizes being $(f_i(t_n))_{i \geq 1}$. Therefore, one can again use Lemma 7.14. Recall that, by Lemma 7.10, for any $\varepsilon > 0$, one can choose $T > 0$ so large that the probability of exploring $\mathscr{C}_{(i)}(\lambda)$ within time $Tn^{\rho-\delta}$ is at least $1 - \varepsilon$. Thus, if $\mathscr{V}_l^b$ denotes the set of blobs explored before time l, then, for any $T > 0$,

$$\sum_{b \in \mathscr{V}_{Tn^{\rho-\delta}}^b} \left| f_b(t_n) - (\nu - 1)|\mathscr{C}_b(t_n)| \right|$$

$$= (1 + o_{\mathbb{P}}(1))Tn^{\rho-\delta} \sum_i \frac{f_i(t_n)}{\sum_i f_i(t_n)} \left| f_i(t_n) - (\nu - 1)|\mathscr{C}_i(t_n)| \right|.$$

Using the Cauchy-Schwarz inequality and Theorem 7.3 it now follows that the above term is $o(n^\rho)$. Therefore (7.49) follows. Finally, to conclude the result for percolated graphs, we use Lemma 7.1. In fact, if $\mathscr{C}_{(i)}^+(\lambda)$ and $\mathscr{C}_{(i)}^-(\lambda)$ denote the i-th largest component of $\mathcal{G}_n(t_c(\lambda) + \varepsilon_n)$ and $\mathcal{G}_n(t_c(\lambda) - \varepsilon_n)$ respectively, then analogously to Lemma 7.17, we can conclude that with high probability

$$\mathscr{C}_{(i)}^-(\lambda) \subseteq \mathscr{C}_{(i)}^p(\lambda) \subseteq \mathscr{C}_{(i)}^+(\lambda), \tag{7.50}$$

for any fixed $i \geq 1$. This completes the proof of Theorem 2.3. $\qquad\square$

Remark 7.18. The fact that the measure can be changed from μ_{fr}^i to $\mu_{\text{ct},i}$ in $n^{-\eta}\mathscr{C}_{(i)}^{\text{fr}}$ follows only from (7.49), which again follows from the entrance boundary conditions. However, the entrance boundary conditions in Theorem 2.10 hold for weight sequences $\boldsymbol{w} = (w_i)_{i \in [n]}$ under rather general assumptions (see Assumption 2.9). Therefore, one could also replace the measure $\mu_{\text{ct},i}$ by $\mu_{w,i}$, where $\mu_{w,i} = \sum_{i \in A} w_i / \sum_{k \in \mathscr{C}_{(i)}(\lambda)} w_i$ and $\boldsymbol{w}$ satisfies Assumption 2.9.

7.7 Graphs conditioned on simplicity: Proof of Theorem 2.4

We will use the following joint construction of the $\text{CM}_n(\boldsymbol{d}, p_n(\lambda))$ and $\text{CM}_n(\boldsymbol{d})$.

Algorithm 7. (S0) Let $\ell_n^p = 2X$, where $X \sim \text{Bin}(\ell_n/2, p_n(\lambda))$. Pick ℓ_n^p many half-edges uniformly at random and color them blue. Color the rest of the half-edges red.

(S1) Pair the blue half-edges using a uniform perfect matching.

(S2) Pair the red half-edges using another independent uniform perfect matching.

If G_I is the graph obtained after (SI) with I=1,2, then (G_1, G_2) is jointly distributed as $(\text{CM}_n(\boldsymbol{d}, p_n(\lambda)), \text{CM}_n(\boldsymbol{d}))$ [32, Lemmas 8.1, 8.2]. Let d_i^p be the number of blue half-edges incident to i and let $\boldsymbol{d}^p = (d_i^p)_{i \in [n]}$. Then, by construction, G_1, conditionally on $\boldsymbol{d}^p$, is distributed as $\text{CM}_n(\boldsymbol{d}^p)$. To complete the proof of Theorem 2.4, consider the exploration

algorithm given by Algorithm 3, now on the graph G_1, conditionally on the blue half-edges selected in Algorithm 7 (S0). The starting vertex is chosen in a size biased manner with sizes proportional to the degrees $\boldsymbol{d}^p$. Let $\mathscr{F}_l$ denote the sigma-algebra generated by the exploration process up to time l. Let $\mathcal{I}_i^n(l)$ denote the indicator that vertex i is discovered upto time l and note that Algorithm 3 will explore the vertices in a size-biased manner with sizes being $\boldsymbol{d}^p$. For convenience, we denote $X = (\mathscr{C}_{(i)}^p(\lambda))_{i \leq K}$ in this section. Consider a bounded continuous function $f : (\mathscr{S}_*)^K \mapsto \mathbb{R}$. Recall from [47, Theorem 1.1] that

$$\liminf_{n \to \infty} \mathbb{P}\left(G_2 \text{ is simple}\right) > 0.$$

Thus, it is enough to show that

$$\mathbb{E}\left[f(X)\mathbb{1}\left\{G_2 \text{ is simple}\right\}\right] - \mathbb{E}\left[f(X)\right]\mathbb{P}\left(G_2 \text{ is simple}\right) \to 0. \tag{7.51}$$

Now, for any $T > 0$, let $\mathcal{A}_{n,T}$ denote the event that X is explored before time Tn^ρ by the exploration algorithm. Using [31, Lemma 13], it follows that

$$\lim_{T \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\mathcal{A}_{n,T}^c\right) = 0. \tag{7.52}$$

