# Geometry of the minimal spanning tree of a random 3-regular graph 

Louigi Addario-Berry ${ }^{1} \cdot$ Sanchayan Sen ${ }^{2}$ (D)

Received: 7 June 2019 / Revised: 30 May 2021 / Accepted: 4 June 2021 / Published online: 12 June 2021 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021


#### Abstract

The global structure of the minimal spanning tree (MST) is expected to be universal for a large class of underlying random discrete structures. However, very little is known about the intrinsic geometry of MSTs of most standard models, and so far the scaling limit of the MST viewed as a metric measure space has only been identified in the case of the complete graph (Addario-Berry et al. in Ann Probab 45(5):3075-3144, 2017). In this work, we show that the MST constructed by assigning i.i.d. continuous edge weights to either the random (simple) 3-regular graph or the 3-regular configuration model on $n$ vertices, endowed with the tree distance scaled by $n^{-1 / 3}$ and the uniform probability measure on the vertices, converges in distribution with respect to Gromov-Hausdorff-Prokhorov topology to a random compact metric measure space. Further, this limiting space has the same law as the scaling limit of the MST of the complete graph identified in Addario-Berry et al. (2017) up to a scaling factor of $6^{1 / 3}$. Our proof relies on a novel argument that proceeds via a comparison between a 3-regular configuration model and the largest component in the critical Erdős-Rényi random graph. The techniques of this paper can be used to establish the scaling limit of the MST in the setting of general random graphs with given degree sequences provided two additional technical conditions are verified.


Keywords Minimal spanning tree • Gromov-Hausdorff distance • Critical percolation • Real tree • Random regular graphs • Graphs with prescribed degree sequence • Configuration model

Mathematics Subject Classification Primary 60C05 - 05C80

[^0]
## Contents

1 Introduction ..... 554
1.1 Organization of the paper ..... 556
1.2 Random graph models ..... 557
2 Main results ..... 558
3 Definitions and various scaling limits ..... 559
3.1 Notation and conventions ..... 559
3.2 Topologies on the space of metric spaces ..... 560
3.3 Scaling limits of component sizes at criticality ..... 562
3.4 Real trees and $\mathbb{R}$-graphs ..... 563
3.5 Geometry of critical random graphs ..... 566
4 Properties of minimal spanning trees ..... 569
4.1 MST and percolation ..... 569
4.2 Cycle-breaking and modified cycle-breaking ..... 570
4.3 Alternate descriptions of the space $\mathscr{M}$ ..... 572
5 Idea of proof ..... 573
6 Proofs of Theorems 2.1, 2.2, 2.3, and 4.8 ..... 575
6.1 GH convergence of the MST of $\mathscr{G}_{n, 3}$ ..... 576
6.2 Proof of Proposition 6.2 ..... 579
6.3 Proof of Proposition 6.3 ..... 579
6.4 Proof of Proposition 6.4 ..... 581
6.5 Proof of Proposition 6.5 ..... 586
6.6 GHP convergence of the MST of $\mathscr{G}_{n, 3}$ ..... 587
6.7 GHP convergence of the MST of $\overline{\mathscr{G}}_{n, 3}$ ..... 595
6.8 Proof of Theorem 2.3 ..... 596
6.9 Proof of Theorem 4.8 ..... 599
7 Discussion ..... 608
Appendix A ..... 612
A. 1 Sketch of proof of Theorem 3.11 ..... 612
A. 2 Sketch of proof of Theorem 3.13 ..... 616
References ..... 617

## 1 Introduction

Consider a finite, connected, and weighted graph $(V, E, w)$, where $(V, E)$ is the underlying graph and $w: E \rightarrow[0, \infty)$ is the weight function. A spanning tree of $(V, E)$ is a tree that is a subgraph of $(V, E)$ with vertex set $V$. A minimal spanning tree (MST) $T$ of $(V, E, w)$ satisfies

$$
\begin{equation*}
\sum_{e \in T} w(e)=\min \left\{\sum_{e \in T^{\prime}} w(e): T^{\prime} \text { is a spanning tree of }(V, E)\right\} . \tag{1.1}
\end{equation*}
$$

The two natural choices for the underlying weighted graph are (i) a deterministic graph (e.g., the complete graph on $n$ vertices or the hypercube) or a random graph (e.g., Erdős-Rényi random graph, random regular graph, or inhomogeneous random graphs [30]) with i.i.d. continuous edge weights assigned to them, and (ii) the complete graph on a finite set of random points in $\mathbb{R}^{d}$ (e.g., $n$ i.i.d. points or a Poisson point process in the unit cube) where the edge weights are some function of the Euclidean length of the edges. The MST in the latter case is sometimes called the Euclidean MST.

The MST is one of the most studied objects in combinatorial optimization and geometric probability and has inspired a large body of work. For an account of law of large numbers and related asymptotics in the Euclidean setting, see e.g., [12,13,17, 19,82]. Central limit theorems (CLT) for the total weight of Euclidean MSTs were first proved by Kesten and Lee [60] and by Alexander [15] in 1996. This was a longstanding open question at the time of its solution. Later certain other CLTs related to MSTs were proved in [62,63]. A question raised in [60] about the convergence rate in the CLT for the total weight of the Euclidean MST was answered in [36].

Studies related to the MST in several other directions were undertaken in [18,28, $77,78,80]$. An account of certain structural and connectivity properties of minimal spanning forests can be found in $[14,16,67,76]$ and the references therein. For an account of the scaling limit of minimal spanning trees in subsets of $\mathbb{Z}^{2}$ with respect to the topology introduced by Aizenman, Burchard, Newman, and Wilson, see, e.g., [7,49].

The MST of $K_{n}$-the complete graph on $[n]:=\{1, \ldots, n\}$ has been studied extensively as well. A celebrated theorem of Frieze [47] shows that under some assumptions on the weight distributions, the total weight of the MST of $K_{n}$ converges in expectation to $\zeta(3)$. Various extensions of this result were proved in [10,21, $46,48,79]$. The central limit theorem for the total weight of the MST of $K_{n}$ constructed using i.i.d. Uniform[0, 1] edge weights was proved in [53].

The global geometric properties of the MST, e.g., the diameter and the typical distance, have also been of considerable interest, but until very recently, there were few rigorous mathematical results on this problem. Frieze and McDiarmid asked a question [45, Research Problem 23] about the 'likely shape of a minimum spanning tree' and the order of the diameter of the MST. In the statistical physics literature, paths in the MST correspond to optimal paths in the so-called strong disorder regime for complex networks. Using empirical observations, it was predicted in [31] (see also [32]) that in the strong disorder regime, the length of optimal paths in complex networks should scale like $n^{1 / 3}$ if the degree distribution of the network has finite third moment, although a rigorous justification of this claim was missing in this work.

An upper bound of the order $n^{1 / 3}$ on the diameter of the MST of $K_{n}$ was proved in [6]: Let $M_{\infty}^{n, \text { er }}$ be $^{1}$ the MST of $K_{n}$ constructed using i.i.d. continuous edge weights, and denote by $\operatorname{diam}\left(M_{\infty}^{n, \text { er }}\right)$ the maximum tree distance between vertices of $M_{\infty}^{n, \text { er }}$. Then $\operatorname{diam}\left(M_{\infty}^{n, e r}\right)=O_{P}\left(n^{1 / 3}\right)$. Nachmias and Peres [74] showed that the diameter of the largest component of the critical Erdős-Rényi random graph is $\Theta_{P}\left(n^{1 / 3}\right)$. There is a natural coupling between MSTs and percolation (see Observation 4.2), which together with the above result gives a matching lower bound :

$$
\begin{equation*}
\operatorname{diam}\left(M_{\infty}^{n, e r}\right)=\Omega_{P}\left(n^{1 / 3}\right) \tag{1.2}
\end{equation*}
$$

Then a stronger result was proved in [5], where the scaling limit of $M_{\infty}^{n, \text { er }}$ viewed as a metric measure space was obtained. We state this result in the following theorem. We refer the reader to Sect. 3.2 for the definition of the Gromov-Hausdorff-Prokhorov topology.

[^1]Theorem 1.1 (Scaling limit of the MST of the complete graph [5]) View $M_{\infty}^{n, e r}$ as a random metric measure space by endowing it with the tree distance and the uniform probability measure on its vertices. Then there exists a random compact metric measure space $\mathscr{M}$ such that

$$
n^{-1 / 3} M_{\infty}^{n, \text { er }} \xrightarrow{d} \mathscr{M}
$$

w.r.t. Gromov-Hausdorff-Prokhorov topology. Further, almost surely, the space $\mathscr{M}$ is a binary real tree and its Minkowski dimension exists and equals 3.

Theorem 1.1 appears to be one of the first scaling limits to be identified for any problem from combinatorial optimization, and so far, the above theorem gives the only result where the metric space scaling limit of the MST has been identified. Several questions about the geometry of $\mathscr{M}$ remain open. For instance, what is the distribution of the typical distance in $\mathscr{M}$ ? More generally, is there a stick-breaking construction of $\mathscr{M}$ ? Is the support of the mass measure $\mu$ on $\mathscr{M}$ the whole of $\mathscr{M}$ ? Since $\mathscr{M}$ is a compact real tree, by [41, Corollary 1.2], the metric space $\mathscr{M}$ (without the measure) is encoded by a random continuous function (see Sect. 3.4). What can we say about the distribution of this function?

The limiting space $\mathscr{M}$ is expected to be a universal object in the following sense: For a wide array of random discrete structures that exhibit mean-field behavior, the MST constructed using i.i.d. continuous edge weights should have a rescaled version of $\mathscr{M}$ as its scaling limit. Examples of such models include the high-dimensional discrete torus, the hypercube, random regular graphs or more generally random graphs with given degree sequence (under finite third moment assumption on the degrees), various models of inhomogeneous random graphs (under appropriate assumptions), boundedsize rules, and the quantum random graph model. See Sect. 7 for a more detailed discussion.

In this work, we take a first step in this broader program of establishing universality of the MST by showing that the above claim is true for the random simple 3-regular graph and the 3-regular configuration model. The core of the largest component of the Erdős-Rényi random graph, in the critical window and also in the barely-supercritical regime up to a certain threshold, can be described by a 3-regular configuration model on a random number of vertices and having random edge lengths (see [56]). This makes the 3-regular case special. We use an indirect approach by exploiting the above coupling between the 3 -regular configuration model and the Erdős-Rényi random graph. However, with two additional technical estimates, our arguments can be extended to establish the scaling limit of the MST of general random graphs with given degree sequences. We refer the reader to Sect. 7 for details.

### 1.1 Organization of the paper

In Sect. 1.2, we describe the random graph models considered in this paper. Section 2 contains precise statements of our main results. We have deferred many definitions to Sect. 3, where we also give the necessary background on results on scaling limits of critical random graph models. The proofs of two results (Theorems 3.11 and 3.13 )
stated in Sect. 3 are outlined in Appendix A. In Sect. 4 we list several properties of MSTs, and describe the connection between MSTs and percolation and the so-called cycle-breaking algorithm. We also state a result (Theorem 4.8) central to our argument. In Sect. 5, we describe some of the ideas used in the proof of Theorem 2.1. The proofs of our main results are given in Sect. 6. In Sect. 7 we discuss the relevance of this work and related open problems.

### 1.2 Random graph models

First we define the classical Erdős-Rényi random graph model. Recall that $K_{n}$ denotes the complete graph on $[n]$.

Definition 1.2 (The Erdős-Rényi process) The Erdős-Rényi process ( $\operatorname{ER}(n, \lambda), \lambda \in$ $\mathbb{R}$ ) is a stochastic process taking values in the space of subgraphs of $K_{n}$ defined as follows: Assign a random variable $U_{i j}$ to each edge $(i, j)$ of $K_{n}$, where $U_{i j}, 1 \leq i<$ $j \leq n$, are i.i.d. Uniform[ 0,1$]$ random variables. $\operatorname{Set} \operatorname{ER}(n, \lambda)$ to be the subgraph of $K_{n}$ whose vertex set is $[n]$ and edge set is $\left\{(i, j): U_{i j} \leq n^{-1}+\lambda n^{-4 / 3}\right\}$.

Remark 1 The Erdős-Rényi process is often defined as a random graph process that is indexed by a parameter $p \in[0,1]$ and takes values in the space of subgraphs of $K_{n}$, where the graph at parameter value $p$ has edge set $\left\{(i, j): U_{i j} \leq p\right\}$. We instead work with the parametrization of Definition 1.2 , as this will be particularly convenient for us.

Now fix a collection of $n$ vertices labeled by $[n]:=\{1,2, \ldots, n\}$ and an associated degree sequence $\mathbf{d}=\left(d_{v}, v \in[n]\right)$ where $\ell_{n}:=\sum_{v \in[n]} d_{v}$ is assumed even. There are two natural constructions resulting in a random graph on $[n]$ with the prescribed degree sequence.

Definition 1.3 (Uniformly distributed simple graphs) Suppose $\mathbf{d}=\left(d_{v}, v \in[n]\right)$ is a given degree sequence. Consider the set of all simple graphs with vertex set [ $n$ ] where vertex $v$ has degree $d_{v}$, and write $\overline{\mathscr{G}}_{n, \mathbf{d}}$ for the random graph having uniform distribution over this set.

When $d_{v}=3$ for all $v \in[n]$, we will denote the corresponding random graph by $\overline{\mathscr{G}}_{n, 3}$. In this case, we assume that $n$ is even.

Recall that a multigraph is a graph where we allow multiple edges and self-loops.
Definition 1.4 (Configuration model $[20,29,73])$ Let $\mathscr{G}_{n, \mathbf{d}}$ be the random multigraph with degree sequence d constructed sequentially as follows: Equip each vertex $v \in[n]$ with $d_{v}$ half-edges or stubs. Initially all half-edges are unpaired, and then sequentially at each step, pick two half-edges uniformly from the set of half-edges that have not yet been paired, and pair them to form a full edge. Repeat till all half-edges have been paired.

When $d_{v}=3$ for all $v \in[n]$, we will denote the corresponding random multigraph by $\mathscr{G}_{n, 3}$. In this case, we assume that $n$ is even.

Note that $\mathscr{G}_{n, \mathbf{d}}$ is not uniformly distributed over the set of multigraphs with degree sequence $\mathbf{d}$. We record the distribution of $\mathscr{G}_{n, \mathbf{d}}$ here for later use. Let $G$ be a multigraph
on vertex set $[n]$ in which there are $x_{i j}$ many edges between $i$ and $j, 1 \leq i<j \leq n$, and vertex $i$ has $x_{i i}$ many loops, so that $d_{i}=x_{i i}+\sum_{j=1}^{n} x_{i j}$ is the total degree of $i$ (note that a loop contributes two to the degree). Let $\ell_{n}=\sum_{i=1}^{n} d_{i}$. Then

$$
\begin{equation*}
\mathbb{P}\left(\mathscr{G}_{n, \mathbf{d}}=G\right)=\frac{1}{\left(\ell_{n}-1\right)!!} \times \frac{\prod_{i \in[n]} d_{i}!}{\prod_{i \in[n]} 2^{x_{i i}} \prod_{1 \leq i \leq j \leq n} x_{i j}!} . \tag{1.3}
\end{equation*}
$$

The proof of (1.3) can be found in [84, Proposition 7.7].

## 2 Main results

In this section we will describe our main results. We first fix some conventions that we will follow throughout this paper.
Convention. (i) For any metric measure space $\boldsymbol{X}=(X, d, \mu)$ and $\alpha>0, \alpha \boldsymbol{X}$ will denote the metric measure space ( $X, \alpha d, \mu$ ), i.e, the space where the metric has been multiplied by $\alpha$ and the measure $\mu$ has remained unchanged. Precise definitions of metric space convergence including the Gromov-Hausdorff-Prokhorov (GHP) topology are deferred to Sect. 3.
(ii) For any finite (not necessarily connected) graph $G$, unless the edge weights are specified, the "MST of $G$ " will mean the (random) minimal spanning tree of the largest component of $G$ obtained by assigning i.i.d. continuous weights to the edges of $G$. It is a standard fact (see Observation 4.1) that the law of the MST constructed using exchangeable edge weights that are almost surely pairwise distinct does not depend on the distribution of the underlying weights. So the above definition of MST of $G$ makes sense.

Recall the definitions of $\overline{\mathscr{G}}_{n, 3}$ and $\mathscr{G}_{n, 3}$ from Sect. 1.2. Our first main result concerns the scaling limit of the MST of $\mathscr{G}_{n, 3}$.

Theorem 2.1 (Scaling limit of the MST of the 3-regular configuration model) For $n$ even, let $M_{n}$ denote the MST of $\mathscr{G}_{n, 3}$. Think of $M_{n}$ as a metric measure space by using the tree distance and the uniform probability measure on the vertices. Let $\mathscr{M}$ be as in Theorem 1.1. Then

$$
n^{-1 / 3} \cdot M_{n} \xrightarrow{d} 6^{1 / 3} \cdot \mathscr{M} \text { as } n \rightarrow \infty
$$

with respect to the Gromov-Hausdorff-Prokhorov topology.
Our next main result concerns the scaling limit of the MST of $\overline{\mathscr{G}}_{n, 3}$.
Theorem 2.2 (Scaling limit of the MST of the simple 3-regular graph) For $n$ even, let $\bar{M}_{n}$ denote the MST of $\overline{\mathscr{G}}_{n, 3}$. Then the result in Theorem 2.1 continues to hold with $\bar{M}_{n}$ in place of $M_{n}$, i.e.,

$$
n^{-1 / 3} \cdot \bar{M}_{n} \xrightarrow{d} 6^{1 / 3} \cdot \mathscr{M} \text { as } n \rightarrow \infty
$$

with respect to the Gromov-Hausdorff-Prokhorov topology.

Remark 2 Let

$$
\begin{equation*}
B_{n}:=\left\{\mathscr{G}_{n, 3} \text { is connected }\right\} . \tag{2.1}
\end{equation*}
$$

By the results of $[43,64]$,

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \mathbb{P}\left(B_{n}\right)=1=\lim _{n \rightarrow \infty} \mathbb{P}\left(\overline{\mathscr{G}}_{n, 3} \text { is connected }\right) \tag{2.2}
\end{equation*}
$$

Thus the conclusions of the two theorems above also hold for $\mathscr{G}_{n, 3}$ and $\overline{\mathscr{G}}_{n, 3}$ conditioned to be connected. Further, the results of Theorems 2.1 and 2.2 remain true if the MST were constructed using exchangeable edge weights that are almost surely pairwise distinct.

Our next result, which is interesting in its own right, is a crucial ingredient in the proofs of the above two theorems.

Theorem 2.3 Almost surely the mass measure $\mu$ on $\mathscr{M}$ is non-atomic, i.e.,

$$
\mathbb{P}(\mu(\{x\})=0 \text { for every } x \in \mathscr{M})=1 .
$$

## 3 Definitions and various scaling limits

### 3.1 Notation and conventions

For any set $A$, we write $|A|$ or $\# A$ for its cardinality and $\mathbb{1}\{A\}$ for the associated indicator function. For any graph $H$, we write $V(H)$ and $E(H)$ for the set of vertices and the set of edges of $H$ respectively. We write $|H|$ for the number of vertices in $H$, i.e., $|H|=|V(H)|$. For any finite connected graph $H=(V, E)$, we write $\operatorname{sp}(H)$ for the number of surplus edges in $H$, i.e.,

$$
\begin{equation*}
\operatorname{sp}(H):=|E|-|V|+1 . \tag{3.1}
\end{equation*}
$$

For any finite multigraph $H=(V, E)$ and $e_{1}, \ldots, e_{k} \in E$, let $H \backslash\left\{e_{1}, \ldots, e_{k}\right\}:=$ $\left(V, E \backslash\left\{e_{1}, \ldots, e_{k}\right\}\right)$. While removing a single edge $e$ we will simply write $H \backslash e$ instead of $H \backslash\{e\}$. Further, denote by Conne $(H)$ the set of all edges $e \in E$ such that $H \backslash e$ is connected. For any finite multigraph $H=(V, E)$ and edges $f_{1}, \ldots, f_{k}$ in the complete graph on $V$, let $H \cup\left\{f_{1}, \ldots, f_{k}\right\}:=\left(V, E \cup\left\{f_{1}, \ldots, f_{k}\right\}\right)$. For two multigraphs $H_{i}=\left(V_{i}, E_{i}\right), i=1,2$, we write $H_{1} \cup H_{2}$ for the multigraph $\left(V_{1} \cup V_{2}, E_{1} \cup E_{2}\right)$. If $H_{2}$ is a connected component of $H_{1}$, then we write $H_{1} \backslash H_{2}$ for the multigraph ( $V_{1} \backslash V_{2}, E_{1} \backslash E_{2}$ ).

For any $u>0, \Gamma_{u}$ will denote a $\operatorname{Gamma}(u, 1)$ random variable. We will write $\Gamma_{u}^{(\alpha)}$, $\alpha \in \Lambda$, to denote i.i.d. $\operatorname{Gamma}(u, 1)$ random variables indexed by the set $\Lambda$.

For any metric space $(X, d)$ and $U \subseteq X$, we define $\operatorname{diam}(U ; X):=\sup \left\{d\left(x_{1}, x_{2}\right):\right.$ $\left.x_{1}, x_{2} \in U\right\}$. We simply write $\operatorname{diam}(U)$ when there is no scope of confusion. For any
$\delta>0$ and $x \in X$, we let $B(x, \delta)=\{y \in X: d(x, y) \leq \delta\}$. For any metric measure space $(X, d, \mu)$, we define $\mathfrak{m}(\delta ; X):=\sup \{\mu(B(x, \delta)): x \in X\}$.

For any tree $\mathbf{t}$ on $[m]$ rooted at $\rho$, we write

$$
\operatorname{ht}(u, \mathbf{t}):=d_{\mathbf{t}}(\rho, u) \text { for } u \in \mathbf{t}, \text { and } \operatorname{ht}(\mathbf{t})=\max _{u \in[m]} \operatorname{ht}(u, \mathbf{t}),
$$

where $d_{\mathbf{t}}(\cdot, \cdot)$ denotes the tree distance on $\mathbf{t}$. If $u \neq \rho$, we write $\overleftarrow{u}^{(1)}$ or simply $\overleftarrow{u}$ for the parent of $u$ in $\mathbf{t}$. If $\overleftarrow{u} \neq \rho$, then $\overleftarrow{u}^{(2)}$ will denote the parent of $\overleftarrow{u}$. Similarly define $\overleftarrow{u}^{(k)}$ for $1 \leq k \leq \operatorname{ht}(u, \mathbf{t})$. We set $\overleftarrow{u}^{(0)}=u$.

For any set $A$ and function $f: A \rightarrow \mathbb{R}$, we let $\|f\|_{\infty}:=\sup _{x \in A}|f(x)|$. We use the standard Landau notation of $o(\cdot), O(\cdot)$ and the corresponding order in probability notation $o_{P}(\cdot), O_{P}(\cdot)$, and $\Theta_{P}(\cdot)$. We use $\xrightarrow{\mathrm{P}}, \xrightarrow{\mathrm{d}}$, and $\xrightarrow{\text { a.s. }}$ to denote convergence in probability, convergence in distribution, and almost sure convergence respectively. We write $\stackrel{\text { d }}{=}$ to mean equality in distribution.

When a graph with edge lengths is viewed as a metric space, the underlying set will be the collection of vertices in the graph joined by line segments (that represent the edges in the graph) of lengths specified by the edge lengths. When not specified, all edge lengths are taken to be one. When a finite connected graph is viewed as a metric measure space, the measure, unless specified otherwise, will be the uniform probability measure on the vertices.

We will work with edge lengths as well as edge weights. To avoid confusion, we make a note here that their roles are completely different. When a graph with edge lengths is viewed as a metric space, the distances are calculated using the edge lengths. In Sect. 4.2, we will define the 'cycle-breaking' process, and edge lengths will be used to perform cycle breaking. On the other hand, edge weights are used to construct the MST (as in (1.1)).

Throughout this paper, $C, C^{\prime}, c, c^{\prime}$ will denote positive universal constants, and their values may change from line to line. Special constants will be indexed as $c_{1}, c_{2}$ etc. We freely omit ceilings and floors when there is little risk of confusion in doing so.

### 3.2 Topologies on the space of metric spaces

We mainly follow [1,5,35,71]. All metric spaces under consideration will be compact. For any compact $(X, d)$ and $A_{1}, A_{2} \subseteq X$, we define the Hausdorff distance between $A_{1}$ and $A_{2}$ to be

$$
d_{H}\left(A_{1}, A_{2}\right):=\inf \left\{\varepsilon>0: A_{1} \subseteq A_{2}^{\varepsilon} \text { and } A_{2} \subseteq A_{1}^{\varepsilon}\right\}
$$

where $A_{1}^{\varepsilon}:=\bigcup_{x \in A_{1}} B(x, \varepsilon)$.
Next we recall the Gromov-Hausdorff distance $d_{\text {GH }}$ between metric spaces. Fix two metric spaces $X_{1}=\left(X_{1}, d_{1}\right)$ and $X_{2}=\left(X_{2}, d_{2}\right)$. For a subset $\Re \subseteq X_{1} \times X_{2}$, the
distortion of $\mathfrak{R}$ is defined as

$$
\operatorname{dis}(\mathfrak{R}):=\sup \left\{\left|d_{1}\left(x_{1}, y_{1}\right)-d_{2}\left(x_{2}, y_{2}\right)\right|:\left(x_{1}, x_{2}\right),\left(y_{1}, y_{2}\right) \in \mathfrak{R}\right\} .
$$

A correspondence $\mathfrak{R}$ between $X_{1}$ and $X_{2}$ is a measurable subset of $X_{1} \times X_{2}$ such that for every $x_{1} \in X_{1}$, there exists at least one $x_{2} \in X_{2}$ such that $\left(x_{1}, x_{2}\right) \in \mathfrak{R}$ and vice-versa. The Gromov-Hausdorff (GH) distance between the two metric spaces ( $X_{1}, d_{1}$ ) and ( $X_{2}, d_{2}$ ) is defined as
$d_{\mathrm{GH}}\left(X_{1}, X_{2}\right)=\frac{1}{2} \inf \left\{\operatorname{dis}(\Re): \mathfrak{R}\right.$ is a correspondence between $X_{1}$ and $\left.X_{2}\right\}$.
Let $\mathfrak{S}_{\mathrm{GH}}$ denote the set of isometry equivalence classes of compact metric spaces endowed with the quotient metric induced by $d_{\mathrm{GH}}$, which we will continue to denote by $d_{\mathrm{GH}}$.

We next define the marked topology; see [71, Sections 6.4 and 6.5] for a more detailed treatment. A marked metric space is a triple $(X, d, Z)$, where $(X, d)$ is a compact metric space and $Z$ is a compact subset of $X$. The isometry classes $[(X, d, Z)]$ of marked spaces are defined in the obvious way, and the set of such isometry classes is denoted by $\mathfrak{S}_{\mathrm{GH}}^{*}$. We put the following metric on $\mathfrak{S}_{\mathrm{GH}}^{*}:$ For $\left[\boldsymbol{X}_{i}\right]=\left[\left(X_{i}, d_{i}, Z_{i}\right)\right] \in$ $\mathfrak{S}_{\mathrm{GH}}^{*}, i=1,2$, define

$$
\begin{equation*}
d_{\mathrm{GH}}^{*}\left(\left[\boldsymbol{X}_{1}\right],\left[\boldsymbol{X}_{2}\right]\right):=\inf _{\phi_{1}, \phi_{2}}\left\{d_{H}\left(\phi_{1}\left(X_{1}\right), \phi_{2}\left(X_{2}\right)\right)+d_{H}\left(\phi_{1}\left(Z_{1}\right), \phi_{2}\left(Z_{2}\right)\right)\right\}, \tag{3.2}
\end{equation*}
$$

where the infimum is taken over all isometric embeddings $\phi_{i}: X_{i} \rightarrow S, i=1,2$, into some metric space $S$. (There is an equivalent definition of the Gromov-Hausdorff distance $d_{\mathrm{GH}}$ that is similar to (3.2); see, e.g., [35, Section 7.3.2].)

The following result is the content of [71, Proposition 9].
Lemma 3.1 (a) The space $\left(\mathfrak{S}_{\mathrm{GH}}^{*}, d_{\mathrm{GH}}^{*}\right)$ is Polish.
(b) A collection $\left\{\left[\left(X_{\alpha}, d_{\alpha}, Z_{\alpha}\right)\right]: \alpha \in \Lambda\right\}$ is relatively compact in $\left(\mathfrak{S}_{\mathrm{GH}}^{*}, d_{\mathrm{GH}}^{*}\right)$ iff $\left\{\left[\left(X_{\alpha}, d_{\alpha}\right)\right]: \alpha \in \Lambda\right\}$ is relatively compact in $\left(\mathfrak{S}_{\mathrm{GH}}, d_{\mathrm{GH}}\right)$, or equivalently, iff the collection of metric spaces $\left\{\left(X_{\alpha}, d_{\alpha}\right): \alpha \in \Lambda\right\}$ is uniformly totally bounded.

To ease notation, we will simply write $\left(X_{1}, d_{1}, Z_{1}\right)$ to denote both the marked metric space and its equivalence class.

A compact metric measure space $(X, d, \mu)$ is a compact metric space $(X, d)$ with an associated finite measure $\mu$ on the Borel sigma algebra of $X$. We will use the Gromov-Hausdorff-Prokhorov (GHP) distance to compare compact metric measure spaces. Given two compact metric measure spaces $\left(X_{1}, d_{1}, \mu_{1}\right)$ and $\left(X_{2}, d_{2}, \mu_{2}\right)$ and a measure $\pi$ on the product space $X_{1} \times X_{2}$, the discrepancy of $\pi$ with respect to $\mu_{1}$ and $\mu_{2}$ is defined as

$$
D\left(\pi ; \mu_{1}, \mu_{2}\right):=\left\|\mu_{1}-\pi_{1}\right\|+\left\|\mu_{2}-\pi_{2}\right\|,
$$

where $\pi_{1}, \pi_{2}$ are the marginals of $\pi$ and $\|\cdot\|$ denotes the total variation of signed measures. Then define the distance $d_{\mathrm{GHP}}\left(X_{1}, X_{2}\right)$ by

$$
d_{\mathrm{GHP}}\left(X_{1}, X_{2}\right):=\inf \left\{\max \left(\frac{1}{2} \operatorname{dis}(\Re), D\left(\pi ; \mu_{1}, \mu_{2}\right), \pi\left(\Re^{c}\right)\right)\right\},
$$

where the infimum is taken over all correspondences $\mathfrak{R}$ and measures $\pi$ on $X_{1} \times X_{2}$.
The function $d_{\mathrm{GHP}}$ is a pseudometric and defines an equivalence relation: $X \sim$ $Y \Leftrightarrow d_{\mathrm{GHP}}(X, Y)=0$. Let $\mathfrak{S}_{\mathrm{GHP}}$ be the set of all equivalence classes of compact metric measure spaces. As before, we continue to denote the quotient metric by $d_{\mathrm{GHP}}$. Then by [1], $\left(\mathfrak{S}_{\mathrm{GHP}}, d_{\mathrm{GHP}}\right)$ is a complete separable metric space. As before, to ease notation, we will continue to use ( $X, d, \mu$ ) to denote both the metric space and the corresponding equivalence class.

Sometimes we will be interested in not just one but an infinite sequence of compact metric measure spaces. Then the relevant space will be $\mathfrak{S}_{\mathrm{GHP}}^{\mathbb{N}}$ equipped with the product topology inherited from $d_{\mathrm{GHP}}$.

### 3.3 Scaling limits of component sizes at criticality

As we will see in the course of our proof, a key step in understanding the geometry of the MST in the supercritical regime is obtaining the metric space scaling limit of the random graph model in the critical window. The starting point for establishing the metric space scaling limit of critical random graph models is understanding the behavior of their component sizes. Aldous [11] studied the maximal components of the Erdős-Rényi random graph in the critical regime and proved the following remarkable result. Recall the notation $\operatorname{sp}(\cdot)$ from (3.1).

Theorem 3.2 [11, Corollary 2] Write $\mathscr{C}_{i}^{\text {n,er }}(\lambda)$ for the $i$-th largest connected component of $\operatorname{ER}(n, \lambda)$. Then there exists a random sequence $\zeta(\lambda)=\left(\left(\xi_{i}(\lambda), N_{i}(\lambda)\right), i \geq 1\right)$ such that as $n \rightarrow \infty$,

$$
\left(\left(n^{-2 / 3}\left|\mathscr{C}_{i}^{n, \text { er }}(\lambda)\right|, s p\left(\mathscr{C}_{i}^{n, \text { er }}(\lambda)\right)\right) ; i \geq 1\right) \xrightarrow{d} \zeta(\lambda)
$$

with respect to product topology.
This convergence in fact holds w.r.t. a stronger topology. We refer the reader to [11] for an explicit description of the limiting sequence $\zeta(\lambda)$. We record here a result about the asymptotic growth of the random variables $\xi_{1}(\lambda)$ and $N_{1}(\lambda)$.

Lemma 3.3 We have, as $\lambda \rightarrow \infty$,

$$
\frac{\xi_{1}(\lambda)}{\lambda} \xrightarrow{d} 2, \quad \text { and } \quad \frac{N_{1}(\lambda)}{\lambda^{3}} \xrightarrow{d} \frac{2}{3} .
$$

The proof of this result can be found in [5, Lemma 5.6]. (See also [4] for the analogue of this result for the multiplicative coalescent in the regime where the scaling limit is a pure-jump process.)

Theorem 3.2 has since been generalized to a number of other random graph models. In the context of graphs with given degree sequence, Nachmias and Peres [75] studied critical percolation on random regular graphs; Riordan [81] analyzed the configuration model with bounded degrees; Joseph [59] considered i.i.d. degrees. A more general result was obtained in [39]. We will state a weaker version of this result next.

For a measure $v$ on $\mathbb{R}$ and $p>0$, write $\sigma_{p}(\nu)=\int_{\mathbb{R}}|x|^{p} \mathrm{~d} \nu$; if $v$ has support $\mathbb{Z}_{\geq 0}$ then $\sigma_{p}(\nu)=\sum_{i \geq 0} i^{p} \nu(i)$. Recall that $v_{n} \rightarrow v$ w.r.t. the Wasserstein distance $W_{p}$ if $v_{n} \rightarrow \nu$ weakly and $\sigma_{q}\left(v_{n}\right) \rightarrow \sigma_{q}(\nu)<\infty$ for all $0 \leq q \leq p$; see [86, Definition 6.8].

Assumption 3.4 Suppose $\mathbf{d}=\mathbf{d}^{(n)}=\left(d_{v}^{(n)}, v \in[n]\right)$ is a degree sequence for each $n \geq 1$, and write $\nu^{n}:=n^{-1} \sum_{v \in[n]} \delta_{d_{v}^{n}}$ for the empirical degree distribution. Assume the following hold as $n \rightarrow \infty$ :
(i) There exists a measure $v$ on $\mathbb{Z}_{\geq 0}$ such that $v^{n} \rightarrow v$ w.r.t. the $W_{3}$ distance.
(ii) The degree sequence is in the critical scaling window, i.e., there exists $\lambda \in \mathbb{R}$ such that

$$
\frac{\sigma_{1}(\nu)}{\left(\sigma_{3}(\nu)-4 \sigma_{1}(\nu)\right)^{2 / 3}} \cdot\left(\frac{\sigma_{2}\left(\nu^{n}\right)}{\sigma_{1}\left(\nu^{n}\right)}-2\right) \cdot n^{1 / 3} \rightarrow \lambda
$$

Note that this assumption implies that $\sigma_{2}(\nu)=2 \sigma_{1}(\nu)$.
Theorem 3.5 [39] Consider a sequence of degree sequences $\mathbf{d}=\mathbf{d}^{(n)}, n \geq 1$, satisfying Assumption 3.4 with limiting empirical distribution v. Let $\mathscr{C}_{i}^{n}$ be the $i$-th largest connected component of $\mathscr{G}_{n, \mathbf{d}}\left(\operatorname{or} \overline{\mathscr{G}}_{n, \mathbf{d}}\right)$. Then as $n \rightarrow \infty$,

$$
\begin{equation*}
\left(\left(\frac{\left(\sigma_{3}(\nu)-4 \sigma_{1}(\nu)\right)^{1 / 3}}{\sigma_{1}(\nu) \cdot n^{2 / 3}} \cdot\left|\mathscr{C}_{i}^{n}\right|, \operatorname{sp}\left(\mathscr{C}_{i}^{n}\right)\right), i \geq 1\right) \xrightarrow{d} \zeta(\lambda) \tag{3.3}
\end{equation*}
$$

with respect to product topology.
This result, in a stronger form, can be found in [39, Theorem 2 and Remark 5]. In [39], the description of the limiting sequence is slightly different. But it is easy to restate the result in the above form using Brownian scaling. In the next section we will use the random sequence $\zeta(\lambda)$ to describe certain metric measure spaces that will appear in our proofs.

### 3.4 Real trees and $\mathbb{R}$-graphs

In this section we will first define real trees and $\mathbb{R}$-graphs and introduce various notions related to them. We will then introduce a family of random $\mathbb{R}$-graphs $\mathscr{H}^{(s)}, s \geq 0$, that act as the building blocks for the scaling limits of various critical random graph models. Using these spaces and the sequence $\zeta(\lambda)$ introduced in Sect. 3.3, we will
define a sequence $\boldsymbol{S}(\lambda)$ of random metric measure spaces; see Construction 3.10. As we will see in the next section, the sequence $\boldsymbol{S}(\lambda)$ describes (up to a multiplicative constant) the scaling limits of the critical random graph models of interest to us.

For any metric space $(X, d)$, a geodesic between $x_{1}, x_{2} \in X$ is an isomeric embedding $f:\left[0, d\left(x_{1}, x_{2}\right)\right] \rightarrow X$ such that $f(0)=x_{1}$ and $f\left(d\left(x_{1}, x_{2}\right)\right)=x_{2} .(X, d)$ is a geodesic space if there is a geodesic between any two points in $X$. An embedded cycle in $X$ is a subset of $X$ that is a homeomorphic image of the unit circle $S^{1}$.

