SCALING LIMITS AND UNIVERSALITY: CRITICAL PERCOLATION ON WEIGHTED GRAPHS CONVERGING TO AN L^3 GRAPHON

JNANESHWAR BASLINGKER[1], SHANKAR BHAMIDI[2], NICOLAS BROUTIN[3], SANCHAYAN SEN[4], AND XUAN WANG[5]

ABSTRACT. We develop a general universality technique for establishing metric scaling limits of critical random discrete structures exhibiting mean-field behavior that requires four ingredients: (i) from the barely subcritical regime to the critical window, components merge approximately like the multiplicative coalescent, (ii) asymptotics of the susceptibility functions are the same as that of the Erdős-Rényi random graph, (iii) asymptotic negligibility of the maximal component size and the diameter in the barely subcritical regime, and (iv) macroscopic averaging of distances between vertices in the barely subcritical regime.

As an application of the general universality theorem, we establish, under some regularity conditions, the critical percolation scaling limit of graphs that converge, in a suitable topology, to an L^3 graphon. In particular, we define a notion of the critical window in this setting. The L^3 assumption ensures that the model is in the Erdős-Rényi universality class and that the scaling limit is Brownian. Our results do not assume any specific functional form for the graphon. As a consequence of our results on graphons, we obtain the metric scaling limit for Aldous-Pittel's RGIV model [9] inside the critical window.

Our universality principle has applications in a number of other problems including in the study of noise sensitivity of critical random graphs [52]. In [10], we use our universality theorem to establish the metric scaling limit of critical bounded size rules. Our method should yield the critical metric scaling limit of Ruciński and Wormald's random graph process with degree restrictions [56] provided an additional technical condition about the barely subcritical behavior of this model can be proved.

CONTENTS

[1]DEPARTMENT OF MATHEMATICS, INDIAN INSTITUTE OF SCIENCE
[2]DEPARTMENT OF STATISTICS AND OR, UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL
[3]SORBONNE UNIVERSITÉ, FRANCE,
[4]DEPARTMENT OF MATHEMATICS, INDIAN INSTITUTE OF SCIENCE
[5]DATABRICKS

1. INTRODUCTION

Over the last two decades, motivated both by questions in fields such as combinatorics and statistical physics as well as the explosion in data on real networks, an array of random graph models have been proposed. Many of these models undergo a phase transition in the size of the largest component. The canonical example of such phenomenon is the Erdős-Rényi random graph $\mathrm{ERRG}(n, t/n)$ where the phase transition occurs at the critical time $t_c(\mathrm{ERRG}) = 1$. One fundamental question in the study of random discrete structures

is understanding the nature of this phase transition, and see how they compare to the phase transition in the Erdős-Rényi random graph.

In the early 2000s, based on simulations, it was predicted in the statistical physics literature [21, 22, 29, 60] that asymptotics in the critical regime are "universal." More precisely, under moment assumptions on the average degree, a wide array of models are conjectured to exhibit the same behavior in the critical regime as the Erdős-Rényi random graph in the sense that for any fixed $i \geq 1$, the size of the i-th maximal component scales like $n^{2/3}$, and the typical disctance in the component scales like $n^{1/3}$. Denoting the i-th maximal component of the Erdős-Rényi random graph ERRG$(n, n^{-1} + \lambda n^{-4/3})$ by $\mathscr{C}_i(\lambda)$, it was proved in [7] that the sequence of rescaled component sizes $(n^{-2/3}|\mathscr{C}_i(\lambda)|; i \geq 1)$ conervges in distribution to a limit that will be defined in Section 2.4. Turning to the intrinsic geometry of the components, it was shown in [4] that rescaling edge lengths in the i-th maximal component $\mathscr{C}_i(\lambda)$ of the Erdős-Rényi random graph ERRG$(n, 1/n + \lambda/n^{4/3})$ and denoting the resulting metric space by $n^{-1/3}\mathscr{C}_i(\lambda)$, one has

$$\left(n^{-1/3}\mathscr{C}_i(\lambda); i \geq 1\right) \xrightarrow{\text{d}} \mathbf{Crit}(\lambda) := (\text{Crit}_i(\lambda); i \geq 1), \tag{1}$$

for a sequence of limiting random fractals that are described in more detail in Section 2. This scaling limit is expected to be the universal scaling limit for a large class of critical random graph models exhibiting mean-field behavior. Further, existing literature suggests that the components of the high-dimensional discrete torus [40, 41, 45] and the hypercube [44] under critical percolation also share the Erdős-Rényi scaling limit.

There are two standard techniques for establishing scaling limits as above: (a) *The exploration technique:* Here, the random graph under consideration is explored, often in a breadth-first or a depth-first manner, and alongside one keeps track of certain associated processes, e.g., the exploration-walk process, the height process, or the contour process. One then proves distributional limits for these associated processes and the metric scaling limit is derived from these limit theorems. (b) *The spanning tree technique:* Here, one first establishes the scaling limit of the rescaled component sizes of the model being considered. Then conditional on the component sizes, one has to find a way to represent each connected component as a well-behaved spanning tree with certain additional random edges added to it. Then the metric scaling limit of each connected component can be derived using the scaling limit of the spanning tree and also incorporating the location of the additional random edges. When applicable, these techniques can be powerful tools in obtaining precise results about the geometry of the model being considered. However, the model being studied has to have some nice properties in order to be amenable to analysis via these methods. For the first method to be applicable, one has to be able to set up a suitable exploration process for which it is possible to obtain the distributional limit in the Skorohod topology along with the assoiated height/contour process. Similarly, for the second technique, each connected component needs to have a spanning tree whose metric scaling limit can be established, which is usually done by connecting its distribution to a well-studied model of random trees. These two methods have been successfully applied to the following two models of random graphs: the inhomogeneous random graph with a rank-one kernel (including the Erdős-Rényi random graph), and the configuration model and (simple) random graphs with a given degree sequence.

See [3, 4, 7, 11–13, 15, 16, 23, 30–33, 39, 48, 53, 55] for various results on the critical scaling limits for these two models proved under different sets of assumptions. In general, it is not clear how the above two methods can be applied to the many other standard models of random discrete structures studied in the literature.

1.1. **Main contributions of this paper.** We prove a general universality principle for systematically approaching the problem of establishing metric scaling limits of critical random discrete structures exhibiting mean-field behavior. Applying this technique requires the verification of four properties: (i) From the barely subcritical regime to the critical window, components merge approximately like the multiplicative coalescent. (ii) The second and the third susceptibility functions at certain suitable points in the barely subcritical regime exhibit the same asymptotics as those of the Erdős-Rényi random graph. (iii) The size of the maximal component and the diameter in the barely subcritical regime is asymptotically negligibile in an appropriate sense. (iv) The average distances between vertices in the barely subcritical regime is well-behaved. Often, these four criteria can be verified for models where it is not clear if the exploration process technique or the spanning tree technique can be used directly, thus allowing for broader applicability.

We apply our general theorem to random graphs obtained under critical percolation on edge-weighted graphs that converge, in a suitable topology, to an L^3 graphon. The theory of graphons is now a standard tool for sampling and studying convergence properties of dense graph sequences, and has found applications in many branches of applied science, especially in statistics and machine learning. The precise definition of the model we are considering is given in Section 1.2.1. The phase transition in this model follows from the results in [17–19]. Note that in all these works, only the transition from the purely subcritical regime to the purely supercritical regime was studied–the near critical behavior was not considered.

As a consequence of our universality principle, we will show that under some regularity conditions, the metric scaling limit in this model is the same as that of the Erdős-Rényi random graph (given in (1)) modulo a constant. To our knowledge, this gives the first result where the existence of a critical window is established and the metric scaling limit is obtained inside the critical window without any assumptions on the specific functional form of the graphon. (As mentioned above, the metric scaling limit for this model was only known in the special case of the rank-one setting where the kernel has a multiplicative form: $W(x, y) = \phi(x)\phi(y), \ x, y \in [0, 1]$.) Here, the L^3 condition ensures that the scaling limit is of a Brownian nature, i.e., the model belongs to the Erdős-Rényi universality class. We expect scaling limits of a different type to arise when the L^3 condition is replaced by an L^p condition where $p \in (2, 3)$; we will revisit this point in Remark 4. As a further application of our result on graphons, we obtain the metric scaling limit of Aldous-Pittel's RGIV model [9] inside the critical window.

The main universality principle of this paper can be applied to several other problems. In [10], we apply our universality theorem to dynamic random graph models including bounded size rules [58] and establish the metric scaling limit inside the critical window. In [52], the authors use this universality result as one of the ingredients to study noise sensitivity of critical random graphs. We expect that our general method can be used to obtain

the critical metric scaling limit of Ruciński and Wormald's random graph process with degree consraints [56]. At present, we are missing a technical condition about the barely subcritical behavior of this model. We will briefly discuss this in Section 7. Further, as shown in [5] (see also [6, 14]), obtaining the metric structure of maximal components at criticality is a key tool in analyzing more complicated objects such as the minimal spanning tree (MST) on the giant component in the supercritical regime. Establishing the critical scaling limits such as those obtained in this paper would be one of the steps in establishing universality for such objects as well.

Organization of the paper: In Section 1.2 we define the specific random graph models considered in this paper. In Section 2, we define mathematical constructs needed to state the results. Then Section 3 describes the general universality result. Section 4 contains results on applications of the universality result to specific models. The proofs of these results are given Sections 5 and 6. Finally, in Section 7, we state a result on the scaling limit of the critical stochastic block model and discuss possible extensions of our results.

1.2. **Models.** In this section, we define the two specific models to which our general universality result (Theorem 3.4 stated below) is applied in this paper.

1.2.1. *Graphons and critical percolation on edge-weighted graphs.* For $p \ge 1$, a Borel measurable function $W : [0, 1]^2 \to [0, \infty)$ is called an L^p graphon if W is symmetric (i.e., $W(x, y) = W(y, x)$ for $x, y \in [0, 1]$), and $\int_0^1 \int_0^1 W(x, y)^p dx dy < \infty$. Clearly, an L^p graphon is also an L^q graphon if $p > q$.