Let X_T denote the random vector consisting of K largest ones among the components explored before time Tn^ρ. Thus,

$$\lim_{T \to \infty} \limsup_{n \to \infty} \mathbb{E}\left[f(X)\mathbb{1}\left\{G_2 \text{ is simple}\right\}\mathbb{1}[\mathcal{A}_{n,T}^c]\right]$$
$$\leq \|f\|_\infty \lim_{T \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\mathcal{A}_{n,T}^c\right) = 0,$$

which implies that

$$\lim_{T \to \infty} \limsup_{n \to \infty} \left|\mathbb{E}\left[f(X)\mathbb{1}\left\{G_2 \text{ is simple}\right\}\right] - \mathbb{E}\left[f(X_T)\mathbb{1}\left\{G_2 \text{ is simple}\right\}\right]\right| = 0. \tag{7.53}$$

Further, let $\mathcal{B}_{n,T}$ denote the event that a vertex v is explored before time Tn^ρ such that v is involved in a self-loop or a multiple edge in G_2. Let v_l denote the *exploring* vertex in the exploration at time l. Without loss of generality, we assume that during the sequential pairing of the half-edges in Algorithm 7 (S2), we first pair the red half-edges associated to $(v_l)_{l \leq Tb_n}$. Let $\ell_n' := \ell_n^p - 2Tb_n - d_1^p + 1$ and $\ell_n'' := (\ell_n - \ell_n^p) - \sum_{i \in [n]} d_i \mathcal{I}_i(Tn^\rho)$. Note that, by Assumption 2.1 (ii), $\sum_{i \in V} d_i = o(n)$ whenever $|V| = o(n)$. Also, using concentration inequalities for the Binomial distribution, $\ell_n^p = p_n(\lambda)\ell_n(1 + o(1))$ almost surely. Thus, we assume that $\ell_n', \ell_n'' \geq c_1 \ell_n$ with probability 1 for some $0 < c_1 < 1$. Note that, uniformly over $l \leq Tb_n$, any blue half-edge of v_l creates a self-loop in G_1 with probability at most $d_{v_l}^p/\ell_n'$, and any red half-edge creates a self-loop with probability at most $(d_{v_l} - d_{v_l}^p)/\ell_n''$. Thus the expected number of self-loops incident to v_l is at most $\mathbb{E}[d_{v_l}^2]/c_1 \ell_n$. Moreover, the expected number of blue-blue multiple edge attached to v_l in G_1, is at most

$$\mathbb{E}[d_{v_l}^p(d_{v_l}^p - 1)]\frac{\sum_{i \in [n]} d_i(d_i - 1)}{c_1 \ell_n(c_1 \ell_n - 1)} \leq \frac{C\mathbb{E}[d_{v_l}^2]}{\ell_n}, \tag{7.54}$$

where we have used Assumption 2.1. While counting the multiple edges incident to v_l in G_2, we have to take care of (i) the creation of a red edge between two vertices having a blue edge and (ii) the creation of two red edges between two vertices. Using identical arguments, the expected number of multiple edges incident to v_l in G_2 is at most $C\mathbb{E}[d_{v_l}^2]/\ell_n$. Therefore,

$$\mathbb{E}\left[\#\{\text{self-loops or multiple edges discovered while } v_l \text{ is exploring}\}\right] \leq \frac{C\mathbb{E}[d_{v_l}^2]}{\ell_n}. \tag{7.55}$$

Thus,

$$\mathbb{P}(\mathcal{B}_{n,T}) \leq \frac{C}{\ell_n} \mathbb{E}\Bigg[\sum_{i \in [n]} d_i^2 \mathcal{I}_i^n(Tn^\rho) \Bigg] = \frac{C}{\ell_n}\Bigg(\mathbb{E}\Bigg[\sum_{i=1}^K d_i^2 \mathcal{I}_i^n(Tn^\rho) \Bigg] + \mathbb{E}\Bigg[\sum_{i=K+1}^n d_i^2 \mathcal{I}_i^n(Tn^\rho) \Bigg] \Bigg).$$

Now, using Assumption 2.1, for every fixed $K \geq 1$,

$$\frac{1}{\ell_n} \mathbb{E}\Bigg[\sum_{i=1}^K d_i^2 \mathcal{I}_i^n(Tn^\rho) \Bigg] \leq \frac{1}{\ell_n} \sum_{i=1}^K d_i^2 \xrightarrow{\mathbb{P}} 0. \tag{7.56}$$

Further, conditionally on Algorithm 7 (S0), the vertices are explored in a size-biased manner with sizes being $(d_i^p/\ell_n^p)_{i \in [n]}$. Therefore, using that $\ell_n^p = p_n(\lambda)\ell_n(1 + o(1))$,

$$\frac{1}{\ell_n} \mathbb{E}_p\Bigg[\sum_{i=K+1}^n d_i^2 \mathcal{I}_i^n(Tn^\rho) \Bigg] \leq \frac{Tn^\rho}{\ell_n} \sum_{i=K+1}^n d_i^2 \mathbb{E}\Bigg[\frac{d_i^p}{\ell_n^p} \Bigg] \leq C\Bigg(n^{-3\alpha} \sum_{i=K+1}^n d_i^3 \Bigg). \tag{7.57}$$