Definition 3.6 (Real trees $[42,61])$ A compact geodesic metric space $(X, d)$ is called a real tree if it has no embedded cycles.

Definition 3.7 ( $\mathbb{R}$-graphs [5]) A compact geodesic metric space $(X, d)$ is called an $\mathbb{R}$-graph if for every $x \in X$, there exists $\varepsilon=\varepsilon(x)>0$ such that $\left(B(x, \varepsilon),\left.d\right|_{B(x, \varepsilon)}\right)$ is a real tree. A measured $\mathbb{R}$-graph is an $\mathbb{R}$-graph with a probability measure on its Borel $\sigma$-algebra.

The core of an $\mathbb{R}$-graph $(X, d)$, denoted by $\operatorname{Core}(X)$, is the union of all the simple arcs having both endpoints in embedded cycles of $X$. If it is non-empty, then (Core $(X), d)$ is an $\mathbb{R}$-graph with no leaves. We define $\operatorname{Conn}(X)$ to be the set of all $x \in X$ such that $x$ belongs to an embedded cycle in $X$.

Clearly, $\operatorname{Conn}(X) \subseteq \operatorname{Core}(X)$. By [5, Theorem 2.7], if $X$ is an $\mathbb{R}$-graph with a non-empty core, then $(\operatorname{Core}(X), d)$ can be represented as $(k(X), e(X)$, len), where $(k(X), e(X))$ is a finite connected multigraph in which all vertices have degree at least 3 and len : $e(X) \rightarrow(0, \infty)$ gives the edge lengths of this multigraph. We denote by $\operatorname{sp}(X)$ the number of surplus edges in $(k(X), e(X))$. On any $\mathbb{R}$-graph $(X, d)$ there exists a unique $\sigma$-finite Borel measure $\ell$, called the length measure, such that if $x_{1}, x_{2} \in X$ and $\left[x_{1}, x_{2}\right]$ is a geodesic path between $x_{1}$ and $x_{2}$ then $\ell\left(\left[x_{1}, x_{2}\right]\right)=d\left(x_{1}, x_{2}\right)$. Further, we define

$$
\begin{equation*}
L(X):=\sum_{e \in e(X)} \operatorname{len}(e)=\ell(\operatorname{Core}(X)) \tag{3.4}
\end{equation*}
$$

Note that $\ell(\operatorname{Conn}(X)) \leq \ell(\operatorname{Core}(X))<\infty$. If $\operatorname{Conn}(X) \neq \emptyset$ (in which case $\ell(\operatorname{Conn}(X))>0$ ), we write $\ell_{\operatorname{Conn}(X)}$ for the restriction of the length measure to Conn $(X)$ normalized to be a probability measure, i.e.,

$$
\ell_{\operatorname{Conn}(X)}(\cdot)=\frac{\ell(\cdot)}{\ell(\operatorname{Conn}(X))} .
$$

Note that any finite connected multigraph with edge lengths, viewed as a metric space, is an $\mathbb{R}$-graph. So the above definitions make sense for any finite connected multigraph $H$. Note the difference between $e(H)$ defined above and $E(H)$-the set of edges in $H$. Note also that in this case, the graph theoretic 2-core of $H$, viewed as a metric space, coincides with the space Core $(H)$ as defined above. We will use Core $(H)$ to denote both the metric space and the graph theoretic 2-core, and the meaning will be clear from the context. Clearly, for any finite connected multigraph $H$ with unit edge
lengths, $L(H)=|E(\operatorname{Core}(H))|$. Further, if $H=(V, E$, len $)$ is a finite connected multigraph with edge lengths, then

$$
\begin{equation*}
\ell(H)=\sum_{e \in E} \operatorname{len}(e) \tag{3.5}
\end{equation*}
$$

We will write $\ell(H)$ to mean the above even when $H$ is not connected.
Functions encoding excursions from zero can be used to construct real trees via a simple procedure. We now describe this construction. An excursion on $[0,1]$ is a continuous function $h \in C([0,1], \mathbb{R})$ with $h(0)=0=h(1)$ and $h(t) \geq 0$ for $t \in(0,1)$. Let $\mathscr{E}_{1}$ be the space of all excursions on the interval [ 0,1$]$. Given an excursion $h \in \mathscr{E}_{1}$, one can construct a real tree as follows. Define a pseudo-metric $d_{h}$ on $[0,1]$ as follows:

$$
d_{h}(s, t):=h(s)+h(t)-2 \inf _{u \in[s, t]} h(u), \text { for } s, t \in[0,1] .
$$

Define the equivalence relation $s \sim t \Leftrightarrow d_{h}(s, t)=0$. Let $[0,1] / \sim$ denote the corresponding quotient space and consider the metric space $\mathscr{T}_{h}:=\left([0,1] / \sim, \bar{d}_{h}\right)$, where $\bar{d}_{h}$ is the metric on the equivalence classes induced by $d_{h}$. Then $\mathscr{T}_{h}$ is a real tree ( $[42,61])$. Let $q_{h}:[0,1] \rightarrow \mathscr{T}_{h}$ be the canonical projection and write $\mu_{\mathscr{T}_{h}}$ for the push-forward of the Lebesgue measure on $[0,1]$ onto $\mathscr{T}_{h}$ via $q_{h}$. Further, we assume that $\mathscr{T}_{h}$ is rooted at $\rho:=q_{h}(0)$. Equipped with $\mu_{\mathscr{T}_{h}}, \mathscr{T}_{h}$ is now a rooted compact metric measure space. Note that by construction, for any $x \in \mathscr{T}_{h}$, the function $h$ is constant on $q_{h}^{-1}(x)$. Thus for each $x \in[0,1]$, we write $\mathrm{ht}(x)=h\left(q_{h}^{-1}(x)\right)$ for the height of this vertex.

The Brownian continuum random tree defined below is a fundamental object in the literature of random real trees.

Definition 3.8 (Aldous's Brownian continuum random tree (CRT) [8]) Let $\mathbf{e}$ be a standard Brownian excursion on $[0,1]$. The real tree $\mathscr{T}_{2 \mathrm{e}}$ is called the Brownian CRT.

It is well-known [8,9] that the associated measure $\mu_{\mathscr{T}_{2 e}}$ (also called the mass measure) is non-atomic and concentrated on the set of leaves of $\mathscr{T}_{2} \mathrm{e}$ almost surely. We will now define a collection of random metric measure spaces $\mathscr{H}^{(s)}, s \geq 2$, using the Brownian CRT. Recall the definition of $\mathscr{G}_{n, 3}$ from Sect. 1.2.
 and $r=3(s-1)$.
(a) Let $\mathscr{K}_{n, 3}$ be distributed as $\mathscr{G}_{n, 3}$ conditioned to be connected. Label its edges arbitrarily as $\left(u_{i}, v_{i}\right), 1 \leq i \leq r$.
(b) Independently of the above, sample $\left(X_{1}, \ldots, X_{r}\right)$ from a Dirichlet $\left(\frac{1}{2}, \ldots, \frac{1}{2}\right)$ distribution.
(c) Independently of the above, sample i.i.d. Brownian CRTs $\mathscr{T}_{1}, \ldots, \mathscr{T}_{r}$. For $1 \leq$ $i \leq r$, let $\rho_{i}$ be the root of $\mathscr{T}_{i}$ and $z_{i}$ be a point in $\mathscr{T}_{i}$ sampled according to its mass measure.
(d) For $1 \leq i \leq r$, construct the metric measure space $\mathscr{T}_{i}^{\prime}$ from $\mathscr{T}_{i}$ by multiplying the distance between each two points by $\sqrt{X_{i}}$ and multiplying the measure of each Borel set by $X_{i}$. Denote the points in $\mathscr{T}_{i}^{\prime}$ that correspond to $\rho_{i}$ and $z_{i}$ by $\rho_{i}^{\prime}$ and $z_{i}^{\prime}$ respectively.
(e) Form a new space $\mathscr{H}^{(s)}$ from $\mathscr{K}_{n, 3}$ by replacing the edge $\left(u_{i}, v_{i}\right)$ by $\mathscr{T}_{i}^{\prime}$ identifying $\rho_{i}^{\prime}$ with $u_{i}$ and $z_{i}^{\prime}$ with $v_{i}, 1 \leq i \leq r$.

This construction of $\mathscr{H}^{(s)}$ was given in [2, Procedure 1]. Note that Core $\left(\mathscr{H}^{(s)}\right)$ is given by the multigraph $\mathscr{K}_{n, 3}$ with associated edge lengths $d_{\mathscr{T}_{i}^{\prime}}\left(\rho_{i}^{\prime}, z_{i}^{\prime}\right)=\sqrt{X_{i}} \cdot d_{\mathscr{T}_{i}}\left(\rho_{i}, z_{i}\right)$, $1 \leq i \leq r$, where $d_{\mathscr{T}_{i}}$ denotes the metric on $\mathscr{T}_{i}$.

For $s=0$, we define the space $\mathscr{H}^{(0)}$ to be the Brownian CRT $\mathscr{T}_{2}$. The explicit construction of the space $\mathscr{H}^{(1)}$ is not relevant to our proof, so we do not include it here, and instead refer the reader to [2, Procedure 1]. Let us also mention here that there are two other constructions of $\mathscr{H}^{(s)}$. In Construction A. 3 below, we describe a 'depth-first construction' of $\mathscr{H}^{(s)}$. This construction was essentially contained in the arguments in [3]. An alternate construction that can be viewed as a 'breadth-first construction' is given in [72, Construction 2.2].

Now, we will define a sequence $S(\lambda)=\left(S_{1}(\lambda), S_{2}(\lambda), \ldots\right)$ of random metric measure spaces. Recall the random sequence $\zeta(\lambda)$ from Theorem 3.5.

Construction 3.10 (The sequence $\boldsymbol{S}(\lambda)$ ) Sample $\boldsymbol{\zeta}(\lambda)=\left(\left(\xi_{i}(\lambda), N_{i}(\lambda)\right), i \geq 1\right)$. For simplicity, write $\xi_{i}=\xi_{i}(\lambda)$, and $N_{i}=N_{i}(\lambda)$. Conditional on $\zeta(\lambda)$, construct the spaces $S_{i}(\lambda)$ independently for $i \geq 1$, where

$$
S_{i}(\lambda) \stackrel{d}{=} \sqrt{\xi_{i}} \cdot \mathscr{H}^{\left(N_{i}\right)} .
$$

Set $\boldsymbol{S}(\lambda)=\left(S_{1}(\lambda), S_{2}(\lambda), \ldots\right)$.
Note that the spaces $\mathscr{H}^{(s)}$ and $S_{i}(\lambda), i \geq 1$, are $\mathbb{R}$-graphs (recall Definition 3.7).

### 3.5 Geometry of critical random graphs

In this section, we will state four results on the geometry and scaling limit of critical random graphs that will be pivotal in our proofs.

Theorem 3.11 (Geometry of uniform connected graphs with a given surplus) Fix an integer $s \geq 2$. Let $\mathscr{H}_{m, s}$ be uniformly distributed over the set of all simple connected graphs on $[m]$ having surplus $s$. Recall the notation $(k(\cdot), e(\cdot)$, len $)$ and $L(\cdot)$ introduced around (3.4). Let $r=3(s-1)$.
(a) We have,

$$
\begin{equation*}
\lim _{m \rightarrow \infty} \mathbb{P}\left(\left(k\left(\mathscr{H}_{m, s}\right), e\left(\mathscr{H}_{m, s}\right)\right) \text { is a 3-regular multigraph }\right)=1 \tag{3.6}
\end{equation*}
$$

In particular,

$$
\begin{equation*}
\lim _{m \rightarrow \infty} \mathbb{P}\left(\left|e\left(\mathscr{H}_{m, s}\right)\right|=r\right)=1 \tag{3.7}
\end{equation*}
$$

Let $e_{1}^{(m)}, \ldots, e_{r}^{(m)}$ (resp. $e_{1}, \ldots, e_{r}$ ) be an enumeration of $\left\{e: e \in e\left(\mathscr{H}_{m, s}\right)\right\}$ (resp. $\left\{e: e \in e\left(\mathscr{H}^{(s)}\right)\right\}$ ). Then as $m \rightarrow \infty$,

$$
\begin{equation*}
\left(\frac{1}{\sqrt{m}} \mathscr{H}_{m, s}, \frac{1}{\sqrt{m}} \cdot\left(\operatorname{len}\left(e_{i}^{(m)}\right), 1 \leq i \leq r\right)\right) \xrightarrow{d}\left(\mathscr{H}^{(s)},\left(\operatorname{len}\left(e_{i}\right), 1 \leq i \leq r\right)\right), \tag{3.8}
\end{equation*}
$$

where the convergence in the first coordinate is with respect to GHP topology. Further, for any $\alpha>0$,

$$
\begin{equation*}
\sup _{m} \mathbb{E}\left[\exp \left(\alpha m^{-1 / 2} L\left(\mathscr{H}_{m, s}\right)\right)\right]<\infty \tag{3.9}
\end{equation*}
$$

As a consequence of (3.8), for every $\varepsilon>0$, there exists $r_{\varepsilon}>0$ such that for all large m,

$$
\begin{equation*}
m^{-1 / 2} L\left(\mathscr{H}_{m, s}\right) \leq 1 / r_{\varepsilon}, \quad \text { and } \quad m^{-1 / 2} \min _{e \in e\left(\mathscr{H}_{m, s}\right)} \operatorname{len}(e) \geq r_{\varepsilon} \tag{3.10}
\end{equation*}
$$

with probability at least $1-\varepsilon$.
(b) Let $V_{i}^{(m)}$ be the set of vertices in $\mathscr{H}_{m, s}$ that are connected to Core $\left(\mathscr{H}_{m, s}\right)$ via $e_{i}^{(m)}$, $1 \leq i \leq r$. (The common endpoints of multiple $e \in e\left(\mathscr{H}_{m, s}\right)$ and their pendant subtrees are assigned to only of the $V_{i}^{(m)}$,s in an arbitrary way.) Recall the real trees $\mathscr{T}_{i}^{\prime}, 1 \leq i \leq r$, from Construction 3.9. Denote the measure on $\mathscr{H}^{(s)}$ by $\mu^{(s)}$. Then as $m \rightarrow \infty$,

$$
\begin{equation*}
\frac{1}{m}\left(\left|V_{i}^{(m)}\right|, 1 \leq i \leq r\right) \xrightarrow{d}\left(\mu^{(s)}\left(\mathscr{T}_{i}^{\prime}\right), 1 \leq i \leq r\right) \sim \operatorname{Dirichlet}\left(\frac{1}{2}, \ldots, \frac{1}{2}\right) . \tag{3.11}
\end{equation*}
$$

(3.6) follows from [56, Theorem 7]. (3.7) follows from (3.6) and the fact $\operatorname{sp}\left(k\left(\mathscr{H}_{m, s}\right)\right)=$ $s$. The rest of the assertions can be proved by following the arguments used in [3]. An outline of the proof is given in Sect. A.1.

Theorem 3.12 (Scaling limit of $\operatorname{ER}(n, \lambda))$ Fix $\lambda \in \mathbb{R}$, and let $\mathscr{C}_{i}^{n, \text { er }}(\lambda)$ denote the $i$-th largest component of $\operatorname{ER}(n, \lambda)$. Then

$$
n^{-1 / 3}\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda), \mathscr{C}_{2}^{n, \text { er }}(\lambda), \ldots\right) \xrightarrow{d} \boldsymbol{S}(\lambda)=\left(S_{1}(\lambda), S_{2}(\lambda), \ldots\right)
$$

with respect to the product topology on $\mathfrak{S}_{\mathrm{GHP}}^{\mathbb{N}}$ as discussed at the end of Sect. 3.2.
This result is the content of [3, Theorem 2]. That the limiting sequence of spaces is same as $\boldsymbol{S}(\boldsymbol{\lambda})$ follows from the discussion around [2, Equation 1].

In [26, Theorem 2.2], the metric space scaling limit of random graphs with a critical degree sequence was established. (See also [22], where a similar result for critical
percolation on the supercritical configuration model was derived as an application of a more general universality principle.) The next result gives a variant of [26, Theorem 2.2]. This result follows from arguments similar to those used in [26]. A sketch of proof is given in Sect. A.2.

Theorem 3.13 Suppose $\left\{\mathbf{d}^{(n)}\right\}_{n \geq 1}$ is a sequence of degree sequences satisfying Assumption 3.4 with limiting empirical distribution v. Further, suppose $f$ : $\{1,2, \ldots\} \rightarrow[0, \infty)$ satisfies $\max \left\{f\left(d_{v}\right): v \in[n]\right\}=o\left(n^{2 / 3}\right)$ and $\sum_{k \geq 1} f(k) v(k)>0$.
(i) Let $\mathscr{C}_{j}^{n}$ be the $j$-th largest component of $\mathscr{G}_{n, \mathbf{d}} . V i e w \mathscr{C}_{j}^{n}$ as a metric space in the usual way; further, assign mass $f\left(d_{v}\right)$ to each $v \in V\left(\mathscr{C}_{j}^{n}\right)$ and normalize it to make it a probability measure. (If $\sum_{v \in \mathscr{C}_{j}^{n}} f\left(d_{v}\right)=0$ then simply take the uniform measure on the vertices.) Denote the resulting metric measure space by $\mathscr{C}_{j}^{n, f}$. Then

$$
n^{-1 / 3}\left(\mathscr{C}_{1}^{n, f}, \mathscr{C}_{2}^{n, f}, \ldots\right) \stackrel{d}{\longrightarrow} \frac{\sigma_{1}(\nu)}{\left(\sigma_{3}(\nu)-4 \sigma_{1}(\nu)\right)^{2 / 3}} \cdot \boldsymbol{S}(\lambda)
$$

with respect to the product topology on $\mathfrak{S}_{\mathrm{GHP}}^{\mathbb{N}}$ jointly with the convergence in (3.3).
(ii) The conclusion of part (i) continues to hold with the same limiting sequence of metric measure spaces if we replace $\mathscr{G}_{n, \mathbf{d}}$ by $\overline{\mathscr{G}}_{n, \mathbf{d}}$.

Next we state a result about the core of the components of a critical graph with given degree sequence.

Theorem 3.14 Suppose $\left\{\mathbf{d}^{(n)}\right\}_{n \geq 1}$ is a sequence of degree sequences satisfying Assumption 3.4 with limiting empirical distribution v. Let $\mathscr{C}_{1}^{n}$ denote the largest component of $\mathscr{G}_{n, \mathbf{d}}$, and write $\mathscr{E}_{1}^{n}=\left|E\left(\mathscr{C}_{1}^{n}\right)\right|$. We will drop the superscript $n$ for convenience. Let $\Gamma_{1}^{(1)}, \Gamma_{1}^{(2)}, \ldots$ be i.i.d. Exponential(1) random variables independent of $\mathscr{G}_{n, \mathbf{d}}$.
(a) Recall the notaion $(k(\cdot), e(\cdot)$, len $)$ and $L(\cdot)$ introduced around (3.4). Then

$$
n^{-1 / 3}\left(L\left(\mathscr{C}_{1}\right), \min _{e \in e\left(\mathscr{C}_{1}\right)} \operatorname{len}(e)\right) \xrightarrow{d} \frac{\sigma_{1}(\nu)}{\left(\sigma_{3}(\nu)-4 \sigma_{1}(\nu)\right)^{2 / 3}} \cdot\left(L\left(S_{1}(\lambda)\right), \min _{e \in e\left(S_{1}(\lambda)\right)} \operatorname{len}(e)\right) .
$$

In particular, for every $\varepsilon>0$, there exists $r_{\varepsilon}>0$ such that for all large $n$,

$$
\operatorname{sp}\left(\mathscr{C}_{1}\right) \leq 1 / r_{\varepsilon}, \quad \frac{L\left(\mathscr{C}_{1}\right)}{n^{1 / 3}} \leq 1 / r_{\varepsilon}, \quad \text { and } \quad \min _{e \in e\left(\mathscr{C}_{1}\right)} \frac{\operatorname{len}(e)}{n^{1 / 3}} \geq r_{\varepsilon}
$$

with probability at least $1-\varepsilon$.
(b) Assign lengths $\Gamma_{1}^{(1)}, \ldots, \Gamma_{1}^{\left(\mathscr{E}_{1}\right)}$ to the edges of $\mathscr{C}_{1}$, and call the resulting graph with edge lengths $\mathscr{C}_{1}^{\text {exp }}$. Then the conclusion in (a) continues to hold with $\mathscr{C}_{1}^{\text {exp }}$ in place of $\mathscr{C}_{1}$.

By Theorem 3.5, $\operatorname{sp}\left(\mathscr{C}_{1}^{n}\right) \xrightarrow{\mathrm{d}} N_{1}(\lambda)$. The other claims in Theorem 3.14(a) follow from the arguments used in the proof of [26, Theorem 2.4]. The claim in (b) can be proved in an identical manner.

## 4 Properties of minimal spanning trees

In this section we discuss various properties of MSTs and give another description of the space $\mathscr{M}$ appearing in Theorem 1.1.

### 4.1 MST and percolation

Suppose $G=(V, E, w)$ is a weighted, connected, and labeled graph. Assume that $w(e) \neq w\left(e^{\prime}\right)$ whenever $e \neq e^{\prime}$. We now state a useful property of the MST.

Lemma 4.1 (Minimax paths property) Let $G=(V, E, w)$ be as above. Then the MST $T$ of $G$ is unique. Further, $T$ has the following property: Any path $\left(x_{0}, \ldots, x_{m}\right)$ with $x_{i} \in V$ and $\left\{x_{i}, x_{i+1}\right\} \in E(T)$ satisfies

$$
\max _{i} w\left(\left\{x_{i}, x_{i+1}\right\}\right) \leq \max _{j} w\left(\left\{x_{j}^{\prime}, x_{j+1}^{\prime}\right\}\right)
$$

for any path $\left(x_{0}^{\prime}, \ldots, x_{m^{\prime}}^{\prime}\right)$ with $\left\{x_{j}^{\prime}, x_{j+1}^{\prime}\right\} \in E$ and $x_{0}=x_{0}^{\prime}$ and $x_{m}=x_{m^{\prime}}^{\prime}$. In words, the maximum edge weight in the path in the MST connecting two given vertices is smallest among all paths in $G$ connecting those two vertices.

Moreover, $T$ is the only spanning tree of $G$ with the above property.
The above lemma is just a restatement of [60, Lemma 2]; see also [14, Proposition 2.1]. We record the following useful observations:

Observation 4.1 Using Lemma 4.1, we see that the MST can be constructed just from the ranks of the different edge weights. Thus the law of the MST constructed using exchangeable edge weights that are almost surely pairwise distinct does not depend on the distribution of the weights.

Observation 4.2 Let $G=(V, E, w)$ be a connected and labeled graph with pairwise distinct edge weights. Let $u \in[0, \infty)$ and $\mathscr{C}$ be a component of the graph $G^{u}=$ $\left(V, E^{u}\right)$, where $E^{u} \subseteq E$ contains only those edges $e$ for which $w(e) \leq u$. Then the restriction of the MST of $(V, E, w)$ to $\mathscr{C}$ is the $\operatorname{MST}$ of $\left(V(\mathscr{C}), E(\mathscr{C}),\left.w\right|_{E(\mathscr{C})}\right)$. This can be argued as follows: If $v, v^{\prime} \in \mathscr{C}$, then there exists a path in $G$ connecting $v$ and $v^{\prime}$ such that all edge weights along this path is at most $u$. By Lemma 4.1, it follows that all edge weights in the path in the MST of $(V, E, w)$ connecting $v$ and $v^{\prime}$ is also smaller than $u$. Thus the restriction of the MST of $(V, E, w)$ to $\mathscr{C}$ is a spanning tree of $\mathscr{C}$. Since the restriction of the MST of $(V, E, w)$ to $\mathscr{C}$ also satisfies the minimax path property, it is the MST of $\mathscr{C}$ (constructed using the restriction of the weight function $w(\cdot)$ to the edges of $\mathscr{C})$. This fact is extremely useful as it can be used to connect the structure of the MST to the geometry of components of the graph under percolation.

Observation 4.3 Let $G=(V, E, w)$ be a connected and labeled graph with pairwise distinct edge weights. Recall the notation Conne(•) from Sect. 3.1. Let $e \in \operatorname{Conne}(G)$ be the edge with the maximum weight among all edges in Conne $(G)$. Then $G^{\prime}=$ $\left(V, E \backslash\{e\}, w^{\prime}\right)$ is connected, where $w^{\prime}$ is the restriction of $w$ to $E \backslash\{e\}$. Further, by Lemma 4.1, $e$ is not contained in the MST of $G$. Thus, the MST of $G^{\prime}$ is the same as the MST of $G$. We can use this algorithm inductively to remove edges until we are left with a tree, and this tree will be the MST of $G$.

### 4.2 Cycle-breaking and modified cycle-breaking

In this section we define two procedures that can be applied to $\mathbb{R}$-graphs and multigraphs. Recall the definitions related to $\mathbb{R}$-graphs from Sect. 3.4.

Definition 4.2 (Cycle-breaking (CB)) Let $X$ be an $\mathbb{R}$-graph. If $X$ has no embedded cycles, then set $\mathrm{CB}(X)=X$. Otherwise, sample $x$ from $\operatorname{Conn}(X)$ using the measure $\ell_{\operatorname{Conn}(X)}$, and set $\mathrm{CB}(X)$ to be the completion of the space $X \backslash\{x\}$ endowed with the intrinsic metric inherited from the metric on $X$. (Thus, $\mathrm{CB}(X)$ is also an $\mathbb{R}$-graph.)

For $k \geq 2$, we inductively define $\mathrm{CB}^{k}(X)$ to be the space $\mathrm{CB}\left(\mathrm{CB}^{k-1}(X)\right)$. (Thus, at the $k$-th step, if $\mathrm{CB}^{k-1}(X)$ has an embedded cycle, then we are using the measure to $\ell_{\mathrm{Conn}\left(\mathrm{CB}^{k-1}(X)\right)}$ to sample a point.)

Note that $\mathrm{CB}^{k}(X)=\mathrm{CB}^{\operatorname{sp}(X)}(X)$ for all $k \geq \operatorname{sp}(X)$, i.e., the spaces $\mathrm{CB}^{k}(X)$ remain the same after all cycles have been cut open. We denote this final space (which is a real tree) by $\mathrm{CB}^{\infty}(X)$.

Next we define a cycle-breaking process for discrete multigraphs. We will use a variation of the above process. More precisely, we will sample edges with replacement. This will turn out to be convenient in our proof.

Definition 4.3 (Cycle-breaking for discrete graphs (CBD)) Let $H=(V, E$, len) be a finite (not necessarily connected) multigraph with edge lengths given by the function len $: E \rightarrow(0, \infty)$. Set $\mathrm{CBD}_{0}(H)=H$. For $k \geq 1$, we inductively define $\mathrm{CBD}_{k}(H)$ as follows: Sample $e_{k}$ from $E$ with probability proportional to len $\left(e_{k}\right)$. If $e_{k}$ is not an edge in $\mathrm{CBD}_{k-1}(H)$, set $\mathrm{CBD}_{k}(H)=\mathrm{CBD}_{k-1}(H)$. Otherwise, if $\mathscr{C}$ is the component of $\mathrm{CBD}_{k-1}(H)$ containing $e_{k}$ and $e_{k} \in \operatorname{Conne}(\mathscr{C})$, then set $\mathrm{CBD}_{k}(H)=\mathrm{CBD}_{k-1}(H) \backslash$ $e_{k}$; and if $e_{k} \notin \operatorname{Conne}(\mathscr{C})$, then sample a point $x$ uniformly on the edge $e_{k}$ and color $x$ red, and set $\mathrm{CBD}_{k}(H)$ to be $\mathrm{CBD}_{k-1}(H)$ with the point $x$ colored red.

Ignoring the colored points, the multigraphs $\mathrm{CBD}_{k}(H)$ are the same (and are all forests) for all large values of $k$. We denote the tree (without any colored points) in this forest with the most number of vertices by $\mathrm{CBD}_{\infty}(H)$.

Suppose $H$ is a finite connected multigraph with edge lengths. Let $f_{1}, \ldots, f_{s}$ be the edges of $H$ that get removed in the process $\left(\operatorname{CBD}_{k}(H), k \geq 1\right)$. Clearly, $s=\operatorname{sp}(H)$. For $1 \leq i \leq s$, let $y_{i}$ be a uniformly sampled point on $f_{i}$. It is easy to see that viewing $H$ as an $\mathbb{R}$-graph, the completion of the space $H \backslash\left\{y_{1}, \ldots, y_{s}\right\}$ has the same distribution as $\mathrm{CB}^{\infty}(H)$. In this coupling, $\mathrm{CBD}_{\infty}(H)$ is a subspace of $\mathrm{CB}^{\infty}(H)$, and

$$
\begin{equation*}
d_{H}\left(\mathrm{CBD}_{\infty}(H), \mathrm{CB}^{\infty}(H)\right) \leq \max _{e \in E} \operatorname{len}(e) \tag{4.1}
\end{equation*}
$$

Further, suppose $G_{1}$ (resp. $G_{2}$ ) is a finite connected graph with edge lengths and $u_{1}$ (resp. $u_{2}$ ) is one of its vertices. Denote by $\left(G_{1}, u_{1}\right) \xrightarrow{a}\left(u_{2}, G_{2}\right)$ the graph obtained by joining $u_{1}$ and $u_{2}$ by an edge of length $a$. Then

$$
\begin{equation*}
\operatorname{CBD}_{\infty}\left(\left(G_{1}, u_{1}\right) \stackrel{a}{-}\left(u_{2}, G_{2}\right)\right) \stackrel{\mathrm{d}}{=}\left(\operatorname{CBD}_{\infty}\left(G_{1}\right), u_{1}\right) \stackrel{a}{-}\left(u_{2}, \operatorname{CBD}_{\infty}\left(G_{2}\right)\right) \tag{4.2}
\end{equation*}
$$

We now record a useful observation that we will use in the proofs. The proof of this result is elementary, so we omit it.

Lemma 4.4 Suppose $H=(V, E$, len $)$ is a finite multigraph with edge lengths.
(a) Assume that len (e), $e \in E$, are exchangeable random variables. For $1 \leq i \leq|E|$, let $\mathscr{E}_{i}$ denote the $i$-th distinct edge sampled in the process $\left(C B D_{k}(H), k \geq 1\right)$. Then for any $j \in\{1, \ldots,|E|-1\}$ and collection of distinct edges $e_{1}, \ldots, e_{j}$, conditional on the event $\left\{\mathscr{E}_{i}=e_{i}\right.$ for $\left.1 \leq i \leq j\right\}, \mathscr{E}_{j+1}$ is uniformly distributed over $E \backslash\left\{e_{1}, \ldots, e_{j}\right\}$.
(b) Assume that $H$ is connected and that len(e), $e \in \operatorname{Conne}(H)$, are exchangeable random variables. For $1 \leq i \leq s p(H)$, let $\mathscr{E}_{i}^{\prime}$ denote the $i$-th edge removed in the process $\left(C B D_{k}(H), k \geq 1\right)$. Consider $j \in\{1, \ldots, s p(H)-1\}$ and a collection of edges $e_{1}, \ldots, e_{j}$ satisfying $e_{i} \in$ Conne $\left(H \backslash\left\{e_{1}, \ldots, e_{i-1}\right\}\right)$ for all $1 \leq i \leq j$. Then conditional on the event $\left\{\mathscr{E}_{i}^{\prime}=e_{i}\right.$ for $\left.1 \leq i \leq j\right\}, \mathscr{E}_{j+1}^{\prime}$ is uniformly distributed over Conne ( $H \backslash\left\{e_{1}, \ldots, e_{j}\right\}$ ).

For any finite multigraph $H=$ ( $V, E$, len) having edge lengths (and possibly points colored red on its edges), we write Shape $[H]$ to denote the multigraph $(V, E)$ (without any red points). We also define $\operatorname{Rem}(H)$ to be the multigraph with edge lengths obtained by removing all edges of $H$ that have at least one red point on them. We now state a lemma that connects cycle-breaking to MSTs.

Lemma 4.5 Suppose $H=(V, E$, len $)$ is a finite connected multigraph with random edge lengths. Assume that len $(e), e \in \operatorname{Conne}(H)$, are exchangeable random variables. Then Shape $\left[C B D_{\infty}(H)\right]$ has the same law as the MST of Shape $[H]$ constructed by assigning exchangeable pairwise distinct weights to the edges in Conne $(H)$ and any arbitrary weights to the other edges.

Note that in the setting of Lemma 4.5, Shape $\left[\mathrm{CBD}_{\infty}(H)\right]$ is not the MST of the weighted graph (Shape $[H], w)$ where $w(e)=\operatorname{len}(e)$, even though they have the same law provided the edge lengths are almost surely pairwise distinct.

Proof of Lemma 4.5 Let $\mathscr{E}_{j}^{\prime}$ be the $j$-th edge removed in the CBD process. Then by Lemma $4.4(\mathrm{~b}), \mathscr{E}_{1}^{\prime}$ is uniformly distributed over Conne $(H)$. In general, conditional on $\mathscr{E}_{i}^{\prime}, 1 \leq i \leq k-1, \mathscr{E}_{k}^{\prime}$ is uniformly distributed over Conne $\left(H \backslash\left\{\mathscr{E}_{1}^{\prime}, \ldots, \mathscr{E}_{k-1}^{\prime}\right\}\right)$.

Now, consider edge weights $(w(e), e \in E)$, such that $w(e), e \in \operatorname{Conne}(H)$, are exchangeable and almost surely pairwise distinct. Then using Observation 4.3, the MST of (Shape[H],w) can be constructed by sequentially removing the edges having maximum weight among all edges whose removal do not disconnect the current graph. By the assumptions on the weights, the edge to be removed at each step is uniformly
distributed over the set of all edges whose removal do not disconnect the current graph. In other words, the sequence of edges removed in the algorithm described in Observation 4.3 has the same law as $\left(\mathscr{E}_{k}^{\prime}, k \geq 1\right)$. This completes the proof.

Recall the notation $k(X), e(X)$, (len $(e), e \in e(X)), \operatorname{sp}(X)$, and $L(X)$ introduced below Definition 3.7. For $r \in(0,1)$ define $\mathscr{A}_{r}$ to be the set of all measured $\mathbb{R}$-graphs $X$ that satisfy

$$
\operatorname{sp}(X)+L(X) \leq 1 / r, \text { and } \min _{e \in e(X)} \operatorname{len}(e) \geq r .
$$

The following theorem will allow us to prove convergence of MSTs from GHP convergence of the underlying graphs.

Theorem 4.6 Fixr $\in(0,1)$. Suppose $(X, d, \mu)$ and $\left(X_{n}, d_{n}, \mu_{n}\right), n \geq 1$, are measured $\mathbb{R}$-graphs in $\mathscr{A}_{r}$ such that $\left(X_{n}, d_{n}, \mu_{n}\right) \rightarrow(X, d, \mu)$ as $n \rightarrow \infty$ w.r.t. GHP topology.
(a) Then $C B^{\infty}\left(X_{n}\right) \xrightarrow{d} C B^{\infty}(X)$ as $n \rightarrow \infty$ w.r.t. GHP topology.
(b) Suppose for each $n \geq 1,\left(X_{n}, d_{n}, \mu_{n}\right)$ is the metric measure space associated with $\left(V_{n}, E_{n}\right.$,len) - a finite connected multigraph with edge lengths. If $\max _{e \in E_{n}}$ len $(e) \rightarrow 0$ as $n \rightarrow \infty$, then $C B D_{\infty}\left(X_{n}\right) \xrightarrow{d} C B^{\infty}(X)$ as $n \rightarrow \infty$ w.r.t. GHP topology.

The result in Theorem 4.6 ((a)) is from [5, Theorem 3.3], while the claim in ((b)) follows from (4.1).

### 4.3 Alternate descriptions of the space $\mathscr{M}$

Recall the construction of the process $\operatorname{ER}(n, \cdot)$ using the random variables $U_{i j}$ from Definition 1.2. Let $\mathscr{C}_{1}^{n, \text { er }}(\lambda)$ be the largest component of $\operatorname{ER}(n, \lambda)$ and let $M_{\lambda}^{n \text {,er }}$ be the MST of $\mathscr{C}_{1}^{n, \text { er }}(\lambda)$ constructed using the random weights $U_{i j},(i, j) \in E\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)$. Then $\lim _{\lambda \rightarrow \infty} M_{\lambda}^{n, \text { er }}=M_{\infty}^{n, \text { er }}$ (in fact $M_{\lambda}^{n, \text { er }}=M_{\infty}^{n, \text { er }}$ for all large $\lambda$ ), where $M_{\infty}^{n, \text { er }}$ is the MST of $K_{n}$ constructed using the random weights $U_{i j}$. Theorem 1.1 says that $n^{-1 / 3} M_{\infty}^{n, \text { er }} \xrightarrow{\mathrm{d}} \mathscr{M}$ as $n \rightarrow \infty$ w.r.t. GHP topology. The natural question to ask here is whether the order in which the limits are taken can be interchanged, i.e., can we first take limit as $n \rightarrow \infty$ for fixed $\lambda$, and then let $\lambda \rightarrow \infty$ ? Now, by [5, Theorem 4.4],

$$
\begin{equation*}
n^{-1 / 3} M_{\lambda}^{n, \text { er }} \xrightarrow{\mathrm{d}} \mathrm{CB}^{\infty}\left(S_{1}(\lambda)\right) \text { as } n \rightarrow \infty \tag{4.3}
\end{equation*}
$$

w.r.t. GHP topology. Then the following theorem answers the above question in the affirmative.

Theorem 4.7 [5, Theorem 4.9] As $\lambda \rightarrow \infty$,

$$
C B^{\infty}\left(S_{1}(\lambda)\right) \xrightarrow{d} \mathscr{M}
$$

with respect to GHP topology.

The space $S_{1}(\lambda)$ has a random number of cycles. The following theorem gives a derandomized version of Theorem 4.7.

Theorem 4.8 Recall the space $\mathscr{H}^{(s)}$ from Construction 3.9. Then

$$
(12 s)^{1 / 6} \cdot C B^{\infty}\left(\mathscr{H}^{(s)}\right) \xrightarrow{d} \mathscr{M} \text { as } s \rightarrow \infty
$$

with respect to GHP topology.
Theorem 4.8 plays a crucial role in our argument. The proof of this result can be read independently of the rest and is deferred to Sect. 6.9.