Now, to any $K \in L^2([0, 1]^2)$ we can associate a bounded linear operator $T_K : L^2[0, 1] \to L^2[0, 1]$ via

$$T_K f(\cdot) := \int_0^1 K(\cdot, y) f(y) dy, \ \ f \in L^2[0, 1]. \tag{2}$$

Further, for $p, q \in [1, \infty]$ and a linear operator $T^\bullet : L^p[0, 1] \to L^q[0, 1]$, we write

$$\|T^\bullet\|_{p,q} = \sup \left\{ \|T^\bullet f\|_q : \|f\|_p = 1 \right\} \tag{3}$$

for the operator norm of $T^\bullet$. We call an L^2 graphon W critical if $\|T_W\|_{2,2} = 1$. The reason for this terminology comes from the criterion for an inhomogeneous random graph to be critical [18, Theorem 6.1].

Consider a sequence of graphs with vertex sets $[n]$ and edge weights $\boldsymbol{\beta}_n = \left(\beta_{ij}^{(n)}; i, j \in [n]\right)$, where $\beta_{ii}^{(n)} = 0$ for $i \in [n]$ and $\beta_{ij}^{(n)} = \beta_{ji}^{(n)} \ge 0$ for $1 \le i \ne j \le n$. Using the weight sequence $\boldsymbol{\beta}_n$, we will formulate a notion (see Conditions 4.1, 4.2, and 4.3 below) for these graphs to converge to a critical L^3 graphon. We write

$$p_{ij}^{(n)} = 1 \wedge (n^{-1} \beta_{ij}^{(n)}) \ \ \text{and} \ \ \widetilde{p}_{ij}^{(n)} = 1 - \exp\left(-\beta_{ij}^{(n)}/n\right), \ \ i \le i \ne j \le n. \tag{4}$$

Let $\mathscr{G}_n^{\mathrm{gphn}} = \mathscr{G}_n^{\mathrm{gphn}}(\boldsymbol{\beta}_n)$ (resp. $\widetilde{\mathscr{G}}_n^{\mathrm{gphn}} = \widetilde{\mathscr{G}}_n^{\mathrm{gphn}}(\boldsymbol{\beta}_n)$) denote the random graph with vertex set $[n]$, where the edges $\{i, j\}$ appear independently with respective probabilities $p_{ij}^{(n)}$ (resp. $\widetilde{p}_{ij}^{(n)}$), $1 \le i < j \le n$. We will use the notation G_n^{gphn} to refer to either of the graphs $\mathscr{G}_n^{\mathrm{gphn}}$ and $\widetilde{\mathscr{G}}_n^{\mathrm{gphn}}$, as all our results will hold for both these graphs. As a consequence of our general universality result (Theorem 3.4), we will obtain the metric scaling limit of this model (Theorem 4.5).

1.2.2. *Aldous and Pittel's RGIV process* [9]*:* The random graph with immigrating vertices (RGIV) process is a dynamic random graph model that evolves as follows: At time 0, the RGIV process is the empty graph. Then vertices immigrate at rate n, and for each pair of existing vertices, an edge connecting them appears at rate $1/n$.

A deterministic version of this process was studied in the physics literature [59], where it was observed that a phase transition occurs in this process at time $\pi/2$; see also [28]. Aldous and Pittel [9] studied the behavior of this process inside the critical window, i.e., at time

$$t_{n,\lambda} = \pi/2 + \lambda n^{-1/3}, \tag{5}$$

and they established the scaling limit of the component sizes. For $\lambda \in \mathbb{R}$, write $\mathscr{G}_n^{\mathrm{rgiv}}(\lambda)$ for the random graph process at time $t_{n,\lambda}$ as given by (5). As an application of Theorem 4.5, we will obtain the metric scaling limit of $\mathscr{G}_n^{\mathrm{rgiv}}(\lambda)$ (Theorem 4.15).

2. PRELIMINARIES

2.1. **GHP convergence of compact metric measure spaces.** We primarily follow [1, 5, 24]. For a compact metric space (X, d) and $A_1, A_2 \subseteq X$, we define the Hausdorff distance between A_1 and A_2 to be

$$d_{\mathrm{H}}(A_1, A_2) := \inf\{\varepsilon > 0 : A_1 \subseteq A_2^\varepsilon \text{ and } A_2 \subseteq A_1^\varepsilon\},$$

where $A_1^\varepsilon := \bigcup_{x \in A_1}\{y \in X : d(y, x) \le \varepsilon\}$. For two metric spaces $X_1 = (X_1, d_1)$ and $X_2 = (X_2, d_2)$ and $\mathscr{R} \subseteq X_1 \times X_2$, define the distortion of $\mathscr{R}$ with respect to X_1 and X_2 as

$$\mathrm{dis}(\mathscr{R}) = \mathrm{dis}(\mathscr{R}; X_1, X_2) := \sup\{|d_1(x_1, y_1) - d_2(x_2, y_2)| : (x_1, x_2), (y_1, y_2) \in \mathscr{R}\}. \tag{6}$$

A correspondence $\mathscr{R}$ 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 $(x_1, x_2) \in \mathscr{R}$ and vice-versa. Let $\mathrm{Corr}(X_1, X_2)$ be the set of all such correspondences. The Gromov-Hausdorff distance between (X_1, d_1) and (X_2, d_2) is defined as

$$d_{\mathrm{GH}}(X_1, X_2) = \frac{1}{2}\inf\{\mathrm{dis}(\mathscr{R}) : \mathscr{R} \in \mathrm{Corr}(X_1, X_2)\}. \tag{7}$$

The Gromov-Hausdorff-Prokhorov distance extends the above to keep track of measures on the corresponding spaces. A compact metric measure space is a compact metric space equipped with a finite measure on its Borel sigma algebra. Let $X_1 = (X_1, d_1, \mu_1)$ and $X_2 = (X_2, d_2, \mu_2)$ be compact metric measure spaces. Suppose π is a measure on the product space $X_1 \times X_2$ with marginals π_1, π_2. The discrepancy of π with respect to X_1 and X_2 is defined as

$$\mathrm{dsc}(\pi) = \mathrm{dsc}(\pi; X_1, X_2) := ||\mu_1 - \pi_1|| + ||\mu_2 - \pi_2||, \tag{8}$$

where $||\cdot||$ denotes the total variation of signed measures. Now define

$$\bar{d}_{\mathrm{GHP}}(X_1, X_2) := \inf\{\max(\mathrm{dis}(\mathscr{R})/2, \mathrm{dsc}(\pi), \pi(\mathscr{R}^c))\}, \tag{9}$$

where the infimum is taken over all correspondences $\mathscr{R}$ and measures π on $X_1 \times X_2$. The function $\bar{d}_{\mathrm{GHP}}$ is a pseudometric and defines an equivalence relation $X_1 \sim X_2$ if and only if $\bar{d}_{\mathrm{GHP}}(X_1, X_2) = 0$ on. Let $\mathscr{S}$ be the space of all equivalence classes of compact metric measure spaces and let d_{GHP} be the induced metric. Then by [1], $(\mathscr{S}, d_{\mathrm{GHP}})$ is a complete

separable metric space. We will continue to use $X = (X, d, \mu)$ to denote both the metric space and the corresponding equivalence class. In this paper we will typically deal with a sequence of infinite sequences of compact metric measure spaces; we will view it as a sequence in $\mathscr{S}^{\mathbb{N}}$. When talking about distributional convergence for a sequence of $\mathscr{S}^{\mathbb{N}}$-valued random objects, it will be understood that the underlying topology on $\mathscr{S}^{\mathbb{N}}$ is the product topology inherited from d_{GHP}.

The scaling operator: We will need to rescale both the metric as well as the associated measures on the components in the critical regime. For $a, b > 0$, let $\mathrm{scl}(a, b) : \mathscr{S} \to \mathscr{S}$ be the scaling operator

$$\mathrm{scl}(a, b)[(X, d, \mu)] := (X, d', \mu'),$$

where $d'(x, y) := a d(x, y)$ for all $x, y \in X$, and $\mu'(A) := b\mu(A)$ for all $A \subseteq X$. For simplicity, write $\mathrm{scl}(a, b)X := \mathrm{scl}(a, b)[(X, d, \mu)]$ and $aX := \mathrm{scl}(a, 1)X$.

2.2. **Notation.** We write $|A|$ or $\#A$ for the cardinality of a set A. For any $n \in \mathbb{Z}_{>0}$, we will write $[n]$ for the set $\{1, 2, \dots, n\}$. For a finite graph G, we write $V(G)$ and $E(G)$ for the set of vertices and the set of edges in G respectively. We write $|G|$ to mean $|V(G)|$, and $\mathrm{spls}(G) := |E(G)| - |G| + 1$ for the number of surplus edges in G. We view a connected component $\mathscr{C}$ of G as a metric space using the graph distance d_G. When the graph G is clear from the context, we will supress the subscript and simply write d for the graph distance. Often in our analysis, associated to a graph G there will be a collection of vertex weights $\mathbf{w} = \{w_v : v \in V(G)\}$. Natural measures on G are (a) *Counting measure:* $\mu_{\mathrm{ct}}(A) := |A|$, for $A \subseteq V(G)$; and (b) *Weighted measure:* $\mu_{\mathbf{w}}(A) := \sum_{v \in A} w_v$, for $A \subseteq V(G)$. For G finite and connected, we will often use G for both the graph and its associated metric measure space. If the measure to be considered is not the counting measure then that will be explicitly mentioned. If nothing is specified, the measure will be understood to be the counting measure. For two graphs G_1 and G_2, we write $G_1 \subseteq G_2$ to mean that G_1 is a subgraph of G_2. For a metric space $(\mathfrak{X}, d)$, write $\mathrm{diam}(\mathfrak{X}) := \sup_{u, v \in \mathfrak{X}} d(u, v)$. For a finite graph G, we let $\mathrm{diam}(G) := \max\{\mathrm{diam}(\mathscr{C}) : \mathscr{C} \text{ connected component of } G\}$.

We will say that a sequence $\mathscr{E}_n$, $n \geq 1$, of events occurs with high probability (whp) if $\mathbb{P}(\mathscr{E}_n) \to 1$. We will freely omit ceilings and floors when there is no confusion in doing so. Throughout this paper, C, C' etc. will denote positive constants whose values may change from line to line; the values of these constants do not depend on the variable (usually 'n' or 'ℓ') used to index sequences. For two sequences of real numbers (a_n) and (b_n) (resp. two sequences of real-valued functions (f_n) and (g_n) defined on a domain D), we will write $a_n \lesssim b_n$ (resp. $f_n \lesssim g_n$) to express the relation $a_n \leq C b_n$ for all n (resp. $f_n(x) \leq C g_n(x)$ for all $x \in D$ and all n), where C is a constant as above.