Now, by Assumption 2.1, the final term in (7.57) tends to zero if we first take $\limsup_{n \to \infty}$ and then take $\lim_{K \to \infty}$. Consequently, for any fixed $T > 0$,

$$\lim_{n \to \infty} \mathbb{P}(\mathcal{B}_{n,T}) = 0. \tag{7.58}$$

Let $\mathcal{E}_{n,T}$ denote the event that no self-loops or multiple edges are attached to the vertices in G_2 that are discovered after time Tn^ρ. Then (7.53) and (7.58) implies that

$$\begin{aligned}
&\lim_{n \to \infty} \mathbb{E}\left[f(X) \mathbb{1}\{\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}\} \right] = \lim_{T \to \infty} \lim_{n \to \infty} \mathbb{E}\left[f(X_T) \mathbb{1}[\mathcal{E}_{n,T}] \right] \\
&= \lim_{T \to \infty} \lim_{n \to \infty} \mathbb{E}\left[f(X_T) \mathbb{P}(\mathcal{E}_{n,T} | \mathscr{F}_{Tn^\rho}) \right] = \lim_{T \to \infty} \lim_{n \to \infty} \mathbb{E}\left[f(X_T) \mathbb{P}(\mathcal{E}_{n,T} | \mathscr{F}_{Tn^\rho}, \mathcal{B}_{n,T}) \right].
\end{aligned} \tag{7.59}$$

Let $\mathcal{G}^*_{Tn^\rho}$ denote the graph obtained from G_2 after removing the vertices discovered upto time Tn^ρ. Then, conditionally on $\mathscr{F}_{Tn^\rho} \cap \mathcal{B}_{n,T}$, $\mathcal{E}_{n,T}$ happens if and only if $\mathcal{G}^*_{Tn^\rho}$ is simple. Also, $\mathcal{G}^*_{Tn^\rho}$ is distributed as a configuration model conditional on its degree sequence, and since only $o(n)$ vertices have been removed, the corresponding ν_n in $\mathcal{G}^*_{Tn^\rho}$ converges in probability to 1. Thus, [43, Theorem 7.12] implies that

$$\mathbb{P}(\mathcal{G}^*_{Tn^\rho} \text{ is simple} | \mathscr{F}_{tn^\rho}) \xrightarrow{\mathbb{P}} \mathrm{e}^{-3/4}, \tag{7.60}$$

and also $\mathbb{P}(G_2 \text{ is simple}) \to \mathrm{e}^{-3/4}$, so that

$$\mathbb{P}(\mathcal{G}^*_{Tn^\rho} \text{ is simple} | \mathscr{F}_{tn^\rho}) - \mathbb{P}(G_2 \text{ is simple}) \xrightarrow{\mathbb{P}} 0. \tag{7.61}$$

Now, using (7.59), (7.51) follows, and the proof of Theorem 2.4 is complete. $\qquad\square$

References

[1] 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. MR-2892951

[2] Louigi Addario-Berry, Nicolas Broutin, Christina Goldschmidt, and Gregory Miermont, *The scaling limit of the minimum spanning tree of the complete graph*, Ann. Probab. **45** (2017), no. 5, 3075–3144. MR-3706739

[3] David Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, Ann. Probab. **25** (1997), no. 2, 812–854. MR-1434128

[4] David Aldous and Vlada Limic, *The entrance boundary of the multiplicative coalescent*, Electron. J. Probab. **3** (1998), no. 3, 1–59. MR-1491528

[5] 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. MR-2063375

[6] David Aldous and Jim Pitman, *A family of random trees with random edge lengths*, Random Struct. Algor. **15** (1999), no. 2, 176–195. MR-1704343

[7] David Aldous and Jim Pitman, *Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent*, Probab. Theory Relat. Fields **118** (2000), no. 4, 455–482. MR-1808372

[8] Noga Alon, Itai Benjamini, and Alan Stacey, *Percolation on finite graphs and isoperimetric inequalities*, Ann. Probab. **32** (2004), no. 3, 1727–1745. MR-2073175

[9] 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. MR-3522292

[10] Shankar Bhamidi, Nicolas Broutin, Sanchayan Sen, and Xuan Wang, *Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős-Rényi random graph*, (2014). arXiv:1411.3417

[11] Shankar Bhamidi, Amarjit Budhiraja, and Xuan Wang, *The augmented multiplicative coalescent and critical dynamic random graph models*, Probab. Theory Relat. Fields **160** (2014), no. 3, 733–796. DOI:10.1007/s00440-013-0540-x

[12] Shankar Bhamidi, Souvik Dhara, Remco van der Hofstad, and Sanchayan Sen, *Global lower mass-bound for critical configuration models in the heavy-tailed regime*, In Preparation.