## 5 Idea of proof

In this section, we outline the proof of the fact that the claimed convergence in Theorem 2.1 holds w.r.t. the GH topology assuming Theorem 4.8. We explain the ideas at a high level, and do our best to avoid getting into the technicalities. For any finite graph $H$, we will write $H^{\text {exp }}$ to denote the graph obtained by assigning i.i.d. Exponential(1) lengths to the edges of $H$. When $H$ is random, the edge lengths are taken to be independent of $H$.

Recall the notation from Construction 3.9. Then

$$
\left(X_{1}, \ldots, X_{r}\right) \stackrel{\mathrm{d}}{=}\left(\Gamma_{1 / 2}^{(1)}, \ldots, \Gamma_{1 / 2}^{(r)}\right) / \Gamma_{r / 2}
$$

where $\Gamma_{1 / 2}^{(j)}, j=1, \ldots, r$, are i.i.d. $\operatorname{Gamma}(1 / 2,1)$ random variables, and $\Gamma_{r / 2}:=$ $\sum_{j=1}^{r} \Gamma_{1 / 2}^{(j)}=\frac{r}{2} \cdot\left(1+o_{P}(1)\right)$ as $s \rightarrow \infty$. Further, it is well-known that typical distance in a Brownian CRT follows a Rayleigh distribution. Consequently, $d_{\mathscr{T}_{i}}\left(\rho_{i}, z_{i}\right)$. $\left(2 \Gamma_{1 / 2}^{(i)}\right)^{1 / 2}, i=1, \ldots, r$, are i.i.d. Exponential(1) random variables. As noted below Construction 3.9, $\operatorname{Core}\left(\mathscr{H}^{(s)}\right)$ can be represented by the multigraph $\mathscr{K}_{n, 3}$ with edge lengths given by $d_{\mathscr{T}_{i}}\left(\rho_{i}, z_{i}\right) \cdot \sqrt{X_{i}}, i=1, \ldots, r$. Hence, Core $\left(\mathscr{H}^{(s)}\right)$ is simply $\gamma_{n}$. $\mathscr{K}_{n, 3}^{\exp }$, where $\gamma_{n}=r^{-1 / 2}\left(1+o_{P}(1)\right)$ as $s \rightarrow \infty$. Now, the space $(12 s)^{1 / 6} \cdot \mathscr{H}^{(s)}$ can be obtained from $(12 s)^{1 / 6}$. Core $\left(\mathscr{H}^{(s)}\right)$ by attaching some random compact trees. As $s \rightarrow \infty$, the maximum diameter of these trees becomes negligible. In other words, the result in Theorem 4.8 continues to hold if we replace $\mathscr{H}^{(s)}$ by Core $\left(\mathscr{H}^{(s)}\right)$, which can in turn be replaced by $r^{-1 / 2} \cdot \mathscr{K}_{n, 3}^{\exp }$. Thus, $(12 s)^{1 / 6} \cdot r^{-1 / 2} \cdot \mathrm{CB}^{\infty}\left(\mathscr{K}_{n, 3}^{\exp }\right) \xrightarrow{\mathrm{d}} \mathscr{M}$, as $n \rightarrow \infty$, with respect to the GH topology. Using the relations $s \sim n / 2$ and $r \sim 3 n / 2$ as $n \rightarrow \infty$, we conclude that $n^{-1 / 3} \cdot \mathrm{CB}^{\infty}\left(\mathscr{K}_{n, 3}^{\exp }\right) \xrightarrow{\mathrm{d}}(0.75)^{1 / 3} \cdot \mathscr{M}$, as $n \rightarrow \infty$, with respect to the GH topology. Now using (4.1) to go from $\mathrm{CB}^{\infty}$ to $\mathrm{CBD}_{\infty}$ and using (2.2), we get

$$
\begin{equation*}
n^{-1 / 3} \cdot \operatorname{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right) \xrightarrow{\mathrm{d}}(0.75)^{1 / 3} \cdot \mathscr{M}, \text { as } n \rightarrow \infty, \tag{5.1}
\end{equation*}
$$

with respect to the GH topology.

Now, observe that for any finite graph $H$, conditional on the event that the number of distinct edges sampled in the process $\left(\mathrm{CBD}_{i}\left(H^{\exp }\right), 1 \leq i \leq T\right)$ is $m$, the collection of these edges has the same distribution as a uniform subset of size $m$ sampled from $E(H)$. From this, one might guess that for an appropriately chosen random $T$, Shape $\left[\operatorname{Rem}\left(\operatorname{CBD}_{T}\left(H^{\text {exp }}\right)\right)\right]$ will have the same distribution as $\operatorname{Perc}(H, p)$ for a deterministic $p$, where $\operatorname{Perc}(H, p)$ denotes the random subgraph of $H$ obtained under percolation with edge retention probability $p$. In fact, we have the following stronger result: Fix $t>0$, and let $R(t)$ be a Poisson $\left(t \cdot \ell\left(H^{\exp }\right)\right.$ ) random variable, where $\ell(\cdot)$ is as in (3.5). Then

$$
\begin{align*}
& \left(\operatorname{Shape}\left[\operatorname{Rem}\left(\operatorname{CBD}_{R(t)}\left(H^{\exp }\right)\right)\right], \operatorname{Rem}\left(\operatorname{CBD}_{R(t)}\left(H^{\exp }\right)\right)\right) \\
& \qquad \stackrel{\mathrm{d}}{=}\left(\operatorname{Perc}\left(H, \frac{1}{1+t}\right), \frac{1}{1+t} \cdot\left(\operatorname{Perc}\left(H, \frac{1}{1+t}\right)\right)^{\exp }\right) . \tag{5.2}
\end{align*}
$$

This is the content of Lemma 6.7 whose proof is rather short.
For fixed $\lambda \in \mathbb{R}$ and $n$ large so that $2 n^{1 / 3}>|\lambda|$, let $t_{n, \lambda}$ be such that $\left(1+t_{n, \lambda}\right)^{-1}=$ $1 / 2+\lambda n^{-1 / 3}$. Write $\mathscr{G}_{n, 3}(\lambda)$ for $\operatorname{Perc}\left(\mathscr{G}_{n, 3}, 1 / 2+\lambda n^{-1 / 3}\right)$. Let $\mathscr{C}_{1}(\lambda)$ denote the largest connected component of $\mathscr{G}_{n, 3}(\lambda)$. Applying (5.2) with $H=\mathscr{G}_{n, 3}$ and $t=$ $t_{n, \lambda}$ will yield the following: Let $\mathfrak{G}_{1}(n, \lambda)$ be the largest connected component of $\operatorname{Rem}\left(\operatorname{CBD}_{R\left(t_{n, \lambda}\right)}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right)$, and set $G_{1}(n, \lambda):=\operatorname{Shape}\left[\mathfrak{G}_{1}(n, \lambda)\right]$. Then

$$
\begin{equation*}
\left(G_{1}(n, \lambda), \mathfrak{G}_{1}(n, \lambda)\right) \stackrel{\mathrm{d}}{=}\left(\mathscr{C}_{1}(\lambda),\left(1 / 2+\lambda n^{-1 / 3}\right) \cdot\left(\mathscr{C}_{1}(\lambda)\right)^{\exp }\right) \tag{5.3}
\end{equation*}
$$

Now, $\mathscr{G}_{n, 3}(\lambda)$, conditional on its degree sequence, is distributed as a configuration model with that degree sequence. Further, it is easy to show that the (random) degree sequence of $\mathscr{G}_{n, 3}(\lambda)$ satisfies Assumption 3.4 with limiting empirical distribution $v=$ Binomial (3, 1/2). Using these observations together with Theorem 3.13, we can show that

$$
\begin{equation*}
n^{-1 / 3} \cdot \mathscr{C}_{1}(\lambda) \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot S_{1}\left((48)^{1 / 3} \cdot \lambda\right) \tag{5.4}
\end{equation*}
$$

w.r.t. the GH topology. Now, consider any self-avoiding path $\pi$ in $\mathscr{C}_{1}(\lambda)$. Since the edge lengths in $\left(\mathscr{C}_{1}(\lambda)\right)^{\text {exp }}$ are i.i.d. Exponential(1) random variables, the length of $\pi$ in $\left(\mathscr{C}_{1}(\lambda)\right)^{\text {exp }}$ will be concentrated around the length of $\pi$ in $\mathscr{C}_{1}(\lambda)$. Thus, leveraging the fact that $\mathscr{C}_{1}(\lambda)$ has only $O_{P}(1)$ many surplus edges, we can show that $d_{\mathrm{GH}}\left(\mathscr{C}_{1}(\lambda),\left(\mathscr{C}_{1}(\lambda)\right)^{\exp }\right)=o_{P}\left(n^{1 / 3}\right)$. Combining this with (5.4), (5.3), Theorem 4.6, and Theorem 3.14 will yield

$$
\begin{gathered}
n^{-1 / 3} \mathrm{CBD}_{\infty}\left(G_{1}(n, \lambda)\right) \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot \mathrm{CB}^{\infty}\left(S_{1}\left((48)^{1 / 3} \cdot \lambda\right)\right), \text { and } \\
n^{-1 / 3} \mathrm{CBD}_{\infty}\left(\mathfrak{G}_{1}(n, \lambda)\right) \xrightarrow{\mathrm{d}}(0.75)^{1 / 3} \cdot \mathrm{CB}^{\infty}\left(S_{1}\left((48)^{1 / 3} \cdot \lambda\right)\right)
\end{gathered}
$$

w.r.t. the GH topology. Using Theorem 4.7, we see that for any $\mathbb{Z}_{>0}$-valued sequence $\lambda_{n}$ that tends to infinity sufficiently slowly,

$$
\begin{align*}
& n^{-1 / 3} \operatorname{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right) \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot \mathscr{M}, \text { and }  \tag{5.5}\\
& n^{-1 / 3} \operatorname{CBD}_{\infty}\left(\mathfrak{G}_{1}\left(n, \lambda_{n}\right)\right) \xrightarrow{\mathrm{d}}(0.75)^{1 / 3} \cdot \mathscr{M} \tag{5.6}
\end{align*}
$$

w.r.t. the GH topology.

Now compare (5.6) with (5.1). Let $B_{n}$ be as in (2.1). Then on the event $B_{n}$ (which, by (2.2), occurs with high probability), $\mathrm{CBD}_{\infty}\left(\mathfrak{G}_{1}\left(n, \lambda_{n}\right)\right)$ is a subspace of $\mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right)$. Since they have the same scaling limit,

$$
\begin{equation*}
n^{-1 / 3} \cdot \mathbb{1}_{B_{n}} \cdot d_{\mathrm{H}}\left(\operatorname{CBD}_{\infty}\left(\mathfrak{G}_{1}\left(n, \lambda_{n}\right)\right), \mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right) \xrightarrow{\mathrm{d}} 0 . \tag{5.7}
\end{equation*}
$$

This follows from a general property of metric spaces; see Proposition 6.5. Suppose, on the event $B_{n}, \mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right)$ is obtained by attaching the trees $\mathfrak{T}_{n, \lambda_{n}}^{(j)}, 1 \leq j \leq k_{n}\left(\lambda_{n}\right)$, each via an edge to a vertex of $\operatorname{CBD}_{\infty}\left(\mathfrak{G}_{1}\left(n, \lambda_{n}\right)\right)$. Set $k_{n}\left(\lambda_{n}\right)=0$ on $B_{n}^{c}$. Then (5.7) is equivalent to the assertion that $n^{-1 / 3} \max _{1 \leq j \leq k_{n}\left(\lambda_{n}\right)} \operatorname{diam}\left(\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right) \xrightarrow{\mathrm{d}} 0$. From this, it is not difficult to argue that

$$
\begin{equation*}
n^{-1 / 3} \max _{1 \leq j \leq k_{n}\left(\lambda_{n}\right)} \operatorname{diam}\left(\operatorname{Shape}\left[\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right]\right) \xrightarrow{\mathrm{d}} 0 . \tag{5.8}
\end{equation*}
$$

Finally, using Lemma4.4(a), the processes $\left(\operatorname{CBD}_{i}\left(\mathscr{G}_{n, 3}^{\exp }\right), i \geq 1\right)$ and $\left(\operatorname{CBD}_{t}\left(\mathscr{G}_{n, 3}\right), t \geq\right.$ 1) can be coupled so that the $j$-th distinct edge sampled is the same in both processes, $1 \leq j \leq 3 n / 2$. In this coupling, on the event $B_{n}, \mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right)$ is $\mathrm{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)$ with the trees Shape $\left[\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right], 1 \leq j \leq k_{n}\left(\lambda_{n}\right)$, attached to its vertices via an edge. This observation together with (2.2), (5.8), (5.5), and Lemma 4.5 shows that

$$
n^{-1 / 3} M_{n} \stackrel{\mathrm{~d}}{=} n^{-1 / 3} \mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right) \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot \mathscr{M}
$$

as $n \rightarrow \infty$ w.r.t. the GH topology.

## 6 Proofs of Theorems 2.1, 2.2, 2.3, and 4.8

We divide the argument into several steps. In Sect. 6.1, we prove a weaker version of Theorem 2.1 that only deals with convergence w.r.t. GH topology. The proof of this result depends on several propositions whose proofs are given in Sections 6.2-6.5. The proof of Theorem 2.1 is then completed in Sect. 6.6. The proof of Theorem 2.2 is given in Sect. 6.7. The proof of Theorem 2.3 is given in Sect. 6.8. Finally, the proof of Theorem 4.8 is given in Sect. 6.9.

### 6.1 GH convergence of the MST of $\mathscr{G}_{n, 3}$

In this section we prove the following weaker version of Theorem 2.1.
Theorem 6.1 Let $M_{n}$ be as in Theorem 2.1. Then $n^{-1 / 3} \cdot M_{n} \xrightarrow{d} 6^{1 / 3} \cdot \mathscr{M}$ with respect to GH topology.

This convergence will be strengthened to GHP convergence in Sect. 6.6. The proof of the above theorem relies on the following four propositions.

Proposition 6.2 For all $r \in \mathbb{N}$, there exists $c>0$ small such that the following holds: Let $G=(V, E)$ be a finite graph with maximum degree at most $r$. Let $\Gamma_{1}^{(e)}, e \in E$, be i.i.d. Exponential(1) random variables. Then for all $m \geq 1$,

$$
\begin{aligned}
& \mathbb{P}\left(G \text { contains a self-avoiding path } P \text { with }|P| \geq m \text { and } \sum_{e \in P} \Gamma_{1}^{(e)} \leq c|P|\right) \\
& \quad \leq|V| \cdot \exp (-m),
\end{aligned}
$$

where $|P|$ denotes the number of edges in the path $P$.
Proposition 6.3 Assign i.i.d. Exponential(1) lengths to the edges of $\mathscr{G}_{n, 3}$ and denote this multigraph with edge lengths by $\mathscr{G}_{n, 3}^{\exp }$. Then

$$
n^{-1 / 3} C B D_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right) \xrightarrow{d} \frac{1}{2} \cdot 6^{1 / 3} \cdot \mathscr{M}=(0.75)^{1 / 3} \cdot \mathscr{M}, \text { as } n \rightarrow \infty
$$

with respect to GH topology.
Remark 3 By Lemma 4.5, $M_{n} \stackrel{\text { d }}{=}$ Shape $\left[\operatorname{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right]$. However, conditional on Shape $\left[\mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right]$, the edge lengths of $\mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right)$ are not exchangeable, which is why Theorem 6.1 cannot be proved by just using Proposition 6.3, and it takes quite a bit of additional work. Note however that Proposition 6.3 implies that

$$
\operatorname{diam}\left(\operatorname{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right)=\Theta_{P}\left(n^{1 / 3}\right)
$$

This observation together with Proposition 6.2 implies that $\operatorname{diam}\left(M_{n}\right)=O_{P}\left(n^{1 / 3}\right)$. As noted before in (1.2) in the case of the complete graph, using Observation 4.2 and Theorem 3.13, it follows that $\operatorname{diam}\left(M_{n}\right)=\Omega_{P}\left(n^{1 / 3}\right)$. Thus, we get that $\operatorname{diam}\left(M_{n}\right)=$ $\Theta_{P}\left(n^{1 / 3}\right)$. By a standard conditioning argument (see (6.48) and (6.49)), this also implies that $\operatorname{diam}\left(\bar{M}_{n}\right)=\Theta_{P}\left(n^{1 / 3}\right)$.

Recall the notation Shape[•] and Rem(•) introduced right before Lemma 4.5. Recall also from (3.5) and the line below the meaning of $\ell(H)$ for finite multigraphs with edge lengths.

Proposition 6.4 Let $S_{1}(\cdot)$ and $\mathscr{G}_{n, 3}^{\exp }$ be as in Construction 3.10 and Proposition 6.3 respectively. For $\lambda \in \mathbb{R}$ satisfying $|\lambda|<n^{1 / 3} / 2$, let $t_{n, \lambda}$ be given by

$$
\frac{1}{1+t_{n, \lambda}}=\frac{1}{2}+\frac{\lambda}{n^{1 / 3}} .
$$

Let $R_{n, \lambda}$ be a Poisson $\left(t_{n, \lambda} \cdot \ell\left(\mathscr{G}_{n, 3}^{\exp }\right)\right)$ random variable. Let $\mathfrak{G}_{1}(n, \lambda)$ be the largest component of $\operatorname{Rem}\left(C B D_{R_{n, \lambda}}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right)$. Let $G_{1}(n, \lambda):=\operatorname{Shape}\left[\mathfrak{G}_{1}(n, \lambda)\right]$. Then for any fixed $\lambda \in \mathbb{R}$,

$$
\begin{align*}
& n^{-1 / 3} C B D_{\infty}\left(G_{1}(n, \lambda)\right) \xrightarrow{d} 6^{1 / 3} \cdot C B^{\infty}\left(S_{1}\left((48)^{1 / 3} \cdot \lambda\right)\right), \quad \text { and }  \tag{6.1}\\
& n^{-1 / 3} C B D_{\infty}\left(\mathfrak{G}_{1}(n, \lambda)\right) \xrightarrow{d}(0.75)^{1 / 3} \cdot C B^{\infty}\left(S_{1}\left((48)^{1 / 3} \cdot \lambda\right)\right) \tag{6.2}
\end{align*}
$$

as $n \rightarrow \infty$ with respect to the GH topology.
Recall the marked topology from Sect. 3.2.
Proposition 6.5 Suppose $\left\{\left(X_{n}^{+}, d_{n}, X_{n}\right)\right\}_{n \geq 1}$ is a sequence of random compact marked metric spaces such that

$$
X_{n}^{+} \xrightarrow{d} Z, \text { and } X_{n} \xrightarrow{d} Z, \text { as } n \rightarrow \infty
$$

with respect to the GH topology for some random compact metric space $Z$. Then $d_{H}\left(X_{n}, X_{n}^{+}\right) \xrightarrow{d} 0$ as $n \rightarrow \infty$.

We first prove Theorem 6.1 assuming the above four propositions. The proofs of Propositions $6.2,6.3,6.4$, and 6.5 will be given in the next four sections. We will make use of the following elementary fact in the proof of Theorem 6.1; we omit its proof.

Lemma 6.6 Suppose $a_{i, j}, i \in \mathbb{Z}_{>0}, j \in \mathbb{Z}_{>0} \cup\{\infty\}$, and $a_{\infty, \infty}$ are elements of some metric space such that $\lim _{j \rightarrow \infty} a_{i, j}=a_{i, \infty}$ for every $i \in \mathbb{Z}_{>0}$, and $\lim _{i \rightarrow \infty} a_{i, \infty}=$ $a_{\infty, \infty}$. Then there exists $a \mathbb{Z}_{>0}$-valued sequence $\left\{i_{j}^{\star}\right\}_{j \in \mathbb{Z}_{>0}}$ with $i_{j}^{\star} \uparrow \infty$ such that for any $\mathbb{Z}_{>0}$-valued sequence $\left\{i_{j}\right\}_{j \in \mathbb{Z}_{>0}}$ satisfying $i_{j} \uparrow \infty$ and $i_{j} \leq i_{j}^{\star}, \lim _{j \rightarrow \infty} a_{i_{j}, j}=$ $a_{\infty, \infty}$.

Proof of Theorem 6.1 Let $B_{n}$ be as in (2.1). Note that on the event $B_{n}$, for any $\lambda \in \mathbb{R}$, the space $\mathrm{CBD}_{R_{n, \lambda}}\left(\mathscr{G}_{n, 3}^{\exp }\right)$ is simply $\mathfrak{G}_{1}(n, \lambda)$ together with some additional connected multigraphs (with edge lengths and red points) each of which is attached to a vertex of $\mathfrak{G}_{1}(n, \lambda)$ via a single edge that has at least one red point on it. Thus, by (4.2), on the event $B_{n}, \mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}^{\exp }\right)$ is $\mathrm{CBD}_{\infty}\left(\mathfrak{G}_{1}(n, \lambda)\right)$ with some additional trees, say $\mathfrak{T}_{n, \lambda}^{(j)}$, $1 \leq j \leq k_{n}(\lambda)$, each of which is attached to a vertex of $\operatorname{CBD}_{\infty}\left(\mathfrak{G}_{1}(n, \lambda)\right)$ via a single edge. Define $k_{n}(\lambda)=0$ on $B_{n}^{c}$ for all $\lambda \in \mathbb{R}$.

Using Lemma 4.4(a), there exists a coupling of the processes $\left(\operatorname{CBD}_{i}\left(\mathscr{G}_{n, 3}^{\exp }\right), i \geq 1\right)$ and $\left(\operatorname{CBD}_{t}\left(\mathscr{G}_{n, 3}\right), t \geq 1\right)$ such that the $j$-th distinct edge sampled is the same in
both processes, $1 \leq j \leq 3 n / 2$. In this coupling, on the event $B_{n}, \mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right)$ is $\operatorname{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$ with Shape $\left[\mathfrak{T}_{n, \lambda}^{(j)}\right], 1 \leq j \leq k_{n}(\lambda)$, attached to its vertices via an edge.

Now using (6.1), (6.2), Theorem 4.7, and Lemma 6.6, it follows that there exists a $\mathbb{Z}_{>0}$-valued sequence $\left\{\lambda_{n}^{\star}\right\}_{n \geq 1}$ with $\lambda_{n}^{\star} \uparrow \infty$ such that for any $\mathbb{Z}_{>0}$-valued sequence $\lambda_{n} \uparrow \infty$ satisfying $\lambda_{n} \leq \lambda_{n}^{\star}$,

$$
\begin{align*}
& n^{-1 / 3} \mathrm{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right) \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot \mathscr{M}, \text { and }  \tag{6.3}\\
& n^{-1 / 3} \mathrm{CBD}_{\infty}\left(\mathfrak{G}_{1}\left(n, \lambda_{n}\right)\right) \xrightarrow{\mathrm{d}}(0.75)^{1 / 3} \cdot \mathscr{M} \tag{6.4}
\end{align*}
$$

with respect to GH topology. Using (6.4) in conjunction with Proposition 6.3, Proposition 6.5, and (2.2), it follows that for any $\mathbb{Z}_{>0}$-valued sequence $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\star}$,

$$
\begin{equation*}
n^{-1 / 3} \max _{1 \leq j \leq k_{n}\left(\lambda_{n}\right)} \operatorname{diam}\left(\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right) \xrightarrow{\mathrm{d}} 0 . \tag{6.5}
\end{equation*}
$$

Denoting the edge lengths of $\mathscr{G}_{n, 3}^{\exp }$ by $\Gamma_{1}^{(e)}, e \in E\left(\mathscr{G}_{n, 3}\right)$, we have, for any $\varepsilon>0$ and any $c>0$,

$$
\begin{aligned}
& \mathbb{P}\left(\max _{1 \leq j \leq k_{n}\left(\lambda_{n}\right)} \operatorname{diam}\left(\operatorname{Shape}\left[\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right]\right) \geq \varepsilon n^{1 / 3}\right) \leq \mathbb{P}\left(\max _{1 \leq j \leq k_{n}\left(\lambda_{n}\right)} \operatorname{diam}\left(\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right) \geq c \varepsilon n^{1 / 3}\right) \\
& \quad+\mathbb{P}\left(\mathscr{G}_{n, 3} \text { contains a self-avoiding path } P \text { with }|P| \geq \varepsilon n^{1 / 3} \text { and } \sum_{e \in P} \Gamma_{1}^{(e)} \leq c \varepsilon n^{1 / 3}\right) .
\end{aligned}
$$

Thus, using Proposition 6.2 together with (6.5), we get

$$
\begin{equation*}
n^{-1 / 3} \max _{1 \leq j \leq k_{n}\left(\lambda_{n}\right)} \operatorname{diam}\left(\text { Shape }\left[\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right]\right) \xrightarrow{\mathrm{d}} 0, \tag{6.6}
\end{equation*}
$$

which in turn shows that

$$
\begin{equation*}
n^{-1 / 3} d_{H}\left(\mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right), \mathrm{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)\right) \xrightarrow{\mathrm{d}} 0 . \tag{6.7}
\end{equation*}
$$

Finally, by Lemma 4.5,

$$
\begin{equation*}
M_{n} \stackrel{\mathrm{~d}}{=} \mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right) . \tag{6.8}
\end{equation*}
$$

The result now follows from (6.3) and (6.7).

### 6.2 Proof of Proposition 6.2

Fix $m \geq 1$ and $k \geq m$. Consider a self-avoiding path $P$ in $G$ with $|P|=k$. Then for any $c>0$ and any $t>0$,

$$
\begin{equation*}
\mathbb{P}\left(\sum_{e \in P} \Gamma_{1}^{(e)} \leq c k\right)=\mathbb{P}(Z \geq k) \leq e^{-t k} \mathbb{E}[\exp (t Z)]=\exp \left(-k\left(t-c\left(e^{t}-1\right)\right)\right) \tag{6.9}
\end{equation*}
$$

where $Z$ is a Poisson $(c k)$ random variable. Hence,

$$
\begin{aligned}
& \mathbb{P}\left(G \text { contains a self-avoiding path } P \text { with }|P| \geq m \text { and } \sum_{e \in P} \Gamma_{1}^{(e)} \leq c|P|\right) \\
& \quad \leq \sum_{k \geq m} \mathbb{P}\left(G \text { contains a self-avoiding path } P \text { with }|P|=k \text { and } \sum_{e \in P} \Gamma_{1}^{(e)} \leq c k\right) \\
& \quad \leq \sum_{k \geq m}|V| r^{k} \exp \left(-k\left(t-c\left(e^{t}-1\right)\right)\right) \\
& \quad=|V| \sum_{k \geq m} \exp \left(-k\left(t-\log r-c\left(e^{t}-1\right)\right)\right) \leq|V| e^{-m},
\end{aligned}
$$

where the second inequality uses (6.9) and the fact that there are at most $|V| r^{k}$ many self-avoiding paths of length $k$ in $G$, and the last step follows if we choose $t$ sufficiently large and $c$ sufficiently small. This completes the proof.

### 6.3 Proof of Proposition 6.3

Recall Construction 3.9. Let $s, n, r, \mathscr{K}_{n, 3}, \mathscr{T}_{1}, \ldots, \mathscr{T}_{r}, \mathscr{T}_{1}^{\prime}, \ldots, \mathscr{T}_{r}^{\prime}, \rho_{i}, z_{i}$, and ( $X_{1}, \ldots, X_{r}$ ) be as in Construction 3.9. Using (2.2), we can assume that $\mathscr{K}_{n, 3}$ and $\mathscr{G}_{n, 3}$ are coupled in a way so that

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \mathbb{P}\left(\mathscr{K}_{n, 3} \neq \mathscr{G}_{n, 3}\right)=0 \tag{6.10}
\end{equation*}
$$

Let $\left\{\Gamma_{1 / 2}^{(j)}\right\}_{1 \leq j \leq r}$ be a sequence of i.i.d. $\operatorname{Gamma}(1 / 2,1)$ random variables. Then

$$
\begin{equation*}
\left(X_{1}, \ldots, X_{r}\right) \stackrel{\mathrm{d}}{=}\left(\Gamma_{1 / 2}^{(1)}, \ldots, \Gamma_{1 / 2}^{(r)}\right) / \Gamma_{r / 2} \tag{6.11}
\end{equation*}
$$

where $\Gamma_{r / 2}=\sum_{j=1}^{r} \Gamma_{1 / 2}^{(j)}$. Note also that

$$
\begin{equation*}
r^{-1} \Gamma_{r / 2} \xrightarrow{\mathrm{~d}} 1 / 2, \text { as } r \rightarrow \infty . \tag{6.12}
\end{equation*}
$$

Let $Y_{i}:=d_{\mathscr{T}_{i}}\left(\rho_{i}, z_{i}\right)$. Then $Y_{i}, 1 \leq i \leq r$, are i.i.d. Rayleigh random variables [8,9] with density $f(y)=y \exp \left(-y^{2} / 2\right), y>0$. Hence $Y_{i}^{2} \stackrel{\mathrm{~d}}{=} 2 \Gamma_{1}$, where $\Gamma_{1} \sim$

Exponential(1). By [68], for $i=1, \ldots, r, Z_{i}^{2}:=Y_{i}^{2} \Gamma_{1 / 2}^{(i)} \stackrel{\mathrm{d}}{=} \Gamma_{1}^{2} / 2$. Hence
$\left\{\sqrt{2} Z_{i}\right\}_{1 \leq i \leq r}$ is an i.i.d. sequence of Exponential(1) random variables.
Let $\mathfrak{Q}_{n, 3}=\left(k\left(\mathscr{H}^{(s)}\right), e\left(\mathscr{H}^{(s)}\right)\right.$, len) be the multigraph with edge lengths that represents Core $\left(\mathscr{H}^{(s)}\right)$. As observed right below Construction 3.9, $\mathfrak{Q}_{n, 3}$ can be constructed by assigning length $Y_{i} \sqrt{X_{i}}=Z_{i} / \sqrt{\Gamma_{r / 2}}$ to the $i$-th edge of $\mathscr{K}_{n, 3}$, $1 \leq i \leq r$. There is a natural coupling between $\mathrm{CB}^{\infty}\left(\mathfrak{Q}_{n, 3}\right)$ and $\mathrm{CB}^{\infty}\left(\mathscr{H}^{(s)}\right)$ in which $\mathrm{CB}^{\infty}\left(\mathscr{H}^{(s)}\right)$ can be obtained by attaching countably many real trees to $\mathrm{CB}^{\infty}\left(\mathfrak{Q}_{n, 3}\right)$, and the diameter of each such real tree is at most $\max _{1 \leq i \leq r} \operatorname{diam}\left(\mathscr{T}_{i}^{\prime}\right)=$ $\max _{1 \leq i \leq r} \sqrt{X_{i}} \cdot \operatorname{diam}\left(\mathscr{T}_{i}\right)$. Thus, in this coupling,

$$
\begin{equation*}
(12 s)^{1 / 6} \cdot d_{H}\left(\mathrm{CB}^{\infty}\left(\mathfrak{Q}_{n, 3}\right), \mathrm{CB}^{\infty}\left(\mathscr{H}^{(s)}\right)\right) \leq(12 s)^{1 / 6} \max _{1 \leq i \leq r} \sqrt{X_{i}} \cdot \operatorname{diam}\left(\mathscr{T}_{i}\right) \xrightarrow{\mathrm{d}} 0, \tag{6.14}
\end{equation*}
$$

where the last step is a consequence of the facts $\max _{1 \leq i \leq r} X_{i}=O_{P}(\log r / r)$ (which can be seen from (6.11) and (6.12)), and $\max _{1 \leq i \leq r} \operatorname{diam}(\mathscr{T})=O_{P}(\sqrt{\log r})$.

Now, in the coupling used in (4.1),

$$
\begin{equation*}
(12 s)^{1 / 6} \cdot d_{H}\left(\mathrm{CBD}_{\infty}\left(\mathfrak{Q}_{n, 3}\right), \mathrm{CB}^{\infty}\left(\mathfrak{Q}_{n, 3}\right)\right) \leq(12 s)^{1 / 6} \cdot \max _{1 \leq i \leq r} Z_{i} / \sqrt{\Gamma_{r / 2}} \xrightarrow{\mathrm{~d}} 0 \tag{6.15}
\end{equation*}
$$

where the last step follows from (6.13) and (6.12). Combining (6.15) with (6.14) and Theorem 4.8 , we see that as $n \rightarrow \infty$,

$$
\begin{equation*}
(12 s)^{1 / 6} \mathrm{CBD}_{\infty}\left(\mathfrak{Q}_{n, 3}\right) \xrightarrow{\mathrm{d}} \mathscr{M} \quad \text { w.r.t. GH topology. } \tag{6.16}
\end{equation*}
$$

Finally, using (6.12) and the relation $r=3(s-1)$, we see that the length of the $i$-th edge in $(12 s)^{1 / 6} \mathfrak{Q}_{n, 3}$ is

$$
(12 s)^{1 / 6} \frac{Z_{i}}{\sqrt{\Gamma_{r / 2}}}=\left(1+o_{P}(1)\right) \cdot\left(\frac{4}{3 n}\right)^{1 / 3} \sqrt{2} \cdot Z_{i}
$$

which together with (6.16) implies that as $n \rightarrow \infty$,

$$
\begin{equation*}
\left(\frac{4}{3 n}\right)^{1 / 3} \cdot \mathrm{CBD}_{\infty}\left(\mathscr{K}_{n, 3}^{\exp }\right) \xrightarrow{\mathrm{d}} \mathscr{M} \text { w.r.t. GH topology, } \tag{6.17}
\end{equation*}
$$

where $\mathscr{K}_{n, 3}^{\exp }$ is the multigraph obtained by assigning lengths $\sqrt{2} Z_{i}, 1 \leq i \leq r$ (which, by (6.13), are i.i.d. Exponential(1) random variables) lengths to the edges of $\mathscr{K}_{n, 3}$. We complete the proof by combining (6.17) and (6.10).

### 6.4 Proof of Proposition 6.4

As in Sect. 5, for any graph $H$ and $p \in[0,1]$, we denote by $\operatorname{Perc}(H, p)$ the random subgraph of $H$ obtained by removing edges of $H$ independently with probability $1-p$. The proof of Proposition 6.4 relies on the following three lemmas.

Lemma 6.7 Let $H$ be a finite multigraph. Let $H^{\exp }$ be the multigraph with edge lengths obtained by assigning i.i.d. Exponential(1) lengths to the edges of H. Fix $t>0$, and let $R(t)$ be a $\operatorname{Poisson}\left(t \cdot \ell\left(H^{\exp }\right)\right)$ random variable. Then

$$
\left(\operatorname{Shape}\left[\operatorname{Rem}\left(C B D_{R(t)}\left(H^{\exp }\right)\right)\right], \operatorname{Rem}\left(\operatorname{CBD}_{R(t)}\left(H^{\exp }\right)\right)\right)
$$

has the same distribution as

$$
\left(\operatorname{Perc}\left(H, \frac{1}{1+t}\right), \frac{1}{1+t} \cdot\left(\operatorname{Perc}\left(H, \frac{1}{1+t}\right)\right)^{\exp }\right)
$$

where the last graph denotes the multigraph obtained by assigning i.i.d. Exponential(1) lengths to the edges of $\operatorname{Perc}(H, 1 /(1+t))$, and then multiplying the lengths by $1 /(1+t)$ (or equivalently, assigning i.i.d.exponential lengths with mean $1 /(1+t)$ to the edges of $\operatorname{Perc}(H, 1 /(1+t)))$.

Next we state two results about the behavior of the configuration model when a uniform subset of edges of given size is removed. Recall the notation $\mathscr{G}_{n, \mathbf{d}}$ from Definition 1.4.
Lemma 6.8 Suppose $\mathbf{d}=\left(d_{1}, \ldots, d_{n}\right)$ is a degree sequence and $\ell_{n}=\sum_{v \in[n]} d_{i}$. Let $m \leq \ell_{n} / 2$ and define $\ell_{n}^{\prime}=\ell_{n}-2 m$. Out of the $\ell_{n} / 2$ edges of $\mathscr{G}_{n, \mathbf{d}}$, sample a subset of $m$ edges uniformly. Let $\mathscr{G}_{n, \mathbf{d}}^{(m)}$ be the graph obtained by removing those $m$ edges. Then

$$
\begin{equation*}
\left(\mathscr{G}_{n, \mathbf{d}}, \mathscr{G}_{n, \mathbf{d}}^{(m)}\right) \stackrel{d}{=}\left(\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}, \mathscr{Q}_{n, \mathbf{d}, m}^{(1)}\right), \tag{6.18}
\end{equation*}
$$

where the pair $\left(\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}, \mathscr{Q}_{n, \mathbf{d}, m}^{(1)}\right)$ is constructed as follows: Start with the vertex set $[n]$ with $d_{i}$ many half-edges attached to vertex $i$. Sample $\ell_{n}^{\prime}$ many half-edges uniformly, and construct $\mathscr{Q}_{n, \mathbf{d}, m}^{(1)}$ by uniformly pairing up those $\ell_{n}^{\prime}$ half-edges. Conditional on this step, uniformly pair the rest of the half-edges to form $\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}$.

Consequently, if $p \in[0,1]$ and $m$ is a Binomial $\left(\ell_{n} / 2,1-p\right)$ random variable independent of $\mathscr{G}_{n, \mathbf{d}}$, then

$$
\begin{equation*}
\mathscr{D}_{n, \mathbf{d}, m}^{(1)} \stackrel{d}{=} \operatorname{Perc}\left(\mathscr{G}_{n, \mathbf{d}}, p\right) \tag{6.19}
\end{equation*}
$$

Equality in both coordinates in (6.18) will be used later in Sect. 6.6. In the proof of Theorem 6.1 we will only need (6.19), which is a consequence of $\mathscr{G}_{n, \mathbf{d}}^{(m)} \stackrel{\mathrm{d}}{=} \mathscr{Q}_{n, \mathbf{d}, m}^{(1)}$, i.e., just the equality of the second coordinates in (6.18). The relation $\mathscr{G}_{n, \mathbf{d}}^{(m)} \stackrel{\text { d }}{=} \mathscr{Q}_{n, \mathbf{d}, m}^{(1)}$ was already observed in [44, Lemmas 3.1 and 3.2]. See also [54] for a related construction.