For two sequences of real numbers (a_n) and (b_n), we will write $a_n \sim b_n$ to mean $a_n/b_n \to 1$ as $n \to \infty$. For two sequences (X_n) and (Y_n) of random variables where X_n and Y_n are defined on the same probability space for each n, we will write $X_n \sim Y_n$ to mean that $X_n/Y_n \xrightarrow{\mathrm{d}} 1$ as $n \to \infty$. (We also write $X \sim \mu$ to mean that the random variable X has law μ; this, however, should not cause any confusion.) We write $X_n = O_P(Y_n)$ if the sequence (X_n/Y_n) is tight. We will write $X_n = o_P(Y_n)$ to mean $X_n/Y_n \xrightarrow{\mathrm{d}} 0$ as $n \to \infty$.

We will denote the uniform probability measures on $[0,1]$ and $[0,1]^2$ by $\mu_\circ$ and $\mu_\square$ respectively. For $x, y \in \mathbb{R}$, we will write $x \wedge y$ and $x \vee y$ for $\min\{x, y\}$ and $\max\{x, y\}$ respectively. For a function f taking values in $\mathbb{R}$, we will denote by f^+ and f^- the functions $f \vee 0$ and $-(f \wedge 0)$ respectively. For $f, g \in L^2([0,1])$, we will write $\langle f, g \rangle$ for $\int_0^1 f(x)g(x)\,dx$.

2.3. **Real trees and quotient spaces.**

For $l > 0$, write $\mathcal{E}_l$ for the space of continuous excursions on $[0, l]$. Fix $h, g \in \mathcal{E}_l$, a locally finite set $\mathcal{P} \subseteq \mathbb{R}_+ \times \mathbb{R}_+$, and set

$$g \cap \mathcal{P} := \big\{ (x, y) \in \mathcal{P} : 0 \le x \le l, \, 0 \le y < g(x) \big\}.$$

Let $\mathcal{T}(h)$ be the real tree encoded by h (see, e.g., [37, 51]) equipped with the push forward of the Lebesgue measure on $[0, l]$. For $(x, y) \in g \cap \mathcal{P}$, let $r(x, y) := \inf\{x' : x' \ge x, \, g(x') \le y\}$. Write $\mathcal{G}(h, g, \mathcal{P})$ for the metric measure space obtained by identifying the pairs of points in $\mathcal{T}(h)$ corresponding to the pairs of points $\{(x, r(x, y)) : (x, y) \in g \cap \mathcal{P}\}$. Thus $\mathcal{G}(h, g, \mathcal{P})$ is obtained by adding a finite number of shortcuts to $\mathcal{T}(h)$.

Tilted Brownian excursions: Let $(\mathbf{e}_l(s), s \in [0, l])$ be a Brownian excursion of length l. Fix $\theta > 0$ and let $\tilde{\mathbf{e}}_l^\theta$ be an $\mathcal{E}_l$-valued random variable such that for any bounded continuous function $f : \mathcal{E}_l \to \mathbb{R}$,

$$\mathbb{E}[f(\tilde{\mathbf{e}}_l^\theta)] = \mathbb{E}\left[f(\mathbf{e}_l) \exp\left(\theta \int_0^l \mathbf{e}_l(s)\,ds \right) \right] \Big/ \mathbb{E}\left[\exp\left(\theta \int_0^l \mathbf{e}_l(s)\,ds \right) \right]. \tag{10}$$

For simplicity, write $\mathbf{e}(\cdot) := \mathbf{e}_1(\cdot)$, $\tilde{\mathbf{e}}^\theta(\cdot) := \tilde{\mathbf{e}}_1^\theta(\cdot)$, and $\tilde{\mathbf{e}}_l(\cdot) := \tilde{\mathbf{e}}_l^1(\cdot)$. In Section 2.5, we will use the random objects defined in this section to describe a construction of the limiting random metric measure spaces of interest in this paper.

2.4. **Multiplicative coalescent and the random graph $\mathcal{G}(\mathbf{x}, q)$.**

A multiplicative coalescent $\big(\mathbf{X}(t); t \in \mathscr{A}\big)$ is a Markov process with state space $l_\downarrow^2 := \{(x_1, x_2, \ldots) : x_1 \ge x_2 \ge \cdots \ge 0, \sum_i x_i^2 < \infty\}$ endowed with the metric inherited from l^2; here, either $\mathscr{A} = \mathbb{R}$ or $\mathscr{A} = [t_0, \infty)$ for some $t_0 \in \mathbb{R}$. Its evolution can be described in words as follows: Fix $\mathbf{x} \in l_\downarrow^2$. Then conditional on $\mathbf{X}(t) = \mathbf{x}$, each pair of clusters i and j merge at rate $x_i x_j$ to form a new cluster of size $x_i + x_j$. While this description makes sense for a finite collection of clusters (i.e., $x_i = 0$ for $i > K$ for some finite K), Aldous [7] showed that it also makes sense for $\mathbf{x} \in l_\downarrow^2$. We refer the reader to [7] for a more detailed study of how the multiplicative coalescent is connected to the evoloution of the Erdős-Rényi random graph.

An object closely related to the multiplicative coalescent is the random graph $\mathcal{G}(\mathbf{x}, q)$: Fix vertex set $[n]$, a collection of positive vertex weights $\mathbf{x} = (x_i, i \in [n])$, and parameter $q > 0$. Construct the random graph $\mathcal{G}(\mathbf{x}, q)$ by placing an edge between $i \ne j \in [n]$ with probability $1 - \exp(-q x_i x_j)$, independently across pairs $\{i, j\}$ with $i \ne j$. For a connected component $\mathscr{C}$ of $\mathcal{G}(\mathbf{x}, q)$, define $\mathrm{mass}(\mathscr{C}) = \sum_{i \in \mathscr{C}} x_i = \mu_\mathbf{x}(\mathscr{C})$ (recall the notation from Section 2.2). Rank the components in terms of their masses and let $\mathscr{C}_i$ be the i-th maximal component.

Condition 2.1. *Consider a sequence of vertex weights $\mathbf{x} = \mathbf{x}^{(n)}$ and a sequence $q = q^{(n)}$. Let $\sigma_k := \sum_{i \in [n]} x_i^k$ for $k = 2, 3$ and $x_{\max} := \max_{i \in [n]} x_i$. Assume that there exists a constant $\lambda \in \mathbb{R}$ such that as $n \to \infty$,*

$$\sigma_3/(\sigma_2)^3 \to 1, \quad q - (\sigma_2)^{-1} \to \lambda, \quad \text{and} \quad x_{\max}/\sigma_2 \to 0.$$

Note that these conditions imply that $\sigma_2 \to 0$. Fix $\lambda \in \mathbb{R}$ and let B be a standard Brownian motion. Define the processes W_λ and $\tilde{W}_\lambda$ via

$$W_\lambda(t) := B(t) + \lambda t - t^2/2, \quad \text{and} \quad \tilde{W}_\lambda(t) := W_\lambda(t) - \inf_{s \in [0,t]} W_\lambda(s), \quad t \geq 0. \tag{11}$$

An excursion of $\tilde{W}_\lambda$ is an interval $(l, r) \subset \mathbb{R}^+$ such that $\tilde{W}_\lambda(l) = \tilde{W}_\lambda(r) = 0$ and $\tilde{W}_\lambda(t) > 0$ for all $t \in (l, r)$. Write $r - l$ for the length of such an excursion. Aldous [7] showed that the lengths of the excursions of $\tilde{W}_\lambda$ can be arranged in decreasing order as $\gamma_1(\lambda) > \gamma_2(\lambda) > \ldots > 0$. Conditional on $\tilde{W}_\lambda$, let $\mathrm{Pois}_i(\lambda)$, $i \geq 1$, be independent random variables where $\mathrm{Pois}_i(\lambda)$ is Poisson distributed with mean equal to the area underneath the excursion with length $\gamma_i(\lambda)$.

Theorem 2.2 ([7]). *Under Condition* 2.1, $(\mathrm{mass}(\mathscr{C}_i); i \geq 1) \xrightarrow{\mathrm{d}} \boldsymbol{\xi}(\lambda) := (\gamma_i(\lambda); i \geq 1)$, *as* $n \to \infty$, *with respect to the topology on* $l^2_\downarrow$. *Further,* $\big((\mathrm{mass}(\mathscr{C}_i), \mathrm{spls}(\mathscr{C}_i)); i \geq 1\big) \xrightarrow{\mathrm{d}} \Xi(\lambda) := \big((\gamma_i(\lambda), \mathrm{Pois}_i(\lambda)); i \geq 1\big)$, *as* $n \to \infty$, *with respect to the product topology.*

Remark 1. In [7, Proposition 4], Aldous only proves the first assertion in Theorem 2.2, and the second assertion in Theorem 2.2 is proved for the special case of the Erdős-Rényi random graph [7, Corollary 2]. However, the second convergence in Theorem 2.2 follows easily from [7, Proposition 10]. ∎

2.5. **Scaling limits of components in critical Erdős-Rényi random graph.** Now we can define the scaling limit of the maximal components in $\mathrm{ERRG}(n, 1/n + \lambda/n^{4/3})$ derived in [4]. Recall the definitions of tilted Brownian excursions and the metric measure space $\mathscr{G}(h, g, \mathscr{P})$ from Section 2.3. Let $\boldsymbol{\xi}(\lambda) = (\gamma_i(\lambda), i \geq 1)$ be as in Theorem 2.2. Conditional on $\boldsymbol{\xi}(\lambda)$, let $\tilde{\mathbf{e}}_{\gamma_i(\lambda)}$, $i \geq 1$, be independent tilted Brownian excursions with $\tilde{\mathbf{e}}_{\gamma_i(\lambda)}$ having length $\gamma_i(\lambda)$. Let $\mathscr{P}_i$, $i \geq 1$, be independent rate one Poisson processes on $\mathbb{R}^2_+$ that are also independent of $\big(\tilde{\mathbf{e}}_{\gamma_i(\lambda)}; i \geq 1\big)$. Now consider the sequence of random metric measure spaces

$$\mathrm{Crit}_i(\lambda) := \mathscr{G}(2\tilde{\mathbf{e}}_{\gamma_i(\lambda)}, \tilde{\mathbf{e}}_{\gamma_i(\lambda)}, \mathscr{P}_i), \qquad i \geq 1. \tag{12}$$