[13] Shankar Bhamidi and Sanchayan Sen, *Geometry of the vacant set left by random walk on random graphs, Wright's constants, and critical random graphs with prescribed degrees*, Random Struct. Algor. (2019). DOI:10.1002/rsa.20880

[14] Shankar Bhamidi, Sanchayan Sen, and Xuan Wang, *Continuum limit of critical inhomogeneous random graphs*, Probab. Theory Relat. Fields **169** (2017), no. 1, 565–641. MR-3704776

[15] 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. MR-3748328

[16] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden, *Scaling limits for critical inhomogeneous random graphs with finite third moments*, Electron. J. Probab. **15** (2010), no. 6, 1682–1702. MR-2735378

[17] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden, *Novel scaling limits for critical inhomogeneous random graphs*, Ann. Probab. **40** (2012), no. 6, 2299–2361. MR-3050505

[18] 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. MR-0595929

[19] Béla Bollobás, *Random Graphs*, Cambridge University Press, Cambridge, 2001. DOI: 10.1017/CBO9780511814068

[20] Béla Bollobás, Svante Janson, and Oliver Riordan, *The phase transition in inhomogeneous random graphs*, Random Struct. Algor. **31** (2007), no. 1, 3–122. MR-2337396

[21] Lidia A Braunstein, Sergey V Buldyrev, Reuven Cohen, Shlomo Havlin, and H Eugene Stanley, *Optimal paths in disordered complex networks*, Phys. Rev. Lett. **91** (2003), no. 16, 168701. DOI: 10.1103/PhysRevLett.91.168701

[22] Lidia A Braunstein, Zhenhua Wu, Yiping Chen, Sergey V Buldyrev, Tomer Kalisky, Sameet Sreenivasan, Reuven Cohen, Eduardo López, Shlomo Havlin, and H Eugene Stanley, *Optimal path and minimal spanning trees in random weighted networks*, Int. J. Bifurc. Chaos **17** (2007), no. 07, 2215–2255. MR-2349740

[23] Tom Britton, Maria Deijfen, and Anders Martin-Löf. *Generating simple random graphs with prescribed degree distribution*, J. Stat. Phys. **124** (2006), vol. 6, 1377–1397. MR-2266448

[24] Nicolas Broutin, Thomas Duquesne, and Minmin Wang, *Limits of multiplicative inhomogeneous random graphs and Lévy trees*, (2018). arXiv:1804.05871

[25] 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. MR-1741774

[26] Fan Chung and Linyuan Lu, *Connected components in random graphs with given expected degree sequences*, Ann. Comb. **6** (2002), no. 2, 125–145. MR-1955514

[27] Fan Chung and Linyuan Lu. *The average distances in random graphs with given expected degrees*, Proc. Natl. Acad. Sci. USA, **99** (2002), no. 25, 15879–82. MR-1944974

[28] Reuven Cohen, Daniel Ben-Avraham, and Shlomo Havlin, *Percolation critical exponents in scale-free networks*, Phys. Rev. E **66** (2002), no. 3, 36113. DOI: 10.1103/PhysRevE.66.036113

[29] Guillaume Conchon-Kerjan and Christina Goldschmidt, *The stable graph: the metric space scaling limits of a critical random graph with i.i.d. power-law degrees*, (2017). arXiv:2002.04954

[30] Souvik Dhara, *PhD Thesis: Critical Percolation on Random Networks with Prescribed Degrees*, Technische Universiteit Eindhoven, (2018). arXiv:1809.03634

[31] Souvik Dhara, Remco van der Hofstad, Johan S. H. van Leeuwaarden, and Sanchayan Sen, *Heavy-tailed configuration models at criticality*, To appear in Ann. Inst. H. Poincaré (B) Probab. Statist. (2016). arXiv:1612.00650

[32] Souvik Dhara, Remco van der Hofstad, Johan S. H. van Leeuwaarden, and Sanchayan Sen, *Critical window for the configuration model: finite third moment degrees*, Electron. J. Probab. **22** (2017), no. 16, 1–33. MR-3622886

[33] Stewart N. Ethier and Thomas G. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley & Sons, Inc., New Jersey, 1986.

[34] Steven N Evans, *Probability and real trees*, École d'Été de Probabilités de Saint-Flour XXXV-2005 (2006), no. 1920. MR-2351587

[35] Nikolaos Fountoulakis, *Percolation on sparse random graphs with given degree sequence*, Internet Math. **4** (2007), no. 1, 329–356. MR-2522948

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

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

[38] 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. MR-2520129

[39] Mikhail Leonidovich Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Translated by Bates, S.M., Birkhäuser Basel, 2007. MR-2307192

[40] Hamed Hatami and Michael Molloy, *The scaling window for a random graph with a given degree sequence*, Random Struct. Algor. **41** (2012), no. 1, 99–123. MR-2943428

[41] Shlomo Havlin, Lidia A. Braunstein, Sergey V. Buldyrev, Reuven Cohen, Tomer Kalisky, Sameet Sreenivasan, and H. Eugene Stanley, *Optimal path in random networks with disorder: A mini review*, Physica A: Statistical Mechanics and its Applications **346** (2005), no. 1-2, 82–92. DOI: 10.1016/j.physa.2004.08.053

[42] Remco van der Hofstad, *Critical behavior in inhomogeneous random graphs*, Random Struct. Algor. **42** (2013), 480–508. MR-3068034

[43] Remco van der Hofstad, *Random Graphs and Complex Networks*, vol. I, Cambridge university press, (2016). MR-3617364

[44] Remco van der Hofstad, *Stochastic Processes on Random Graphs*, Lecture notes for the 47th Summer School in Probability Saint-Flour 2017, (2017). *In preparation.*

[45] Remco van der Hofstad, Svante Janson, and Malwina Luczak, *Component structure of the configuration model: barely supercritical case*, Random Struct. Algor. (2016). DOI: 10.1002/rsa.20837

[46] Svante Janson, *On percolation in random graphs with given vertex degrees*, Electron. J. Probab. **14** (2009), 87–118. MR-2471661