Lemma 6.9 Suppose $m=m(n)$ satisfies

$$
\begin{equation*}
n^{-2 / 3}(3 n-4 m) \rightarrow \lambda_{0} \text { as } n \rightarrow \infty \tag{6.20}
\end{equation*}
$$

for some $\lambda_{0} \in \mathbb{R}$. Let $\mathscr{G}_{n, 3}^{(m)}$ be the graph obtained by removing a uniform subset of $m$ edges from $\mathscr{G}_{n, 3}$. Let $\mathbf{d}^{\prime}=\mathbf{d}^{\prime}(n):=\left(d_{1}^{\prime}, \ldots, d_{n}^{\prime}\right)$, where $d_{v}^{\prime}$ is the (random) degree of $v$ in $\mathscr{G}_{n, 3}^{(m)}$. Let $v$ be the $\operatorname{Binomial}(3,1 / 2)$ distribution. Then as $n \rightarrow \infty$,

$$
\begin{array}{r}
v_{n}^{\prime}(i):=\frac{1}{n} \#\left\{v \in[n]: d_{v}^{\prime}=i\right\} \xrightarrow{d} v(i), i=0,1,2,3 \\
\frac{1}{n} \sum_{v \in[n]}\left(d_{v}^{\prime}\right)^{3} \xrightarrow{d} \sum_{i=0}^{3} i^{3} v(i), \text { and } n^{1 / 3}\left(\frac{\sum_{v \in[n]}\left(d_{v}^{\prime}\right)^{2}}{\sum_{v \in[n]} d_{v}^{\prime}}-2\right) \xrightarrow{d} \frac{\lambda_{0}}{3} . \tag{6.22}
\end{array}
$$

Remark 4 Loosely speaking, Lemma 6.9 says that the sequence $\left\{\mathbf{d}^{\prime}(n)\right\}_{n=2,4, \ldots}$ of random degree sequences satisfies Assumption 3.4 in probability. This in particular implies that results for configuration models that only require Assumption 3.4 (e.g., Theorem 3.13) apply directly to the random graphs $\mathscr{G}_{n, 3}^{(m)}$. This can be argued as follows: By Skorohod representation theorem, we can construct $\boldsymbol{v}_{n}^{\prime \prime}:=\left(v_{n}^{\prime \prime}(i) ; i=0,1,2,3\right)$, $n=2,4,6, \ldots$, on the same probability space such that $\boldsymbol{v}_{n}^{\prime \prime} \stackrel{\mathrm{d}}{=}\left(v_{n}^{\prime}(i) ; i=0,1,2,3\right)$ for $n=2,4,6, \ldots$, and that

$$
v_{n}^{\prime \prime}(i) \xrightarrow{\text { a.s. }} v(i), i=0,1,2,3, \quad \text { and } \quad n^{1 / 3}\left(\frac{\sum_{i=0}^{3} i^{2} v_{n}^{\prime \prime}(i)}{\sum_{i=0}^{3} i v_{n}^{\prime \prime}(i)}-2\right) \xrightarrow{\text { a.s. }} \frac{\lambda_{0}}{3}
$$

in this space. We can further assume that $\pi_{2}, \pi_{4}, \pi_{6}, \ldots$ are also defined on this space, where (a) $\pi_{n}$ is a uniform permutation of $n$ elements for $n=2,4, \ldots$; (b) $\pi_{2}, \pi_{4}, \pi_{6}, \ldots$ are independent; and (c) $\left(\pi_{n} ; n=2,4, \ldots\right)$ is independent of $\left(\boldsymbol{v}_{n}^{\prime \prime} ; n=2,4, \ldots\right)$. For $n=2,4,6, \ldots$, let $\mathbf{d}^{\prime \prime}(n)$ be the random sequence of length $n$ obtained by applying $\pi_{n}$ to the sequence $(0, \ldots, 0,1, \ldots, 1,2, \ldots, 2,3, \ldots, 3)$ with $i$ appearing $n v_{n}^{\prime \prime}(i)$ many times, $i=0,1,2,3$. Then $\mathbf{d}^{\prime \prime}(n) \stackrel{\text { d }}{=} \mathbf{d}^{\prime}(n)$ for $n=2,4,6, \ldots$, and further, in this space, the convergences in (6.21) and (6.22) hold almost surely if $\mathbf{d}^{\prime}(n)$ is replaced by $\mathbf{d}^{\prime \prime}(n)$. Conditional on $\left(\mathbf{d}^{\prime \prime}(n) ; n=2,4, \ldots\right)$, construct $H_{2}, H_{4}, \ldots$, where $H_{n}$ is distributed as a configuration model with degree sequence $\mathbf{d}^{\prime \prime}(n)$. (The exact way in which $H_{2}, H_{4}, \ldots$ are coupled is not important here. For definiteness, let us take them to be independent conditional on ( $\left.\mathbf{d}^{\prime \prime}(n) ; n=2,4, \ldots\right)$.) Then Theorem 3.13 applies to the sequence of random graphs $H_{n}$. Now, from the equality in the second coordinate in (6.18), we see that $\mathscr{G}_{n, 3}^{(m)}$, conditional on $\mathbf{d}^{\prime}(n)$, is distributed as a configuration model with degree sequence $\mathbf{d}^{\prime}(n)$. Hence, $\mathscr{G}_{n, 3}^{(m)} \stackrel{\text { d }}{=} H_{n}$, and consequently, Theorem 3.13 applies to the random graphs $\mathscr{G}_{n, 3}^{(m)}$.

We now prove Proposition 6.4 assuming the above three lemmas.

Completing the proof of Proposition 6.4 We first note that if $v$ is the $\operatorname{Binomial}(3,1 / 2)$ distribution, then

$$
\begin{equation*}
\sigma_{1}(v)=3 / 2, \quad \sigma_{2}(v)=3, \quad \text { and } \sigma_{3}(v)=27 / 4 \tag{6.23}
\end{equation*}
$$

Next, by Lemma 6.7,

$$
\begin{equation*}
\operatorname{Shape}\left[\operatorname{Rem}\left(\operatorname{CBD}_{R_{n, \lambda}}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right)\right] \stackrel{\mathrm{d}}{=} \operatorname{Perc}\left(\mathscr{G}_{n, 3}, 1 / 2+\lambda n^{-1 / 3}\right) \tag{6.24}
\end{equation*}
$$

Now, for any $p \in[0,1]$, the number of edges removed from $\mathscr{G}_{n, 3}$ to construct $\operatorname{Perc}\left(\mathscr{G}_{n, 3}, p\right)$ is a $\operatorname{Binomial}(3 n / 2,1-p)$ random variable. In particular, when $p=$ $1 / 2+\lambda n^{-1 / 3}$, the number of edges removed is

$$
m=\frac{3 n}{2}\left(\frac{1}{2}-\frac{\lambda}{n^{1 / 3}}\right)+O_{P}(\sqrt{n})
$$

which satisfies (6.20) with $\lambda_{0}=6 \lambda$. Further, conditional on $m, \operatorname{Perc}\left(\mathscr{G}_{n, 3}, 1 / 2+\right.$ $\lambda n^{-1 / 3}$ ) is distributed as $\mathscr{G}_{n, 3}^{(m)}$, where the notation is as in Lemma 6.9. Thus, by Lemma 6.9, the (random) degree sequence of $\operatorname{Perc}\left(\mathscr{G}_{n, 3}, 1 / 2+\lambda n^{-1 / 3}\right)$ satisfies (6.21) and (6.22) with limiting parameter $\lambda_{0} / 3=2 \lambda$. Finally, using (6.19), it follows that $\operatorname{Perc}\left(\mathscr{G}_{n, 3}, 1 / 2+\lambda n^{-1 / 3}\right)$, conditional on its degree sequence, is distributed as a configuration model with that degree sequence. Hence by Theorem 3.5, Theorem 3.13, (6.23), and (6.24),

$$
\begin{align*}
& \mathscr{V}=\Theta_{P}\left(n^{2 / 3}\right), \mathscr{E}=\Theta_{P}\left(n^{2 / 3}\right), \mathscr{S}=O_{P}(1), \mathscr{D}=\Theta_{P}\left(n^{1 / 3}\right), \text { and }(6.25) \\
& n^{-1 / 3} \cdot G_{1}(n, \lambda) \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot S_{1}\left((48)^{1 / 3} \cdot \lambda\right) \text { w.r.t. GH topology, } \tag{6.26}
\end{align*}
$$

where $\mathscr{V}=\left|G_{1}(n, \lambda)\right|, \mathscr{E}=\left|E\left(G_{1}(n, \lambda)\right)\right|, \mathscr{S}=\operatorname{sp}\left(G_{1}(n, \lambda)\right)$, and $\mathscr{D}$ denotes the diameter of $G_{1}(n, \lambda)$. By Lemma 6.7, conditional on $G_{1}(n, \lambda)$, the lengths of the edges of $\mathfrak{G}_{1}(n, \lambda)$ are

$$
\begin{equation*}
\left(1 / 2+\lambda n^{-1 / 3}\right) \cdot\left(\Gamma_{1}^{(1)}, \ldots, \Gamma_{1}^{(\mathscr{E})}\right) \tag{6.27}
\end{equation*}
$$

where $\Gamma_{1}^{(1)}, \ldots, \Gamma_{1}^{(\mathscr{E})}$ are i.i.d. Exponential(1) random variables.
Now it is easy to see that for any two vertices in $G_{1}(n, \lambda)$, there are at most $2^{\mathscr{S}}$ many self-avoiding paths connecting them, and the length of any such self-avoiding path is at most $6(\mathscr{S}+1) \mathscr{D}$. For any such self-avoiding path $P$ and any $\eta>0$, by standard concentration inequalities,

$$
\begin{equation*}
\mathbb{P}_{G_{1}}\left(\left|\sum_{j \in P} \Gamma_{1}^{(j)}-1\right| \geq(6(\mathscr{S}+1) \mathscr{D})^{1 / 2+\eta}\right) \leq \exp \left(-c(6(\mathscr{S}+1) \mathscr{D})^{2 \eta}\right), \tag{6.28}
\end{equation*}
$$

where $\mathbb{P}_{G_{1}}$ denotes probability conditional on $G_{1}(n, \lambda)$. Let $G_{1}^{\exp }(n, \lambda)$ be the graph with edge lengths obtained by assigning lengths $\Gamma_{1}^{(1)}, \ldots, \Gamma_{1}^{(\mathscr{E})}$ to the edges of $G_{1}(n, \lambda)$. Then by (6.28),
$\mathbb{P}_{G_{1}}\left(d_{\mathrm{GH}}\left(G_{1}^{\exp }(n, \lambda), G_{1}(n, \lambda)\right) \geq(6(\mathscr{S}+1) \mathscr{D})^{1 / 2+\eta}\right) \leq \mathscr{V}^{2} \cdot 2^{\mathscr{S}} \exp \left(-c(6(\mathscr{S}+1) \mathscr{D})^{2 \eta}\right)$.
Thus, by (6.25) and (6.26), $n^{-1 / 3} \cdot G_{1}^{\exp }(n, \lambda) \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot S_{1}\left((48)^{1 / 3} \cdot \lambda\right)$ w.r.t. GH topology, which together with (6.27) implies

$$
\begin{equation*}
n^{-1 / 3} \cdot \mathfrak{G}_{1}(n, \lambda) \xrightarrow{\mathrm{d}} \frac{1}{2} \cdot 6^{1 / 3} \cdot S_{1}\left((48)^{1 / 3} \cdot \lambda\right) \tag{6.29}
\end{equation*}
$$

w.r.t. GH topology. The claim now follows from (6.26) and (6.29) by using Theorem 4.6 and Theorem 3.14.

The rest of this section is devoted to the proofs of Lemmas 6.7, 6.8, and 6.9 .
Proof of Lemma 6.7 Let $|E(H)|=r$. Run two independent Poisson point processes (PPP)-a 'red' PPP and a 'blue' PPP, with intensities $t$ and 1 respectively. Let $X_{1}<$ $\ldots<X_{r}$ be the locations of the first $r$ blue points. Enumerate the edges of $H$ in any way, and assign length ( $X_{i}-X_{i-1}$ ) to the $i$-th edge, $i=1, \ldots, r$, where $X_{0}=0$. Call the resulting graph with edge lengths $H_{1}$. Let $\widetilde{R}$ be the number of red points in [ $0, X_{r}$ ]. Identifying the $i$-th edge of $H_{1}$ with the interval [ $X_{i-1}, X_{i}$ ], $i=1, \ldots, r$, place a red point on $H_{1}$ corresponding to the location of each of the $\widetilde{R}$ red points in [ $0, X_{r}$ ]. Call the resulting graph with red points $H_{2}$.

Now, note that $H_{1} \stackrel{\mathrm{~d}}{=} H^{\text {exp }}$. Next, conditional on the blue PPP, $\widetilde{R}$ follows a $\operatorname{Poisson}\left(t X_{r}\right) \equiv \operatorname{Poisson}\left(t \cdot \ell\left(H_{1}\right)\right)$ distribution. Thus,

$$
\left(H_{1}, \widetilde{R}\right) \stackrel{\mathrm{d}}{=}\left(H^{\exp }, R(t)\right)
$$

Finally, conditional on the blue PPP and $\widetilde{R}$, the locations of the red points in $\left[0, X_{r}\right]$ are i.i.d. Uniform $\left[0, X_{r}\right]$ random variables, which implies that

$$
\begin{equation*}
\operatorname{Rem}\left(\operatorname{CBD}_{R(t)}\left(H^{\exp }\right)\right) \stackrel{\mathrm{d}}{=} \operatorname{Rem}\left(H_{2}\right) \tag{6.30}
\end{equation*}
$$

Now $\operatorname{Rem}\left(H_{2}\right)$ can be generated in the following alternate way: Sample independent random variables $Z_{1}, \ldots, Z_{r}$, where $Z_{i} \sim \operatorname{Poisson}\left(t\left(X_{i}-X_{i-1}\right)\right), i=1, \ldots, r$. (Here $Z_{i}$ corresponds to the number of red points in $\left[X_{i-1}, X_{i}\right]$.) Remove the $i$-th edge of $H$ iff $Z_{i} \geq 1$, and assign independent lengths $Y_{i}$ to the remaining edges, where $Y_{i} \stackrel{\mathrm{~d}}{=}\left(\left(X_{i}-X_{i-1}\right) \mid Z_{i}=0\right)$.

Combining (6.30) with the facts that $\mathbb{1}_{Z_{i}=0}, i \geq 1$, are i.i.d. Bernoulli $(1 /(1+t))$ random variables, and $\left(\left(X_{i}-X_{i-1}\right) \mid Z_{i}=0\right)$ has an exponential distribution with mean $1 /(1+t)$, it follows that

$$
\operatorname{Rem}\left(\operatorname{CBD}_{R(t)}\left(H^{\exp }\right)\right) \stackrel{\mathrm{d}}{=} \frac{1}{1+t} \cdot\left(\operatorname{Perc}\left(H, \frac{1}{1+t}\right)\right)^{\exp }
$$

Now the result follows immediately.
Proof of Lemma 6.8 Let $G$ be a graph on [ $n$ ] with degree sequence $\mathbf{d}$. Let $G^{\prime}$ be a subgraph of $G$. Let $\mathbf{d}^{\prime}=\left(d_{1}^{\prime}, \ldots, d_{n}^{\prime}\right)$ be the degree sequence of $G^{\prime}$. Let $x_{i j}$ (resp. $x_{i j}^{\prime}$ ) be the number of edges between $i$ and $j$ in $G$ (resp. $G^{\prime}$ ), $i \neq j$, and let $x_{i i}$ (resp. $x_{i i}^{\prime}$ ) denote the number of loops attached to vertex $i$ in $G$ (resp. $G^{\prime}$ ). Using (1.3) it follows that

$$
\begin{equation*}
\mathbb{P}\left(\mathscr{G}_{n, \mathbf{d}}=G, \mathscr{G}_{n, \mathbf{d}}^{(m)}=G^{\prime}\right)=\frac{1}{\left(\ell_{n}-1\right)!!} \times \frac{\prod_{i \in[n]} d_{i}!}{\prod_{i \in[n]} 2^{x_{i i}} \prod_{i \leq j} x_{i j}!} \times \frac{\prod_{i \leq j}\binom{x_{i j}}{x_{i j}^{\prime}}}{\binom{\ell_{n} / 2}{m}} \tag{6.31}
\end{equation*}
$$

and

$$
\begin{align*}
\mathbb{P}\left(\mathscr{Q}_{n, \mathbf{d}, m}^{(1)}=G^{\prime}, \mathscr{Q}_{n, \mathbf{d}, m}^{(2)}=G\right)= & \frac{\prod_{i \in[n]}\binom{d_{i}}{d_{i}^{\prime}}}{\binom{\ell_{n}}{\ell_{n}^{\prime}}} \times \frac{1}{\left(\ell_{n}^{\prime}-1\right)!!} \times \frac{\prod_{i \in[n]} d_{i}^{\prime}!}{\prod_{i \in[n]} 2^{x_{i i}^{\prime}} \prod_{i \leq j} x_{i j}^{\prime}!} \\
& \times \frac{1}{\left(\ell_{n}-\ell_{n}^{\prime}-1\right)!!} \times \frac{\prod_{i \in[n]}\left(d_{i}-d_{i}^{\prime}\right)!}{\prod_{i \in[n]} 2^{x_{i i}-x_{i i}^{\prime}} \prod_{i \leq j}\left(x_{i j}-x_{i j}^{\prime}\right)!} . \tag{6.32}
\end{align*}
$$

A direct computation shows that the right sides of (6.31) and (6.32) are equal. This completes the proof.

Proof of Lemma 6.9 We use the alternate construction of $\mathscr{G}_{n, 3}^{(m)}$ from Lemma 6.8. For each $v \in[n]$, let $f_{v, i}$ denote the $i$-th half edge attached to $v, i=1,2,3$. Let $E_{v, i}$ denote the event that $f_{v, i}$ is one of the $3 n-2 m$ selected half edges. Then

$$
\mathbb{P}\left(E_{v, i}\right)=(3 n-2 m) / 3 n, \text { for } 1 \leq i \leq 3 .
$$

and

$$
\mathbb{P}\left(E_{v_{1}, i_{1}} \cap E_{v_{2}, i_{2}}\right)=\frac{(3 n-2 m)(3 n-2 m-1)}{3 n(3 n-1)}, \text { whenever }\left(v_{1}, i_{1}\right) \neq\left(v_{2}, i_{2}\right) .
$$

Since $d_{v}^{\prime}=\sum_{i=1}^{3} \mathbb{1}\left\{E_{v, i}\right\}$,
$\mathbb{E}\left[\sum_{v \in[n]} d_{v}^{\prime 2}\right]=n \cdot \mathbb{E}\left[d_{1}^{\prime 2}\right]=n\left[3 \times \frac{(3 n-2 m)}{3 n}+6 \times \frac{(3 n-2 m)(3 n-2 m-1)}{3 n(3 n-1)}\right]$.

Using this relation, (6.20), and the fact that $\sum_{v \in[n]} d_{v}^{\prime}=3 n-2 m$, a direct computation shows that

$$
\begin{equation*}
\lim _{n \rightarrow \infty} n^{1 / 3}\left(\frac{\mathbb{E}\left[\sum_{v \in[n]} d_{v}^{\prime 2}\right]}{\sum_{v \in[n]} d_{v}^{\prime}}-2\right)=\frac{\lambda_{0}}{3} . \tag{6.33}
\end{equation*}
$$

Now it is straightforward to check that for any four distinct pairs $\left(v_{j}, i_{j}\right), 1 \leq j \leq 4$, each of the quantities $\operatorname{Cov}\left(\mathbb{1}\left\{E_{v_{1}, i_{1}}\right\}, \mathbb{1}\left\{E_{v_{2}, i_{2}}\right\}\right), \operatorname{Cov}\left(\mathbb{1}\left\{E_{v_{1}, i_{1}} \cap E_{v_{2}, i_{2}}\right\}, \mathbb{1}\left\{E_{v_{3}, i_{3}}\right\}\right)$, and $\operatorname{Cov}\left(\mathbb{1}\left\{E_{v_{1}, i_{1}} \cap E_{v_{2}, i_{2}}\right\}, \mathbb{1}\left\{E_{v_{3}, i_{3}} \cap E_{v_{4}, i_{4}}\right\}\right)$ is negative. Thus, for any $v_{1} \neq v_{2}$, $\operatorname{Cov}\left(d_{v_{1}}^{\prime 2}, d_{v_{2}}^{\prime 2}\right)<0$, which implies that

$$
\operatorname{Var}\left(\sum_{v \in[n]} d_{v}^{\prime 2}\right) \leq \sum_{v \in[n]} \operatorname{Var}\left(d_{v}^{\prime 2}\right)=O(n) .
$$

This combined with (6.33) proves the second convergence in (6.22).
Next, for $v \in[n]$ and $k=0,1,2,3$,

$$
\mathbb{P}\left(d_{v}^{\prime}=k\right)=\binom{3}{k}\binom{3 n-3}{2 m+k-3} /\binom{3 n}{2 m}
$$

which together with (6.20) yields

$$
\lim _{n \rightarrow \infty} \frac{1}{n} \mathbb{E}\left[\#\left\{v \in[n]: d_{v}^{\prime}=k\right\}\right]=\lim _{n \rightarrow \infty} \mathbb{P}\left(d_{1}^{\prime}=k\right)=v(k)
$$

A little computation will show that $\operatorname{Var}\left[\#\left\{v \in[n]: d_{v}^{\prime}=k\right\}\right]=O(n)$ for $k=$ $0,1,2,3$. This proves (6.21). Finally, the first convergence in (6.22) follows from (6.21). This completes the proof.

### 6.5 Proof of Proposition 6.5

We will use the following lemmas in the proof:
Lemma 6.10 Suppose $Y_{1}$ and $Y_{2}$ are two real valued random variables defined on the same probability space such that $Y_{1} \leq Y_{2}$ almost surely. Suppose further that $Y_{1} \stackrel{d}{=} Y_{2}$. Then $Y_{1}=Y_{2}$ almost surely.

This is an elementary lemma, and we omit the proof.
Lemma 6.11 Suppose $\left\{\left(Z_{n}^{+}, d_{n}, Z_{n}\right)\right\}_{n \geq 1}$ is a sequence in $\mathfrak{S}_{\mathrm{GH}}^{*}$ satisfying $\left(Z_{n}^{+}, d_{n}, Z_{n}\right) \rightarrow\left(Z_{0}^{+}, d, Z_{0}\right)$ for some marked space $\left(Z_{0}^{+}, d, Z_{0}\right)$. Then

$$
d_{H}\left(Z_{n}^{+}, Z_{n}\right) \rightarrow d_{H}\left(Z_{0}^{+}, Z_{0}\right) .
$$

Proof For any isometric embeddings $\phi_{n}: Z_{n}^{+} \rightarrow Z^{\star}$ and $\psi_{n}: Z_{0}^{+} \rightarrow Z^{\star}$ into some common space $Z^{\star}$, we have

$$
d_{H}\left(Z_{n}^{+}, Z_{n}\right) \leq d_{H}\left(\phi_{n}\left(Z_{n}^{+}\right), \psi_{n}\left(Z_{0}^{+}\right)\right)+d_{H}\left(Z_{0}^{+}, Z_{0}\right)+d_{H}\left(\psi_{n}\left(Z_{0}\right), \phi_{n}\left(Z_{n}\right)\right)
$$

Using symmetry, we see that

$$
\begin{equation*}
\left|d_{H}\left(Z_{n}^{+}, Z_{n}\right)-d_{H}\left(Z_{0}^{+}, Z_{0}\right)\right| \leq d_{H}\left(\phi_{n}\left(Z_{n}^{+}\right), \psi_{n}\left(Z_{0}^{+}\right)\right)+d_{H}\left(\psi_{n}\left(Z_{0}\right), \phi_{n}\left(Z_{n}\right)\right) \tag{6.34}
\end{equation*}
$$

Using the fact $\left(Z_{n}^{+}, d_{n}, Z_{n}\right) \rightarrow\left(Z_{0}^{+}, d, Z_{0}\right)$, we can choose $\phi_{n}, \psi_{n}$ in a way so that the right side of (6.34) goes to zero as $n \rightarrow \infty$.

We will now complete the proof of Proposition 6.5. For any compact metric space ( $X, d$ ) and $\delta>0$, let $N_{\delta}(X)$ be the minimum number of closed $\delta$ balls needed to cover $X$.

Since $X_{n}^{+} \xrightarrow{\mathrm{d}} Z$, the sequence $\left\{\left(X_{n}^{+}, d_{n}\right)\right\}_{n \geq 1}$ is relatively compact w.r.t. GH topology. Using Lemma 3.1(b), the sequence $\left\{\left(X_{n}^{+}, d_{n}, X_{n}\right)\right\}_{n \geq 1}$ is relatively compact w.r.t. the marked topology. Thus, there exists a subsequence $\left\{n_{k}\right\}_{k \geq 1}$ and a random marked space $\left(Z_{0}^{+}, d, Z_{0}\right)$ such that

$$
\left(X_{n_{k}}^{+}, d_{n_{k}}, X_{n_{k}}\right) \xrightarrow{\mathrm{d}}\left(Z_{0}^{+}, d, Z_{0}\right)
$$

as $k \rightarrow \infty$ with respect to the marked topology. Since $X_{n}^{+} \xrightarrow{\mathrm{d}} Z$ and $X_{n} \xrightarrow{\mathrm{~d}} Z$, we must have $Z_{0}^{+} \stackrel{\mathrm{d}}{=} Z \stackrel{\mathrm{~d}}{=} Z_{0}$ as compact metric spaces. In particular, for all $\varepsilon>0$,

$$
\begin{equation*}
N_{\varepsilon}\left(Z_{0}^{+}\right) \stackrel{\mathrm{d}}{=} N_{\varepsilon}\left(Z_{0}\right) \tag{6.35}
\end{equation*}
$$

Since $Z_{0}$ is a closed subset of $Z_{0}^{+}$, for every $\varepsilon>0, N_{\varepsilon}\left(Z_{0}\right) \leq N_{\varepsilon}\left(Z_{0}^{+}\right)$almost surely. Then it follows from (6.35) and Lemma 6.10 that

$$
\mathbb{P}\left(N_{\varepsilon}\left(Z_{0}^{+}\right)=N_{\varepsilon}\left(Z_{0}\right)\right)=1
$$

for every $\varepsilon>0$. This implies that $\mathbb{P}\left(d_{H}\left(Z_{0}^{+}, Z_{0}\right)=0\right)=1$. Thus, using Lemma 6.11, we conclude that $d_{H}\left(X_{n_{k}}^{+}, X_{n_{k}}\right) \xrightarrow{\mathrm{P}} 0$.

Now for any subsequence $\left\{m_{\ell}\right\}_{\ell \geq 1}$, using the above argument, we can extract a further subsequence $\left\{m_{\ell_{k}}\right\}_{k \geq 1}$ such that $d_{H}\left(X_{m_{\ell_{k}}}^{+}, X_{m_{\ell_{k}}}\right) \xrightarrow{\mathrm{P}} 0$ as $k \rightarrow \infty$. Thus the claim follows.

### 6.6 GHP convergence of the MST of $\mathscr{G}_{n, 3}$

In this section we improve the convergence in Theorem 6.1 to GHP convergence, thus completing the proof of Theorem 2.1. Let $\mathfrak{G}_{1}(n, \lambda)$ and $G_{1}(n, \lambda)$ be as in the
statement of Proposition 6.4, and let $k_{n}(\lambda)$ and $\mathfrak{T}_{n, \lambda}^{(j)}, 1 \leq j \leq k_{n}(\lambda)$, be as in the proof of Theorem 6.1. For $v \in \mathfrak{G}_{1}(n, \lambda)$, let $d_{v, \lambda}$ be the degree of $v$ in $\mathfrak{G}_{1}(n, \lambda)$, and define $d_{v, \lambda}^{\text {avail }}:=3-d_{v, \lambda}$. Thus, $d_{v, \lambda}^{\text {avail }}$ denotes the number of distinct edges sampled in the process $\left(\mathrm{CBD}_{i}\left(\mathscr{G}_{n, 3}^{\exp }\right), 1 \leq i \leq R_{n, \lambda}\right)$ that were incident to $v$, and one can picture this degree deficiency as 'available half-edges' attached to $v$.

On the event $B_{n}, \operatorname{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right)$ is $\operatorname{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$ with Shape $\left[\mathfrak{T}_{n, \lambda}^{(j)}\right], 1 \leq j \leq$ $k_{n}(\lambda)$, attached to its vertices via a single edge; let $T_{n, \lambda}^{(i)}(v), 1 \leq i \leq r_{v, \lambda}$, be the trees (arranged following some deterministic rule) attached to $v \in \operatorname{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$. Clearly, $0 \leq r_{v, \lambda} \leq d_{v, \lambda}^{\text {avail }}$. Thus, the collection of trees $T_{n, \lambda}^{(i)}(v), 1 \leq i \leq r_{v, \lambda}$, $v \in \operatorname{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$, is simply Shape $\left[\mathfrak{T}_{n, \lambda}^{(j)}\right], 1 \leq j \leq k_{n}(\lambda)$, in some order. Recall from the proof of Theorem 6.1 that we define $k_{n}(\lambda)=0$ for all $\lambda \in \mathbb{R}$ on $B_{n}^{c}$. Accordingly, we set $r_{v, \lambda}=0$ for all $v \in \operatorname{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$ on the event $B_{n}^{c}$.

Construct the spaces $\mathfrak{M}_{n, \lambda}^{\text {attach }}$ and $\mathfrak{M}_{n, \lambda}^{\text {avail }}$ by endowing $\operatorname{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$ with the tree distance and respectively assigning mass

$$
\begin{align*}
& p_{v, \lambda}^{\text {attach }}:=\left\{\begin{array}{l}
1 /\left|G_{1}(n, \lambda)\right|, \text { on } B_{n}^{c}, \\
\frac{1}{n}\left(1+\sum_{i=1}^{r_{v, \lambda}}\left|T_{n, \lambda}^{(i)}(v)\right|\right), \text { on } B_{n},
\end{array}\right. \\
& \text { and } p_{v, \lambda}^{\text {avail }}:=\left\{\begin{array}{l}
1 /\left|G_{1}(n, \lambda)\right|, \text { if } \sum_{u \in G_{1}(n, \lambda)} d_{u, \lambda}^{\text {avail }}=0, \\
d_{v, \lambda}^{\text {avail }} /\left(\sum_{u \in G_{1}(n, \lambda)} d_{u, \lambda}^{\text {avail }}\right), \text { otherwise, }
\end{array}\right. \tag{6.36}
\end{align*}
$$

to $v \in \operatorname{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$. Note that $\sum_{v \in G_{1}(n, \lambda)} p_{v, \lambda}^{\text {attach }}=\sum_{v \in G_{1}(n, \lambda)} p_{v, \lambda}^{\text {avail }}=1$. Note also that the first and the third asymptotics in (6.25) imply that $\mathbb{P}\left(\sum_{u \in G_{1}(n, \lambda)} d_{u, \lambda}^{\text {avail }}=\right.$ $0) \rightarrow 0$ as $n \rightarrow \infty$. Thus, the value of $p_{v, \lambda}^{\text {avail }}$ on the event $\left\{\sum_{u \in G_{1}(n, \lambda)} d_{u, \lambda}^{\text {avail }} \geq 1\right\}$ is the one relevant for distributional asymptotics of $\mathfrak{M}_{n, \lambda}^{\text {avail }}$. Similarly, using (2.2), only the value of $p_{v, \lambda}^{\text {attach }}$ on $B_{n}$ is relevant for the asymptotic behavior of $\mathfrak{M}_{n, \lambda}^{\text {attach }}$.

Throughout Sect. 6.6, all sequences $\left\{\lambda_{n}\right\}_{n \geq 1}$ will be $\mathbb{Z}_{>0}$-valued sequences, and we will not mention this explicitly.

Lemma 6.12 Let $\lambda_{n}^{\star}$ be as in the proof of Theorem 6.1. Then for all $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\star}$,

$$
n^{-1 / 3} d_{\mathrm{GHP}}\left(C B D_{\infty}\left(\mathscr{G}_{n, 3}\right), \mathfrak{M}_{n, \lambda_{n}}^{\text {attach }}\right) \xrightarrow{P} 0 .
$$

Lemma 6.13 There exists a sequence $\lambda_{n}^{\dagger} \uparrow \infty$ such that for all $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\dagger}$,

$$
n^{-1 / 3} \mathfrak{M}_{n, \lambda_{n}}^{\text {avail }} \xrightarrow{d} 6^{1 / 3} \cdot \mathscr{M} \quad \text { w.r.t. the GHP topology. }
$$

Lemma 6.14 There exists a sequence $\lambda_{n}^{\circ} \uparrow \infty$ such that for all $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\circ}$,

$$
n^{-1 / 3} d_{\mathrm{GHP}}\left(\mathfrak{M}_{n, \lambda_{n}}^{\text {attach }}, \mathfrak{M}_{n, \lambda_{n}}^{\text {avail }}\right) \xrightarrow{P} 0 .
$$

Completing the proof of Theorem 2.1 The result follows upon combining Lemmas 6.12, $6.13,6.14$, and (6.8).

Proof of Lemma 6.12 On the event $B_{n}$, define the correspondence $C$ between $\mathfrak{M}_{n, \lambda_{n}}^{\text {attach }}$ and $\mathrm{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right)$ as follows:

$$
C:=\left\{(v, u): v \in \mathfrak{M}_{n, \lambda_{n}}^{\text {attach }} \text { and } u \in\{v\} \cup\left(\bigcup_{i=1}^{r_{v, \lambda_{n}}} T_{n, \lambda_{n}}^{(i)}(v)\right)\right\} .
$$

Let $\pi$ be a measure on $\mathfrak{M}_{n, \lambda_{n}}^{\text {attach }} \times \operatorname{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right)$ given by $\pi(\{(v, u)\})=1 / n$ for $(v, u) \in C$. Then with this choice of $C$ and $\pi$, the claim follows immediately if we use (6.6).

Proof of Lemma 6.13 Assign mass $p_{v, \lambda}^{\text {avail }}$ to $v \in G_{1}(n, \lambda)$ and call the resulting metric measure space $G_{1}^{\text {avail }}(n, \lambda)$. Using Theorem 3.13 with $f(k)=3-k, k=0, \ldots, 3$, and the arguments used to prove (6.26), we see that $n^{-1 / 3} G_{1}^{\text {avail }}(n, \lambda) \xrightarrow{\mathrm{d}} 6^{1 / 3}$. $S_{1}\left((48)^{1 / 3} \cdot \lambda\right)$ w.r.t. GHP topology. Using Theorems 3.14 and 4.6 , it follows that for each $\lambda \in \mathbb{R}$,

$$
n^{-1 / 3} \mathfrak{M}_{n, \lambda}^{\text {avail }} \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot \mathrm{CB}^{\infty}\left(S_{1}\left((48)^{1 / 3} \cdot \lambda\right)\right) \quad \text { as } n \rightarrow \infty
$$

w.r.t. GHP topology. The claim now follows from Theorem 4.7 and Lemma 6.6.

To prove Lemma 6.14 we will make use of Lemma 6.15 stated below. Let $\lambda_{n}^{\star}$ be as in the proof of Theorem 6.1.

Lemma 6.15 There exists a sequence $\lambda_{n}^{\diamond} \uparrow \infty$ such that
(i) $\lambda_{n}^{\diamond} \leq \lambda_{n}^{\star}$,
(ii) $\mathbb{P}\left(\left|\bar{G}_{1}\left(n, \lambda_{n}^{\diamond}\right)\right|>n / 2\right) \rightarrow 0$, and
(iii) for any $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\diamond}$, the following holds: For every $n$, fix an enumeration $v_{1}, v_{2}, \ldots$ of the vertices of $G_{1}\left(n, \lambda_{n}\right)$ measurable w.r.t. the $\sigma$-field generated by $C B D_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)$, and define

$$
Z_{n}:=\max _{1 \leq j_{1} \leq j_{2} \leq\left|G_{1}\left(n, \lambda_{n}\right)\right|}\left|\sum_{s=j_{1}}^{j_{2}} \sum_{i=1}^{r_{v_{s}, \lambda_{n}}} \frac{\left|T_{n, \lambda_{n}}^{(i)}\left(v_{s}\right)\right|}{n-\left|G_{1}\left(n, \lambda_{n}\right)\right|}-\sum_{s=j_{1}}^{j_{2}} p_{v_{s}, \lambda_{n}}^{\text {avail }}\right|,
$$

where $p_{v, \lambda}^{\text {avail }}$ is as defined in (6.36). Then $Z_{n} \xrightarrow{P} 0$.
We first prove Lemma 6.14 assuming Lemma 6.15.
Proof of Lemma 6.14 On the event $B_{n}$, construct $\mathfrak{M}_{n, \lambda}^{\text {modi }}$ by endowing $\mathrm{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$ with the tree distance and assigning mass

$$
p_{v, \lambda}^{\text {modi }}:=\frac{\sum_{i=1}^{r_{v, \lambda}}\left|T_{n, \lambda}^{(i)}(v)\right|}{n-\left|G_{1}(n, \lambda)\right|}
$$

to $v \in \mathrm{CBD}_{\infty}\left(G_{1}(n, \lambda)\right)$. On $B_{n}^{c}$, set $\mathfrak{M}_{n, \lambda}^{\text {modi }}=\mathfrak{M}_{n, \lambda}^{\text {attach }}$. As observed in (6.25), $\left|G_{1}(n, \lambda)\right|=\Theta_{P}\left(n^{2 / 3}\right)$. Thus,

$$
\sum_{v \in G_{1}(n, \lambda)}\left|p_{v, \lambda}^{\text {modi }}-p_{v, \lambda}^{\text {attach }}\right|=O_{P}\left(n^{-1 / 3}\right)
$$

It follows that for each $\lambda \in \mathbb{R}, d_{\mathrm{GHP}}\left(\mathfrak{M}_{n, \lambda}^{\text {modi }}, \mathfrak{M}_{n, \lambda}^{\text {attach }}\right) \xrightarrow{\mathrm{P}} 0$ as $n \rightarrow \infty$. Thus, we can choose a sequence $\lambda_{n}^{\oplus} \uparrow \infty$ such that $\mathbb{P}\left(\left|G_{1}\left(n, \lambda_{n}^{\oplus}\right)\right|>n / 2\right) \rightarrow 0$, and further, for all $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\oplus}$,

$$
\begin{equation*}
d_{\mathrm{GHP}}\left(\mathfrak{M}_{n, \lambda_{n}}^{\text {modi }}, \mathfrak{M}_{n, \lambda_{n}}^{\text {attach }}\right) \xrightarrow{\mathrm{P}} 0 . \tag{6.37}
\end{equation*}
$$

Set $\lambda_{n}^{\circ}:=\min \left\{\lambda_{n}^{\oplus}, \lambda_{n}^{\dagger}, \lambda_{n}^{\diamond}\right\}$, where $\lambda_{n}^{\dagger}$ (resp. $\lambda_{n}^{\diamond}$ ) is as in Lemma 6.13 (resp. Lemma 6.15). Fix a sequence $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\circ}$.