For the rest of the paper we let

$$\mathbf{Crit}(\lambda) := (\mathrm{Crit}_i(\lambda), i \geq 1). \tag{13}$$

3. RESULTS: UNIVERSALITY

Our plan is to extend Theorem 2.2 in two stages. In the first stage, we consider the random graph $\mathscr{G}(\mathbf{x}, q)$, and for $i \geq 1$, we view the component $\mathscr{C}_i$ as a metric measure space endowed with the measure $\mu_{\mathbf{x}}$. Then [16, Theorem 7.3] implies that under Condition 2.1 and Condition 3.1 stated below, the sequence $(\mathscr{C}_i, i \geq 1)$ properly rescaled converges to $\mathbf{Crit}(\lambda)$ as in (13). In the second stage of the extension of Theorem 2.2, we replace each vertex $i \in [n]$ in the graph $\mathscr{G}(\mathbf{x}, q)$ with a metric measure space (M_i, d_i, μ_i); we refer to the spaces (M_i, d_i, μ_i), $i \in [n]$, as 'blobs.' In Theorem 3.4, we show that the metric measure space $\bar{\mathscr{C}}_i$ now associated with $\mathscr{C}_i$, under Conditions 2.1 and 3.3, converges to $\mathrm{Crit}_i(\lambda)$ after proper rescaling, owing to macroscopic averaging of distances within blobs.

Before moving on to the statements of these results, let us explain the strategy to apply this general theorem to specific random graph models. Consider a random graph process that undergoes a phase transition in the size of the largest component, and let $\mathbb{G}_n$ denote

the random graph at a point inside the critical window. We couple $\mathbb{G}_n$ with a suitable barely subcritical random graph model $\mathbb{G}_n^\bullet$, and we take the components of $\mathbb{G}_n^\bullet$ to be the 'blobs.' This can be seen as a kind of renormalization technique. We then study the connections that form in going from $\mathbb{G}_n^\bullet$ to $\mathbb{G}_n$, and we show that the connectivity pattern between blobs can be *approximately* described through the graph $\mathscr{G}(\mathbf{x}, q)$ for a suitable q, where x_i is a model dependent functional of the i-th blob. We emphasize that at this stage, we are viewing each blob as a single vertex ignoring their internal structure. Thus, $\mathscr{G}(\mathbf{x}, q)$ should be thought of as approximately describing the blob-level superstructure owing to edges created while going from $\mathbb{G}_n^\bullet$ to $\mathbb{G}_n$. Now, once we verify that Condition 3.3 is satisfied by the blobs, Theorem 3.4 applies to the random graph $\overline{\mathbb{G}}_n$ obtained by replacing the vertices of $\mathscr{G}(\mathbf{x}, q)$ by the blobs (where the internal structure of the blobs have been incorporated). Then the final step is to show that $\mathbb{G}_n$ is well-approximated by $\overline{\mathbb{G}}_n$.

3.1. **Blob-level superstructure.** Recall the random graph $\mathscr{G}(\mathbf{x}, q)$ from Section 2.4, and the notation used in Condition 2.1. View $(\mathscr{C}_i, i \geq 1)$ as metric measure spaces equipped with the graph distance and the weighted measure $\mu_{\mathbf{x}}$.

Condition 3.1. *Assume that there exist $\eta_0 \in (0, \infty)$ and $r_0 \in (0, \infty)$ such that $x_{\max}/\sigma_2^{3/2+\eta_0} \to 0$ and $\sigma_2^{r_0}/x_{\min} \to 0$ as $n \to \infty$, where $x_{\min} := \min_{i \in [n]} x_i$.*

Theorem 3.2. *Under Conditions 2.1 and 3.1, $\left(\mathrm{scl}(\sigma_2, 1) \mathscr{C}_i^{(n)}, i \geq 1 \right) \xrightarrow{\mathrm{d}} \mathbf{Crit}(\lambda)$ as $n \to \infty$.*

Remark 2. Theorem 3.2 can be derived as a special case of Theorem 3.4 stated below. However, we state Theorem 3.2 in the above form as this explains the motivation behind us proceeding in two stages as discussed above. Note also that the scaling limit of the critical Erdős-Rényi random graph $\mathrm{ERRG}(n, n^{-1} + \lambda n^{-4/3})$ derived in [4] can be recovered from Theorem 3.2 by taking $q = n^{1/3} + \lambda$ and $x_i = n^{-2/3}$ for $i \in [n]$ (which results in $\sigma_2 = n^{-1/3}$). ∎

3.2. **Incorporating the internal structure of the blobs.** We need the following ingredients:

(a) **Blob level superstructure:** A simple finite graph $\mathscr{G}$ with vertex set $[n]$ and vertex weight sequence $\mathbf{x} := (x_i, i \in [n])$.

(b) **Blobs:** A family of compact metric measure spaces $\mathbf{M} := \{(M_i, d_i, \mu_i) : i \in [n]\}$, where μ_i is a probability measure for all i.

(c) **Blob to blob junction points:** A collection of points $\mathbf{X} := \{X_{i,j} : i \in [n], j \in [n]\}$ such that $X_{i,j} \in M_i$ for all i, j.

Using these three ingredients, we can define a metric measure space $\Gamma(\mathscr{G}, \mathbf{x}, \mathbf{M}, \mathbf{X}) = (\bar{M}, \bar{d}, \bar{\mu})$ as follows: Let $\bar{M} := \bigsqcup_{i \in [n]} M_i$. Define the measure $\bar{\mu}$ as

$$\bar{\mu}(A) = \sum_{i \in [n]} x_i \mu_i(A \cap M_i), \text{ for } A \subseteq \bar{M}. \tag{14}$$

The metric $\bar{d}$ is obtained by using the intra-blob distance functions $(d_i; i \in [n])$ together with the graph distance on $\mathscr{G}$ by putting an edge of length one between the pairs of vertices

$\{\{X_{i,j}, X_{j,i}\} : \{i, j\}$ is an edge in $\mathcal{G}\}$. Thus, for $x, y \in \bar{M}$ with $x \in M_{j_1}$ and $y \in M_{j_2}$,

$$\bar{d}(x, y) = \inf\left\{k + d_{j_1}(x, X_{j_1, i_1}) + \sum_{\ell=1}^{k-1} d_{i_\ell}(X_{i_\ell, i_{\ell-1}}, X_{i_\ell, i_{\ell+1}}) + d_{j_2}(X_{j_2, i_{k-1}}, y)\right\}, \quad (15)$$

where the infimum is taken over $k \geq 1$ and all paths $(i_0, i_1, \ldots, i_{k-1}, i_k)$ in $\mathcal{G}$ with $i_0 = j_1$ and $i_k = j_2$. (Here, the infimum of an empty set is understood to be $+\infty$. This corresponds to the case where j_1 and j_2 do not belong to the same compoenent of $\mathcal{G}$.)

The above is a deterministic procedure for creating a new metric measure space. Now assume that we are provided with a parameter sequence $q^{(n)}$, weight sequence $\mathbf{x}^{(n)} := (x_i^{(n)}, i \in [n])$, and the family of metric measure spaces $\mathbf{M}^{(n)} := \left((M_i^{(n)}, d_i^{(n)}, \mu_i^{(n)}), i \in [n]\right)$. As before, we will suppress the dependence on n. Let $\mathcal{G}(\mathbf{x}, q)$ be the random graph defined in Section 2.4 constructed using the weight sequence $\mathbf{x}$ and the parameter q. Let $(\mathcal{C}_i, i \geq 1)$ denote the connected components of $\mathcal{G}(\mathbf{x}, q)$ ranked in decreasing order of their masses. Let $X_{i,j}$, $i, j \in [n]$, be independent random variables (that are also independent of $\mathcal{G}(\mathbf{x}, q)$) such that for each fixed i, $X_{i,j}$, $j \in [n]$, are i.i.d. M_i-valued random variables that are μ_i distributed. Let $\mathbf{X} = (X_{i,j}, i, j \in [n])$.

Let $\bar{\mathcal{G}}(\mathbf{x}, q, \mathbf{M}) := \Gamma(\mathcal{G}, \mathbf{x}, \mathbf{M}, \mathbf{X})$ be the (random) compact metric measure space constructed as above. Let $\bar{\mathcal{C}}_i$ be the component in $\bar{\mathcal{G}}(\mathbf{x}, q, \mathbf{M})$ corresponding to the i-th maximal component $\mathcal{C}_i$ in $\mathcal{G}(\mathbf{x}, q)$, viewed as a compact metric measure space as explained in (14) and (15). Define the quantities

$$u_i := \int_{M_i \times M_i} d_i(x, y) \mu_i(dx) \mu_i(dy), \qquad i \in [n], \quad (16)$$

$$\tau := \sum_{i \in [n]} x_i^2 u_i \text{ and } \operatorname{diam}_{\max} := \max_{i \in [n]} \operatorname{diam}(M_i). \quad (17)$$

Note that u_i is the expectation of the distance between two blob-to-blob junction points in blob i and τ / σ_2 is the weighted average of these "typical" distances.