[47] Svante Janson, *The probability that a random multigraph is simple*, Comb. Probab. Comp. **18** (2009), no. 1-2, 205–225. MR-2497380

[48] Svante Janson, *Susceptibility of random graphs with given vertex degrees*, J. Combin. **1** (2010), no. 3-4, 357–387. MR-2799217

[49] Svante Janson, Donald E Knuth, Tomasz Łuczak, and Boris Pittel, *The birth of the giant component*, Random Struct. Algor. **4** (1993), no. 3, 233–358. MR-1220220

[50] Svante Janson and Malwina J. Luczak, *A new approach to the giant component problem*, Random Struct. Algor. **34** (2009), no. 2, 197–216. MR-2490288

[51] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski, *Random Graphs.* Wiley, New York, 2000. DOI: 10.1002/9781118032718

[52] Adrien Joseph, *The component sizes of a critical random graph with given degree sequence*, Ann. Appl. Probab. **24** (2014), no. 6, 2560–2594. MR-3262511

[53] M. Kang and T. G. Seierstad, *The critical phase for random graphs with a given degree sequence*, Comb. Probab. Comp. **17** (2008), no. 01, 67–86. MR-2376424

[54] Jean François Le Gall, *Random trees and applications*, Probab. Surveys **2** (2005), 245–311. DOI: 10.1214/154957805100000140

[55] Robert H Martin, *Nonlinear operators and differential equations in Banach spaces*, Krieger Publishing Co., Inc., Melbourne, FL, 1986.

[56] 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. MR-1370952

[57] Asaf Nachmias and Yuval Peres, *Critical percolation on random regular graphs*, Random Struct. Algor. **36** (2010), no. 2, 111–148. MR-2583058

[58] Asaf Nachmias and Yuval Peres, *The critical random graph, with martingales*, Israel J. Math. **176** (2010), no. 1, 29–41. MR-2653185

[59] Jim Pitman, *Random mappings, forests, and subsets associated with Abel-Cayley-Hurwitz multinomial expansions*, Séminaire Lotharingien de Combinatoire **46** (2001), 02. MR-1877634

[60] Oliver Riordan, *The phase transition in the configuration model*, Comb. Probab. Comp. **21** (2012), 265–299. MR-2900063

[61] L. C. G. Rogers and David Williams, *Diffusions, Markov processes, and Martingales*, 2nd ed., vol. 1, John Wiley & Sons, Ltd., Chichester, (1994). MR-1780932

[62] Ward Whitt, *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*, Springer-Verlag, New York, (2002). DOI: 10.1007/b97479

[63] Nicholas C Wormald, *Differential equations for random processes and random graphs*, Ann. Appl. Probab. (1995), 1217–1235. MR-1384372

A Rescaling the excursions: Proof of Proposition 4.3

Note that due to the difference in the choice of $p_n(\lambda)$ in [31, Assumption 2] and this paper, λ must be replaced by $\lambda\nu$. Let $\mathcal{E}(\cdot)$ denote the operator that maps a process to its ordered vector of excursion lengths, and $\mathcal{A}(\cdot)$ maps a process to the vector of areas under those excursions. Let us use $\mathrm{Exp}(b)$ as a generic notation to write an exponential random variable with rate b. Now,

$$
\frac{1}{\sqrt{\nu}}\mathcal{E}\left(\sum_{i\geq 1} \frac{\theta_i}{\sqrt{\nu}}\left(\mathbb{1}\left\{ \mathrm{Exp}(\theta_i/(\mu\sqrt{\nu})) \leq t \right\} - (\theta_i/(\mu\sqrt{\nu}))t \right) + \lambda\nu t \right)
$$

$$
\stackrel{d}{=} \frac{1}{\nu}\mathcal{E}\left(\sum_{i\geq 1} \frac{\theta_i}{\sqrt{\nu}}\left(\mathbb{1}\left\{ \mathrm{Exp}(\theta_i/(\mu\nu)) \leq u \right\} - (\theta_i/(\mu\nu))u \right) + \lambda u\sqrt{\nu} \right) \qquad \text{(A.1)}
$$

$$
\stackrel{d}{=} \frac{1}{\nu}\mathcal{E}\left(\sum_{i\geq 1} \frac{\theta_i}{\mu\nu}\left(\mathbb{1}\left\{ \mathrm{Exp}(\theta_i/(\mu\nu)) \leq u \right\} - (\theta_i/(\mu\nu))u \right) + \frac{\lambda}{\mu}u \right),
$$

where the last step follows by rescaling the space by $\mu\sqrt{\nu}$ and noting that the rescaling of space does not affect excursion lengths. Again,

$$
\mathcal{A}\left(\sum_{i\geq 1}\frac{\theta_i}{\mu\sqrt{\nu}}\left(\mathbb{1}\left\{\mathrm{Exp}(\theta_i/(\mu\sqrt{\nu}))\leq t\right\}-(\theta_i/(\mu\sqrt{\nu}))t\right)+\frac{\lambda\nu}{\mu}t\right)
$$
$$
\stackrel{d}{=}\mathcal{A}\left(\sum_{i\geq 1}\frac{\theta_i}{\mu\nu}\left(\mathbb{1}\left\{\mathrm{Exp}(\theta_i/(\mu\nu))\leq u\right\}-(\theta_i/(\mu\nu))t\right)+\frac{\lambda}{\mu}u\right),
\tag{A.2}
$$

which is obtained by rescaling both the space and time by $\sqrt{\nu}$. Thus, the proof follows.