Fix $\delta>0$. Let $N_{\delta}^{(n)}$ be the minimum number of closed $\delta n^{1 / 3}$ balls needed to cover $\mathrm{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)$. By Lemma 6.13, $\left\{N_{\delta}^{(n)}\right\}_{n \geq 1}$ is tight. Write $\mathbb{V}_{n}$ for the set of vertices of $\mathrm{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)$ and $d_{\infty}$ for the tree distance in $\mathrm{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)$. Let $A_{1}, \ldots, A_{N_{\delta}^{(n)}}$ be a partition of $\mathbb{V}_{n}$ such that for $1 \leq j \leq N_{\delta}^{(n)}, d_{\infty}\left(v, v^{\prime}\right) \leq 2 \delta n^{1 / 3}$ if $v, v^{\prime} \in A_{j}$.

Let $v_{1}, v_{2}, \ldots$ be an enumeration of $\mathbb{V}_{n}$ such that for each $j \leq N_{\delta}^{(n)}$, all vertices $v \in A_{j}$ appear successively. Note that this enumeration is measurable w.r.t. the $\sigma$-field generated by $\operatorname{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)$. By Lemma 6.15,

$$
\begin{equation*}
\max _{1 \leq j \leq N_{\delta}^{(n)}}\left|\sum_{v \in A_{j}}\left(p_{v, \lambda_{n}}^{\text {modi }}-p_{v, \lambda_{n}}^{\text {avail }}\right)\right| \xrightarrow{\mathrm{P}} 0 . \tag{6.38}
\end{equation*}
$$

Let $\mu_{n}^{\text {modi }}$ be the measure on $\mathfrak{X}_{n}:=\left\{A_{1}, \ldots, A_{N_{\delta}^{(n)}}\right\}$ given by $\mu_{n}^{\text {modi }}\left(A_{j}\right)=$ $\sum_{v \in A_{j}} p_{v, \lambda_{n}}^{\text {modi }}$. Define $\mu_{n}^{\text {avail }}$ on $\mathfrak{X}_{n}$ analogously. Then the total variation distance between $\mu_{n}^{\text {modi }}$ and $\mu_{n}^{\text {avail }}$ satisfies

$$
d_{\mathrm{TV}}\left(\mu_{n}^{\text {modi }}, \mu_{n}^{\text {avail }}\right) \leq \frac{1}{2} \times N_{\delta}^{(n)} \times \max _{1 \leq j \leq N_{\delta}^{(n)}}\left|\sum_{v \in A_{j}}\left(p_{v, \lambda_{n}}^{\text {modi }}-p_{v, \lambda_{n}}^{\text {avail }}\right)\right| \xrightarrow{\mathrm{P}} 0,
$$

where the last step uses (6.38) and the fact that $\left\{N_{\delta}^{(n)}\right\}_{n \geq 1}$ is tight. Thus, for each $n$, we can construct $\mathfrak{X}_{n}$-valued random variables $X_{n}^{\text {modi }}$ and $X_{n}^{\text {avail }}$ distributed as $\mu_{n}^{\text {modi }}$ and $\mu_{n}^{\text {avail }}$ respectively such that $\mathbb{P}\left(X_{n}^{\text {modi }} \neq X_{n}^{\text {avail }}\right) \xrightarrow{\mathrm{P}} 0$. Using $X_{n}^{\text {modi }}$ and $X_{n}^{\text {avail }}$, there is a natural way to construct $\mathbb{V}_{n}$-valued random variables $Y_{n}^{\text {modi }}$ and $Y_{n}^{\text {avail }}$ such that $\mathbb{P}\left(Y_{n}^{\text {modi }}=v\right)=p_{v, \lambda_{n}}^{\text {modi }}$, and $\mathbb{P}\left(Y_{n}^{\text {avail }}=v\right)=p_{v, \lambda_{n}}^{\text {avail }}$ for all $v \in \mathbb{V}_{n}$, and further,

$$
\mathbb{P}\left(d_{\infty}\left(Y_{n}^{\text {modi }}, Y_{n}^{\text {avail }}\right)>2 \delta n^{1 / 3}\right) \xrightarrow{\mathrm{P}} 0 .
$$

Since $\delta>0$ was arbitrary, we get $n^{-1 / 3} d_{\mathrm{GHP}}\left(\mathfrak{M}_{n, \lambda_{n}}^{\text {modi }}, \mathfrak{M}_{n, \lambda_{n}}^{\text {avail }}\right) \xrightarrow{\mathrm{P}} 0$, which combined with (6.37) completes the proof.

The proof of Lemma 6.15 relies on the next two lemmas.
Lemma 6.16 There exist universal constants $c_{1}, c_{2}>0$ such that for any $m \geq 1$ and probability vector $\boldsymbol{p}:=\left(p_{1}, \ldots, p_{m}\right)$,

$$
\mathbb{P}\left(\max _{j \in[m]}\left|\sum_{i=1}^{j} p_{\pi(i)}-\frac{j}{m}\right| \geq x \sigma(\boldsymbol{p})\right) \leq \exp \left(-c_{1} x \log \log x\right), \text { for } x \geq c_{2},
$$

where $\pi$ is a uniform permutation on $[m]$, and $\sigma(\boldsymbol{p}):=\sqrt{p_{1}^{2}+\ldots+p_{m}^{2}}$. Consequently, using the relation $\sigma(\boldsymbol{p}) \leq \max _{j} \sqrt{p_{j}}$, we get, for $x \geq c_{2}$,

$$
\begin{equation*}
\mathbb{P}\left(\max _{j_{1}<j_{2}}\left|\sum_{i=j_{1}+1}^{j_{2}} p_{\pi(i)}-\frac{j_{2}-j_{1}}{m}\right| \geq 2 x \cdot \max _{j} \sqrt{p_{j}}\right) \leq 2 \exp \left(-c_{1} x \log \log x\right) \tag{6.39}
\end{equation*}
$$

This result gives a quantitative concentration inequality for the partial sums of exchangeable random variables. The result can be found in the above form in [26, Lemma 7.5], but was essentially already contained in [24, Lemma 4.9].

Lemma 6.17 (i) Fix $\lambda \in \mathbb{R}$. For every $v \in G_{1}(n, \lambda)$, append ( $d_{v, \lambda}^{\text {avail }}-r_{v, \lambda}$ ) many zeros to the sequence $\left(\left|T_{n, \lambda}^{(i)}(v)\right|, 1 \leq i \leq r_{v, \lambda}\right)$ and let $\left(\alpha_{n, \lambda}^{(i)}(v), 1 \leq i \leq d_{v, \lambda}^{\text {avail }}\right)$ be a uniform permutation of the resulting sequence; use independent permutations for different $v \in G_{1}(n, \lambda)$ that are also independent of all the other random variables being considered. Then conditional on Shape $\left[\operatorname{Rem}\left(C B D_{R_{n, \lambda}}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right)\right]$ and $C B D_{\infty}\left(G_{1}(n, \lambda)\right)$, the family

$$
\left(\alpha_{n, \lambda}^{(i)}(v) ; 1 \leq i \leq d_{v, \lambda}^{\text {avail }}, v \in G_{1}(n, \lambda)\right)
$$

of random variables is exchangeable.
(ii) Let $\lambda_{n}^{\star}$ be as in the proof of Theorem 6.1. Then for any $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\star}$,

$$
\begin{equation*}
\max \left\{\frac{\left|T_{n, \lambda_{n}}^{(i)}(v)\right|}{n}: 1 \leq i \leq r_{v, \lambda_{n}}, v \in C B D_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)\right\} \xrightarrow{P} 0 \tag{6.40}
\end{equation*}
$$

Proof of Lemma 6.15 By (6.25), $\left|G_{1}(n, \lambda)\right|=\Theta_{P}\left(n^{2 / 3}\right)$. So, in particular, for every $\lambda \in \mathbb{R}, \mathbb{P}\left(\left|G_{1}(n, \lambda)\right|>n / 2\right) \rightarrow 0$. Hence, we can choose $\lambda_{n}^{\diamond} \uparrow \infty$ slowly enough such that $\mathbb{P}\left(\left|G_{1}\left(n, \lambda_{n}^{\diamond}\right)\right|>n / 2\right) \rightarrow 0$ as $n \rightarrow \infty$. We can further take $\lambda_{n}^{\diamond} \leq \lambda_{n}^{\star}$.

Fix $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \lambda_{n}^{\diamond}$. Let $v_{1}, v_{2}, \ldots$ be an enumeration of the vertices of $G_{1}\left(n, \lambda_{n}\right)$ measurable w.r.t. the $\sigma$-field generated by $\operatorname{CBD}_{\infty}\left(G_{1}\left(n, \lambda_{n}\right)\right)$. Define

$$
p_{n, \lambda_{n}}^{(i)}\left(v_{s}\right):=\frac{\alpha_{n, \lambda_{n}}^{(i)}\left(v_{s}\right)}{n-\left|G_{1}\left(n, \lambda_{n}\right)\right|}, \quad 1 \leq i \leq d_{v, \lambda_{n}}^{\text {avail }}, 1 \leq s \leq\left|G_{1}\left(n, \lambda_{n}\right)\right| \text {, }
$$

where $0 / 0$ is interpreted as 1 . Since $\lambda_{n} \leq \lambda_{n}^{\diamond}, \mathbb{P}\left(n-\left|G_{1}\left(n, \lambda_{n}\right)\right| \geq n / 2\right) \rightarrow 1$ as $n \rightarrow \infty$ by our choice of $\lambda_{n}^{\diamond}$. Thus, using Lemma 6.17 (ii) and the fact that $\lambda_{n} \leq \lambda_{n}^{\diamond} \leq \lambda_{n}^{\star}$,

$$
\begin{equation*}
\max \left\{p_{n, \lambda_{n}}^{(i)}\left(v_{s}\right): 1 \leq i \leq d_{v, \lambda_{n}}^{\text {avail }}, 1 \leq s \leq\left|G_{1}\left(n, \lambda_{n}\right)\right|\right\} \xrightarrow{\mathrm{P}} 0 . \tag{6.41}
\end{equation*}
$$

Now, for any $1 \leq j_{1} \leq j_{2} \leq\left|G_{1}\left(n, \lambda_{n}\right)\right|$,

$$
\sum_{s=j_{1}}^{j_{2}} \sum_{i=1}^{r_{v s}, \lambda_{n}}\left|T_{n, \lambda_{n}}^{(i)}\left(v_{s}\right)\right|=\sum_{s=j_{1}}^{j_{2}} \sum_{i=1}^{d_{v, ~}^{\text {avail }}} \alpha_{n, \lambda_{n}}^{(i)}\left(v_{s}\right)
$$

and in particular, on the event $B_{n}$,

$$
\sum_{s=1}^{\left|G_{1}\left(n, \lambda_{n}\right)\right|} \sum_{i=1}^{d_{v s, \lambda_{n}}^{\text {avail }}} \alpha_{n, \lambda_{n}}^{(i)}\left(v_{s}\right)=n-\left|G_{1}\left(n, \lambda_{n}\right)\right|
$$

Thus, on the event $B_{n} \cap\left\{\left|G_{1}\left(n, \lambda_{n}\right)\right|<n\right\},\left(p_{n, \lambda_{n}}^{(i)}\left(v_{s}\right) ; 1 \leq i \leq d_{v, \lambda_{n}}^{\text {avail }}, 1 \leq s \leq\right.$ $\left.\left|G_{1}\left(n, \lambda_{n}\right)\right|\right)$ is a probability vector. By (2.2), $\mathbb{P}\left(B_{n} \cap\left\{\left|G_{1}\left(n, \lambda_{n}\right)\right|<n\right\}\right) \rightarrow 1$ as $n \rightarrow \infty$. Thus, the desired result follows from Lemma 6.17 (i), (6.39), and (6.41).
Proof of Lemma 6.17(i) Consider a finite (non-random) graph $H$ and $t>0$, and let $H^{\text {exp }}$ and $R(t)$ be as in the statement of Lemma 6.7. Then conditional on Shape $\left[\operatorname{Rem}\left(\mathrm{CBD}_{R(t)}\left(H^{\exp }\right)\right)\right]=H_{0}$, the order in which the edges in $E(H) \backslash E\left(H_{0}\right)$ were sampled for the first time in the CBD process is a uniform permutation on $E(H) \backslash E\left(H_{0}\right)$.

Using the above observation, Lemma 6.7, and Lemma 6.8, we can generate Shape $\left[\operatorname{Rem}\left(\operatorname{CBD}_{R_{n, \lambda}}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right)\right], \mathscr{G}_{n, 3}$, and $\left(\alpha_{n, \lambda}^{(i)}(v) ; 1 \leq i \leq d_{v, \lambda}^{\text {avail }}, v \in G_{1}(n, \lambda)\right)$ jointly as follows:
(a) Sample a $\operatorname{Binomial}\left(3 n / 2,1 / 2-\lambda n^{-1 / 3}\right)$ random variable. For simplicity, we denote the realization by $m$.
(b) Conditional on step (a), sample $\mathscr{Q}_{n, \mathbf{3}, m}^{(1)}$ as in Lemma 6.8, where $\mathbf{3}=(3, \ldots, 3)$. By (6.19), $\mathscr{Q}_{n, 3, m}^{(1)} \stackrel{\mathrm{d}}{=} \operatorname{Perc}\left(\mathscr{G}_{n, 3}, 1 / 2+\lambda n^{-1 / 3}\right)$. Hence, using Lemma 6.7,

$$
\begin{equation*}
\mathscr{Q}_{n, \mathbf{3}, m}^{(1)} \stackrel{\mathrm{d}}{=} \operatorname{Shape}\left[\operatorname{Rem}\left(\operatorname{CBD}_{R_{n, \lambda}}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right)\right] \tag{6.42}
\end{equation*}
$$

Thus, the largest component of $\mathscr{Q}_{n, 3, m}^{(1)}$, say $\mathscr{C}_{1}$, has the same law as $G_{1}(n, \lambda)$. Let $d_{v}$ be the degree of $v \in \mathscr{C}_{1}$. Then each $v \in \mathscr{C}_{1}$ has $3-d_{v}$ many 'available' half-edges; we denote them by $f_{v, i}, 1 \leq i \leq 3-d_{v}, v \in \mathscr{C}_{1}$.
(c) Conditional on steps (a) and (b), generate $\mathscr{Q}_{n, \mathbf{3}, m}^{(2)}$ as in Lemma 6.8. From (6.18) it follows that

$$
\begin{equation*}
\mathscr{Q}_{n, \mathbf{3}, m}^{(2)} \stackrel{\mathrm{d}}{=} \mathscr{G}_{n, 3} \tag{6.43}
\end{equation*}
$$

jointly with the equality in distribution in (6.42). Let $\mathscr{E}_{1}, \ldots, \mathscr{E}_{m}$ be the edges that are in $\mathscr{Q}_{n, \mathbf{3}, m}^{(2)}$ but not in $\mathscr{Q}_{n, \mathbf{3}, m}^{(1)}$. For $v \in \mathscr{C}_{1}$ and $1 \leq i \leq 3-d_{v}$, let $\widetilde{f}_{v, i}$ denote the edge that was formed by pairing $f_{v, i}$ with another half-edge.
(d) If $\mathscr{Q}_{n, \mathbf{3}, m}^{(2)}$ is not connected, go to the next step. If $\mathscr{Q}_{n, \mathbf{3}, m}^{(2)}$ is connected, let $\pi$ be a uniform permutation of $m$ elements independent of steps (a), (b), and (c) above. Consider the edges $\mathscr{E}_{\pi(1)}, \ldots, \mathscr{E}_{\pi(m)}$ sequentially in this order, and at each step, remove the edge being considered from $\mathscr{Q}_{n, 3, m}^{(2)}$ if its removal does not disconnect the current graph. Denote the resulting graph by $\mathscr{Q}$. Then $\mathscr{Q}$ has the same law as Shape $\left[\operatorname{CBD}_{R_{n, \lambda}}\left(\mathscr{G}_{n, 3}^{\exp }\right)\right]$.
(e) If $\mathscr{Q}_{n, \mathbf{3}, m}^{(2)}$ is not connected, define $Q^{(i)}(v)$ to be the empty graph for $1 \leq i \leq 3-d_{v}$, $v \in \mathscr{C}_{1}$. If $\mathscr{Q}_{n, 3, m}^{(2)}$ is connected, then note that $\mathscr{Q}$ as constructed in (d) is simply $\mathscr{C}_{1}$ together with some connected multigraphs each of which is connected to a vertex of $\mathscr{C}_{1}$ by a single edge; for $v \in \mathscr{C}_{1}$ and $1 \leq i \leq 3-d_{v}$, set $Q^{(i)}(v)$ to be the connected multigraph that is connected to $v$ via $\widetilde{f}_{v, i}$, with the convention that $Q^{(i)}(v)$ is the empty graph if $\widetilde{f}_{v, i}$ was removed in step (d). Then

$$
\left(\operatorname{sort}\left(\left|Q^{(i)}(v)\right|, 1 \leq i \leq 3-d_{v}\right) ; v \in \mathscr{C}_{1}\right) \stackrel{\mathrm{d}}{=}\left(\operatorname{sort}\left(\alpha_{n, \lambda}^{(i)}(v), 1 \leq i \leq d_{v, \lambda}^{\text {avail }}\right) ; v \in G_{1}(n, \lambda)\right)
$$

jointly with the distributional equalities in (6.42) and (6.43), where sort(•) arranges the entries of a finite sequence in decreasing order.
Conditional on steps (a) and (b) above, the rest of the procedure is symmetric with respect to the available half-edges attached to the vertices of $\mathscr{C}_{1}$. Hence, conditional on $\mathscr{Q}_{n, \mathbf{3}, m}^{(1)}$, the family $\left(\left|Q^{(i)}(v)\right| ; 1 \leq i \leq 3-d_{v}, v \in \mathscr{C}_{1}\right)$ is exchangeable. Now, the conditional distribution of $\left(\left|Q^{(i)}(v)\right| ; 1 \leq i \leq 3-d_{v}, v \in \mathscr{C}_{1}\right)$ given $\mathscr{Q}_{n, \mathbf{3}, m}^{(1)}$ is equal to the conditional distribution of the same sequence given $\mathscr{Q}_{n, 3, m}^{(1)}$ and $\mathrm{CBD}_{\infty}\left(\mathscr{C}_{1}\right)$. Thus, $\left(\left|Q^{(i)}(v)\right| ; 1 \leq i \leq 3-d_{v}, v \in \mathscr{C}_{1}\right)$ is an exchangeable family given $\mathscr{Q}_{n, 3, m}^{(1)}$ and $\mathrm{CBD}_{\infty}\left(\mathscr{C}_{1}\right)$. Thus the claim follows.

We need the following result before proceeding to the proof of Lemma 6.17((ii)). Recall the notation $\mathfrak{m}(\cdot ; \cdot)$ from Sect. 3.1.

Lemma 6.18 Suppose $\left(Z_{n}, d_{n}, \mu_{n}\right) \rightarrow(Z, d, \mu)$ as $n \rightarrow \infty$ in $\mathfrak{S}_{\mathrm{GHP}}$. If $\mu$ is nonatomic, then

$$
\lim _{\varepsilon \downarrow 0} \limsup _{n \rightarrow \infty} \mathfrak{m}\left(\varepsilon, Z_{n}\right)=0 .
$$

Proof Using the convergence $\left(Z_{n}, d_{n}, \mu_{n}\right) \rightarrow(Z, d, \mu)$, it is easy to see that for every $\varepsilon>0$ and sufficiently large $n$,

$$
\mathfrak{m}\left(\varepsilon, Z_{n}\right) \leq \mathfrak{m}(2 \varepsilon, Z)+\varepsilon
$$

It follows easily from the compactness of $(Z, d)$ and the non-atomicity of $\mu$ that

$$
\lim _{\varepsilon \downarrow 0} \mathfrak{m}(2 \varepsilon, Z)=0
$$

Thus the claim follows.

We now continue with
Proof of Lemma 6.17((ii)) Recall the notation from Construction 3.9. Let $\Gamma_{r / 2}$ and $\mathscr{K}_{n, 3}^{\exp }$ be as in the proof of Proposition 6.3. As observed in the proof of Proposition 6.3, there is a natural isometric embedding of $\mathscr{K}_{n, 3}^{\exp }$ into $\sqrt{2 \Gamma_{r / 2}} \cdot \mathscr{H}^{(s)}$. In this embedding, $\sqrt{2 \Gamma_{r / 2}} \cdot \mathscr{H}^{(s)}$ can be obtained by attaching countably many real trees to $\mathscr{K}_{n, 3}^{\exp }$. Let $\left(\mathscr{K}_{n, 3}^{\exp }, \mu_{n}\right)$ be the measured $\mathbb{R}$-graph derived by endowing $\mathscr{K}_{n, 3}^{\exp }$ by the measure obtained by projecting the measure from $\sqrt{2 \Gamma_{r / 2}} \cdot \mathscr{H}^{(s)}$ onto the attachment points in $\mathscr{K}_{n, 3}^{\text {exp }}$. Thus, the $\mu_{n}$ measure of the $i$-th edge of $\mathscr{K}_{n, 3}^{\text {exp }}$ is $X_{i}, 1 \leq i \leq r$, where $\left(X_{1}, \ldots, X_{r}\right) \sim \operatorname{Dirichlet}(1 / 2, \ldots, 1 / 2)$ as in Construction 3.9.

Arguing as in (6.14), it is easy to show that as $n \rightarrow \infty$,

$$
n^{-1 / 3} d_{\mathrm{GHP}}\left(\mathrm{CB}^{\infty}\left(\sqrt{2 \Gamma_{r / 2}} \cdot \mathscr{H}^{(s)}\right), \mathrm{CB}^{\infty}\left(\mathscr{K}_{n, 3}^{\exp }\right)\right) \xrightarrow{\mathrm{P}} 0 .
$$

Combining this with Theorem 4.8 and (6.12), we get

$$
\begin{equation*}
n^{-1 / 3} \mathrm{CB}^{\infty}\left(\mathscr{K}_{n, 3}^{\exp }\right) \xrightarrow{\mathrm{d}}(0.75)^{1 / 3} \cdot \mathscr{M} \text { w.r.t. GHP topology. } \tag{6.44}
\end{equation*}
$$

Suppose $\mathscr{G}_{n, 3}$ and $\mathscr{K}_{n, 3}$ are coupled as in (6.10). On the event $\left\{\mathscr{K}_{n, 3}=\mathscr{G}_{n, 3}\right\}$, the tree $\mathrm{CB}^{\infty}\left(\mathscr{K}_{n, 3}^{\text {exp }}\right)$ can be obtained by
(i) attaching each of the trees $\mathfrak{T}_{n, \lambda_{n}}^{(j)}, 1 \leq j \leq k_{n}\left(\lambda_{n}\right)$, to $\operatorname{CBD}_{\infty}\left(\mathfrak{G}_{1}\left(n, \lambda_{n}\right)\right)$ via a single edge, and then
(ii) attaching some additional line segments to the space thus obtained.
(Recall that in the CB process edges are cut open, while in the CBD process edges are removed. Because of this difference these additional line segments need to be attached.) Thus, using (6.44), Theorem 2.3, (6.5), together with Lemma 6.18, we see that

$$
\begin{equation*}
\max _{1 \leq j \leq k_{n}\left(\lambda_{n}\right)} \mathbb{1}_{\left\{\mathscr{Y}_{n, 3}=\mathscr{K}_{n, 3}\right\}} \times \mu_{n}\left(\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right) \xrightarrow{\mathrm{P}} 0 . \tag{6.45}
\end{equation*}
$$

Consider the tree among $\mathfrak{T}_{n, \lambda_{n}}^{(j)}, 1 \leq j \leq k_{n}\left(\lambda_{n}\right)$, that has the maximum number of edges (pick any one if there is more that one such tree), and let $f_{1}, \ldots, f_{\mathscr{E}_{\text {max }}}$ be an enumeration of its edges. On the event $\left\{\mathscr{G}_{n, 3}=\mathscr{K}_{n, 3}\right\}$,

$$
\begin{equation*}
\max _{1 \leq j \leq k_{n}\left(\lambda_{n}\right)} \mu_{n}\left(\mathfrak{T}_{n, \lambda_{n}}^{(j)}\right) \geq \sum_{j=1}^{\mathscr{E}_{\text {max }}} \mu_{n}\left(f_{j}\right) \geq \sum_{i=1}^{\mathscr{E}_{\text {max }}} X_{(i)} \tag{6.46}
\end{equation*}
$$

where $X_{(1)}<\ldots<X_{(r)}$ are the order statistics corresponding to $\left(X_{1}, \ldots, X_{r}\right)$. Now for any $\varepsilon \in(0,1)$,

$$
\begin{equation*}
\sum_{i=1}^{\varepsilon r} X_{(i)} \xrightarrow{\mathrm{P}} 2 \cdot \mathbb{E}\left[\Gamma_{1 / 2} \mathbb{1}_{\left\{\Gamma_{1 / 2} \leq Q_{\varepsilon}\right\}}\right], \tag{6.47}
\end{equation*}
$$

where $\Gamma_{1 / 2} \sim \operatorname{Gamma}(1 / 2,1)$ and $\mathbb{P}\left(\Gamma_{1 / 2} \leq Q_{\varepsilon}\right)=\varepsilon$. It now follows from (6.45), (6.46) and (6.47) that

$$
\mathscr{E}_{\max } / r \xrightarrow{\mathrm{P}} 0,
$$

which in turn implies (6.40). This completes the proof.

### 6.7 GHP convergence of the MST of $\overline{\mathscr{G}}_{n, 3}$

In this section we prove Theorem 2.2. We first state two fundamental results about the configuration model and uniform simple graphs with prescribed degree.
(a) From (1.3) (see also [29,70]), it follows that conditional on being simple, the configuration model has the same distribution as $\overline{\mathscr{G}}_{n, \mathbf{d}}$, i.e.,

$$
\begin{equation*}
\mathbb{P}\left(\overline{\mathscr{G}}_{n, \mathbf{d}} \in \cdot\right)=\mathbb{P}\left(\mathscr{G}_{n, \mathbf{d}} \in \cdot \mid \mathscr{G}_{n, \mathbf{d}} \text { is simple }\right) . \tag{6.48}
\end{equation*}
$$

(b) By [55, Theorem 1.1], if $\sum_{v \in[n]} d_{v} \rightarrow \infty$ and $\sum_{v \in[n]} d_{v}^{2}=O\left(\sum_{v \in[n]} d_{v}\right)$, then

$$
\begin{equation*}
\liminf _{n \rightarrow \infty} \mathbb{P}\left(\mathscr{G}_{n, \mathbf{d}} \text { is simple }\right)>0 \tag{6.49}
\end{equation*}
$$

Let $\mathfrak{M}_{n, \lambda}^{\text {avail }}$ and $\mathfrak{M}_{n, \lambda}^{\text {attach }}$ be as defined around (6.36). Define the spaces $\overline{\mathfrak{M}}_{n, \lambda}^{\text {avail }}$ and $\overline{\mathfrak{M}}_{n, \lambda}^{\text {attach }}$ analogously for $\overline{\mathscr{G}}_{n, 3}$. Using (6.48) and (6.49), it follows that the analogues of Lemma 6.12 and Lemma 6.14 hold for $\overline{\mathscr{G}}_{n, 3}$ : Fix $\delta>0$ and consider a $\mathbb{Z}_{>0}$-valued sequence $\lambda_{n} \uparrow \infty$ with $\lambda_{n} \leq \min \left\{\lambda_{n}^{\star}, \lambda_{n}^{\circ}\right\}$. Then

$$
\begin{aligned}
& \mathbb{P}\left(d_{\mathrm{GHP}}\left(\mathrm{CBD}_{\infty}\left(\overline{\mathscr{G}}_{n, 3}\right), \overline{\mathfrak{M}}_{n, \lambda_{n}}^{\text {attach }}\right)>\delta n^{1 / 3}\right) \\
& \quad=\mathbb{P}\left(d_{\mathrm{GHP}}\left(\operatorname{CBD}_{\infty}\left(\mathscr{G}_{n, 3}\right), \mathfrak{M}_{n, \lambda_{n}}^{\text {attach }}\right)>\delta n^{1 / 3} \mid \mathscr{G}_{n, 3} \text { is simple }\right) \rightarrow 0,
\end{aligned}
$$

as $n \rightarrow \infty$. Similarly

$$
n^{-1 / 3} d_{\mathrm{GHP}}\left(\overline{\mathfrak{M}}_{n, \lambda_{n}}^{\text {attach }}, \overline{\mathfrak{M}}_{n, \lambda_{n}}^{\text {avail }}\right) \xrightarrow{\mathrm{P}} 0 .
$$

To complete the proof, we have to show that the analogue of Lemma 6.13 remains true for $\overline{\mathfrak{M}}_{n, \lambda_{n}}^{\text {avail }}$. Thus, it suffices to prove that for each fixed $\lambda \in \mathbb{R}$,

$$
n^{-1 / 3} \cdot \overline{\mathfrak{M}}_{n, \lambda}^{\text {avail }} \xrightarrow{\mathrm{d}} 6^{1 / 3} \cdot \mathrm{CB}^{\infty}\left(S_{1}\left((48)^{1 / 3} \cdot \lambda\right)\right), \quad \text { as } n \rightarrow \infty
$$

w.r.t. GHP topology. Let $f: \mathfrak{S}_{\mathrm{GHP}} \rightarrow \mathbb{R}$ be bounded continuous. Then it suffices to show that as $n \rightarrow \infty$,

$$
\mathbb{E}\left[f\left(n^{-1 / 3} \mathfrak{M}_{n, \lambda}^{\text {avail }} \mid \mathscr{G}_{n, 3} \text { is simple }\right)\right]-\mathbb{E}\left[f\left(n^{-1 / 3} \mathfrak{M}_{n, \lambda}^{\text {avail }}\right)\right] \rightarrow 0
$$

or equivalently

$$
\mathbb{E}\left[f\left(n^{-1 / 3} \mathfrak{M}_{n, \lambda}^{\text {avail }} \cdot \mathbb{1}\left\{\mathscr{G}_{n, 3} \text { is simple }\right\}\right)\right]-\mathbb{E}\left[f\left(n^{-1 / 3} \mathfrak{M}_{n, \lambda}^{\text {avail }}\right)\right] \times \mathbb{P}\left(\mathscr{G}_{n, 3} \text { is simple }\right) \rightarrow 0
$$

This can be proved by using techniques similar to the ones used in the proof of [40, Theorem 3]; see the argument given in [40, Section 7]. We omit the details.

### 6.8 Proof of Theorem 2.3

Let $M_{\lambda}^{n, \text { er }}$ be as defined at the beginning of Sect. 4.3. Using Observation 4.2, $M_{\lambda}^{n, \text { er }}$ is a subtree of $M_{\infty}^{n, \text { er }}$. Consider the forest obtained from $M_{\infty}^{n, \text { er }}$ by deleting all edges in $M_{\lambda}^{n, \text { er }}$, and for every $v \in V\left(M_{\lambda}^{n, \text { er }}\right)$, let $T_{v, \lambda}^{n, \text { er }}$ be the tree in this forest that contains $v$. Define $p_{v, \lambda}^{n, \text { er }}:=\left|T_{v, \lambda}^{n, \text { er }}\right| / n$. We now state two lemmas that will be needed in the proof.

Lemma 6.19 For every $\lambda \in \mathbb{R}$, conditional on $\operatorname{ER}(n, \lambda)$, the family $\left(p_{v, \lambda}^{n, \text { er }} ; v \in\right.$ $\left.V\left(M_{\lambda}^{n, \text { er }}\right)\right)$ of random variables is exchangeable.

Remark 5 At the beginning of [5, Section 4.4], it is stated that $\left(\left|T_{v, \lambda}^{n, \text { er }}\right|, v \in V\left(M_{\lambda}^{n, \text { er }}\right)\right)$ is exchangeable conditional on $V\left(M_{\lambda}^{n, \text { er }}\right)$. (Here, we have translated the claim in [5] using our notation.) However, in [5, Page 3114], the vertices of $M_{\lambda}^{n \text {,er }}$ are relabeled so that the vertices in each element of a given $\operatorname{cover}\left(B_{\lambda}^{n, i}, 1 \leq i \leq N_{\varepsilon}^{n}\right)$ of $M_{\lambda}^{n, \text { er }}$ by small-diameter sets appear successively. This labeling contains some information about the relative positioning of the vertices in the tree, because if two vertices are close in this arrangement of the vertices, then their tree distance is likely to be small as well. In other words, the symmetry between the roles of a pair of vertices appearing consecutively and a pair of vertices that are far from each other in this arrangement does not follow directly. Exchangeability of $\left(\left|T_{v, \lambda}^{n, \text { er }}\right|, v \in V\left(M_{\lambda}^{n, \text { er }}\right)\right)$ is needed conditional on this labeling. Thus, the result implicitly being used in the proof of [5, Proposition 4.8] is the following: $\left(\left|T_{v, \lambda}^{n, \text { er }}\right|, v \in V\left(M_{\lambda}^{n, \text { er }}\right)\right)$ is exchangeable conditional on $M_{\lambda}^{n, \text { er }}$. This stronger form of exchangeability follows from Lemma 6.19.

Proof of Lemma 6.19 We outline the proof briefly. Using Lemma 4.1, conditional on the graph $\operatorname{ER}(n, \lambda), M_{\infty}^{n, \text { er }}$ can be generated as follows:
(i) Let $\mathscr{E}_{\text {out }}$ denote the set of edges of the complete graph $K_{n}$ whose endpoints are in two different components of $\operatorname{ER}(n, \lambda)$. Let $\mathscr{E}_{\text {in }}$ denote the set of edges of $\operatorname{ER}(n, \lambda)$. Construct the graph $\operatorname{ER}(n, \lambda) \cup \mathscr{E}_{\text {out }}$. Assign i.i.d. Uniform $\left[n^{-1}+\right.$ $\left.\lambda n^{-4 / 3}, 1\right]$ weights to the edges in $\mathscr{E}_{\text {out }}$, and independently of this, assign i.i.d. Uniform $\left[0, n^{-1}+\lambda n^{-4 / 3}\right]$ weights to the edges in $\mathscr{E}_{\text {in }}$. Denote the weight assigned to an edge $e$ by $w_{e}$.
(ii) From the graph $\operatorname{ER}(n, \lambda) \cup \mathscr{E}_{\text {out }}$, delete all edges $e \in \mathscr{E}_{\text {out }}$ that are part of a cycle $\pi$ (with no repeated edges) in $\operatorname{ER}(n, \lambda) \cup \mathscr{E}_{\text {out }}$ and $w_{e}$ is the maximum among all edge weights in $\pi$.
(iii) For each $i \geq 1$, delete all edges $e \in E\left(\mathscr{C}_{i}^{n, \text { er }}(\lambda)\right)$ that are part of a cycle $\pi$ (with no repeated edges) in $\mathscr{C}_{i}^{n, \text { er }}(\lambda)$ and $w_{e}$ is the maximum among all edge weights in $\pi$.

The marginal distribution of the resulting tree will be the same as that of $M_{\infty}^{n, \text { er }}$. Consider two distinct vertices $v_{1}, v_{2} \in V\left(\mathscr{C}_{1}^{n, e r}(\lambda)\right)$. If we interchange the values $w_{\left\{v_{1}, u\right\}}$ and $w_{\left\{v_{2}, u\right\}}$ for every vertex $u \notin V\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)$, then it is easy to check that in the above procedure, the set of edges removed in step (iii) remains the same, and the set of edges in $\mathscr{E}_{\text {out }}$ that are not incident to $v_{1}$ or $v_{2}$ and are removed in step (ii) remains the same. Further, if $\left\{v_{1}, u\right\} \in \mathscr{E}_{\text {out }}$ was removed in step (ii) before the interchange of edge weights, then the edge $\left\{v_{2}, u\right\}$ will be removed in step (ii) after the interchange and vice versa. Consequently, the values of $p_{v_{1}, \lambda}^{n, \text { er }}$ and $p_{v_{2}, \lambda}^{n, \text { er }}$ would be swapped as a result of the interchange of edge weights. This shows that conditional on $\operatorname{ER}(n, \lambda)$, the law of $\left(p_{v, \lambda}^{n, \text { er }}, v \in V\left(M_{\lambda}^{n, \text { er }}\right)\right)$ is invariant under transpositions. We can repeat the same argument with any permutation of $V\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)$ to get the claimed exchangeability.