Condition 3.3. *There exist $\eta_0 \in (0, \infty)$ and $r_0 \in (0, \infty)$ such that* (a) *Condition* 3.1 *holds, and* (b) *as $n \to \infty$, $(\operatorname{diam}_{\max} \cdot \sigma_2^{3/2 - \eta_0})/(\tau + \sigma_2) \to 0$ and $\sigma_2 x_{\max} \operatorname{diam}_{\max}/\tau \to 0$.*

Theorem 3.4. *Under Conditions* 2.1 *and* 3.3,

$$\left(\operatorname{scl}\left(\frac{\sigma_2^2}{\sigma_2 + \tau}, 1\right) \bar{\mathcal{C}}_i ; i \geq 1\right) \xrightarrow{\mathrm{d}} \mathbf{Crit}(\lambda), \text{ as } n \to \infty.$$

Remark 3. Depending on whether $\lim_{n \to \infty} \tau / \sigma_2 \in [0, \infty)$ or equals ∞, the above theorem deals with different scales. If each blob is a fixed connected graph G, then $\lim_{n \to \infty} \tau / \sigma_2 \in [0, \infty)$. (The critical Erdős-Rényi random graph corresponds to the case where G is a single vertex.) In the applications below, $\sigma_2 \sim n^{\delta - 1/3}$ while $\tau \sim n^{2\delta - 1/3}$ which corresponds to the case where the above limit equals ∞. ∎

4. RESULTS FOR THE ASSOCIATED RANDOM GRAPH MODELS

Consider a sequence of edge weights $\boldsymbol{\beta}_n = \left(\beta_{ij}^{(n)}; i, j \in [n]\right)$, $n \geq 1$, as in Section 1.2.1. Let $Q_{ij}^{(n)} = \left((i-1)/n, i/n\right) \times \left((j-1)/n, j/n\right)$, $1 \leq i, j \leq n$. Let $W_n : [0, 1]^2 \to [0, \infty)$ be given by

$$W_n(x, y) = \begin{cases} \beta_{ij}^{(n)}, & \text{if } (x, y) \in Q_{ij}^{(n)}, \ 1 \leq i, j \leq n, \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

4.1. **The main result for graphons.** Recall the notation introduced in Section 1.2.1.

Condition 4.1. *The sequence* $(W_n; n \geq 1)$ *is uniformly bounded in* $L^3([0,1]^2)$, *i.e.,*

$$\sup_{n \geq 1} \int_{[0,1]^2} W_n^3 = \sup_{n \geq 1} \frac{1}{n^2} \sum_{i,j \in [n]} (\beta_{ij}^{(n)})^3 < \infty. \qquad (19)$$

Further, there exists an L^3 *graphon* W *that is critical (i.e.,* $\|T_W\|_{2,2} = 1$) *and a symmetric* $H \in L^2([0,1]^2)$ *such that*

$$\left\| T_{H-n^{1/3}(W_n-W)} \right\|_{2,2} = \left\| T_H - n^{1/3}(T_{W_n} - T_W) \right\|_{2,2} \to 0 \quad \text{as } n \to \infty. \qquad (20)$$

Let $D_n := \bigcup_{i=1}^n Q_{ii}^{(n)}$. Then it is easy to show that $n^{1/3} \|T_{W \cdot \mathbb{1}_{D_n}}\|_{2,2} \to 0$ as $n \to \infty$ for any $W \in L^3([0,1]^2)$. Further, $\|T_{H \cdot \mathbb{1}_{D_n}}\|_{2,2} \leq \left(\int_{D_n} H^2\right)^{1/2} \to 0$ as $n \to \infty$ for any $H \in L^2([0,1]^2)$. Consequently, the following relatively easier to check condition implies (20):

$$\int_{[0,1]^2 \setminus D_n} \left[H(x,y) - n^{1/3}\left(W_n(x,y) - W(x,y)\right) \right]^2 dx dy \to 0 \quad \text{as } n \to \infty. \qquad (21)$$

We will need some additional regularity conditions on the sequence $(\boldsymbol{\beta}_n)$.

Condition 4.2. *There exists* $\theta_1 \in (0, 1/8)$ *such that*

$$\max_{i \in [n]} \frac{1}{n} \sum_{j \in [n]} \left(\beta_{ij}^{(n)}\right)^{3/2} = o(n^{\theta_1}),$$

or equivalently, there exists $\theta_0 \in (0, 1/12)$ *such that*

$$\max_{x \in [0,1]} \left(\int_0^1 W_n^{3/2}(x,y) \, dy \right)^{2/3} = o(n^{\theta_0}).$$

Condition 4.3. *Let* θ_0 *be as in Condition 4.2. Then there exist* $\delta_0 \in (0, 1/3 - 2\theta_0)$ *and* $\varpi_0 \in [0, 2\delta_0 + 1/3)$ *such that the following hold: For each* $n \geq 1$, *there exists* $B_n \subseteq [n]$ *satisfying*

$$|B_n| = o(n^{\varpi_0}), \quad \sum_{i \in B_n} \sum_{j \in [n]} \left(\beta_{ij}^{(n)}\right)^2 = o(n^{4/3}), \quad \text{and} \qquad (22)$$

$$\#\left\{(i,j) : i, j \in [n] \setminus B_n \text{ and } \beta_{ij}^{(n)} \leq n^{-\delta_0}\right\} = o(n^{1+\delta_0}). \qquad (23)$$

Note that Condition 4.3 is trivially satisfied if there exists $\delta_0 \in (0, 1/3 - 2\theta_0)$ such that for all large n, $\beta_{ij}^{(n)} > n^{-\delta_0}$ for $1 \leq i \neq j \leq n$.

Remark 4. Before proceeding to the statements of our results, let us explain Conditions 4.1, 4.2, and 4.3. We first observe two simple facts: While studying the metric space structure of the graphs, the labels of the vertices do not play any special role. Hence, for a sequence of edge weights $(\boldsymbol{\beta}_n)$ satisfying Conditions 4.2 and 4.3, if we can find permutations π_n on $[n]$ such that Condition 4.1 is satisfied by the sequence of weight sequences $\left(\beta_{\pi_n(i),\pi_n(j)}^{(n)}; i, j \in [n]\right)$, then the conclusion in Theorem 4.5 stated below continues to hold. Next, note that if W is not a critical graphon, then we simply consider the random graph on $[n]$ where the edges $\{i, j\}$ appear independently with respective probabilities $1 \wedge (c\beta_{ij}^{(n)}/n)$ and $c = 1/\|T_W\|_{2,2}$. So, the assumption $\|T_W\|_{2,2} = 1$ enables us to avoid having this normalizing constant all throughout the paper.

The assumption that W is an L^3 graphon and the condition (19) are the analogues of the 'finite third moment condition' imposed on the degree distribution while studying the

configuration model and uniform (simple) graphs with given degree sequence; it is known [11, 12, 15, 30, 31, 33] that for the latter two models, under the finite third moment condition, one gets a Brownian scaling limit inside the critical window, whereas under finite p-th moment with $p \in (2, 3)$ but infinite third moment condition, the scaling limit inside the critical window is of a different nature. We expect to observe a similar phenomenon in the model studied in this paper when we replace (19) by L^p boundedness with $p \in (2, 3)$, and when the limiting graphon is L^p but not L^3. We will pursue this in future work.

Now, the Erdős-Rényi random graph ERRG$(n, t_n/n)$ exhibits critical behavior if there exists $\lambda \in \mathbb{R}$ such that $n^{1/3}(t_n - 1) \to \lambda$ as $n \to \infty$. Analogously, (20) can be written as

$$n^{1/3}(W_n - W) \to H \quad \text{as} \quad n \to \infty, \tag{24}$$

where the convergence is in the sense of (20). This describes the critical window for percolation on weighted graphs converging to an L^3 graphon. Here, the critical graphon W plays the role of the critical threshold $t_c = 1$ for the Erdős-Rényi random graph, and the function $H \in L^2([0, 1]^2)$ replaces λ–the location inside the critical window.

There is a huge body of work surrounding the convergence of graphons. Two important topologies with respect to which convergence is studied are the ones induced by the cut norm and the cut metric [20, 38, 47]. It is easy to see that (20) implies that the convergence in (24) holds with respect to the cut norm, and thus the latter induces a weaker topology than the one considered in (20); see [47, Lemma E.6] (this result is stated for bounded graphons; however, this particular implication does not require the boundedness assumption). It is not clear to us if convergence in the cut norm in (24) would, in general, be a strong enough assumption to deduce the metric scaling limit from. Note however that by [47, Lemma E.6], in the special case where $H \in L^\infty([0, 1]^2)$ and $\sup_{n,x,y} n^{1/3}|W_n(x, y) - W(x, y)| < \infty$, convergence in (24) with respect to the cut norm is equivalent to (20).

Condition 4.2 is similar in spirit to the assumption in [18, Display (3.11)]. We work with the smaller exponent '3/2' as opposed to the exponent '2' in [18, Display (3.11)]. Also, we allow the possibility $\sup_{x \in [0,1]} \int_0^1 W^{3/2}(x, y) dy = \infty$ provided $\max_{x \in [0,1]} \int_0^1 W_n^{3/2}(x, y) dy$ grows in a controlled way.

Condition 4.3 is a technical assumption that allows us to couple the random graph $\mathscr{G}_n^{\text{gphn}}$ with the multiplicative coalescent and apply Theorem 3.4. In applications, B_n is to be taken as the set of vertices $i \in [n]$ such that there are too many $j \in [n]$ for which $\beta_{ij}^{(n)}$ is too small. The assumption in (22) will ensure that the vertices in B_n do not have a significant influence on the intrinsic geometry of the maximal components in $\mathscr{G}_n^{\text{gphn}}$. The condition (23) says that there are not many pairs (i, j) with both i and j from outside of B_n such that $\beta_{ij}^{(n)}$ is too small. ∎

Since W is a symmetric function in $L^3([0, 1]^2) \subseteq L^2([0, 1]^2)$, T_W is a compact and self-adjoint operator on $L^2([0, 1])$. Now, $\|T_W\|_{2,2} = 1$ by assumption, and consequently, either $+1$ or -1 is an eigenvalue of T_W. In the present setting, we have the following in addition:

Lemma 4.4. *The following hold under Conditions* 4.1, 4.2, *and* 4.3:

(a) $\|T_W\|_{2,2} = 1$ *is an eigenvalue of* T_W.

(b) *The eigenspace corresponding to* 1 *is of the form* $\{c\psi : c \in \mathbb{R}\}$ *for an eigenfunction* ψ *that satisfies* $\psi(x) > 0$ *for* $\mu_\circ$ *a.e.* x *and* $\|\psi\|_2 = 1$. *(Clearly, such a* ψ *is unique.) Further,* $\int_0^1 \psi^3 < \infty$.

(c) *There exists* $\Delta_W \in (0,1]$ *such that* $|\lambda| \leq 1 - \Delta_W$ *for each eigenvalue* $\lambda \neq 1$ *of* T_W. *In other words,* -1 *is not an eigenvalue of* T_W.