B Barely subcritical exploration process: Proofs of Lemmas 6.5 and 6.6

Proof of Lemma 6.3. Recall the representation of $\bar{S}_n^j(t)$. It is enough to show that

$$
\sup_{t\in[0,T]}n^{-\alpha}\left|\sum_{i\in[n]}d_i'\left(\mathcal{I}_i^n(tn^{\alpha+\delta})-\frac{d_i'}{\ell_n'}tn^{\alpha+\delta}\right)\right|=\sup_{t\in[0,T]}n^{-\alpha}|M_n(tn^{\alpha+\delta})|\stackrel{\mathbb{P}}{\to}0.
\tag{B.1}
$$

Fix any $T>0$ and define $\ell_n'(T)=\ell_n'-2Tn^{\alpha+\delta}-1$, and $M_n'(l)=\sum_{i\in[n]}d_i'(\mathcal{I}_i^n(l)-(d_i/\ell_n'(T))l)$. Note that

$$
\sup_{t\in[0,T]}n^{-\alpha}|M_n(tn^{\alpha+\delta})-M_n'(tn^{\alpha+\delta})|\leq Tn^{\delta}\frac{(2Tn^{\alpha+\delta}-1)\sum_{i\in[n]}d_i'^2}{\ell_n'(T)^2}=o_{\mathbb{P}}(1),
\tag{B.2}
$$

and thus the proof reduces to showing that

$$
\sup_{t\in[0,T]}n^{-\alpha}|M_n'(tn^{\alpha+\delta})|\stackrel{\mathbb{P}}{\to}0.
\tag{B.3}
$$

Note that, uniformly over $l\leq Tn^{\alpha+\delta}$,

$$
\mathbb{P}\left(\mathcal{I}_i^n(l+1)=1\big|\mathscr{F}_l\right)\leq\frac{d_i'}{\ell_n'(T)}\quad\text{on the set }\{\mathcal{I}_i^n(l)=0\}.
\tag{B.4}
$$

Therefore,

$$
\mathbb{E}\left[M_n'(l+1)-M_n'(l)\big|\mathscr{F}_l\right]
$$
$$
=\mathbb{E}\left[\sum_{i\in[n]}n^{-\alpha}d_i'\left(\mathcal{I}_i^n(l+1)-\mathcal{I}_i^n(l)-\frac{d_i'}{\ell_n'(T)}\right)\bigg|\mathscr{F}_l\right]
$$
$$
=\sum_{i\in[n]}n^{-\alpha}d_i'\left(\mathbb{E}\left[\mathcal{I}_i^n(l+1)\big|\mathscr{F}_l\right]\mathbb{1}\left\{\mathcal{I}_i^n(l)=0\right\}-\frac{d_i'}{\ell_n'(T)}\right)\leq 0.
$$

Thus $(M_n'(l))_{l=1}^{Tn^{\alpha+\delta}}$ is a super-martingale. Further, uniformly for all $l\leq Tn^{\alpha+\delta}$,

$$
\mathbb{P}\left(\mathcal{I}_i^n(l)=0\right)\leq\left(1-\frac{d_i'}{\ell_n'}\right)^l.
\tag{B.5}
$$

Thus, Assumption 2 gives

$$
n^{-\alpha}\big|\mathbb{E}[M_n'(l)]\big|\leq n^{-\alpha}\sum_{i\in[n]}d_i'\left(1-\left(1-\frac{d_i'}{\ell_n'}\right)^l-\frac{d_i'}{\ell_n'}l\right)+n^{-\alpha}l\sum_{i\in[n]}d_i'^2\left(\frac{1}{\ell_n'(T)}-\frac{1}{\ell_n'}\right)
$$
$$
\leq\frac{l^2}{2\ell_n'^2n^{\alpha}}\sum_{i\in[n]}d_i'^3+o(1)=o(1),
$$

where we have used the fact that $n^{-\alpha}l\sum_{i\in[n]}d_i'^2(1/\ell_n'(T)-1/\ell_n')=O(n^{2\rho+1-\alpha-2})=O(n^{(\tau-4)/(\tau-1)})$, uniformly for $l\leq Tn^{\alpha+\delta}$ and, in the last step, that fact that $\delta<\eta$. Therefore, uniformly over $l\leq Tn^{\alpha+\delta}$,

$$\lim_{n\to\infty}\left|\mathbb{E}[M_n'(l)]\right|=0. \tag{B.6}$$

Now, note that for any $(x_1,x_2,\dots)$, $0\leq a+b\leq x_i$ and $a,b>0$ one has $\prod_{i=1}^R(1-a/x_i)(1-b/x_i)\geq\prod_{i=1}^R(1-(a+b)/x_i)$. Thus, for all $l\geq1$ and $i\neq j$,

$$\mathbb{P}\left(\mathcal{I}_i^n(l)=0,\mathcal{I}_j^n(l)=0\right)\leq\mathbb{P}\left(\mathcal{I}_i^n(l)=0\right)\mathbb{P}\left(\mathcal{I}_j^n(l)=0\right) \tag{B.7}$$

and therefore $\mathcal{I}_i^n(l)$ and $\mathcal{I}_j^n(l)$ are negatively correlated. Observe also that, uniformly over $l\leq Tb_n$,

$$\mathrm{Var}\left(\mathcal{I}_i^n(l)\right)\leq\mathbb{P}\left(\mathcal{I}_i^n(l)=1\right)\leq\sum_{l_1=1}^l\mathbb{P}\left(\text{vertex }i\text{ is first discovered at stage }l_1\right)\leq\frac{ld_i'}{\ell_n'(T)}. \tag{B.8}$$