Lemma 6.20 (Lemma 4.11 of [5]) Let $\Delta_{n, \lambda}:=\max _{v \in V\left(M_{\lambda}^{n, \text { er }}\right)} p_{v, \lambda}^{n, \text { er }}$. Then for every $\delta>0$,

$$
\limsup _{\lambda \rightarrow \infty} \limsup _{n \rightarrow \infty} \mathbb{P}\left(\Delta_{n, \lambda}>\delta\right)=0
$$

We are now ready to prove Theorem 2.3. Observe the following facts:
(a) Fix $s \geq 2$ and let $r=3(s-1)$. Let $e_{1}, \ldots, e_{r}$ be an enumeration of $e\left(\mathscr{H}^{(s)}\right)$. Then

$$
\left(\operatorname{len}\left(e_{i}\right), 1 \leq i \leq r\right) \stackrel{\mathrm{d}}{=}\left(Y_{i} \cdot\left(\frac{\Gamma_{1 / 2}^{(i)}}{\sum_{j=1}^{r} \Gamma_{1 / 2}^{(j)}}\right)^{1 / 2}, 1 \leq i \leq r\right),
$$

where $Y_{i}, 1 \leq i \leq r$, are i.i.d. Rayleigh random variables independent of $\Gamma_{1 / 2}^{(i)}$, $1 \leq i \leq r$, which are i.i.d. $\operatorname{Gamma}(1 / 2,1)$ random variables. As observed in (6.13), $\sqrt{2} \cdot Y_{i} \sqrt{\Gamma_{1 / 2}^{(i)}}, 1 \leq i \leq r$, are i.i.d. Exponential(1) random variables. Thus, for all $\delta>0$,

$$
\lim _{s \rightarrow \infty} \mathbb{P}\left(\min _{e \in e\left(\mathscr{H}^{(s)}\right)} \operatorname{len}(e) \geq s^{-\frac{3}{2}-\delta}\right)=1
$$

(b) By (3.8), for any $s \geq 2$,

$$
\frac{1}{\sqrt{m}} \min _{e \in e\left(\mathscr{H}_{m, s}\right)} \operatorname{len}(e) \xrightarrow{\mathrm{d}} \min _{e \in e\left(\mathscr{H}^{(s)}\right)} \operatorname{len}(e) \text { as } m \rightarrow \infty .
$$

(c) $\mathscr{C}_{1}^{n, \text { er }}(\lambda)$ can be generated as follows:
(i) Sample $\left|\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right|$ and $\operatorname{sp}\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)$. Denote the realizations by $m$ and $s$ respectively.
(ii) Conditional on the previous step, generate $\mathscr{H}_{m, s}$ and set this graph to be $\mathscr{C}_{1}^{n, \text { er }}(\lambda)$.
(d) By Lemma 3.3, $\mathbb{P}\left(\xi_{1}(\lambda) \leq \lambda\right)+\mathbb{P}\left(N_{1}(\lambda)<\lambda^{3} / 2\right) \rightarrow 0$ as $\lambda \rightarrow \infty$.
(e) By Theorem 3.2, $\left(n^{-2 / 3}\left|\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right|, \operatorname{sp}\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)\right) \xrightarrow{\text { d }}\left(\xi_{1}(\lambda), N_{1}(\lambda)\right)$ as $n \rightarrow$ $\infty$.

Combining the above, we see that

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \mathbb{P}\left(\min _{e \in e\left(\mathscr{C}_{1}^{n, \mathrm{er}}(\lambda)\right)} \operatorname{len}(e) \leq n^{1 / 3} / \lambda^{5}\right)=: \varepsilon_{1}(\lambda) \rightarrow 0, \text { as } \lambda \rightarrow \infty \tag{6.50}
\end{equation*}
$$

Using the convergences $\operatorname{sp}\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right) \xrightarrow{\mathrm{d}} N_{1}(\lambda)$ as $n \rightarrow \infty$ and $N_{1}(\lambda) / \lambda^{3} \xrightarrow{\mathrm{P}}$ $2 / 3$ as $\lambda \rightarrow \infty$ together with (3.6), we see that

$$
\begin{align*}
& \lim _{n \rightarrow \infty} \mathbb{P}\left(k\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right) \text { is not a 3-regular multigraph }\right) \\
& \quad=\mathbb{P}\left(N_{1}(\lambda) \leq 1\right)=: \varepsilon_{2}(\lambda) \rightarrow 0, \text { as } \lambda \rightarrow \infty \tag{6.51}
\end{align*}
$$

Let $e_{i}^{n, \text { er }}(\lambda), 1 \leq i \leq\left|e\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)\right|$, be an enumeration of $e\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)$. Let $V_{i}^{n, \text { er }}(\lambda)$ be the set of vertices in $\mathscr{C}_{1}^{n, \text { er }}(\lambda)$ that are connected to Core $\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)$ via $e_{i}^{n, \text { er }}(\lambda)$. (As before, the common endpoints of multiple $e \in e\left(\mathscr{C}_{1}^{n, \mathrm{er}}(\lambda)\right)$ and their pendant subtrees are assigned to only of the $V_{i}^{n, e r}(\lambda)$ 's in an arbitrary way.) From (3.11) and arguments as above,

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \mathbb{P}\left(\max _{i} \frac{\left|V_{i}^{n, \text { er }}(\lambda)\right|}{\left|\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right|} \geq(\log \lambda)^{2} / \lambda^{3}\right)=: \varepsilon_{3}(\lambda) \rightarrow 0, \text { as } \lambda \rightarrow \infty . \tag{6.52}
\end{equation*}
$$

Denote the complements of the events in (6.50), (6.51), and (6.52) by $E_{n, \lambda}^{(1)}, E_{n, \lambda}^{(2)}$, and $E_{n, \lambda}^{(3)}$ respectively. Let $E_{n, \lambda}:=\cap_{j=1}^{3} E_{n, \lambda}^{(j)}$. Note that on the event $E_{n, \lambda}$, any $\mathscr{U} \subseteq$ $V\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)$ with $\operatorname{diam}\left(\mathscr{U} ; \mathscr{C}_{1}^{n, \text { er }}(\lambda)\right) \leq n^{1 / 3} /\left(2 \lambda^{5}\right)$ can intersect $V_{i}^{n, \text { er }}(\lambda)$ for at most three values of $i$.

Let $\mathscr{Z}_{\lambda}^{n, \text { er }}$ be the graph obtained by attaching, for each $v \in \mathscr{C}_{1}^{n, \text { er }}(\lambda)$, the tree $T_{v, \lambda}^{n, \text { er }}$ to $\mathscr{\mathscr { C }}_{1}^{n, \text { er }}(\lambda)$ via identification of the vertices labeled $v$. Consider $\tilde{\mathscr{U}} \subseteq[n]$ with $\operatorname{diam}\left(\tilde{\mathscr{U}} ; \mathscr{Z}_{\lambda}^{n, \text { er }}\right) \leq n^{1 / 3} /\left(2 \lambda^{5}\right)$. Let

$$
\mathscr{U}=\left\{v \in \mathscr{C}_{1}^{n, \text { er }}(\lambda): T_{v, \lambda}^{n, \text { er }} \cap \tilde{\mathscr{U}} \neq \emptyset\right\} .
$$

Then $\operatorname{diam}\left(\mathscr{U} ; \mathscr{C}_{1}^{n, \mathrm{er}}(\lambda)\right) \leq n^{1 / 3} /\left(2 \lambda^{5}\right)$. Consequently, on the event $E_{n, \lambda}$,

$$
\begin{align*}
\frac{1}{n} \cdot|\tilde{\mathscr{U}}| & \leq \sum_{v \in \mathscr{U}} \frac{\left|T_{v, \lambda}^{n, \mathrm{er}}\right|}{n} \leq 3 \times \max _{i}\left(\sum_{v \in V_{i}^{n, \mathrm{er}}(\lambda)} p_{v, \lambda}^{n, \mathrm{er}}\right) \\
& \left.\leq\left. 3 \cdot \max _{i}\right|_{v \in V_{i}^{n, \mathrm{er}}(\lambda)} p_{v, \lambda}^{n, \text { er }}-\frac{\left|V_{i}^{n, \mathrm{er}}(\lambda)\right|}{\mid \mathscr{C}_{1}^{n, \mathrm{er}}(\lambda)} \right\rvert\,+3(\log \lambda)^{2} \frac{1}{\lambda^{3}} . \tag{6.53}
\end{align*}
$$

Arrange the vertices in $\mathscr{C}_{1}^{n, \text { er }}(\lambda)$ in a sequence so that for each $i$, the vertices in $V_{i}^{n, \text { er }}(\lambda)$ appear consecutively. This arrangement is measurable w.r.t. the sigma field generated by $\operatorname{ER}(n, \lambda)$. By Lemma 6.19, conditional on this arrangement, $\left\{p_{v, \lambda}^{n, \text { er }}\right\}_{v \in V\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda)\right)}$ is an exchangeable sequence. Using (6.39) with $x=\Delta_{n, \lambda}^{-1 / 4}$, we see that

$$
\begin{align*}
& \varepsilon_{4}^{(n)}(\lambda):=\mathbb{P}\left(\max _{i}\left|\sum_{v \in V_{i}^{n, \mathrm{er}}(\lambda)} p_{v, \lambda}^{n, \text { er }}-\frac{\left|V_{i}^{n, \mathrm{er}}(\lambda)\right|}{\mid \mathscr{C}_{1}^{n, \mathrm{er}}(\lambda)}\right| \geq 2 \Delta_{n, \lambda}^{1 / 4}\right) \\
& \quad \leq \mathbb{P}\left(\Delta_{n, \lambda}^{-1 / 4} \leq c_{2}\right)+2 \cdot \mathbb{E}\left[\exp \left(-c \Delta_{n, \lambda}^{-1 / 4} \log \log \Delta_{n, \lambda}^{-1 / 4}\right)\right] \tag{6.54}
\end{align*}
$$

where $c_{2}$ is as in Lemma 6.16. Combining (6.53) and (6.54), we see that

$$
\mathbb{P}\left(\mathfrak{m}\left(\left(2 \lambda^{5}\right)^{-1} ; n^{-1 / 3} \mathscr{Z}_{\lambda}^{n, \mathrm{er}}\right) \geq 6 \Delta_{n, \lambda}^{1 / 4}+3 \cdot(\log \lambda)^{2} \cdot \lambda^{-3}\right) \leq \mathbb{P}\left(E_{n, \lambda}^{c}\right)+\varepsilon_{4}^{(n)}(\lambda)
$$

Since $M_{\infty}^{n, \text { er }}$ is a subtree of $\mathscr{Z}_{\lambda}^{n, \text { er }}, \mathfrak{m}\left(\delta ; M_{\infty}^{n, \text { er }}\right) \leq \mathfrak{m}\left(\delta ; \mathscr{Z}_{\lambda}^{n, \text { er }}\right)$ for every $\delta>0$. Thus, using Theorem 1.1, we conclude that for every $\eta>0$ and $\lambda>1$,

$$
\begin{align*}
& \mathbb{P}\left(\mathfrak{m}\left(\left(4 \lambda^{5}\right)^{-1} ; \mathscr{M}\right) \geq \eta+3 \cdot(\log \lambda)^{2} \cdot \lambda^{-3}\right) \\
& \quad \leq \limsup _{n \rightarrow \infty} \mathbb{P}\left(6 \Delta_{n, \lambda}^{1 / 4}>\eta\right)+\sum_{i=1}^{3} \varepsilon_{i}(\lambda)+\limsup _{n \rightarrow \infty} \varepsilon_{4}^{(n)}(\lambda) \tag{6.55}
\end{align*}
$$

The result follows upon using Lemma 6.20 and noting that the right side of (6.55) tends to zero as $\lambda \rightarrow \infty$.

### 6.9 Proof of Theorem 4.8

For $p \in(0,1)$ and $m \in \mathbb{N}$, let $G_{p}^{m}$ be distributed as follows: For any connected graph $H$ on [ $m$ ] having $r$ edges,

$$
\mathbb{P}\left(G_{p}^{m}=H\right) \propto p^{r}(1-p)^{-r}
$$

That is, $G_{p}^{m}$ is an Erdős-Rényi random graph conditioned to be connected. We start with the following lemma:

Lemma 6.21 If pm ${ }^{3 / 2} \leq 1$, then

$$
\begin{equation*}
\mathbb{P}\left(s p\left(G_{p}^{m}\right) \geq 2\right) \leq C p^{2} m^{3} \tag{6.56}
\end{equation*}
$$

for some universal constant $C$. Consequently, for any $\varepsilon>0$ there exists $\lambda(\varepsilon)>0$ such that for all $\lambda \geq \lambda(\varepsilon)$,

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \mathbb{P}\left(\operatorname{sp}\left(\mathscr{C}_{i}^{n, \text { er }}(\lambda)\right) \geq 2 \text { for some } i \geq 2\right) \leq \varepsilon \tag{6.57}
\end{equation*}
$$

Proof Suppose $\mathbf{t}$ is a rooted tree on $[m]$. For $v \in[m]$ define

$$
\begin{align*}
& R\left(\overleftarrow{v}^{(k)}, v, \mathbf{t}\right):=\left\{u \in[m]: \overleftarrow{u}=\overleftarrow{v}^{(k)} \text { and } u>\leftarrow_{v}^{(k-1)}\right\}, \text { for } 1 \leq k \leq \operatorname{ht}(v, \mathbf{t}), \text { and }  \tag{6.58}\\
& R(v, \mathbf{t}):=\bigcup_{k=1}^{\operatorname{ht}(v, \mathbf{t})} R\left(\overleftarrow{v}^{(k)}, v, \mathbf{t}\right) . \tag{6.59}
\end{align*}
$$

Let $g(\mathbf{t}):=\sum_{v \in[m]}|R(v, \mathbf{t})|$ and $\operatorname{MR(t)}=\max _{v \in[m]}|R(v, \mathbf{t})|$. Let $T_{m}$ denote a uniform rooted tree on $[m]$ and let $\widetilde{T}_{m}$ be distributed as

$$
\begin{equation*}
\mathbb{P}\left(\widetilde{T}_{m}=\mathbf{t}\right)=\frac{(1-p)^{-g(\mathbf{t})} \mathbb{P}\left(T_{m}=\mathbf{t}\right)}{\mathbb{E}\left[(1-p)^{-g\left(T_{m}\right)}\right]} \tag{6.60}
\end{equation*}
$$

Then by [3, Proposition 8], $G_{p}^{m}$ has the same law as the random graph obtained from $\widetilde{T}_{m}$ by placing an edge with probability $p$ independently between every pair of vertices $v$ and $u$, where $v \in[m]$ and $u \in R\left(v, \widetilde{T}_{m}\right)$, and then forgetting the identity of the root of $\widetilde{T}_{m}$. In particular, $\operatorname{sp}\left(G_{p}^{m}\right)$ is distributed as $\operatorname{Binomial}(N, p)$ where $N \stackrel{\text { d }}{=} g\left(\widetilde{T}_{m}\right)$. Hence

$$
\begin{equation*}
\mathbb{P}\left(\operatorname{sp}\left(G_{p}^{m}\right) \geq 2\right) \leq p^{2} \mathbb{E}\left[g\left(\widetilde{T}_{m}\right)^{2}\right] \leq p^{2} m^{2} \mathbb{E}\left[M R\left(\widetilde{T}_{m}\right)^{2}\right] \tag{6.61}
\end{equation*}
$$

It follows from (6.60) that for any $x>0$,

$$
\begin{aligned}
\mathbb{P}\left(M R\left(\widetilde{T}_{m}\right) \geq x \sqrt{m}\right) & \leq \mathbb{E}\left[(1-p)^{-g\left(T_{m}\right)} \mathbb{1}_{\left\{M R\left(T_{m}\right) \geq x \sqrt{m}\right\}}\right] \\
& \leq\left[\mathbb{P}\left(M R\left(T_{m}\right) \geq x \sqrt{m}\right)\right]^{\frac{1}{2}}\left[\mathbb{E}(1-p)^{-2 g\left(T_{m}\right)}\right]^{\frac{1}{2}} \leq C e^{-C^{\prime} x^{2}} e^{C^{\prime \prime} p^{2} m^{3}},
\end{aligned}
$$

where the last step uses [3, Lemmas 13 and 14]. Using the fact that $p^{2} m^{3} \leq 1$, it follows that $\mathbb{E}\left[M R\left(\widetilde{T}_{m}\right)^{2}\right] \leq C m$. This in conjunction with (6.61) yields (6.56).

Next note that by [57, Theorem A.1], for any $\varepsilon \in(0,1)$, there exists $\lambda(\varepsilon)>0$ such that for all $\lambda \geq \lambda(\varepsilon)$,

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \mathbb{P}\left(\sum_{i \geq 2}\left|\mathscr{C}_{i}^{n, \text { er }}(\lambda)\right|^{2} \geq \varepsilon n^{4 / 3}\right) \leq \varepsilon \tag{6.62}
\end{equation*}
$$

Denote the event on the left side of (6.62) by $F(n, \lambda ; \varepsilon)$. Now conditional on the component sizes of $\operatorname{ER}(n, \lambda)$ being equal to $m_{1}, m_{2}, \ldots, m_{r}$, the components are distributed as $G_{p}^{m_{1}}, \ldots, G_{p}^{m_{r}}$ with $p=n^{-1}+\lambda n^{-4 / 3}$. Further, for any $\varepsilon \in(0,1 / 4)$ and large $n, p m_{i}^{3 / 2} \leq 1$ if $m_{i}^{2} \leq \varepsilon n^{4 / 3}$. Hence, using (6.56),

$$
\begin{align*}
& \mathbb{P}\left(\operatorname{sp}\left(\mathscr{C}_{i}^{n, \mathrm{er}}(\lambda)\right) \geq 2 \text { for some } i \geq 2\right) \\
& \quad \leq \mathbb{P}(F(n, \lambda ; \varepsilon))+C \mathbb{E}\left[\mathbb{1}\left\{F(n, \lambda ; \varepsilon)^{c}\right\} \cdot n^{-2} \sum_{i \geq 2}\left|\mathscr{C}_{i}^{n, \mathrm{er}}(\lambda)\right|^{3}\right] \leq \mathbb{P}(F(n, \lambda ; \varepsilon))+C \varepsilon . \tag{6.63}
\end{align*}
$$

(6.63) together with (6.62) yields (6.57).

Our next lemma roughly states that inside the critical window, the number of surplus edges in the largest component of the Erdős-Rényi random graph takes all large integer values with high probability, and during this time of the evolution, every other component is either a tree or is unicyclic.

Lemma 6.22 For every $\varepsilon>0$, there exists $s_{\varepsilon} \in \mathbb{N}$ such that for all integers $s \geq s_{\varepsilon}$,

$$
\begin{equation*}
\liminf _{n} \mathbb{P}\left(\mathscr{A}\left[s_{\varepsilon}, s\right]\right) \geq 1-\varepsilon \tag{6.64}
\end{equation*}
$$

where $\mathscr{A}\left[s_{\varepsilon}, s\right]$ denotes the event that there exist $\lambda_{1} \leq \lambda_{2}$ such that in the interval $\left[\lambda_{1}, \lambda_{2}\right]$, the process $\operatorname{sp}\left(\mathscr{C}_{1}^{n, \text { er }}(\cdot)\right)$ assumes all values in $\left\{s_{\varepsilon}, s_{\varepsilon}+1, \ldots, s\right\}$, and $\operatorname{sp}\left(\mathscr{C}_{i}^{n, \text { er }}(\lambda)\right) \leq 1$ for all $i \geq 2$ and $\lambda \in\left[\lambda_{1}, \lambda_{2}\right]$.

Proof For $k \geq 1$, define

$$
\beta(k)=\frac{k(k+1)}{(k+1 / 6)(k+5 / 6)} .
$$

We say that the "leader changes in $\operatorname{ER}(n, \cdot)$ after time $\lambda$ " if there exists $\lambda^{\prime}>\lambda$ such that the component in $\operatorname{ER}\left(n, \lambda^{\prime}\right)$ containing $\mathscr{C}_{1}^{n, \text { er }}(\lambda)$ is not $\mathscr{C}_{1}^{n, \text { er }}\left(\lambda^{\prime}\right)$. Fix $\eta>0$ and choose $\lambda(\eta)$ large such that the following hold:

$$
\begin{align*}
& \limsup _{n \rightarrow \infty} \mathbb{P}(\text { The leader does not change in } \operatorname{ER}(n, \cdot) \text { after time } \lambda(\eta)) \geq 1-\eta \\
& \prod_{j \geq 0} \beta\left(\left[\lambda(\eta)^{3} / 3\right]+j\right) \geq 1-\eta  \tag{6.65}\\
& \limsup _{n \rightarrow \infty} \mathbb{P}\left(\operatorname{sp}\left(\mathscr{C}_{i}^{n, \text { er }}(\lambda(\eta))\right) \leq 1 \text { for all } i \geq 2\right) \geq 1-\eta, \text { and }  \tag{6.67}\\
& \limsup _{n \rightarrow \infty} \mathbb{P}\left(\lambda(\eta)^{3} / 3<\operatorname{sp}\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda(\eta))\right)<\lambda(\eta)^{3}\right) \geq 1-\eta \tag{6.68}
\end{align*}
$$

(6.65) uses [65, Theorem 7] (see also [4]). (6.66) uses the fact that $\prod_{k \geq 1} \beta(k)>0$. (6.67) uses (6.57). (6.68) uses Lemma 3.3 and Theorem 3.2.

Let $F_{k}^{n}$ denote the event that there exists $\lambda \in \mathbb{R}$ such that the random graph $\operatorname{ER}(n, \lambda)$ has exactly one component with surplus $k+1$ and the surplus of every other component is at most one. Then by (6.67) and (6.68),

$$
\begin{equation*}
\liminf _{n \rightarrow \infty} \mathbb{P}\left(\bigcup_{k=\left[\lambda(\eta)^{3} / 3\right]}^{\left[\lambda(\eta)^{3}\right]} F_{k}^{n}\right) \geq 1-2 \eta \tag{6.69}
\end{equation*}
$$

By [58, Theorem 5.28], for $\left[\lambda(\eta)^{3} / 3\right] \leq k \leq\left[\lambda(\eta)^{3}\right]$ and any $s \geq\left[\lambda(\eta)^{3}\right]+1$,

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \mathbb{P}\left(\bigcap_{j=k+1}^{s} F_{j}^{n} \mid F_{k}^{n} \backslash\left(\bigcup_{j=\left[\lambda(\eta)^{3} / 3\right]}^{k-1} F_{j}^{n}\right)\right)=\prod_{j=k+1}^{s} \beta(j) . \tag{6.70}
\end{equation*}
$$

Note that

$$
\begin{aligned}
& \mathbb{P}\left(\bigcap_{j=\left[\lambda(\eta)^{3}\right]}^{s} F_{j}^{n}\right) \geq \sum_{k=\left[\lambda(\eta)^{3} / 3\right]}^{\left[\lambda(\eta)^{3}\right]} \mathbb{P}\left(\left(\bigcap_{j=k}^{s} F_{j}^{n}\right) \backslash\left(\bigcup_{j=\left[\lambda(\eta)^{3} / 3\right]}^{k-1} F_{j}^{n}\right)\right) \\
& \geq \sum_{k=\left[\lambda(\eta)^{3} / 3\right]}^{\left[\lambda(\eta)^{3}\right]} \mathbb{P}\left(\bigcap_{j=k+1}^{s} F_{j}^{n} \mid F_{k}^{n} \backslash\left(\bigcup_{j=\left[\lambda(\eta)^{3} / 3\right]}^{k-1} F_{j}^{n}\right)\right) \cdot \mathbb{P}\left(F_{k}^{n} \backslash\left(\bigcup_{j=\left[\lambda(\eta)^{3} / 3\right]}^{k-1} F_{j}^{n}\right)\right) \\
& \geq \mathbb{P}\left(\bigcup_{k=\left[\lambda(\eta)^{3} / 3\right]}^{\left[\lambda(\eta)^{3}\right]} F_{k}^{n}\right) \times \min _{\star} \mathbb{P}\left(\bigcap_{j=k+1}^{s} F_{j}^{n} \mid F_{k}^{n} \backslash\left(\bigcup_{j=\left[\lambda(\eta)^{3} / 3\right]}^{k-1} F_{j}^{n}\right)\right),
\end{aligned}
$$

where $\min _{\star}$ is minimum taken over $\left[\lambda(\eta)^{3} / 3\right] \leq k \leq\left[\lambda(\eta)^{3}\right]$. Thus, (6.69), (6.70), and (6.66) give

$$
\liminf _{n \rightarrow \infty} \mathbb{P}\left(\bigcap_{j=\left[\lambda(\eta)^{3}\right]}^{s} F_{j}^{n}\right) \geq(1-2 \eta)(1-\eta) .
$$

Combining this with (6.67) and (6.68), we see that

$$
\begin{aligned}
& \liminf _{n \rightarrow \infty} \mathbb{P}\left(\left\{\bigcap_{j=\left[\lambda(\eta)^{3}\right]}^{s} F_{j}^{n}\right\} \bigcap\left\{\operatorname{sp}\left(\mathscr{C}_{i}^{n, \text { er }}(\lambda(\eta))\right) \leq 1 \text { for all } i \geq 2\right\}\right. \\
& \left.\bigcap\left\{\lambda(\eta)^{3} / 3<\operatorname{sp}\left(\mathscr{C}_{1}^{n, \text { er }}(\lambda(\eta))\right)<\lambda(\eta)^{3}\right\}\right)>1-5 \eta
\end{aligned}
$$

Thus, for all large $n$, in the process $(\operatorname{ER}(n, \lambda), \lambda \geq \lambda(\eta))$, with probability at least $1-5 \eta$, the surplus of the component containing $\mathscr{C}_{1}^{n, \text { er }}(\lambda(\eta))$ assumes all values in $\left\{\left[\lambda(\eta)^{3}\right]+1, \ldots, s\right\}$, and during this part of the evolution, the surplus of the other components remains at most one. By (6.65), the component containing $\mathscr{C}_{1}^{n, \text { er }}(\lambda(\eta))$ remains the largest component after time $\lambda(\eta)$ with probability at least $1-\eta$. Thus, (6.64) follows if we take $s_{\varepsilon}=\left[\lambda(\eta)^{3}\right]+1$ with $\eta=\varepsilon / 6$.

Let $\mathscr{H}_{n, s}$ be as in Theorem 3.11 and let $L(\cdot)$ be as in (3.4). Define $\widetilde{\mathscr{H}}_{n, s}$ and $\widetilde{\mathscr{H}}^{(s)}$ via

$$
\mathbb{E}\left[f\left(\widetilde{\mathscr{H}}_{n, s}\right)\right]=\frac{\mathbb{E}\left[f\left(\mathscr{H}_{n, s}\right) L\left(\mathscr{H}_{n, s}\right)\right]}{\mathbb{E}\left[L\left(\mathscr{H}_{n, s}\right)\right]} \text { and } \mathbb{E}\left[f\left(\widetilde{\mathscr{H}}^{(s)}\right)\right]=\frac{\mathbb{E}\left[f\left(\mathscr{H}^{(s)}\right) L\left(\mathscr{H}^{(s)}\right)\right]}{\mathbb{E}\left[L\left(\mathscr{H}^{(s)}\right)\right]}
$$

for every bounded measurable $f: \mathfrak{S}_{\mathrm{GHP}} \rightarrow \mathbb{R}$. For $s \geq 2$ define

$$
\begin{align*}
& \tau_{s}=\inf \left\{\lambda: \operatorname{sp}\left(\mathscr{C}^{\star}\right)=s \text { for some component } \mathscr{C}^{\star} \text { of } \operatorname{ER}(n, \lambda)\right. \\
& \quad \text { and } \operatorname{sp}(\mathscr{C}) \leq 1 \text { for every other component } \mathscr{C} \text { of } \operatorname{ER}(n, \lambda)\} . \tag{6.71}
\end{align*}
$$

If $\tau_{s}<\infty$, define $\mathscr{C}_{\tau_{s}, \star}^{n, \star}$ to be the (unique) component of $\operatorname{ER}\left(n, \tau_{s}\right)$ with $\operatorname{sp}\left(\mathscr{C}_{\tau_{s}}^{n, \star}\right)=s$. If $\tau_{s}=\infty$, define $\mathscr{C}_{\tau_{s}}^{n, \star}$ to be the one-point space.

Lemma 6.23 Fix $s \geq 2$. Let $U_{i j}, 1 \leq i<j \leq n$, be the i.i.d. Uniform[0, 1] random variables used in the construction of $\mathrm{ER}(n, \cdot)$. Let $\widetilde{\mathscr{H}}_{n, s}$ be independent of $\left(U_{i j}, 1 \leq\right.$ $i<j \leq n$ ). Define $M_{\tau_{s}}^{n, \star}$ to be the MST of $\mathscr{C}_{\tau_{s}}^{n, \star}$ constructed using the weights $U_{i j}$ if $\tau_{s}<\infty$ and $\left|\mathscr{C}_{\tau_{s}}^{n, \star}\right| \geq \log n$, and set $M_{\tau_{s}}^{n, \star}=C B D_{\infty}\left(\widetilde{\mathscr{H}}_{n, s}\right)$ otherwise. Then as $n \rightarrow \infty$,

$$
\left(\left|M_{\tau_{s}}^{n, \star}\right|\right)^{-1 / 2} M_{\tau_{s}}^{n, \star} \xrightarrow{d} C B^{\infty}\left(\widetilde{\mathscr{H}}^{(s)}\right) \quad \text { w.r.t. GHP topology. }
$$

Proof For convenience, we will assume that the random vector ( $U_{i j}, 1 \leq i<j \leq n$ ) is given by the identity map on the canonical probability space $[0,1]^{\binom{n}{2}}$ endowed with the $\binom{n}{2}$-fold product of the uniform measure on $[0,1]$.

For any subgraph $H$ of the complete graph on [ $n$ ], define the event

$$
F_{H}:=\left\{\tau_{s}<\infty, \operatorname{ER}\left(n, \tau_{s}\right) \backslash \mathscr{C}_{\tau_{s}}^{n, \star}=H\right\} .
$$

Fix any $H$ with $\mathbb{P}\left(F_{H}\right)>0$. Then $\mathbb{P}\left(F_{H} \cap\left\{\mathscr{C}_{\tau_{s}}^{n, \star}=H_{1}\right\}\right)>0$ for any connected graph $H_{1}$ with

$$
\begin{equation*}
V\left(H_{1}\right)=[n] \backslash V(H), \text { and } \operatorname{sp}\left(H_{1}\right)=s \tag{6.72}
\end{equation*}
$$

Now for any $H_{1}$ satisfying (6.72), the realizations ( $u_{i j} ; 1 \leq i<j \leq n$ ) of the random variables $U_{i j}$ for which $F_{H} \cap\left\{\mathscr{C}_{\tau_{s}}^{n, \star}=H_{1}\right\}$ holds are given by

$$
\begin{aligned}
F_{H} \cap\left\{\mathscr{C}_{\tau_{s}}^{n, \star}=H_{1}\right\} & =\left\{\max \left\{u_{i j}:(i, j) \in E\left(H_{1}\right) \cup E(H)\right\}\right. \\
& =\max \left\{u_{i j}:(i, j) \in E\left(\operatorname{Core}\left(H_{1}\right)\right)\right\} \\
& \left.<\min \left\{u_{i j}:(i, j) \notin E\left(H_{1}\right) \cup E(H)\right\}\right\},
\end{aligned}
$$

and for any such realization $\left(u_{i j}\right)$, we have $\left(u_{\pi(i, j)} ; 1 \leq i<j \leq n\right) \in F_{H} \cap\left\{\mathscr{C}_{\tau_{s}}^{n, \star}=\right.$ $\left.H_{1}\right\}$ for any permutation $\pi$ of $\{(i, j): 1 \leq i<j \leq n\}$ satisfying $\pi(i, j)=(i, j)$ for all $(i, j) \notin E\left(\operatorname{Core}\left(H_{1}\right)\right)$. Hence, conditional on $\tau_{s}<\infty$ and $\operatorname{ER}\left(n, \tau_{s}\right)$, the random
variables $U_{i, j},(i, j) \in \operatorname{Core}\left(\mathscr{C}_{\tau_{s}}^{n, \star}\right)$, are exchangeable. Using Lemma 4.5, we see that the following equality of conditional distributions hold for any $m \geq \log n$ :

$$
\begin{equation*}
\left(M _ { \tau _ { s } } ^ { n , \star } | \tau _ { s } < \infty , | \mathscr { C } _ { \tau _ { s } } ^ { n , \star } | = m ) \stackrel { \mathrm { d } } { = } \left(\operatorname{CBD}_{\infty}\left(\mathscr{C}_{\tau_{s}}^{n, \star}\right)\left|\tau_{s}<\infty,\left|\mathscr{C}_{\tau_{s}}^{n, \star}\right|=m\right)\right.\right. \tag{6.73}
\end{equation*}
$$

Next, for any two graphs $G_{1}, G_{2}$ on $[n]$, write $\mathbb{P}^{\text {er }}\left(G_{1}, G_{2}\right)$ to denote the probability that there exist $\lambda_{1} \leq \lambda_{2}$ such that $\operatorname{ER}\left(n, \lambda_{1}\right)=G_{1}$ and $\operatorname{ER}\left(n, \lambda_{2}\right)=G_{2}$. Thus, if $G_{1}$ is a subgraph of $G_{2}$, then

$$
\begin{equation*}
\mathbb{P}^{\mathrm{er}}\left(G_{1}, G_{2}\right)=\frac{1}{N!} \cdot\left|E\left(G_{1}\right)\right|!\cdot\left(\left|E\left(G_{2}\right)\right|-\left|E\left(G_{1}\right)\right|\right)!\cdot\left(N-\left|E\left(G_{2}\right)\right|\right)!, \tag{6.74}
\end{equation*}
$$

where $N=\binom{n}{2}$. Now for any $H_{1}$ satisfying (6.72),

$$
\begin{aligned}
& \mathbb{P}\left(\mathscr{C}_{\tau_{s}}^{n, \star}=H_{1} \mid F_{H}\right)=\frac{1}{\mathbb{P}\left(F_{H}\right)} \mathbb{P}\left(\tau_{s}<\infty, \operatorname{ER}\left(n, \tau_{s}\right)=H_{1} \cup H\right) \\
& \quad=\frac{1}{\mathbb{P}\left(F_{H}\right)} \sum_{e \in E\left(\operatorname{Core}\left(H_{1}\right)\right)} \mathbb{P}^{\operatorname{er}}\left(\left(H_{1} \backslash e\right) \cup H, H_{1} \cup H\right) \propto\left|E\left(\operatorname{Core}\left(H_{1}\right)\right)\right|=L\left(H_{1}\right),
\end{aligned}
$$

where in the penultimate step we have used (6.74) to deduce that the summands are the same for any $H_{1}$ satisfying (6.72). Thus, for any $m \geq \log n$, the conditional distribution of $\mathscr{C}_{\tau_{s}}^{n, \star}$ given $\tau_{s}<\infty$ and $\left|\mathscr{C}_{\tau_{s}}^{n, \star}\right|=m$ satisfies

$$
\begin{equation*}
\left(\mathscr{C}_{\tau_{s}}^{n, \star}\left|\tau_{s}<\infty,\left|\mathscr{C}_{\tau_{s}}^{n, \star}\right|=m\right) \stackrel{\mathrm{d}}{=} \widetilde{\mathscr{H}}_{m, s}\right. \tag{6.75}
\end{equation*}
$$

Now for any bounded continuous $f: \mathfrak{S}_{\mathrm{GHP}} \rightarrow \mathbb{R}$,
$\lim _{m \rightarrow \infty} \mathbb{E}\left[f\left(\frac{1}{\sqrt{m}} \widetilde{\mathscr{H}}_{m, s}\right)\right]=\lim _{m \rightarrow \infty} \frac{\mathbb{E}\left[f\left(\frac{1}{\sqrt{m}} \mathscr{H}_{m, s}\right) L\left(\mathscr{H}_{m, s}\right)\right]}{\mathbb{E}\left[L\left(\mathscr{H}_{m, s}\right)\right]}=\frac{\mathbb{E}\left[f\left(\mathscr{H}^{(s)}\right) L\left(\mathscr{H}^{(s)}\right)\right]}{\mathbb{E}\left[L\left(\mathscr{H}^{(s)}\right)\right]}=\mathbb{E}\left[f\left(\widetilde{\mathscr{H}}^{(s)}\right)\right]$, where the second step uses (3.8) and (3.9). Hence $m^{-1 / 2} \widetilde{\mathscr{H}}_{m, s} \xrightarrow{\mathrm{~d}} \widetilde{\mathscr{H}}^{(s)}$ as $m \rightarrow \infty$ w.r.t. GHP topology. Using Theorem 4.6 and (3.10), it follows that as $m \rightarrow \infty$,

$$
m^{-1 / 2} \mathrm{CBD}_{\infty}\left(\widetilde{\mathscr{H}}_{m, s}\right) \xrightarrow{\mathrm{d}} \mathrm{CB}^{\infty}\left(\widetilde{\mathscr{H}}^{(s)}\right)
$$

w.r.t. GHP topology. Now the claim follows from (6.73) and (6.75).