We defer the proof of Lemma 4.4 to Section 6.1. Define

$$\alpha := \frac{1}{(\int_0^1 \psi)^2}, \quad \chi := \frac{\int_0^1 \psi^3}{(\int_0^1 \psi)^3}, \quad \text{and } \zeta := \frac{\langle \psi, T_H \psi \rangle}{(\int_0^1 \psi)^2}. \tag{25}$$

Theorem 4.5. *For* $i \geq 1$, *let* $\mathscr{C}_i(G_n^{gphn})$ *be the* i-*th largest component in* G_n^{gphn}, *and view it as a metric measure space using the counting measure. Under Conditions* 4.1, 4.2, *and* 4.3,

$$\left(scl\left(\frac{\chi^{2/3}}{\alpha n^{1/3}}, \frac{\chi^{1/3}}{n^{2/3}} \right) \mathscr{C}_i(G_n^{gphn}), \ i \geq 1 \right) \xrightarrow{d} \mathbf{Crit}\left(\frac{\zeta}{\chi^{2/3}} \right) \quad as \quad n \to \infty. \tag{26}$$

Theorem 4.5 implies the distributional convergence of the scaled sizes of the maximal components. However, it does not immediately show that the number of cycles in each maximal component forms a tight sequence, since there may be cycles of length $o(n^{1/3})$ that do not show up in the GHP limit. The next result shows the aforementioned tightness, and in fact, its proof will show that there are no cycles of length $o(n^{1/3})$ in each maximal component whp. Let $\Xi(\lambda)$ be as in the statement of Theorem 2.2.

Theorem 4.6. *Under Conditions* 4.1, 4.2, *and* 4.3,

$$\left(\left(n^{-2/3} \chi^{1/3} |\mathscr{C}_i(G_n^{gphn})|, \ spls\left(\mathscr{C}_i(G_n^{gphn}) \right) \right); \ i \geq 1 \right) \xrightarrow{d} \Xi\left(\chi^{-2/3} \zeta \right) \quad as \quad n \to \infty, \tag{27}$$

with respect to the product topology.

Remark 5. From our proofs, it will follow that Theorem 4.6 remains true if we weaken Conditions 4.2 and 4.3 to allow $\theta_0 \in (0, 1/6)$ (equivalently, $\theta_1 \in (0, 1/4)$) and $\delta_0 \in (0, 1/3 - \theta_0)$. The stronger assumptions in Conditions 4.2 and 4.3 are really needed for the metric scaling limit in Theorem 4.5. However, we state both results under the same set of conditions for simplicity.

Remark 6. It is known that $\gamma_1(\lambda)/\lambda \xrightarrow{d} 2$ as $\lambda \to \infty$; see [5, Lemma 5.6] for a proof (see also [2] for the analogue of this result for the multiplicative coalescent in the regime where the scaling limit is a pure-jump process). Thus, in the special case where $H = \lambda W$ for some $\lambda \in \mathbb{R}$, (27) implies that

$$\lim_{\lambda \to \infty} \lambda^{-1} \left(\lim_{n \to \infty} n^{-2/3} |\mathscr{C}_i(G_n^{gphn})| \right) = 2 \left(\int_0^1 \psi \right) \cdot \left(\int_0^1 \psi^3 \right)^{-1}, \tag{28}$$

where both limits are to be interpreted in the distribution sense. The analogue of (28) in the purely supercritical regime was proved in [18, Theorem 3.17 (ii)]. (The constant c_0 appearing in [18, Display (3.12)] is missing in (28) since we assume $\|T_W\|_{2,2} = 1$. Similarly, the factor $\int \psi^2$ in [18, Display (3.12)] is absent in (28) because we use the normalization $\|\psi\|_2 = 1$.) Thus, (28) generalizes [18, Theorem 3.17 (ii)] to the right side of the critical window.

4.2. **Some special cases of edge weights.** We discuss four special cases of edge weights below:

4.2.1. *The stochastic block model.* Theorem 4.5 can be applied to the important special case of the stochastic block model (SBM). However, for the SBM, instead of working with compact operators, one can just work with finite dimensional matrices. So, we can rework the proof of Theorem 4.5 in the simpler setting of matrices, and prove a result for the SBM that is slightly more general than the one we get as a corollary of Theorem 4.5. We will state this result without proof in Section 7.

4.2.2. *The edge weights $W(i/n, j/n)$.* Suppose $K \in L^3([0,1]^2)$ is a nonnegative, symmetric kernel, and assume that

$$W = \|T_K\|_{2,2}^{-1} \cdot K. \tag{29}$$

When K has reasonable continuity properties, a natural class of edge weights to study would be $W(i/n, j/n), 1 \le i \ne j \le n$. Incorporating the location in the critical window, we consider the more general setting where the edge weights are given by

$$\beta_{ii}^{(n)} = 0 \;\; \text{for} \;\; i \in [n] \;\; \text{and} \;\; \beta_{ij}^{(n)} = \left(W\left(\tfrac{i}{n}, \tfrac{j}{n}\right) + n^{-1/3} H\left(\tfrac{i}{n}, \tfrac{j}{n}\right)\right) \vee 0 \;\; \text{for} \;\; 1 \le i \ne j \le n. \tag{30}$$

For simplicity, we assume that either $H = \lambda W$ for some $\lambda \in \mathbb{R}$, or H is continuous on $[0,1]^2$.

Clearly, if K is bounded away from zero and Hölder(γ) on $[0,1]^2$ with $\gamma > 1/3$, then Conditions 4.1, 4.2, and 4.3 are satisfied by the edge weights in (30), and consequently, Theorems 4.5 and 4.6 are applicable in such cases. Some specific examples where Theorems 4.5 and 4.6 are applicable are

$$\eta + (x \vee y)^a, \;\; \text{or} \;\; \eta \vee (x \vee y)^a, \;\; \text{or} \;\; (\eta + x + y)^a, \;\; \text{or} \;\; \eta + (x + y)^a, \;\; \text{for any} \;\; a, \eta > 0. \tag{31}$$

(Note that the above Hölder condition is not met in some of these examples.)

Now, suppose K is either unbounded, or is not bounded away from zero, or both. We briefly discuss applications of Theorems 4.5 and 4.6 to such graphons via a few simple examples. Consider the family of unbounded kernels

$$K(x, y) = (x \vee y)^{-a}, \;\; \text{or} \;\; (x + y)^{-a}, \;\; \text{or} \;\; \left(x^a + y^a\right)^{-1}, \;\; 0 < a < 2/3. \tag{32}$$

See Sections 3.6, 16.1, 16.3, and 16.6 of [18] for a survey of the works related to inhomogeneous random graphs with kernels that are functions of $(x \vee y)$. Note that the kernels in (32) are not in $L^3([0,1]^2)$ if $a \ge 2/3$. It can be easily checked that Conditions 4.1, 4.2, and 4.3 are satisfied by the edge weights in (30) when K is as in (32), and consequently, Theorems 4.5 and 4.6 are applicable in these cases. Next, take the unbounded kernels $|x - y|^{-a}$, which are in $L^3([0,1]^2)$ if $0 < a < 1/3$. Then Theorems 4.5 and 4.6 apply when

$$K(x, y) := |x - y|^{-a}, \;\; 0 < a < 1/6. \tag{33}$$

Some examples of kernels K that are not bounded away from zero to which Theorems 4.5 and 4.6 are applicable are

$$K(x, y) = (x \wedge y)^a, \;\; a > 0, \tag{34}$$

$$K(x, y) = (x \vee y)^a, \;\; 0 \le a < 1, \;\; \text{and} \tag{35}$$

$$K(x, y) = (x + y)^a, \;\; 0 \le a < 1. \tag{36}$$

Now, unlike (31) or (34), Theorem 4.5 cannot be used directly to get the scaling limit for $K(x, y) = (x \vee y)^a$ or $(x + y)^a$ when $a \geq 1$ and the edge weights are as in (30). However, Theorem 4.5 is applicable to these kernels when the edge weights are $\beta_{ij}^{(n)} \vee n^{-\delta}$ where $\beta_{ij}^{(n)}$ is as in (30) and $\delta \in (1/6, 1/3)$. Thus, to get the metric scaling limit in such cases, it would suffice to establish the scaling limit of the maximal component sizes and their surplus edges. Take, for example, the kernel $x \vee y$. Note that $\|T_{x \vee y}\|_{2,2} =: z_0$ is the unique positive solution to $\tanh(1/\sqrt{z}) = \sqrt{z}$, and the corresponding (one-dimensional) eigenspace is spanned by

$$\tilde{\psi}(x) = \frac{\sqrt{2}}{\cosh(1/z_0^{1/2})} \cdot \cosh(x/z_0^{1/2}), \quad x \in [0, 1]. \tag{37}$$

Recall the definition of $\Xi(\lambda)$ from the statement of Theorem 2.2.

Conjecture 4.7. *Fix $\lambda \in \mathbb{R}$, and consider G_n^{gphn} constructed using the edge weights*

$$\tilde{\beta}_{ij}^{(n)} = \frac{1}{z_0}\Big(1 + \frac{\lambda}{n^{1/3}}\Big) \cdot \Big(\frac{i}{n} \vee \frac{j}{n}\Big), \quad 1 \leq i \neq j \leq n. \tag{38}$$

Then

$$\Big(\Big(\frac{\tilde{\chi}^{1/3}}{n^{2/3}}\big|\mathscr{C}_i(G_n^{gphn})\big|, \; \mathrm{spls}\big(\mathscr{C}_i(G_n^{gphn})\big)\Big); \; i \geq 1\Big) \xrightarrow{\mathrm{d}} \Xi\Big(\frac{\tilde{\zeta}}{\tilde{\chi}^{2/3}}\Big) \quad as \quad n \to \infty \tag{39}$$

with respect to the product topology, where $\tilde{\chi} = (\int_0^1 \tilde{\psi}^3)(\int_0^1 \tilde{\psi})^{-3}$ and $\tilde{\zeta} = \lambda(\int_0^1 \tilde{\psi})^{-2}$ with $\tilde{\psi}$ as in (37).

Now, for any $\delta \in (1/6, 1/3)$, Theorem 4.5 and Theorem 4.6 show that the convergences in (26) and (27) hold when the edge weights are $\tilde{\beta}_{ij}^{(n)} \vee n^{-\delta}$, where $\tilde{\beta}_{ij}^{(n)}$ is as in (38). This can be used in conjuction with the convergence in (39) and a coupling argument to get the metric scaling limit when the edge weights are as in (38).