Therefore, using the negative correlation in (B.7), uniformly over $l\leq Tn^{\alpha+\delta}$,

$$n^{-2\alpha}\mathrm{Var}\left(M_n'(l)\right)\leq\frac{l}{\ell_n'(T)n^{2\alpha}}\sum_{i\in[n]}d_i'^3=o(1). \tag{B.9}$$

Now we can use the super-martingale inequality [61, Lemma 2.54.5] stating that for any super-martingale $(M(t))_{t\geq0}$, with $M(0)=0$,

$$\varepsilon\mathbb{P}\left(\sup_{s\leq t}|M(s)|>3\varepsilon\right)\leq3\mathbb{E}\left[|M(t)|\right]\leq3\left(\left|\mathbb{E}\left[M(t)\right]\right|+\sqrt{\mathrm{Var}\left(M(t)\right)}\right). \tag{B.10}$$

Thus (B.3) follows using (B.6), (B.9), and (B.10). $\qquad\square$

Proof of Lemma 6.4. Fix any $T>0$ and recall that $\ell_n(T)=\ell_n'-2Tn^{\alpha+\delta}-1$. Denote $W(l)=\sum_{i\in[n]}w_i\mathcal{I}_i^n(l)$. Firstly, observe that

$$\mathbb{E}[W(l+1)-W(l)|\mathscr{F}_l]=\sum_{i\in[n]}w_i\mathbb{E}\left[\mathcal{I}_i^n(l+1)\big|\mathscr{F}_l\right]\mathbb{1}\left\{\mathcal{I}_i^n(l)=0\right\}\leq\frac{\sum_{i\in[n]}d_i'w_i}{\ell_n'(T)},$$

uniformly over $l\leq Tn^{\alpha+\delta}$. Therefore, $(\tilde{W}(l))_{l=1}^{Tn^{\alpha+\delta}}$ is a super-martingale, where $\tilde{W}(l)=W(l)-(\sum_{i\in[n]}d_i'w_i/\ell_n')l$. Again, the goal is to use (B.10). Using (B.5), we can show that $\left|\mathbb{E}[\tilde{W}(l)]\right|=o(n^{\alpha+\delta})$, uniformly over $l\leq Tn^{\alpha+\delta}$. Also, using (B.7) and (B.8) and Assumption 2.9, $\mathrm{var}(\tilde{W}(l))\leq\sum_{i\in[n]}w_i^2\mathrm{var}(\mathcal{I}_i^n(l))=o(n^{2(\alpha+\delta)})$, uniformly over $l\leq Tn^{\alpha+\delta}$. Finally, using (B.10), we conclude the proof. $\qquad\square$

C Barely subcritical exploration process: Proof of Fact 6.3

Consider exploring $\mathrm{CM}_n(\boldsymbol{d}')$ by Algorithm 3 with V_n^* being the starting vertex. 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 upto time l. Thus,

$$S_n(0)=d_{V_n^*}',\quad\text{and}\quad S_n(l)=S_n(l-1)+(d_{(l)}'-2), \tag{C.1}$$

and when $\boldsymbol{S}_n$ hits zero, then $\mathscr{C}'(V_n^*)$ has been explored. Using the Doob-Meyer decomposition, one can write

$$S_n(l)=S_n(0)+M_n(l)+A_n(l), \tag{C.2}$$

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

$$A_n(l) = \sum_{j=1}^{l} \mathbb{E}\big[d'_{(j)} - 2 | \mathscr{F}_{j-1}\big], \qquad \langle M_n \rangle(l) = \sum_{j=1}^{l} \mathrm{Var}\big(d'_{(j)} | \mathscr{F}_{j-1}\big). \tag{C.3}$$

Let $t_n = n^{\alpha + \delta + c_0}$, where we choose $c_0 > 0$ (sufficiently small) such that $\alpha + \delta + c_0 < 1 - \delta$. Such a choice of c_0 is always possible since $\alpha + 2\delta < \alpha + 2\eta = 2 - 3\alpha < 1$ as $3\alpha > 1$. We will show that, for all sufficiently large n,

$$A_n(t_n) \leq -\frac{\lambda_0}{2} t_n n^{-\delta} \quad \text{almost surely,} \tag{C.4}$$

where λ_0 is given by (2.9), and for any $\varepsilon > 0$,

$$\mathbb{P}(M_n(t_n) > \varepsilon t_n n^{-\delta}) \leq C' e^{-C'' \varepsilon^2 n^{\varepsilon_0}}, \tag{C.5}$$

where $C', C'' > 0$ are constants. Note that $S_n(0) = d_{V_n^*} \leq Cn^\alpha = o(t_n n^{-\delta})$. Thus, if $M_n(t_n) \leq \varepsilon t_n n^{-\delta}$ and (C.4) holds, then $S_n(t_n) < 0$ and hence $\mathscr{C}'(V_n^*)$ is explored before time t_n. This in turn implies that $\sum_{i \in \mathscr{C}'(V_n^*)} d'_i \leq 2t_n$, since one edge is explored per step. Therefore it is enough to prove (C.4) and (C.5). Let $\mathscr{V}_j$ denote the set of vertices explored up to time j. Recall that, by the definition of ν'_n in (2.9), $\frac{1}{\ell'_n} \sum_{i \in [n]} d'^2_i - 2 = \lambda_0 n^{-\delta} + o(n^{-\delta})$. Then, uniformly over $j \leq t_n$,