Proof of Theorem 4.8 Fix $0<\varepsilon<1 / 2$. For $s \geq 3$, define $\lambda_{s}$ by the relation $2 \lambda_{s}^{3}=3 s$. Define

$$
\begin{equation*}
E_{s}:=\left\{\mathscr{A}[s, s] \bigcap\left\{\text { the leader does not change in } \operatorname{ER}(n, \cdot) \text { after time } \lambda_{s} / 2\right\}\right\} \tag{6.76}
\end{equation*}
$$

where $\mathscr{A}[\cdot, \cdot]$ is as in Lemma 6.22. Using (6.64) and (6.65), choose $s_{1}$ large so that

$$
\limsup _{n \rightarrow \infty} \mathbb{P}\left(E_{s}^{c}\right) \leq \varepsilon \text { for all } s \geq s_{1} .
$$

Let $M_{\lambda}^{n, \text { er }}$ denote the MST of $\mathscr{C}_{1}^{n, \text { er }}(\lambda)$ constructed using the same i.i.d. Uniform $[0,1]$ random variables $U_{i j}$ used to construct the process $\operatorname{ER}(n, \cdot)$. If the leader does not change after time $\lambda$, then using Observation 4.2 , we see that $M_{\lambda}^{n, \text { er }}$ is a subtree of $M_{\lambda^{\prime}}^{n, \text { er }}$ for any $\lambda^{\prime}>\lambda$. Thus, using [5, Lemma 4.5], we can choose $s_{2}$ large enough so that

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \mathbb{P}\left(d_{H}\left(M_{\lambda}^{n, \text { er }}, M_{\lambda^{\prime}}^{n, \text { er }}\right)>\varepsilon n^{1 / 3}\right) \leq \varepsilon \text { for all } \lambda, \lambda^{\prime} \geq \lambda_{s_{2}} \tag{6.77}
\end{equation*}
$$

Next, define $\bar{\lambda}_{s}=\lambda_{s}(1+\varepsilon)$ and $\underline{\lambda}_{s}=\lambda_{s}(1-\varepsilon)$, and using Lemma 3.3, Theorem 3.2, (6.57), and (6.76), choose $s_{3}$ large enough so that

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \mathbb{P}\left(F_{s, \varepsilon}^{c}\right) \leq 2 \varepsilon \text { for all } s \geq s_{3}, \tag{6.78}
\end{equation*}
$$

where

$$
\begin{gathered}
F_{s, \varepsilon}:=E_{s} \bigcap\left\{\left|\mathscr{C}_{1}^{n, \text { er }}\left(\underline{\lambda}_{s}\right)\right| \geq 2 \underline{\lambda}_{s}(1-\varepsilon) n^{2 / 3} \quad \text { and } \quad\left|\mathscr{C}_{1}^{n, \text { er }}\left(\bar{\lambda}_{s}\right)\right| \leq 2 \bar{\lambda}_{s}(1+\varepsilon) n^{2 / 3}\right\} \\
\bigcap\left\{2 \leq \operatorname{sp}\left(\mathscr{C}_{1}^{n, \text { er }}\left(\underline{\lambda}_{s}\right)\right) \leq s-1 \quad \text { and } \quad \operatorname{sp}\left(\mathscr{C}_{1}^{n, \text { er }}\left(\bar{\lambda}_{s}\right)\right) \geq s+1\right\} \\
\bigcap\left\{\forall i \geq 2, \operatorname{sp}\left(\mathscr{C}_{i}^{n, \text { er }}\left(\underline{\lambda}_{s}\right)\right) \leq 1 \quad \text { and } \quad \operatorname{sp}\left(\mathscr{C}_{i}^{n, \text { er }}\left(\bar{\lambda}_{s}\right)\right) \leq 1\right\}
\end{gathered}
$$

Set $s_{0}:=\max \left\{s_{2}, s_{3}\right\}$. From now on, we will only consider $s \geq s_{0}$.
Let $\tau_{s}$ be as in (6.71). If $\tau_{s}<\infty$, let $M_{\tau_{s}}^{n, \text { er }}$ be the MST of $\mathscr{C}_{1}^{n, \text { er }}\left(\tau_{s}\right)$ constructed using the edge weights $U_{i j}$. If $\tau_{s}=\infty$, set $\mathscr{C}_{1}^{n, \text { er }}\left(\tau_{s}\right)$ to be the complete graph $K_{n}$, and let $M_{\tau_{s}}^{n, \text { er }}=M_{\infty}^{n, \text { er }}$-the MST of $K_{n}$ constructed using the edge weights $U_{i j}$. Note that on the event $F_{s, \varepsilon}, \underline{\lambda}_{s}<\tau_{s}<\bar{\lambda}_{s}, \mathscr{C}_{1}^{n, \text { er }}\left(\tau_{s}\right)=\mathscr{C}_{\tau_{s}}^{n, \star}$, and $M_{\tau_{s}}^{n, \text { er }}=M_{\tau_{s}}^{n, \star}$, where the notation is as in Lemma 6.23. Thus, writing $\mathscr{L}(\cdot)$ and $d_{\mathrm{PR}}(\cdot, \cdot)$ to denote the law of a random metric measure space and the Prokhorov distance between two measures respectively, it follows from Lemma 6.23 that

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} d_{\mathrm{PR}}\left(\mathscr{L}\left(\frac{(12 s)^{1 / 6}}{\left(\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)\right|\right)^{1 / 2}} M_{\tau_{s}}^{n, \text { er }}\right), \mathscr{L}\left((12 s)^{1 / 6} \cdot \mathrm{CB}^{\infty}\left(\widetilde{\mathscr{H}}^{(s)}\right)\right)\right) \leq 2 \varepsilon \tag{6.79}
\end{equation*}
$$

Next note that on $F_{s, \varepsilon}, M_{\hat{\lambda}_{s}}^{n, \text { er }} \subseteq M_{\tau_{s}}^{n, \text { er }} \subseteq M_{\bar{\lambda}_{s}}^{n \text {,er }}$. On $F_{s, \varepsilon}$, for every $i \in M_{\tau_{s}}^{n \text {,er }}$, let $V_{i}:=\left\{j \in M_{\bar{\lambda}_{s}}^{n, \text { er }}:\right.$ the path connecting $j$ and $i$ in $M_{\bar{\lambda}_{s}}^{n, \text { er }}$ intersects $M_{\tau_{s}}^{n, \text { er }}$ only at $\left.i\right\}$.

Note that $i \in V_{i}$. Let $C$ be the correspondence between $M_{\tau_{s}}^{n, \text { er }}$ and $M_{\bar{\lambda}_{s}}^{n, \text { er }}$ given by $C=\left\{(i, j): \quad i \in M_{\tau_{s}}^{n, \text { er }}, j \in V_{i}\right\}$. Define a measure $\pi$ on $M_{\tau_{s}}^{n, \text { er }} \times M_{\bar{\lambda}_{s}}^{n, \text { er }}$ via $\pi\{(i, j)\}=1 /\left|\mathscr{C}_{1}^{n, \text { er }}\left(\bar{\lambda}_{s}\right)\right|$ for $(i, j) \in C$. Then on the event $F_{s, \varepsilon}$,

$$
\operatorname{dis}(C) \leq \frac{1}{2} d_{\mathrm{H}}\left(M_{\underline{\lambda}_{s}}^{n, \mathrm{er}}, M_{\bar{\lambda}_{s}}^{n, \text { er }}\right), \text { and } \pi\left(C^{c}\right)=0
$$

Further, writing $\mu_{1}$ and $\mu_{2}$ for the uniform probability measures on $M_{\tau_{s}}^{n, \text { er }}$ and $M_{\bar{\lambda}_{s}}^{n \text {,er }}$ respectively, on the event $F_{s, \varepsilon}$,

$$
\begin{align*}
D\left(\pi ; \mu_{1}, \mu_{2}\right) & \leq \sum_{i \in \mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)}\left|\frac{1}{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)\right|}-\frac{\left|V_{i}\right|}{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|}\right| \\
& \leq \sum_{i \in \mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)}\left(\frac{1}{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)\right|}-\frac{1}{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|}\right)+\sum_{i \in \mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)} \frac{\left|V_{i}\right|-1}{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|} \\
& \leq \sum_{i \in \mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)}\left(\frac{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|-\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|}{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\underline{\lambda}_{s}\right)\right| \cdot\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|}\right)+\frac{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|-\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\lambda_{s}\right)\right|}{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|} \\
& \leq 2 \times \frac{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\bar{\lambda}_{s}\right)\right|-\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\underline{\lambda}_{s}\right)\right|}{\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\underline{\lambda}_{s}\right)\right|} \leq 2 \frac{(1+\varepsilon)^{2}-(1-\varepsilon)^{2}}{(1-\varepsilon)^{2}} \leq 32 \varepsilon, \tag{6.80}
\end{align*}
$$

where the last step uses $\varepsilon<1 / 2$. By (6.77), (6.78), and (6.80), it follows that

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \mathbb{P}\left(d_{\mathrm{GHP}}\left(n^{-\frac{1}{3}} \cdot M_{\tau_{s}}^{n, \text { er }}, n^{-\frac{1}{3}} \cdot M_{\bar{\lambda}_{s}}^{n, \text { er }}\right)>32 \varepsilon\right) \leq 3 \varepsilon \tag{6.81}
\end{equation*}
$$

Note that

$$
\begin{equation*}
\left\{2 \underline{\lambda}_{s}(1-\varepsilon) n^{2 / 3} \leq\left|\mathscr{C}_{1}^{n, \text { er }}\left(\tau_{s}\right)\right| \leq 2 \bar{\lambda}_{s}(1+\varepsilon) n^{2 / 3}\right\} \supseteq F_{s, \varepsilon} \tag{6.82}
\end{equation*}
$$

Hence, on the event $F_{s, \varepsilon}$,

$$
\begin{align*}
d_{\mathrm{GHP}}\left(\frac{(12 s)^{1 / 6}}{\left(\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)\right|\right)^{1 / 2}} M_{\tau_{s}}^{n, \text { er }}, \frac{1}{n^{1 / 3}} M_{\tau_{s}}^{n, \text { er }}\right) & \leq \operatorname{diam}\left(M_{\tau_{s}}^{n, \text { er }}\right)\left|\frac{(12 s)^{1 / 6}}{\left(\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)\right|\right)^{1 / 2}}-\frac{1}{n^{1 / 3}}\right| \\
& \leq \frac{1}{n^{1 / 3}} \times \operatorname{diam}\left(M_{\infty}^{n, \mathrm{er}}\right) \times 3 \varepsilon, \tag{6.83}
\end{align*}
$$

where the last step uses (6.82) and the relation $2 \lambda_{s}^{3}=3 s$. By Theorem 1.1, the sequence of random variables $\left(n^{-1 / 3} \operatorname{diam}\left(M_{\infty}^{n, \text { er }}\right) ; n \geq 1\right)$ is tight. It thus follows from (6.78),
(6.81), and (6.83) that

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \mathbb{P}\left(d_{\mathrm{GHP}}\left(\frac{(12 s)^{1 / 6}}{\left(\left|\mathscr{C}_{1}^{n, \mathrm{er}}\left(\tau_{s}\right)\right|\right)^{1 / 2}} M_{\tau_{s}}^{n, \text { er }}, \frac{1}{n^{1 / 3}} M_{\bar{\lambda}_{s}}^{n, \text { er }}\right) \geq \sqrt{\varepsilon}\right)=: \delta_{\varepsilon} \tag{6.84}
\end{equation*}
$$

satisfies $\delta_{\varepsilon} \downarrow 0$ as $\varepsilon \downarrow 0$.
By (4.3), $n^{-1 / 3} M_{\bar{\lambda}_{s}}^{n} \xrightarrow{\mathrm{e}} \mathrm{CB}^{\infty}\left(S_{1}\left(\bar{\lambda}_{s}\right)\right)$ as $n \rightarrow \infty$ w.r.t. GHP topology. Combining this with (6.84) and (6.79), we see that

$$
\begin{equation*}
d_{\mathrm{PR}}\left(\mathscr{L}\left((12 s)^{1 / 6} \cdot \mathrm{CB}^{\infty}\left(\widetilde{\mathscr{H}}^{(s)}\right)\right), \mathscr{L}\left(\mathrm{CB}^{\infty}\left(S_{1}\left(\bar{\lambda}_{s}\right)\right)\right)\right) \leq 2 \varepsilon+\delta_{\varepsilon}+\sqrt{\varepsilon} \tag{6.85}
\end{equation*}
$$

Finally, by Theorem 4.7, $\mathrm{CB}^{\infty}\left(S_{1}\left(\bar{\lambda}_{s}\right)\right) \xrightarrow{\mathrm{d}} \mathscr{M}$ as $s \rightarrow \infty$ w.r.t. GHP topology. Combining this observation with (6.85) implies that

$$
(12 s)^{1 / 6} \cdot \mathrm{CB}^{\infty}\left(\widetilde{\mathscr{H}}^{(s)}\right) \xrightarrow{\mathrm{d}} \mathscr{M} \text { as } s \rightarrow \infty
$$

w.r.t. GHP topology. Now the proof is completed by using Lemma 6.24 stated below.

Lemma 6.24 For any bounded measurable $f: \mathfrak{S}_{\mathrm{GHP}} \rightarrow \mathbb{R}$,
$\mathbb{E}\left[f\left((12 s)^{1 / 6} \cdot C B^{\infty}\left(\widetilde{\mathscr{H}}^{(s)}\right)\right)\right]-\mathbb{E}\left[f\left((12 s)^{1 / 6} \cdot C B^{\infty}\left(\mathscr{H}^{(s)}\right)\right)\right] \rightarrow 0$ as $s \rightarrow \infty$.
Proof Let $r=3(s-1)$. Let $\left(X_{1}, \ldots, X_{r}\right)$ be as in Construction 3.9 and $Y_{i}, Z_{i}$, $1 \leq i \leq r$, and $\Gamma_{r / 2}$ be as in the proof of Proposition 6.3. Then $\Gamma_{r / 2} \sim \operatorname{Gamma}(r / 2,1)$, and as observed in (6.13), $\sqrt{2} Z_{i}, 1 \leq i \leq r$, are i.i.d. Exponential(1) random variables. Hence

$$
\begin{equation*}
L\left(\mathscr{H}^{(s)}\right)=\sum_{i=1}^{r} Y_{i} \sqrt{X_{i}}=\sum_{i=1}^{r} \frac{Z_{i}}{\sqrt{\Gamma_{r / 2}}}=\sqrt{r} \cdot\left(1+o_{P}(1)\right) . \tag{6.86}
\end{equation*}
$$

Further, for any $s \geq 3$,

$$
\begin{equation*}
\mathbb{E}\left[\frac{L\left(\mathscr{H}^{(s)}\right)^{2}}{r}\right] \leq \frac{1}{r} \cdot\left(\mathbb{E}\left[\left(\sum_{i=1}^{r} Z_{i}\right)^{4}\right]\right)^{1 / 2} \cdot\left(\mathbb{E}\left[\Gamma_{r / 2}^{-2}\right]\right)^{1 / 2} \leq C \tag{6.87}
\end{equation*}
$$

for a universal constant $C$. It follows from (6.86) and (6.87) that

$$
\begin{equation*}
\lim _{s \rightarrow \infty} \frac{1}{\sqrt{r}} \cdot \mathbb{E}\left[L\left(\mathscr{H}^{(s)}\right)\right]=1 \tag{6.88}
\end{equation*}
$$

which in turn implies that $r^{-1 / 2}\left(L\left(\mathscr{H}^{(s)}\right)-\mathbb{E}\left[L\left(\mathscr{H}^{(s)}\right)\right]\right) \xrightarrow{\mathrm{P}} 0$ as $s \rightarrow \infty$. Now

$$
\mathbb{E}\left[\left(L\left(\mathscr{H}^{(s)}\right)-\mathbb{E}\left[L\left(\mathscr{H}^{(s)}\right)\right]\right)^{2}\right] \leq \mathbb{E}\left[L\left(\mathscr{H}^{(s)}\right)^{2}\right] \leq C r
$$

by using (6.87). It thus follows that

$$
\begin{equation*}
\lim _{s \rightarrow \infty} r^{-1 / 2} \mathbb{E}\left|L\left(\mathscr{H}^{(s)}\right)-\mathbb{E}\left(L\left(\mathscr{H}^{(s)}\right)\right)\right|=0 . \tag{6.89}
\end{equation*}
$$

Hence

$$
\begin{aligned}
& \left|\mathbb{E}\left(f\left((12 s)^{1 / 6} \cdot \mathrm{CB}^{\infty}\left(\widetilde{\mathscr{H}}^{(s)}\right)\right)\right)-\mathbb{E}\left(f\left((12 s)^{1 / 6} \cdot \mathrm{CB}^{\infty}\left(\mathscr{H}^{(s)}\right)\right)\right)\right| \\
& \quad \leq\|f\|_{\infty} \cdot \frac{\mathbb{E}\left|L\left(\mathscr{H}^{(s)}\right)-\mathbb{E}\left(L\left(\mathscr{H}^{(s)}\right)\right)\right|}{\mathbb{E}\left[L\left(\mathscr{H}^{(s)}\right)\right]} \rightarrow 0
\end{aligned}
$$

as $s \rightarrow \infty$, where the last step follows from (6.88) and (6.89).

## 7 Discussion

Here we briefly discuss universality of the scaling limit of the MST and related open problems.
(a) Universality of MST scaling limit for models exhibiting mean-field behavior: The geometry of the MST of an underlying discrete structure is closely related to the geometry of the structure under critical percolation. The behavior under critical percolation of several models exhibiting mean-field behavior is well-understood. In [3], the metric space scaling limit of the critical Erdős-Rényi random graph was established. Soon after this work, an abstract universality principle was developed in [22,27] which was used to establish Erdős-Rényi type scaling limits for a wide array of critical random graph models including the configuration model under critical percolation, various models of inhomogeneous random graphs, and the Bohman-Frieze process. In [26], the metric space scaling limit of random graphs with critical degree sequence having finite third moment was established. Further, existing literature suggests that the components of the high-dimensional discrete torus $[51,52,85]$ and the hypercube [83] under critical percolation, and the critical quantum random graph model [38] also share the Erdős-Rényi scaling limit. It is believed that the scaling limit of the MST of each of these models exists and has the same law as $\mathscr{M}$ up to a scaling factor.

We briefly discuss here how such a result might be proved for general random graphs with given degree sequences.

Assumption 7.1 Suppose $\mathbf{d}=\mathbf{d}^{(n)}=\left(d_{v}^{(n)}, v \in[n]\right)$ is a degree sequence for each $n \geq 1$, and write $v^{n}:=n^{-1} \sum_{v \in[n]} \delta_{d_{v}^{n}}$ for the empirical degree distribution. Assume the following hold:
(i) There exists a measure $v$ on $\mathbb{Z}_{\geq 0}$ such that $\nu^{n} \rightarrow \nu$ as $n \rightarrow \infty$ w.r.t. the $W_{3}$ distance.
(ii) We have $\nu^{n}(0)=0$ for all $n \geq 1, \lim _{n \rightarrow \infty} n^{1 / 2} \cdot v^{n}(1)=0$, and $v(2)=0$.

Assumption 7.1 ((ii)) ensures that $\mathbb{P}\left(\mathfrak{B}_{n}\right) \rightarrow 1$, where $\mathfrak{B}_{n}$ denotes the event that $\mathscr{G}_{n, \mathbf{d}}$ is connected [43, Theorem 2.2]. Note that Assumption 7.1 ((ii)) in particular implies that $\sigma_{2}(\nu) \geq 3 \sigma_{1}(\nu)$, which is stronger than the condition for supercriticality of $\mathscr{G}_{n, \mathbf{d}}$, namely, $\sigma_{2}(\nu)>2 \sigma_{1}(\nu)$.
Conjecture 7.2 Let $M^{\mathbf{d}}$ (resp. $\bar{M}^{\mathbf{d}}$ ) denote the MST of $\mathscr{G}_{n, \mathbf{d}}$ (resp. $\overline{\mathscr{G}}_{n, \mathbf{d}}$ ), and let $\mathscr{M}$ be as in Theorem 1.1. Then under Assumption 7.1,

$$
\begin{equation*}
n^{-1 / 3} \cdot M^{\mathbf{d}} \xrightarrow{d} \beta(\nu) \cdot \mathscr{M} \quad \text { as } n \rightarrow \infty \tag{7.1}
\end{equation*}
$$

with respect to the GHP topology, where the constant $\beta(v)$ is given by the following prescription: Let $D$ and $Y$ be random variables such that $D$ has law $v$ and, conditional on $D, Y$ is $\operatorname{Binomial}(D, p)$-distributed,
where $p=\sigma_{1}(v) /\left(\sigma_{2}(v)-\sigma_{1}(v)\right)$. Then

$$
\beta(\nu):=\mathbb{E}[Y] \cdot\left(\mathbb{E}\left[Y^{3}\right]-4 \cdot \mathbb{E}[Y]\right)^{-2 / 3}
$$

Further, (7.1) continues to hold if we replace $M^{\mathbf{d}}$ by $\bar{M}^{\mathbf{d}}$.
In the context of the 3-regular configuration model, $v=\delta_{\{3\}}$. A simple calculation shows that $\beta\left(\delta_{\{3\}}\right)=6^{1 / 3}$, which is exactly the constant in Theorem 2.1. Let $U_{e}$, $e \in E\left(\mathscr{G}_{n, \mathbf{d}}\right)$, be i.i.d. Uniform[0,1] random variables conditional on $\mathscr{G}_{n, \mathbf{d}}$. Suppose $M^{\mathbf{d}}$ is constructed using these edge weights. For $\lambda \geq 0$, let

$$
\mathfrak{p}_{\lambda}^{n}:=\frac{\sigma_{1}\left(\nu^{n}\right)}{\sigma_{2}\left(\nu^{n}\right)-\sigma_{1}\left(\nu^{n}\right)}+\frac{\lambda}{n^{1 / 3}}
$$

Let $\mathscr{G}_{n, \mathbf{d}}(\lambda)$ be the graph with vertex set $[n]$ and edge set $\left\{e \in E\left(\mathscr{G}_{n, \mathbf{d}}\right): U_{e} \leq \mathfrak{p}_{\lambda}^{n}\right\}$. Write $\mathscr{C}_{1}^{\mathbf{d}}(\lambda)$ for the largest connected component of $\mathscr{G}_{n, \mathbf{d}}(\lambda)$, and let $M_{\lambda}^{\mathbf{d}}$ denote the MST of $\mathscr{C}_{1}^{\mathbf{d}}(\lambda)$ constructed using the edge weights $U_{e}, e \in E\left(\mathscr{C}_{1}^{\mathbf{d}}(\lambda)\right)$.

Using Observation 4.2, on the event $\mathfrak{B}_{n}, M_{\lambda}^{\mathbf{d}}$ is the restriction of $M^{\mathbf{d}}$ to $\mathscr{C}_{1}^{\mathbf{d}}(\lambda)$. On $\mathfrak{B}_{n}$, consider the forest obtained by removing from $M^{\mathbf{d}}$ the vertices in $M_{\lambda}^{\mathbf{d}}$ and all edges incident to the vertices in $M_{\lambda}^{\mathbf{d}}$; let $\mathfrak{T}_{i, \lambda}^{\mathbf{d}}, i=1, \ldots, k_{\lambda}^{\mathbf{d}}$, denote the trees in this forest, and set $X_{\lambda}^{\mathbf{d}}:=\max _{i}\left|\mathfrak{T}_{i, \lambda}^{\mathbf{d}}\right|$. On $\mathfrak{B}_{n}^{c}$, set $X_{\lambda}^{\mathbf{d}}=0$.

To prove Conjecture 7.2, it is enough to prove the following two estimates: For all $\varepsilon>0$,

$$
\begin{align*}
& \lim _{\lambda \rightarrow \infty} \limsup _{n \rightarrow \infty} \mathbb{P}\left(\mathfrak{B}_{n} \cap\left\{d_{\mathrm{H}}\left(M_{\lambda}^{\mathbf{d}}, M^{\mathbf{d}}\right)>\varepsilon n^{1 / 3}\right\}\right)=0, \quad \text { and }  \tag{7.2}\\
& \lim _{\lambda \rightarrow \infty} \limsup _{n \rightarrow \infty} \mathbb{P}\left(X_{\lambda}^{\mathbf{d}}>\varepsilon n\right)=0 . \tag{7.3}
\end{align*}
$$

Note that on the event $\mathfrak{B}_{n}, M_{\lambda}^{\mathbf{d}}$ is a subtree of $M^{\mathbf{d}}$, so it makes sense to measure the Hausdorff distance between them; moreover, this distance is bounded from above
by the maximum of the diameters of the trees $\mathfrak{T}_{i, \lambda}^{\mathbf{d}}, i=1, \ldots, k_{\lambda}^{\mathbf{d}}$. In the setting of the complete graph, results analogous to (7.2) and (7.3) were established in [6] and [5, Lemma 4.11] respectively. For inhomogeneous random graphs in the heavy-tailed regime, the analogue of (7.2) is proved in [25, Theorem 6.1].

The bound in (7.2) together with Lemma 4.5, Theorems 3.13, 3.14, and 4.6 would imply that (7.1) holds with respect to the GH topology. This could be strengthened to GHP convergence with the help of the following exchangeability result: For each $v \in V\left(\mathscr{C}_{1}^{\mathbf{d}}(\lambda)\right)$, let $d_{v, \lambda}^{(n)}$ denote the degree of $v$ in $\mathscr{C}_{1}^{\mathbf{d}}(\lambda)$, and define $d_{v, \lambda}^{\text {avail }}:=d_{v}^{(n)}-d_{v, \lambda}^{(n)}$. On the event $\mathfrak{B}_{n}$, for every $v \in V\left(M_{\lambda}^{\mathbf{d}}\right)$, let $\mathfrak{T}_{i, \lambda}^{\mathbf{d}}(v), 1 \leq i \leq r_{v, \lambda}$, be the trees among $\mathfrak{T}_{i, \lambda}^{\mathbf{d}}, 1 \leq i \leq k_{\lambda}^{\mathbf{d}}$, that are attached to $v$ in $M^{\mathbf{d}}$. For every $v \in V\left(M_{\lambda}^{\mathbf{d}}\right)$, append ( $d_{v, \lambda}^{\text {avail }}-r_{v, \lambda}$ ) many zeros to the sequence $\left(\left|\mathfrak{T}_{i, \lambda}^{\mathbf{d}}(v)\right|, 1 \leq i \leq r_{v, \lambda}\right)$ and let $\left(\alpha_{\mathbf{d}, \lambda}^{(i)}(v), 1 \leq i \leq d_{v, \lambda}^{\text {avail }}\right)$ be a uniform permutation of the resulting sequence; use independent permutations for different $v \in V\left(M_{\lambda}^{\mathbf{d}}\right)$ that are also independent of all the other random variables being considered. On $\mathfrak{B}_{n}^{c}$, set $\alpha_{\mathbf{d}, \lambda}^{(i)}(v):=0$ for $v \in V\left(\mathscr{C}_{1}^{\mathbf{d}}(\lambda)\right)$ and $i=1, \ldots, d_{v, \lambda}^{\text {avail }}$. Then conditional on $\mathscr{G}_{n, \mathbf{d}}(\lambda)$ and $M_{\lambda}^{\mathbf{d}}$, the family

$$
\begin{equation*}
\left(\alpha_{\mathbf{d}, \lambda}^{(i)}(v) ; 1 \leq i \leq d_{v, \lambda}^{\text {avail }}, v \in \mathscr{C}_{1}^{\mathbf{d}}(\lambda)\right) \text { of random variables is exchangeable. } \tag{7.4}
\end{equation*}
$$

The proof of (7.4) is similar to those of Lemma 6.17 (i) and Lemma 6.19. We outline the argument here for the readers' convenience. We can generate $\mathscr{G}_{n, \mathbf{d}}(\lambda), M_{\lambda}^{\mathbf{d}}$, and $\left(\alpha_{\mathbf{d}, \lambda}^{(i)}(v) ; 1 \leq i \leq d_{v, \lambda}^{\text {avail }}, v \in \mathscr{C}_{1}^{\mathbf{d}}(\lambda)\right)$ jointly as follows:
(a) Sample a $\operatorname{Binomial}\left(\sum_{v} d_{v}^{(n)} / 2,1-\mathfrak{p}_{\lambda}^{n}\right)$ random variable. For simplicity, we denote the realization by $m$.
(b) Consider the vertex set [ $n$ ] with $d_{v}^{(n)}$ many half-edges attached to the vertex $v$. Conditional on step (a), sample $\mathscr{Q}_{n, \mathbf{d}, m}^{(1)}$ as in Lemma 6.8. By (6.19),

$$
\begin{equation*}
\mathscr{D}_{n, \mathbf{d}, m}^{(1)} \stackrel{\mathrm{d}}{=} \mathscr{G}_{n, \mathbf{d}}(\lambda) . \tag{7.5}
\end{equation*}
$$

Thus, the largest component of $\mathscr{Q}_{n, \mathbf{d}, m}^{(1)}$, say $\mathscr{C}_{1}^{\bullet}$, has the same law as $\mathscr{C}_{1}^{\mathbf{d}}(\lambda)$. Let $d_{v}^{\bullet}$ be the degree of $v \in \mathscr{C}_{1}^{\bullet}$. Then each $v \in \mathscr{C}_{1}^{\bullet}$ has $d_{v}^{(n)}-d_{v}^{\bullet}$ many 'available' half-edges; we denote them by $f_{v, i}, 1 \leq i \leq d_{v}^{(n)}-d_{v}^{\bullet}, v \in \mathscr{C}_{1}^{\bullet}$.
(c) Assign i.i.d. Uniform[0, $\left.\mathfrak{p}_{\lambda}^{n}\right]$ weights to the edges of $\mathscr{Q}_{n, \mathbf{d}, m}^{(1)}$. Let $M_{1}^{\bullet}$ be the MST of $\mathscr{C}_{1}^{\bullet}$ constructed using these edge weights. Then

$$
\begin{equation*}
M_{1}^{\bullet} \stackrel{\mathrm{d}}{=} M_{\lambda}^{\mathrm{d}} \tag{7.6}
\end{equation*}
$$

jointly with the equality in distribution in (7.5).
(d) Conditional on steps (a), (b), and (c), sample $\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}$ as in Lemma 6.8, i.e., by uniformly pairing the previously unpaired half-edges attached to the vertices in $\mathscr{Q}_{n, \mathbf{d}, m}^{(1)}$. Then $\mathscr{Q}_{n, \mathbf{d}, m}^{(2)} \stackrel{\text { d }}{=} \mathscr{G}_{n, \mathbf{d}}$ jointly with (7.5) and (7.6). Let $\mathscr{E}_{\text {new }}:=E\left(\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}\right) \backslash$ $E\left(\mathscr{Q}_{n, \mathbf{d}, m}^{(1)}\right)$.
(e) If $\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}$ is not connected, go to the next step. If $\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}$ is connected, assign i.i.d. Uniform $\left[p_{\lambda}^{n}, 1\right]$ weights to the edges in $\mathscr{E}_{\text {new }}$. Let $w_{e}$ denote the weight assigned to the edge $e$. Delete all edges $e \in \mathscr{E}_{\text {new }}$ for which there exists a cycle $\pi$ (with no repeated edges) in $\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}$ such that $e$ is an edge in $\pi$ and $w_{e}$ is the maximum among all edge weights in $\pi$. This will yield a connected multigraph which we denote by $\mathscr{Q}$.
(f) If $\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}$ is not connected, define $Q^{(i)}(v)$ to be the empty graph for $1 \leq i \leq$ $d_{v}^{(n)}-d_{v}^{\bullet}, v \in \mathscr{C}_{1}^{\bullet}$. If $\mathscr{Q}_{n, \mathbf{d}, m}^{(2)}$ is connected, then note that $\mathscr{Q}$ as constructed in (e) is simply $\mathscr{C}_{1}^{\bullet}$ together with some connected multigraphs each of which is connected to a vertex of $\mathscr{C}_{1}^{\bullet}$ by a single edge; for $v \in \mathscr{C}_{1}^{\bullet}$ and $1 \leq i \leq d_{v}^{(n)}-d_{v}^{\bullet}$, set $Q^{(i)}(v)$ to be the connected multigraph that is connected to $v$ via the edge formed by pairing $f_{v, i}$ with another half edge in step (d), with the convention that $Q^{(i)}(v)$ is the empty graph if this edge was removed in step (e). Then

$$
\left(\operatorname{sort}\left(\left|Q^{(i)}(v)\right|, 1 \leq i \leq d_{v}^{(n)}-d_{v}^{\bullet}\right) ; v \in \mathscr{C}_{1}^{\mathbf{0}}\right) \stackrel{\mathrm{d}}{=}\left(\operatorname{sort}\left(\alpha_{\mathbf{d}, \lambda}^{(i)}(v), 1 \leq i \leq d_{v, \lambda}^{\text {avail }}\right) ; v \in \mathscr{C}_{1}^{\mathbf{d}}(\lambda)\right)
$$

jointly with the distributional equalities in (7.5) and (7.6), where sort(•) arranges the entries of a finite sequence in decreasing order.

Conditional on steps (a), (b), and (c) above, the rest of the procedure is symmetric with respect to the available half-edges attached to the vertices of $\mathscr{C}_{1}^{\bullet}$. Hence, conditional on $\mathscr{Q}_{n, \mathbf{d}, m}^{(1)}$ and $M_{1}^{\bullet}$, the family $\left(\left|Q^{(i)}(v)\right| ; 1 \leq i \leq d_{v}^{(n)}-d_{v}^{\bullet}, v \in \mathscr{C}_{1}^{\bullet}\right)$ is exchangeable. Thus, (7.4) follows.

The GH convergence in (7.1) can be lifted to GHP convergence using (7.3), (7.4), and arguments similar to the ones used in this paper. To carry out this argument, one would need to consider the metric measure space obtained by assigning mass $d_{v, \lambda}^{\text {avail }}$ to each $v \in V\left(\mathscr{C}_{1}^{\mathbf{d}}(\lambda)\right)$ and normalizing it to make it a probability measure. This does not quite fit into the framework of Theorem 3.13 (i) where the same function $f$ is used for all vertices. This was done to keep the statement of that theorem simple. However, the proof of Theorem 3.13 outlined in Sect. A. 2 goes through without any change for the measure being considered in this setting.
(b) MST scaling limit in the heavy-tailed regime: This regime seems more interesting. Consider scale-free random graphs on $n$ vertices where the tail of the empirical degree distribution $v_{n}$ asymptotically decays like $v_{n}([x, \infty)) \sim x^{1-\tau}$ for some $\tau \in(3,4)$. (In particular, the degree distribution asymptotically has infinite third moment and finite second moment.) It is predicted [31,32] that typical distance on the MST of such graphs scale like $n^{\frac{\tau-3}{\tau-1}}$. In this regime, the scaling limit at criticality was first established in [24] for inhomogeneous random graphs, and in [23] for random graphs with given degree sequences. The recent preprint [33] studies scaling limits of critical inhomogeneous random graphs in greater generality. The works [37,50] study scaling limits of critical random graphs with i.i.d. heavy-tailed degree sequences and alternate constructions of the limiting spaces. Very recently in [25], the scaling limit of the MST on the giant component in a supercritical inhomogeneous random graph with tail expoenent $\tau \in(3,4)$ has been established. Almost surely, the limiting space in [25] is compact, every point in this space either has degree one (leaf), or two, or
infinity (hub), both the set of leaves and the set of hubs are dense in this space, and the Minkowski dimension of this space equals $(\tau-1) /(\tau-3)$. We expect this space to be the candidate for the scaling limit of the MST of a wide array of heavy-tailed random graphs under some general assumptions.

Acknowledgements The authors thank two anonymous referees whose careful reading and detailed comments led to significant improvements in the paper. LAB was supported in part by an NSERC Discovery Grant and Discovery Accelerator Supplement, and by an FRQNT Team Grant, during the preparation of this research. SS was partially supported by the Infosys Foundation, Bangalore and by MATRICS grant MTR/2019/000745 from SERB.

## Appendix A

Our aim in this section is to briefly describe the ideas needed to prove Theorems 3.11 and 3.13 .

## A. 1 Sketch of proof of Theorem 3.11

Suppose $\mathbf{t}$ is a rooted tree with vertices labeled by $[m]$ and let $R(\cdot, \mathbf{t})$ be as in (6.59). For $s \geq 1$, let

$$
\begin{aligned}
A_{s}(\mathbf{t}):=\left\{\left(v_{1}, u_{1}, \ldots, v_{s}, u_{s}\right):\right. & 1 \leq v_{1} \leq \ldots \leq v_{s} \leq m, u_{i} \in R\left(v_{i}, \mathbf{t}\right) \\
& \text { if } \left.i<j \text { and } v_{i}=v_{j} \text { then } u_{i}<u_{j}\right\}
\end{aligned}
$$

Note that $s!\times\left|A_{s}(\mathbf{t})\right| \leq\left|A_{1}(\mathbf{t})\right|^{s}$. Let $T_{m}$ denote a uniform rooted labeled tree on $[m]$, and let $\bar{T}_{m}$ be distributed as

$$
\mathbb{P}\left(\bar{T}_{m}=\mathbf{t}\right)=\frac{\mathbb{P}\left(T_{m}=\mathbf{t}\right) \cdot\left|A_{s}(\mathbf{t})\right|}{\mathbb{E}\left(\left|A_{s}\left(T_{m}\right)\right|\right)}
$$

Then we have the following decomposition of $\mathscr{H}_{m, s}$ :
Theorem A. 1 Fix $s \geq$ 1. Sample $\bar{T}_{m}$, and conditional on the realization, sample $\left(\bar{v}_{1, m}, \bar{u}_{1, m}, \ldots, \bar{v}_{s, m}, \bar{u}_{s, m}\right)$ from $A_{s}\left(\bar{T}_{m}\right)$ uniformly. Place an edge between $\bar{v}_{i, m}$ and $\bar{u}_{i, m}$ for $1 \leq i \leq s$, and then forget about the root of $\bar{T}_{m}$. Call the resulting graph $\overline{\mathscr{H}}_{m, s}$. Then $\overline{\mathscr{H}}_{m, s} \stackrel{d}{=} \mathscr{H}_{m, s}$.

This can be seen as follows: Consider a simple, connected, rooted graph $G$ on $[m]$ with $\operatorname{sp}(G)=s$. Let $\mathbf{t}$ be the tree constructed by following a depth-first exploration of $G$ starting at its root, and let $v_{i}, u_{i}, 1 \leq i \leq s$, be the endpoints of the $s$ edges that need to be added to $\mathbf{t}$ to recover $G$. We can arrange $v_{1}, u_{1}, \ldots, v_{s}, u_{s}$ in a unique way so that the resulting sequence becomes an element of $A_{s}(\mathbf{t})$. It thus follows that the set of simple, connected, rooted graphs on $[m]$ having $s$ surplus edges is in bijective
correspondence with the set

$$
\begin{equation*}
\left\{\left(\mathbf{t}, v_{1}, u_{1}, \ldots, v_{s}, u_{s}\right): \mathbf{t} \text { rooted tree on }[m],\left(v_{1}, u_{1}, \ldots, v_{s}, u_{s}\right) \in A_{s}(\mathbf{t})\right\} . \tag{A.1}
\end{equation*}
$$

Then one can show that if we root $\mathscr{H}_{m, s}$ at a uniform vertex, then its corresponding element in the set (A.1) will be distributed as $\left(\bar{T}_{m}, \bar{v}_{1, m}, \bar{u}_{1, m}, \ldots, \bar{v}_{s, m}, \bar{u}_{s, m}\right)$. We omit the details as similar ideas have already been used in [3,26,27].