Finally, as examples of kernels that are neither bounded, nor bounded away from zero consider

$$K(x, y) = \frac{1}{(x \vee y)^a} - 1, \quad \text{or} \quad \Big(\frac{1}{x \vee y} - 1\Big)^a, \quad \text{or} \quad \Big(\frac{2}{x + y} - 1\Big)^a, \quad 0 < a < 2/3. \tag{40}$$

As discussed in [18, Section 16.3], the first two kernels in (40) with $a = 1$ correspond to the CHKNS model [26] and a related process due to Durrett [35]. Note also that the kernels in (40) are not in $L^3([0, 1]^2)$ when $a \geq 2/3$. Here, again it is routine to verify that Conditions 4.1, 4.2, and 4.3 are satisfied.

4.2.3. *Edge weights sampled using uniform random variables.* For a graphon W, a natural collection of edge weights that is widely studied in the statistics literature is $W(U_i, U_j)$, $1 \leq i \neq j \leq n$, where $U_1, \ldots, U_n$ are i.i.d. Uniform[0, 1] random variables. As explained at the beginning of Remark 4, a permutation of the vertex labels does not make any difference, so we can instead take the edge weights as $W(V_i^{(n)}, V_j^{(n)})$, $1 \leq i \neq j \leq n$, where $V_1^{(n)} \leq \cdots \leq V_n^{(n)}$ are $U_1, \ldots, U_n$ arranged in nondecreasing order. We will work in a slightly more general setting and take the edge weights as

$$\beta_{ii}^{(n)} = 0 \;\; \text{for} \;\; i \in [n] \;\; \text{and} \;\; \beta_{ij}^{(n)} = \Big(W\big(V_i^{(n)}, V_j^{(n)}\big) + \frac{1}{n^{1/3}} \cdot H\big(V_i^{(n)}, V_j^{(n)}\big)\Big) \vee 0, \;\; 1 \leq i \neq j \leq n, \tag{41}$$

where $H : [0, 1]^2 \to \mathbb{R}$ is symmetric and Borel measurable.

4.2.4. *Edge weights that provide the best L^2-approximation.* For $f \in L^1([0,1]^2)$, let $f_{Q_{ij}^{(n)}} f :=$
$n^2 \int_{Q_{ij}^{(n)}} f$ denote the average of f over $Q_{ij}^{(n)}$, $i, j \in [n]$. Then, in view of Condition 4.1 and
(21), another natural choice for the edge weights would be

$$\beta_{ii}^{(n)} = 0 \quad \text{for } i \in [n] \quad \text{and} \quad \beta_{ij}^{(n)} = \left(\fint_{Q_{ij}^{(n)}} W + \frac{1}{n^{1/3}} \fint_{Q_{ij}^{(n)}} H \right) \vee 0 \quad \text{for } 1 \le i \ne j \le n, \quad (42)$$

since for given W and H, the function W_n as given by (18) corresponding to the edge
weights in (42) gives the best L^2-approximation for $(W + n^{-1/3} H)$ among all functions that
are zero on D_n, and constant and nonnegative on $Q_{ij}^{(n)}$ for $1 \le i \ne j \le n$.

In the next two sections, we formulate regularity conditions for W and H under which
Theorems 4.5 and 4.6 are applicable to the collections of edge weights in (41) and (42). In
these regularity assumptions, we eschew complete generality in favor of relatively simpler
conditions.

4.3. **Special case: the edge weights in** (41). For $\varepsilon > 0$ and $K : [0,1]^2 \to [0,\infty)$, let $\mathrm{Osc}_{\varepsilon;K} :$
$[0,1]^2 \to [0,\infty]$ be given by

$$\mathrm{Osc}_{\varepsilon;K}(x,y) := \sup \left\{ |K(x',y') - K(x,y)| : (x',y') \in [0,1]^2 \ \text{and} \ |x - x'| + |y - y'| \le |\log \varepsilon| \sqrt{\varepsilon} \right\}.$$

Condition 4.8. *Assume that W is an L^3 graphon satisfying the following:*

(a) *For each $\varepsilon > 0$, there exists a Borel $\mathscr{A}_\varepsilon \subseteq [0,1]^2$ such that*

$$\int_{\widetilde{\mathscr{A}_\varepsilon}} W^2(x,y) dx dy + \int_{[0,1]^2 \setminus \mathscr{A}_\varepsilon} \mathrm{Osc}_{\varepsilon;W}^2(x,y) dx dy = o(\varepsilon^{2/3}) \ \text{as } \varepsilon \to 0,$$

where $\widetilde{\mathscr{A}_\varepsilon}$ is the set of points in $[0,1]^2$ such that there is at least one point in $\mathscr{A}_\varepsilon$ at distance at most $(|\log \varepsilon| \sqrt{\varepsilon})$ from them.

(b) *There exists $p_1 > 8$ such that*

$$\int_0^1 \left(\int_0^1 W^{3/2}(x,y) dy \right)^{p_1} dx < \infty.$$

(c) *There exist $\delta_0 \in (0, 1/3 - 4/(3p_1))$ and $\varpi_0 \in [0, 2\delta_0 + 1/3)$ such that the following hold:*
For each $\varepsilon > 0$, there exists a Borel $\mathscr{B}_\varepsilon \subseteq [0,1]$ satisfying

$$\mu_\circ(\mathscr{B}_\varepsilon) = o(\varepsilon^{1-\varpi_0}), \quad \int_{\mathscr{B}_\varepsilon \times [0,1]} W^2(x,y) dx dy = o(\varepsilon^{2/3}), \quad \text{and} \quad (43)$$

$$\mu_\square \left(\left\{ (x,y) \in ([0,1] \setminus \mathscr{B}_\varepsilon)^2 : W(x,y) \le \varepsilon^{\delta_0} \right\} \right) = o(\varepsilon^{1-\delta_0}) \quad \text{as } \varepsilon \to 0. \quad (44)$$

Note that if $\inf_{x,y \in [0,1]} W(x,y) > 0$, then Condition 4.8 (c) is satisfied trivially with $\mathscr{B}_\varepsilon = \emptyset$
for all $\varepsilon > 0$, any $\delta_0 \in (0, 1/3 - 4/(3p_1))$, and any $\varpi_0 \in [0, 2\delta_0 + 1/3)$. Condition 4.8 (a)
is only needed while verifying that the convergence in (20) holds in distribution when
$H \in L^p([0,1]^2)$ for some $p > 2$ and the edge weights are given by (41); see the proof of
Theorem 4.9 given in Section 6.12. For certain kernels that do not satisfy Condition 4.8 (a),
it may still be possible to directly verify that (20) holds in distribution; see e.g., the discus-
sion below (48). We expect that Condition 4.8 (a) can be substantially weakened, and we
will come back to this point in Section 7.

Theorem 4.9. *Consider G_n^{gphn} constructed using edge weights of the form* (41). *If W is a
critical L^3 graphon that satisfies Condition 4.8 and either*

(a) $H = \lambda W$ *for some* $\lambda \in \mathbb{R}$, *or*

(b) $H^+ \in L^3([0,1]^2)$ *and* $H^- \in L^q([0,1]^2)$ *where* $q = q(\delta_0) := 3(1 - \delta_0)/(1 - 3\delta_0)$,

then the convergences in (26) *and* (27) *hold with the same constants as in* (25).

If further $\inf_{x,y\in[0,1]} W(x, y) > 0$, *then the convergences in* (26) *and* (27) *hold with the same constants as in* (25) *for any* $H \in L^3([0,1]^2)$.

The next result follows easily from Theorem 4.9.

Corollary 4.10. *Suppose* $K : [0,1]^2 \to [0,\infty)$ *is a symmetric, Borel measurable function that is bounded, bounded away from zero (i.e.,* $\inf_{x,y\in[0,1]} K(x, y) > 0$*), and satisfies*

$$\int_{[0,1]^2} \mathrm{Osc}^2_{\varepsilon;K}(x, y)\,dx\,dy = o(\varepsilon^{2/3}) \ as \ \varepsilon \to 0. \tag{45}$$

Let $W = \|T_K\|_{2,2}^{-1} \cdot K$, *and consider* G_n^{gphn} *constructed using edge weights of the form* (41), *where* $H \in L^3([0,1]^2)$. *Then the convergences in* (26) *and* (27) *hold with the same constants as in* (25).

In particular, Corollary 4.10 is applicable to any K that is Hölder(γ) on $[0,1]^2$ with $\gamma > 2/3$ and is bounded away from zero. However, (45) may hold even when the above Hölder condition does not. For example, Corollary 4.10 applies to the kernels in (31).

Now, as we did in Section 4.2.2, consider W as in (29), where $K : [0,1]^2 \to [0,\infty)$ is a symmetric function in $L^3([0,1]^2)$ that is either unbounded, or is not bounded away from zero, or both. We briefly discuss applications of Theorem 4.9 to such graphons via a few examples.

Consider the family of unbounded kernels

$$K(x, y) = (x \vee y)^{-a}, \quad 0 < a < 1/3. \tag{46}$$

Note that the kernel $(x \vee y)^{-a}$ is unbounded and in $L^3([0,1]^2)$ iff $0 < a < 2/3$. However, Theorem 4.9 applies only when $0 < a < 1/3$. To see that Theorem 4.9 can be applied when K is as in (46), it is enough to verify that K satisfies Condition 4.8. To this end, choose $b \in \big((3(1 - a))^{-1}, 1/2\big)$, and for each $\varepsilon \in (0, 1)$, set

$$\mathscr{A}_\varepsilon = [0, \varepsilon^b] \times [0, \varepsilon^b] \ \ \text{and} \ \ \mathscr{B}_\varepsilon = \emptyset.$$

Then Condition 4.8 is satisfied with these choices of $\mathscr{A}_\varepsilon, \mathscr{B}_\varepsilon$, and any $p_1 > 8$, $0 < \delta_0 < 1/3 - 4/(3p_1)$ and $\varpi_0 \in [0, 2\delta_0 + 1/3]$. Similar calculations will show that Condition 4.8 is also satisfied when working with the kernels

$$(x + y)^{-a}, \ \ 0 < a < 1/3, \tag{47}$$

and consequently, Theorem 4.9 is applicable.