$$\begin{aligned}
\mathbb{E}\big[d'_{(j)} - 2 | \mathscr{F}_{j-1}\big] = \mathbb{E}\big[d'_{(j)} | \mathscr{F}_{j-1}\big] - 2 &\leq \frac{\sum_{i \notin \mathscr{V}_{j-1}} d'^2_i}{\ell'_n - 2t_n + 1} - 2 \\
&\leq \frac{\sum_{i \in [n]} d'^2_i}{\ell'_n - 2t_n + 1} - 2 = \frac{\ell_n(2 - \lambda_0 n^{-\delta} + o(n^{-\delta}))}{\ell'_n - 2t_n + 1} - 2 \\
&= -\lambda_0 n^{-\delta} + O(t_n/n) + o(n^{-\delta}) \leq -\frac{\lambda_0}{2} n^{-\delta},
\end{aligned} \tag{C.6}$$

for all sufficiently large n, where in the final step we have used the fact that $t_n = o(n^{1-\delta})$. Thus (C.4) follows.

To prove (C.5), we use Freedman's inequality [36, 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 $a, b > 0$,

$$\mathbb{P}(Y(k) \geq a, \text{ and } \langle Y \rangle(k) \leq b) \leq \exp\left(-\frac{a^2}{2(Ra + b)}\right). \tag{C.7}$$

Note that, uniformly over $j \leq t_n$,

$$\mathrm{Var}\big(d'_{(j)} | \mathscr{F}_{j-1}\big) \leq \mathbb{E}[d'^2_{(j)} | \mathscr{F}_{j-1}] = \frac{\sum_{j \notin \mathscr{V}_{j-1}} d'^3_j}{\ell'_n - 2t_n + 1} \leq \frac{\sum_{j \in [n]} d'^3_j}{\ell'_n - 2t_n + 1} \leq Cn^{3\alpha - 1}, \tag{C.8}$$

so that, almost surely,

$$\langle M_n \rangle(t_n) \leq Ct_n n^{3\alpha - 1}. \tag{C.9}$$

Also, $d_{(j)} \leq Cn^\alpha$ almost surely. Thus, applying (C.7) with $a = \varepsilon t_n n^{-\delta}$, $b = Ct_n n^{3\alpha - 1}$ and $R = Cn^\alpha$, and also using (C.9), it follows that

$$\mathbb{P}(M_n(t_n) > \varepsilon t_n n^{-\delta}) \leq \exp\left(-C' \frac{\varepsilon^2 t_n^2 n^{-2\delta}}{2(t_n n^{3\alpha - 1} + \varepsilon n^\alpha t_n n^{-\delta})}\right) \leq C' e^{-C'' \varepsilon n^{\varepsilon_0}}, \tag{C.10}$$

where in the last step we have used the fact that $\alpha - \delta > \alpha - \eta = 3\alpha - 1$ and $t_n = n^{\alpha + \delta + \varepsilon_0}$. Thus the proof of (C.5) follows. $\qquad \square$

D Limit of exploration process: Proof sketch for Proposition 7.8

The proof of Proposition 7.8 can be carried out using similar ideas as [31, Theorem 8]. The key idea to prove Proposition 7.8 is that the scaling limit is governed by the vertices having large degrees only. More precisely, for any $\varepsilon > 0$ and $T > 0$,

$$\lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\sup_{t \leq T} n^{-\alpha} \left| \sum_{i > K} a_i \left(\mathcal{I}_i^n(tn^{\rho-\delta}) - \frac{a_i}{\ell_n^a} tn^{\rho-\delta} \right) \right| > \varepsilon \right) = 0. \tag{D.1}$$

This can be proved using martingale estimates, see [31, Section 4]. Thus, if one considers the truncated sum

$$\sum_{i \leq K} a_i \left(\mathcal{I}_i^n(l) - \frac{a_i}{\ell_n^a} l \right) + (\nu_n(\boldsymbol{a}) - 1) \, l,$$

with the first K (fixed) terms it is enough to show that the iterated limit of the truncated process (first taking $\lim_{n \to \infty}$ and then $\lim_{K \to \infty}$) converges to $\mathbf{S}$ with respect to the Skorohod J_1 topology. Now, using the fact that $a_i / \sum_i a_i \xrightarrow{\mathbb{P}} \theta_i/(\mu\nu)$, and the fact that the vertices are explored in a size-biased manner with sizes being $(a_i)_{i \geq 1}$, it follows that (see [31, Lemma 9]), for each fixed $K \geq 1$,

$$\left(\mathcal{I}_i^n(tn^{\rho-\delta}) \right)_{i \in [K], t \geq 0} \xrightarrow{d} \left(\mathbb{1}\left\{ \mathrm{Exp}(\theta_i/(\mu\nu)) \leq t \right\} \right)_{i \in [K], t \geq 0}. \tag{D.2}$$

This concludes the proof of Proposition 7.8.

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 http://www.vtex.lt/en/ejms.html
[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: http://www.imstat.org/publications/open.htm