For any tree $\mathbf{t}$ on $[m]$ rooted at $\rho$, endow the children of each vertex in $\mathbf{t}$ with the linear order induced by their labels. Let $\rho=w_{0}, w_{1}, \ldots, w_{m-1}$ be the vertices of $\mathbf{t}$ in order of appearance in a depth-first exploration of $\mathbf{t}$ using the above order. Let $\mathrm{Ht}_{\mathbf{t}}:[0, m] \rightarrow \mathbb{R}$ be the height function of $\mathbf{t}$ given by $\mathrm{Ht}_{\mathbf{t}}(m)=0$, and

$$
\operatorname{Ht}_{\mathbf{t}}(x)=\operatorname{ht}\left(w_{\lfloor x\rfloor}, \mathbf{t}\right), \quad x \in[0, m)
$$

The following lemma is a collection of some standard results about $T_{m}$ :
Lemma A. 2 (i) The following convergences hold:

$$
\begin{align*}
& m^{-1 / 2} \mathrm{Ht}_{T_{m}}(m \cdot) \xrightarrow{d} 2 \mathbf{e}(\cdot), \quad \text { and } \\
& m^{-1 / 2} \max _{v \in[m]}|2| R\left(v, T_{m}\right)\left|-\operatorname{ht}\left(v, T_{m}\right)\right| \xrightarrow{P} 0 \tag{A.2}
\end{align*}
$$

where the convergence in (A.2) is w.r.t. the Skorohod $J_{1}$ topology.
(ii) For all $m \geq 1, \mathbb{P}\left(\operatorname{ht}\left(T_{m}\right) \geq x \sqrt{m}\right) \leq c x^{3} \exp \left(-x^{2} / 2\right)$.
(iii) For all $x \geq 0$ and $m \geq 1$,

$$
\mathbb{P}\left(\max _{v \in[m]}\left|R\left(v, T_{m}\right)\right| \geq x \sqrt{m}\right) \leq c \exp \left(-c^{\prime} x^{2}\right)
$$

Using the bounds $\left|A_{s}\left(T_{m}\right)\right| \times s!\leq\left|A_{1}\left(T_{m}\right)\right|^{s}$ and $\left|A_{1}\left(T_{m}\right)\right| \leq m \cdot \max _{v \in[m]}\left|R\left(v, T_{m}\right)\right|$, we further have

$$
\mathbb{P}\left(\left|A_{s}\left(T_{m}\right)\right| \geq x m^{3 s / 2}\right) \leq c \exp \left(-c^{\prime} x^{2 / s}\right)
$$

for any $s \geq 1, x \geq 0$, and $m \geq 1$.
(iv) For any $s \geq 1, m^{-3 s / 2}\left(\left|A_{1}\left(T_{m}\right)\right|^{s}-\left|A_{s}\left(T_{m}\right)\right| \times s!\right) \xrightarrow{P} 0$.

Lemma A.2((i)) follows from [69]. ((ii)) follows from [66, Corollary 1]. ((iii)) is the content of [3, Lemma 13]. The proof of ((iv)) is similar to that of [26, Lemma 7.3 (iii)].

Sketch of proof of (3.9) In view of Theorem A.1, $s \cdot\left(\mathrm{ht}\left(\bar{T}_{m}\right)+1\right)$ dominates $L\left(\mathscr{H}_{m, s}\right)$ stochastically for any $s \geq 1$. Thus, (3.9) follows from Lemma A. 2 ((ii)) and ((iii)).

To prove the other assertions in Theorem 3.11 it will be convenient to work with two slightly different spaces $\mathscr{H}_{m, s}^{\circ}$ and $\mathscr{H}_{m, s}^{\dagger}$ which we define next. Recall the notation
$R(\cdot, \cdot, \cdot)$ from (6.58). Sample $T_{m}^{\circ}$ according to distribution

$$
\begin{equation*}
\mathbb{P}\left(T_{m}^{\circ}=\mathbf{t}\right)=\frac{\mathbb{P}\left(T_{m}=\mathbf{t}\right) \cdot\left|A_{1}(\mathbf{t})\right|^{s}}{\mathbb{E}\left[\left|A_{1}\left(T_{m}\right)\right|^{s}\right]}, \quad \mathbf{t} \text { rooted tree on }[m] . \tag{A.3}
\end{equation*}
$$

Conditional on $T_{m}^{\circ}$, sample an i.i.d. sequence of triples $\left(v_{i, m}^{\circ}, u_{i, m}^{\circ}, f_{i, m}^{\circ}\right), 1 \leq i \leq s$, where

$$
\begin{aligned}
& \mathbb{P}\left(v_{i, m}^{\circ}=v \mid T_{m}^{\circ}\right)=\left|R\left(v, T_{m}^{\circ}\right)\right| /\left|A_{1}\left(T_{m}^{\circ}\right)\right|, \quad v \in[m], \\
& \left.\mathbb{P}\left(u_{i, m}^{\circ}=u \mid T_{m}^{\circ}, v_{i, m}^{\circ}\right)=\left|R\left(u, v_{i, m}^{\circ}, T_{m}^{\circ}\right)\right| /\left|R\left(v, T_{m}^{\circ}\right)\right|, \quad u \in\left\{\overleftarrow{\left(\overleftarrow{v i}_{\circ}^{\circ}(k)\right.}\right): 1 \leq k \leq \operatorname{ht}\left(v_{i, m}^{\circ}\right)\right\}, \quad \text { and } \\
& \mathbb{P}\left(f_{i, m}^{\circ}=f \mid T_{m}^{\circ}, v_{i, m}^{\circ}, u_{i, m}^{\circ}\right)=1 /\left|R\left(u_{i, m}^{\circ}, v_{i, m}^{\circ}, T_{m}^{\circ}\right)\right|, \quad f \in R\left(u_{i, m}^{\circ}, v_{i, m}^{\circ}, T_{m}^{\circ}\right) .
\end{aligned}
$$

Let $\mathscr{H}_{m, s}^{\dagger}$ (resp. $\mathscr{H}_{m, s}^{\circ}$ ) be the space obtained by adding an edge between $v_{i, m}^{\circ}$ and $f_{i, m}^{\circ}$ (resp. between $v_{i, m}^{\circ}$ and $u_{i, m}^{\circ}$ ) for $1 \leq i \leq s$, and then forgetting about the root of $T_{m}^{\circ}$. It follows from Lemma A. 2 ((iii)) and ((iv)) that the total variation distance between the laws of $\overline{\mathscr{H}}_{m, s}$ (as defined in Theorem A.1) and $\mathscr{H}_{m, s}^{\dagger}$ tends to zero as $m \rightarrow \infty$. It thus follows from Theorem A. 1 that there exists a coupling of $\mathscr{H}_{m, s}$ and $\mathscr{H}_{m, s}^{\dagger}$ such that

$$
\begin{equation*}
\mathbb{P}\left(\mathscr{H}_{m, s} \neq \mathscr{H}_{m, s}^{\dagger}\right) \rightarrow 0, \quad \text { as } \quad m \rightarrow \infty \tag{A.4}
\end{equation*}
$$

We will now recall an alternate construction of $\mathscr{H}^{(s)}$ which is essentially given in [3]; see also the discussion below [2, Equation (1)]. We first introduce some notation. For any $f:[0,1] \rightarrow \mathbb{R}, x \in[0,1]$, and $h>0$, let

$$
\operatorname{prev}(x, h ; f)=\sup \{y \in[0, x): f(y)=h\}, \quad \text { and } \quad \operatorname{next}(x, h ; f)=\inf \{y \in(x, 1]: f(y)<h\},
$$

where $\sup \}=-\infty$ and $\inf \}=\infty$ by convention.
Construction A. 3 (Alternate construction of $\mathscr{H}^{(s)}$ ) Fix an integer $s \geq 2$.
(a) Sample $\mathbf{e}^{\circ}$ with law given by

$$
\mathbb{E}\left[f\left(\mathbf{e}^{\circ}\right)\right]=\frac{\mathbb{E}\left[f(\mathbf{e})\left(\int_{0}^{1} \mathbf{e}(t) d t\right)^{s}\right]}{\mathbb{E}\left[\left(\int_{0}^{1} \mathbf{e}(t) d t\right)^{s}\right]}
$$

(b) Conditional on $\mathbf{e}^{\circ}$, sample i.i.d. points $y_{1}^{\circ}, \ldots, y_{s}^{\circ}$ having density $\mathbf{e}^{\circ}(y) / \int_{0}^{1} \mathbf{e}^{\circ}(t) d t$.
(c) Conditional on the above, sample $h_{1}^{\circ}, \ldots, h_{s}^{\circ}$ independently, where $h_{i}^{\circ} \sim$ $\operatorname{Unif}\left[0, \mathbf{e}^{\circ}\left(y_{i}^{\circ}\right)\right] . \operatorname{Set} x_{i}^{\circ}=\operatorname{prev}\left(y_{i}^{\circ}, h_{i}^{\circ} ; \mathbf{e}^{\circ}\right)$.
(d) Form the quotient space $\mathscr{T}^{\circ} / \sim$, where $\sim$ is the equivalence relation under which $q_{\mathrm{e}^{\circ}}\left(x_{i}^{\circ}\right) \sim q_{\mathrm{e}^{\circ}}\left(y_{i}^{\circ}\right), 1 \leq i \leq s$.

Then $\mathscr{H}^{(s)} \stackrel{d}{=} 2 \cdot\left(\mathscr{T}^{\circ} / \sim\right)$.

Now observe that $\mathscr{H}_{m, s}^{\circ}$ has a similar alternate construction: First sample $T_{m}^{\circ}$ as in (A.3). Let $w_{0}, \ldots, w_{m-1}$ be the vertices of $T_{m}^{\circ}$ in order of appearance in a depth-first exploration of $T_{m}^{\circ}$. Let $\mathrm{Ht}^{\circ}$ be the height function of $T_{m}^{\circ}$. Conditional on $T_{m}^{\circ}$, sample i.i.d. random variables $y_{1, m}^{\circ}, \ldots, y_{s, m}^{\circ}$, where

$$
\mathbb{P}\left(y_{i, m}^{\circ}=j \mid T_{m}^{\circ}\right)=\left|R\left(w_{j}, T_{m}^{\circ}\right)\right| /\left|A_{1}\left(T_{m}^{\circ}\right)\right|, \quad 1 \leq j \leq m-1 .
$$

Conditional on the above, sample $h_{1, m}^{\circ}, \ldots, h_{s, m}^{\circ}$ independently via

$$
\begin{aligned}
& \mathbb{P}\left(h_{i, m}^{\circ}=\mathrm{Ht}^{\circ}\left(y_{i, m}^{\circ}\right)-k \mid T_{m}^{\circ}, y_{1, m}^{\circ}, \ldots, y_{s, m}^{\circ}\right) \\
& \quad=\frac{\left|R\left(\overleftarrow{v}(k), v, T_{m}^{\circ}\right)\right|}{\left|R\left(v, T_{m}^{\circ}\right)\right|}, \quad 1 \leq k \leq \operatorname{Ht}^{\circ}\left(y_{i, m}^{\circ}\right)
\end{aligned}
$$

where $v=w_{y_{i, m}^{\circ}}$. Let $x_{i, m}^{\circ}=\operatorname{prev}\left(y_{i, m}^{\circ}, h_{i, m}^{\circ} ; \mathrm{Ht}^{\circ}\right)-1$. Then $\mathscr{H}_{m, s}^{\circ}$ has the same distribution as the space obtained by placing an edge in $T_{m}^{\circ}$ between $w_{y_{i, m}^{\circ}}$ and $w_{x_{i, m}^{\circ}}$ for $1 \leq i \leq s$.
Sketch of proof of (3.8) Using Lemma A. 2 ((i)) and ((iii)), it can be shown that the following convergences hold jointly:

$$
\begin{equation*}
\frac{1}{\sqrt{m}} \mathrm{Ht}^{\circ}(m \cdot) \xrightarrow{\mathrm{d}} 2 \mathbf{e}^{\circ}(\cdot), \quad \text { and }\left(\frac{x_{i, m}^{\circ}}{m}, \frac{y_{i, m}^{\circ}}{m}, \frac{h_{i, m}^{\circ}}{\sqrt{m}}\right) \xrightarrow{\mathrm{d}}\left(x_{i}^{\circ}, y_{i}^{\circ}, 2 h_{i}^{\circ}\right), \quad 1 \leq i \leq s \tag{A.5}
\end{equation*}
$$

as $m \rightarrow \infty$. Using Construction A. 3 and the above alternate construction of $\mathscr{H}_{m, s}^{\circ}$, it is now routine to prove the assertion in (3.8) for $\mathscr{H}_{m, s}^{\circ}$, from which it follows that the same is true for $\mathscr{H}_{m, s}^{\dagger}$. The desired result now follows from (A.4).

Let $y_{(i), m}^{\circ}\left(\right.$ resp. $\left.y_{(i)}^{\circ}\right), 1 \leq i \leq s$, be $y_{i, m}^{\circ}\left(\right.$ resp. $\left.y_{i}^{\circ}\right), 1 \leq i \leq s$, arranged in an increasing order. For $1 \leq i \leq s-1$ define $z_{i, m}^{\circ}$ and $z_{i}^{\circ}$ via

$$
\begin{aligned}
& z_{i, m}^{\circ}=\min \left\{t \in\left[y_{(i), m}^{\circ}, y_{(i+1), m}^{\circ}\right]: \operatorname{Ht}^{\circ}(t)=\min \left\{\operatorname{Ht}^{\circ}(a): y_{(i), m}^{\circ} \leq a \leq y_{(i+1), m}^{\circ}\right\}\right\}, \quad \text { and } \\
& \mathbf{e}^{\circ}\left(z_{i}^{\circ}\right)=\inf \left\{\mathbf{e}^{\circ}(t): y_{(i)}^{\circ} \leq t \leq y_{(i+1)}^{\circ}\right\} .
\end{aligned}
$$

Further, define

$$
\begin{aligned}
& x_{i, m}^{\circ,+}=\operatorname{next}\left(x_{i, m}^{\circ}, h_{i, m}^{\circ}+1 ; \mathrm{Ht}^{\circ}\right), \quad x_{i}^{\circ,+}=\operatorname{next}\left(x_{i}^{\circ}, h_{i}^{\circ} ; \mathbf{e}^{\circ}\right), \quad 1 \leq i \leq s, \\
& z_{i, m}^{\circ,-}=\operatorname{prev}\left(z_{i, m}^{\circ}, \operatorname{Ht}^{\circ}\left(z_{i, m}^{\circ}\right)-1 ; \mathrm{Ht}^{\circ}\right)-1, z_{i}^{\circ,-}=\operatorname{prev}\left(z_{i}^{\circ}, \mathbf{e}^{\circ}\left(z_{i}^{\circ}\right) ; \mathbf{e}^{\circ}\right), \quad 1 \leq i \leq s-1, \\
& z_{i, m}^{\circ,+}=\operatorname{next}\left(z_{i, m}^{\circ}, \mathrm{Ht}^{\circ}\left(z_{i, m}^{\circ}\right) ; \mathrm{Ht}^{\circ}\right)-1, \quad z_{i}^{\alpha^{\circ}++}=\operatorname{next}\left(z_{i}^{\circ}, \mathbf{e}^{\circ}\left(z_{i}^{\circ}\right) ; \mathbf{e}^{\circ}\right), \quad 1 \leq i \leq s-1 .
\end{aligned}
$$

Sketch of proof of (3.11) From (A.5) it follows that the following convergence holds jointly with the convergence in (A.5): As $m \rightarrow \infty$,

$$
\begin{equation*}
\frac{x_{i, m}^{\circ,+}}{m} \xrightarrow{\mathrm{~d}} x_{i}^{\circ,+}, 1 \leq i \leq s, \text { and } \frac{1}{m}\left(z_{i, m}^{\circ}, z_{i, m}^{\circ,-}, z_{i, m}^{\circ,+}\right) \xrightarrow{\mathrm{d}}\left(z_{i}^{\circ}, z_{i}^{\circ,-}, z_{i}^{\circ,+}\right), 1 \leq i \leq s-1 \tag{A.6}
\end{equation*}
$$

Arrange $x_{i}^{\circ}, x_{i}^{\circ,+}, y_{i}^{\circ}, 1 \leq i \leq s$, and $z_{i}^{\circ}, z_{i}^{\circ,-}, z_{i}^{\circ,+}, 1 \leq i \leq s-1$, (resp. $x_{i, m}^{\circ}, x_{i, m}^{\circ,+}, y_{i, m}^{\circ}, 1 \leq i \leq s$, and $\left.z_{i, m}^{\circ}, z_{i, m}^{\circ,-}, z_{i, m}^{\circ,+}, 1 \leq i \leq s-1\right)$ in increasing order as $a_{1}, \ldots, a_{6 s-3}$ (resp. as $a_{1, m}, \ldots, a_{6 s-3, m}$ ). Let

$$
\Delta_{j}=a_{j+1}-a_{j}, \quad \text { and } \quad \Delta_{j, m}=a_{j+1, m}-a_{j, m}, \quad 1 \leq j \leq 6 s-4
$$

Then it follows from (A.6) and the second convergence in (A.5) that

$$
\begin{equation*}
\left(\Delta_{j, m}, 1 \leq j \leq 6 s-4\right) \xrightarrow{\mathrm{d}}\left(\Delta_{j}, 1 \leq j \leq 6 s-4\right), \quad \text { as } \quad m \rightarrow \infty \tag{A.7}
\end{equation*}
$$

jointly with (A.5) and (A.6).
Recall the notation used in (3.11), and note that there exists a partition $\mathscr{P}=$ $\left\{\mathscr{P}_{1}, \ldots, \mathscr{P}_{r}\right\}$ of $[6 s-4]$ that depends only on the realizations of $\mathbf{e}^{\circ}$ and $x_{i}^{\circ}, y_{i}^{\circ}$, $1 \leq i \leq s$, such that

$$
\begin{equation*}
\left(\mu^{(s)}\left(\mathscr{T}_{i}^{\prime}\right), 1 \leq i \leq r\right) \stackrel{\mathrm{d}}{=}\left(\sum_{j \in \mathscr{P}_{i}} \Delta_{j}, 1 \leq i \leq r\right) \tag{A.8}
\end{equation*}
$$

Further, it follows from (A.5) that for large $m$, the vector consisting of the numbers of vertices in $\mathscr{H}_{m, s}^{\circ}$ that are connected to the different elements of $e\left(\mathscr{H}_{m, s}^{\circ}\right)$ is given by $\left(\sum_{j \in \mathscr{P}_{i}} \Delta_{j, m}, 1 \leq i \leq r\right)$, where the common endpoints of multiple $e \in e\left(\mathscr{H}_{m, s}^{\circ}\right)$ and the vertices in their pendant subtrees have been accounted for in $\sum_{j \in \mathscr{P}_{i}} \Delta_{j, m}$ for exactly one value of $i$ in a specific way. Using (A.7) and (A.8), we get the analogue of (3.11) for $\mathscr{H}_{m, s}^{\circ}$ for the above specific way of assigning the common endpoints of multiple $e \in e\left(\mathscr{H}_{m, s}^{\circ}\right)$ and the vertices in their pendant subtrees to the different terms $\sum_{j \in \mathscr{P}_{i}} \Delta_{j, m}$.

This together with (A.4) would complete the proof if we can show that the sizes of the pendant subtrees of the common endpoints of multiple $e \in e\left(\mathscr{H}_{m, s}^{\circ}\right)$ are asymptotically negligible. This negligibility claim follows from the following facts:
(A) $Y_{i, m}=o_{P}(m), 1 \leq i \leq s$, where $Y_{i, m}$ denotes the number of descendants of $v_{i, m}^{\circ}$ in $T_{m}^{\circ}$.
(B) $X_{i, m}=o_{P}(m), 1 \leq i \leq s$, where $X_{i, m}$ denotes the number of descendants of $u_{i, m}^{\circ}$ in $T_{m}^{\circ}$ that are not in the subtree that contains $v_{i, m}^{\circ}$.
(C) For every $\varepsilon>0$,
$\mathbb{P}\left(\exists v \in T_{m}^{\circ}: v\right.$ has at least three subtrees in $T_{m}^{\circ}$ each of size $\left.\geq \varepsilon m\right) \rightarrow 0$ as $m \rightarrow \infty$.
(A) and (C) follow from (A.5) and the facts that $q_{\mathrm{e}^{\circ}}\left(y_{i}^{\circ}\right)$ is almost surely a leaf in $\mathscr{T}_{\mathrm{e}}{ }^{\circ}$ and that $\mathscr{T}_{\mathrm{e}^{\circ}}$ is almost surely binary. The proof of $(\mathrm{B})$ is also routine.

## A. 2 Sketch of proof of Theorem 3.13

Assume that for each $m \geq 1, \boldsymbol{k}^{(m)}=\left(k_{i}^{(m)}, i \geq 0\right)$, where $k_{i}^{(m)}$ are nonnegative integers satisfying $\sum_{i \geq 0} k_{i}^{(m)}=m$ and $\sum_{i \geq 0} i k_{i}^{(m)}=m-1$. Then there exist trees on $m$ vertices
in which for each $i \geq 0$, there are exactly $k_{i}^{(m)}$ many vertices with $i$ many children. We call $\boldsymbol{k}^{(m)}$ the child sequence of such a tree. Assumption 3.4 gives the criterion for graphs with given degree sequences to be critical. The following assumption gives the analogous criterion for plane trees with given child sequences.

Assumption A. 4 There exists a pmf ( $p_{0}, p_{1}, \ldots$ ) with

$$
p_{0}>0, \quad \sum_{i \geq 1} i p_{i}=1, \quad \text { and } \sum_{i \geq 1} i^{2} p_{i}<\infty
$$

such that

$$
\frac{k_{i}^{(m)}}{m} \rightarrow p_{i} \text { for } i \geq 0, \text { and } \frac{1}{m} \sum_{i \geq 0} i^{2} k_{i}^{(m)} \rightarrow \sum_{i \geq 1} i^{2} p_{i}
$$

We will write $\sigma^{2}=\sum_{i} i^{2} p_{i}-1$ for the variance associated with the $\operatorname{pmf}\left(p_{0}, p_{1}, \ldots\right)$.
Let $\mathbb{T}_{\boldsymbol{k}^{(m)}}$ be the set of plane trees with child sequence $\boldsymbol{k}^{(m)}$. Let $\mathscr{T}_{\boldsymbol{k}^{(m)}}$ be a uniform element of $\mathbb{T}_{\boldsymbol{k}^{(m)}}$ endowed with the tree distance and the uniform probability measure on $m$ vertices and viewed as a metric measure space. Broutin and Marckert [34] showed that under Assumption A.4, $\sigma m^{-1 / 2} \mathscr{T}_{\boldsymbol{k}^{(m)}} \xrightarrow{\text { d }} \mathscr{T}_{2}$ e w.r.t. GHP topology. The following variant of this result follows from [26, Lemma 7.4 and Lemma 7.6]:

Lemma A. 5 Suppose $\boldsymbol{k}^{(m)}$ satisfies Assumption A.4. Further, suppose $f_{m}:\{0,1, \ldots\} \rightarrow$ $[0,1]$ is such that

$$
\sum_{i \geq 0} k_{i}^{(m)} f_{m}(i)=1, \quad \text { and } \quad \lim _{m \rightarrow \infty} \max _{i: k_{i}^{(m)}>0} f_{m}(i)=0
$$

Let $\mathscr{T}_{\boldsymbol{k}^{(m)}}^{f_{m}}$ be a uniform element of $\mathbb{T}_{\boldsymbol{k}^{(m)}}$ endowed with the tree distance and the measure that assigns probability $f_{m}(i)$ to any node that has $i$ children, $i \geq 0$. Then

$$
\sigma m^{-1 / 2} \cdot \mathscr{T}_{k^{(m)}}^{f_{m}} \xrightarrow{d} \mathscr{T}_{2} \mathbf{e} \quad \text { w.r.t. GHP topology. }
$$

Now we can prove Theorem 3.13 using the above lemma and the techniques used in the proof of [26, Theorem 2.2].

## References

1. Abraham, R., Delmas, J.-F., Hoscheit, P.: A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. Electron. J. Probab. 18(14), 1-21 (2013)
2. Addario-Berry, L., Broutin, N., Goldschmidt, C.: Critical random graphs: limiting constructions and distributional properties. Electron. J. Probab. 15(25), 741-775 (2010)
3. Addario-Berry, L., Broutin, N., Goldschmidt, C.: The continuum limit of critical random graphs. Probab. Theory Relat. Fields 152(3-4), 367-406 (2012)
4. Addario-Berry, L., Bhamidi, S., Sen, S.: A probabilistic approach to the leader problem in random graphs. arXiv preprint arXiv:1703.09908 (2017)
5. Addario-Berry, L., Broutin, N., Goldschmidt, C., Miermont, G.: The scaling limit of the minimum spanning tree of the complete graph. Ann. Probab. 45(5), 3075-3144 (2017)
6. Addario-Berry, L., Broutin, N., Reed, B.: Critical random graphs and the structure of a minimum spanning tree. Random Struct. Algorithms 35(3), 323-347 (2009)
7. Aizenman, M., Burchard, A., Newman, C. M., Wilson, D. B.: Scaling limits for minimal and random spanning trees in two dimensions. Random Struct. Algorithms 15(3-4), 319-367 (1999). Statistical physics methods in discrete probability, combinatorics, and theoretical computer science (Princeton, NJ, 1997)
8. Aldous, D.: The continuum random tree. I. Ann. Probab. 19, 1-28 (1991)
9. Aldous, D.: The continuum random tree. III. Ann. Probab. 21, 248-289 (1993)
10. Aldous, D.: A random tree model associated with random graphs. Random Struct. Algorithms 1(4), 383-402 (1990)
11. Aldous, D.: Brownian excursions, critical random graphs and the multiplicative coalescent. Ann. Probab. 25(2), 812-854 (1997)
12. Aldous, D., Steele, J.M.: Asymptotics for Euclidean minimal spanning trees on random points. Probab. Theory Relat. Fields 92(2), 247-258 (1992)
13. Alexander, K.S.: Rates of convergence of means for distance-minimizing subadditive Euclidean functionals. Ann. Appl. Probab. 4(3), 902-922 (1994)
14. Alexander, K.S.: Percolation and minimal spanning forests in infinite graphs. Ann. Probab. 23(1), 87-104 (1995)
15. Alexander, K.S.: The RSW theorem for continuum percolation and the CLT for Euclidean minimal spanning trees. Ann. Appl. Probab. 6(2), 466-494 (1996)
16. Alexander, K.S., Molchanov, S.A.: Percolation of level sets for two-dimensional random fields with lattice symmetry. J. Statist. Phys. 77(3-4), 627-643 (1994)
17. Avram, F., Bertsimas, D.: The minimum spanning tree constant in geometrical probability and under the independent model: a unified approach. Ann. Appl. Probab. 2(1), 113-130 (1992)
18. Bai, Z.D., Lee, S., Penrose, M.D.: Rooted edges of a minimal directed spanning tree on random points. Adv. in Appl. Probab. 38(1), 1-30 (2006)
19. Beardwood, J., Halton, J.H., Hammersley, J.M.: The shortest path through many points. Proc. Cambr. Philos. Soc. 55, 299-327 (1959)
20. Bender, E.A., Canfield, E.R.: The asymptotic number of labeled graphs with given degree sequences. J. Comb. Theory Ser. A 24(3), 296-307 (1978)
21. Beveridge, A., Frieze, A., McDiarmid, C.: Random minimum length spanning trees in regular graphs. Combinatorica 18(3), 311-333 (1998)
22. Bhamidi, S., Broutin, N., Sen, S., Wang, X.: Scaling limits of random graph models at criticality: universality and the basin of attraction of the Erdös-Rényi random graph. arXiv preprint arXiv:1411.3417 (2014)
23. Bhamidi, S., Dhara, S., Hofstad, R. v. d., Sen, S.: Universality for critical heavy-tailed network models: metric structure of maximal components, arXiv preprint arXiv:1703.07145 (2017)
24. Bhamidi, S., van der Hofstad, R., Sen, S.: The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. Probab. Theory Relat. Fields 170(1-2), 387-474 (2018)
25. Bhamidi, S., Sen, S.: Geometry of the minimal spanning tree in the heavy-tailed regime: new universality classes. arXiv preprint arXiv:2009.10696 (2020)
26. Bhamidi, S., Sen, S.: Geometry of the vacant set left by random walk on random graphs, Wright's constants, and critical random graphs with prescribed degrees. Random Struct. Algorithms 56(3), 676-721 (2020)
27. Bhamidi, S., Sen, S., Wang, X.: Continuum limit of critical inhomogeneous random graphs. Probab. Theory Relat. Fields $\mathbf{1 6 9}(1-2)$, 565-641 (2017)
28. Bhatt, A.G., Roy, R.: On a random directed spanning tree. Adv. Appl. Probab. 36(1), 19-42 (2004)
29. Bollobás, B.: Random Graphs, Cambridge Studies in Advanced Mathematics, vol. 73, 2nd edn. Cambridge University Press, Cambridge (2001)
30. Bollobás, B., Janson, S., Riordan, O.: The phase transition in inhomogeneous random graphs. Random Struct. Algorithms 31(1), 3-122 (2007)
31. Braunstein, L.A., Buldyrev, S.V., Cohen, R., Havlin, S., Stanley, H.E.: Optimal paths in disordered complex networks. Phys. Rev. Lett. 91(16), 168701 (2003)
32. Braunstein, L.A., Wu, Z., Chen, Y., Buldyrev, S.V., Kalisky, T., Sreenivasan, S., Cohen, R., Lopez, E., Havlin, S., Stanley, H.E.: Optimal path and minimal spanning trees in random weighted networks. Int. J. Bifurc. Chaos 17(07), 2215-2255 (2007)
33. Broutin, N., Duquesne, T., Wang, M.: Limits of multiplicative inhomogeneous random graphs and Lévy trees. arXiv preprint arXiv:1804.05871 (2018)
34. Broutin, N., Marckert, J.-F.: Asymptotics of trees with a prescribed degree sequence and applications. Random Struct. Algorithms 44(3), 290-316 (2014)
35. Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry, Graduate Studies in Mathematics, vol. 33. American Mathematical Society, Providence (2001)
36. Chatterjee, S., Sen, S.: Minimal spanning trees and Stein's method. Ann. Appl. Probab. 27(3), 15881645 (2017)
37. Conchon-Kerjan, G., Goldschmidt, C.: The stable graph: the metric space scaling limit of a critical randomgraph with IID power-law degrees. arXiv preprint arXiv:2002.04954 (2020)
38. Dembo, A., Levit, A., Vadlamani, S.: Component sizes for large quantum Erdös-Rényi graph near criticality. Ann. Probab. 47(2), 1185-1219 (2019)
39. Dhara, S., van der Hofstad, R., van Leeuwaarden, J. S. H., Sen, S.: Critical window for the configuration model: finite third moment degrees. Electron. J. Probab. 22, Paper No. 16, 33 (2017)
40. Dhara, S., van derHofstad, R., van Leeuwaarden, J.S.H., Sen, S.: Heavy-tailed configuration models at criticality. Ann. Inst. Henri Poincaré Probab. Stat. 56(3), 1515-1558 (2020)
41. Duquesne, T.: The coding of compact real trees by real valued functions. arXiv preprint math (2006)
42. Evans, S.N.: Probability and Real Trees. Lecture Notes in Mathematics, vol. 1920. Springer, Berlin (2008)
43. Federico, L., van derHofstad, R.: Critical window for connectivity in the configuration model. Combin. Probab. Comput. 1-21 (2017)
44. Fountoulakis, N.: Percolation on sparse random graphs with given degree sequence. Internet Math. 4(4), 329-356 (2007)
45. Frieze, A., McDiarmid, C.: Algorithmic theory of random graphs. Random Struct. Algorithms 10(1-2), 5-42 (1997)
46. Frieze, A., Ruszinkó, M., Thoma, L.: A note on random minimum length spanning trees. Electron. J. Combin. 7(1), R41 (2000)
47. Frieze, A.M.: On the value of a random minimum spanning tree problem. Discrete Appl. Math. 10(1), 47-56 (1985)
48. Frieze, A.M., McDiarmid, C.J.: On random minimum length spanning trees. Combinatorica 9(4), 363-374 (1989)
49. Garban, C., Pete, G., Schramm, O.: The scaling limits of the minimal spanning tree and invasion percolation in the plane. Ann. Probab. 46(6), 3501-3557 (2018)
50. Goldschmidt, C., Haas, B., Sénizergues, D.: Stable graphs: distributions and line-breaking construction. arXiv preprint arXiv:1811.06940 (2018)
51. Heydenreich, M., van der Hofstad, R.: Random graph asymptotics on high-dimensional tori. Commun. Math. Phys. 270(2), 335-358 (2007)
52. Heydenreich, M., van der Hofstad, R.: Random graph asymptotics on high-dimensional tori II: volume, diameter and mixing time. Probab. Theory Relat. Fields 149(3-4), 397-415 (2011)
53. Janson, S.: The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph. Random Struct. Algorithms 7(4), 337-355 (1995)
54. Janson, S.: On percolation in random graphs with given vertex degrees. Electron. J. Probab. 14, 86-118 (2009)
55. Janson, S.: The probability that a random multigraph is simple. Comb. Probab. Comput. 18(1-2), 205-225 (2009)
56. Janson, S., Knuth, D.E., Łuczak, T., Pittel, B.: The birth of the giant component. Random Struct. Algorithms 4(3), 233-358 (1993)
57. Janson, S., Luczak, M.J.: Susceptibility in subcritical random graphs. J. Math. Phys. 49(12), 125207 (2008)
58. Janson, S., Luczak, T., Rucinski, A.: Randomgraphs, vol. 45. Wiley, New York (2000)
59. Joseph, A.: The component sizes of a critical random graph with given degree sequence. Ann. Appl. Probab. 24(6), 2560-2594 (2014)
60. Kesten, H., Lee, S.: The central limit theorem for weighted minimal spanning trees on random points. Ann. Appl. Probab. 6(2), 495-527 (1996)
61. Le Gall, J.-F.: Random trees and applications. Probab. Surv. 2, 245-311 (2005)
62. Lee, S.: The central limit theorem for Euclidean minimal spanning trees. I. Ann. Appl. Probab. 7(4), 996-1020 (1997)
63. Lee, S.: The central limit theorem for Euclidean minimal spanning trees II. Adv. Appl. Probab. 31(4), 969-984 (1999)
64. Łuczak, T.: Sparse random graphs with a given degree sequence. In: Proceedings of the Symposium on Random Graphs, Poznan, pp. 165-182 (1989)
65. Łuczak, T.: Component behavior near the critical point of the random graph process. Random Struct. Algorithms 1(3), 287-310 (1990)
66. Łuczak, T.: The number of trees with large diameter. J. Austr. Math. Soc. 58(3), 298-311 (1995)
67. Lyons, R., Peres, Y., Schramm, O.: Minimal spanning forests. Ann. Probab. 34(5), 1665-1692 (2006)
68. Malik, H.J.: Exact distribution of the product of independent generalized gamma variables with the same shape parameter. Ann. Math. Stat. 39(5), 1751-1752 (1968)
69. Marckert, J.-F., Mokkadem, A.: The depth first processes of Galton-Watson trees converge to the same Brownian excursion. Ann. Probab. 31, 1655-1678 (2003)
70. McDiarmid, C.: Concentration. Probabilistic methods for algorithmic discrete mathematics, pp. 195248 (1998)
71. Miermont, G.: Tessellations of random maps of arbitrary genus. Ann. Sci. Éc. Norm. Supér. (4) 42(5), 725-781 (2009)
72. Miermont, G., Sen, S.: On breadth-first constructions of scaling limits of random graphs and random unicellular maps. arXiv preprint arXiv:1908.04403 (2019)
73. Molloy, M., Reed, B.: A critical point for random graphs with a given degree sequence. Random Struct. Algorithms 6(2-3), 161-180 (1995)
74. Nachmias, A., Peres, Y.: Critical random graphs: diameter and mixing time. Ann. Probab. 36(4), 1267-1286 (2008)
75. Nachmias, A., Peres, Y.: Critical percolation on random regular graphs. Random Struct. Algorithms 36, 111-148 (2010)
76. Newman, C., Tassion, V., Wu, W.: Critical percolation and the minimal spanning tree in slabs. Commun. Pure Appl. Math. 70(11), 2084-2120 (2017)
77. Penrose, M.D.: The random minimal spanning tree in high dimensions. Ann. Probab. 24(4), 1903-1925 (1996)
78. Penrose, M.D.: The longest edge of the random minimal spanning tree. Ann. Appl. Probab. 7(2), 340-361 (1997)
79. Penrose, M.D.: Random minimal spanning tree and percolation on the n-cube. Random Struct. Algorithms 12(1), 63-82 (1998)
80. Penrose, M.D., Wade, A.R.: Random minimal directed spanning trees and Dickman-type distributions. Adv. Appl. Probab. 36(3), 691-714 (2004)
81. Riordan, O.: The phase transition in the configuration model. Comb. Probab. Comput. 21(1-2), 265299 (2012)
82. Steele, J.M.: Growth rates of Euclidean minimal spanning trees with power weighted edges. Ann. Probab. 16(4), 1767-1787 (1988)
83. van der Hofstad, R., Nachmias, A.: Hypercube percolation. Preprint. J. Eur. Math. Soc. (2012)
84. van der Hofstad, R.: Random Graphs and Complex Networks, vol. I. Cambridge University Press, Cambridge (2017)
85. van der Hofstad, R., Sapozhnikov, A.: Cycle structure of percolation on high-dimensional tori. Ann. Inst. Henri Poincaré Probab. Stat. 50(3), 999-1027 (2014)
86. Villani, C.: Optimal transport, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer, Berlin, Old and new (2009)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.


[^0]:    Sanchayan Sen
    sanchayan.sen1@gmail.com
    Louigi Addario-Berry
    louigi.addario@mcgill.ca
    1 Department of Mathematics and Statistics, McGill University, Montreal, Canada
    2 Department of Mathematics, Indian Institute of Science, Bangalore, India

[^1]:    1 Here, the superscript 'er' is being used to refer to the Erdős-Rényi random graph. The reason behind using this notation will become clear in Sect. 4.3.