Now, let us look at two examples where K is not bounded away from zero. Consider the family of kernels

$$K(x, y) = (x \wedge y)^a, \ \ a > 0. \tag{48}$$

The kernel $x \wedge y$ is particularly important to us as it is related to the RGIV model; see the proof of Theorem 4.15 stated below. First, consider the case $a > 1/6$. To see that the kernels

in (48) satisfy Condition 4.8 when $a > 1/6$, we choose

$$\delta_0 \in \left(\frac{2a}{3(2a+1)}, \frac{1}{3} \right), \quad \varpi_0 \in \left((1 - \delta_0/a) \vee 0, \, 2\delta_0 + 1/3 \right),$$

and for each $\varepsilon \in (0,1)$, set $\mathscr{A}_\varepsilon = \emptyset$ and $\mathscr{B}_\varepsilon = [0, \varepsilon^{\delta_0/a}]$. Then Condition 4.8 is satisfied with these choices of $\mathscr{A}_\varepsilon, \mathscr{B}_\varepsilon, \delta_0, \varpi_0$, and any $p_1 > 8$ such that $\delta_0 < 1/3 - 4/(3p_1)$. When $a \in (0, 1/6]$, Condition 4.8 (a) is not satisfied. However, as mentioned right below Condition 4.8, for the kernels in (48), we can directly verify that (20) holds if H is as in the statement of Theorem 4.9. In fact, some straightforward calculation will show that convergence in distribution in (20) follows in this case if we can show that

$$n^{-1} \sum_{i \in [n]} \left((V_i^{(n)})^a - i^a/n^a \right)^2 = o_P(n^{-2/3}),$$

and this can be proved easily. Thus, the conclusion of Theorem 4.9 remains true when K is as in (48).

Next, we consider the kernels in (35). Fix any $\delta_0 \in \left(a/(2+a), 1/3 \right)$. Then Condition 4.8 is satisfied with this choice of δ_0, $\mathscr{A}_\varepsilon = \emptyset, \mathscr{B}_\varepsilon = \emptyset$, and any $\varpi_0 \in [0, 2\delta_0 + 1/3)$ and any $p_1 > 8$ such that $\delta_0 < 1/3 - 4/(3p_1)$. Similar calculations show that Theorem 4.9 is applicable to the kernels in (36). Further, the analogue of Conjecture 4.7 and the discussion around Conjecture 4.7 are valid here as well.

We now look at a few examples of kernels that are both unbounded and not bounded away from zero. Consider

$$K(x,y) = \frac{1}{(x \vee y)^a} - 1, \quad 0 < a < 1/3. \tag{49}$$

Kernels of this form are in $L^3([0,1]^2)$ when $0 < a < 2/3$. As was the case in (46) and (47), Theorem 4.9 applies to such kernels when $0 < a < 1/3$. To verify Condition 4.8 for the kernels in (49), choose $b \in \left((3(1-a))^{-1}, 1/2 \right)$, $\delta_0 \in (2/9, 1/3)$, $\varpi_0 \in (1 - \delta_0, 2\delta_0 + 1/3)$, and let

$$\mathscr{A}_\varepsilon = [0, \varepsilon^b] \times [0, \varepsilon^b] \quad \text{and} \quad \mathscr{B}_\varepsilon = \left[1 - 2\varepsilon^{\delta_0}/a, 1 \right].$$

Then Condition 4.8 is satisfied with these choices of $\mathscr{A}_\varepsilon, \mathscr{B}_\varepsilon, \delta_0, \varpi_0$, and any $p_1 > 8$ such that $\delta_0 < 1/3 - 4/(3p_1)$. Similarly, Theorem 4.9 is applicable when

$$K(x,y) = \left(\frac{1}{x \vee y} - 1 \right)^a \quad \text{or} \quad \left(\frac{2}{x+y} - 1 \right)^a, \quad 0 < a < 1/3.$$

4.4. **Special case: the edge weights in** (42)**.** Let us now turn to the edge weights in (42), and formulate sufficient conditions for Theorem 4.5 to be applicable.

Condition 4.11. *Assume that W is an L^3 graphon satisfying the following:*

(a) *We have,*

$$\sum_{1 \le i < j \le n} \int_{Q_{ij}^{(n)}} \left(W - \fint_{Q_{ij}^{(n)}} W \right)^2 = o(n^{-2/3}). \tag{50}$$

(b) *Condition 4.8 (b) holds.*

(c) *Let p_1 be as in* (b). *Then there exist $\delta_0 \in (0, 1/3 - 4/(3p_1))$ and $\varpi_0 \in [0, 2\delta_0 + 1/3)$ such that the following hold: For each n there exists $B_n \subseteq [n]$ such that*

$$|B_n| = o(n^{\varpi_0}), \quad \int_{[0,1]\times\mathscr{B}_{1/n}} W^2(x,y)\,dx\,dy = o(n^{-2/3}), \quad and \tag{51}$$

$$\mu_\square\big(\{(x,y) \in ([0,1]\setminus\mathscr{B}_{1/n})^2 : W(x,y) \le n^{-\delta_0}\}\big) = o(n^{\delta_0-1}) \quad as\ n \to \infty, \tag{52}$$

where $\mathscr{B}_{1/n} := \bigcup_{i\in B_n} ((i-1)/n, i/n)$.

Clearly, if W is Hölder(γ) on $[0,1]^2$ with $\gamma > 1/3$, then W satisfies (50). Further, if W has weak partial derivatives $D_x W$ and $D_y W$ on $(0,1)^2$, then by the Poincaré inequality for functions in Sobolev spaces,

$$\int_{Q_{ij}^{(n)}} \left(W - \fint_{Q_{ij}^{(n)}} W\right)^2 \le C n^{-2} \int_{Q_{ij}^{(n)}} \left[(D_x W)^2 + (D_y W)^2\right],$$

where C is a universal constant. (See [36, Chapter 5.8.1, Theorem 2] where this inequality is stated for balls. It also holds for square regions, as squares satisfy the so-called cone condition.) Finally, we have the trivial bound: $\int_{Q_{ij}^{(n)}} \left(W - \fint_{Q_{ij}^{(n)}} W\right)^2 \le \int_{Q_{ij}^{(n)}} W^2$. We collect these observations in the next lemma.

Lemma 4.12. *An L^3 graphon W satisfies* (50) *if either*

(a) *W is Hölder(γ) on $[0,1]^2$ with $\gamma > 1/3$, or if*
(b) *W has weak partial derivatives $D_x W$ and $D_y W$ on $(0,1)^2$ and*

$$\sum_{1\le i < j \le n} \min\left\{\int_{Q_{ij}^{(n)}} W^2, \frac{1}{n^2}\int_{Q_{ij}^{(n)}} \left[(D_x W)^2 + (D_y W)^2\right]\right\} = o(n^{-2/3}). \tag{53}$$

Theorem 4.13. *Consider G_n^{gphn} constructed using edge weights of the form* (42). *If W is a critical L^3 graphon satisfying Condition* 4.11 *and either*

(a) *$H = \lambda W$ for some $\lambda \in \mathbb{R}$, or*
(b) *$H^+ \in L^3([0,1]^2)$ and $H^- \in L^q([0,1]^2)$ where $q = q(\delta_0) := 3(1-\delta_0)/(1-3\delta_0)$,*

then the convergences in (26) *and* (27) *hold with the same constants as in* (25).

If further $\inf_{x,y\in[0,1]} W(x,y) > 0$, then the convergences in (26) *and* (27) *hold with the same constants as in* (25) *for any $H \in L^3([0,1]^2)$.*

The following result is immediate from Theorem 4.13:

Corollary 4.14. *Suppose W is a critical L^3 graphon that satisfies $\inf_{x,y\in[0,1]} W(x,y) > 0$. Assume further that at least one of the following two conditions hold:*

(a) *W is Hölder(γ) on $[0,1]^2$ with $\gamma > 1/3$.*
(b) *W satisfies Condition* 4.8 (b) *and the condition in Lemma* 4.12 (b).

Consider G_n^{gphn} constructed using edge weights of the form (42), *where $H \in L^3([0,1]^2)$. Then the convergences in* (26) *and* (27) *hold with the same constants as in* (25).

For some specific examples, note that Theorem 4.13 is applicable to W of the form (29), where K is any of the kernels in (31), (32), (33), (34), (35), (36), and (40) (here, we can use Lemma 4.12 (b) to verify that (50) is satisfied).

4.5. **The metric scaling limit of Aldous and Pittel's RGIV process.** We now state the result on the metric scaling limit of the RGIV process inside the critical window; recall the notation from Section 1.2.2.

Theorem 4.15. *Fix $\lambda \in \mathbb{R}$. For $i \geq 1$, let $\mathscr{C}_i\big(\mathscr{G}_n^{rgiv}(\lambda)\big)$ be the i-th largest component in $\mathscr{G}_n^{rgiv}(\lambda)$, and view it as a metric measure space using the counting measure. Then*

$$\left(\text{scl}\left(\frac{a_1^{rgiv}}{n^{1/3}}, \frac{a_2^{rgiv}}{n^{2/3}} \right) \mathscr{C}_i\big(\mathscr{G}_n^{rgiv}(\lambda)\big), i \geq 1 \right) \xrightarrow{\text{d}} \mathbf{Crit}\big(a_3^{rgiv} \cdot \lambda\big) \quad as \quad n \to \infty, \tag{54}$$

where

$$a_1^{rgiv} = \frac{2^{8/3}}{3^{2/3} \cdot \pi}, \quad a_2^{rgiv} = \frac{2^{1/3}}{3^{1/3}}, \quad and \quad a_3^{rgiv} = \frac{3^{2/3}}{2^{2/3}}.$$

5. PROOFS: UNIVERSALITY

In this section we prove Theorem 3.4. Sections 5.1–5.3 collect some preliminary results related to the model. Section 5.4 studies the scaling limit of a related model of a *connected* random graph. Building on this result, Section 5.5 completes the proof.

5.1. **Size-biased random order.** Given a collection of positive weights $\mathbf{x} = (x_i, i \in [n])$, a *size-biased random order* of $[n]$ is a random permutation $(v(1), v(2), \ldots, v(n))$ such that $\mathbb{P}(v(1) = k) \propto x_k$ for $k \in [n]$, and for $2 \leq i \leq n$ and $k \in [n] \setminus \{v(1), \ldots, v(i-1)\}$,

$$\mathbb{P}\big(v(i) = k | v(1), v(2), \ldots, v(i-1)\big) \propto x_k. \tag{55}$$

One construction of such an ordering is as follows: Generate independent exponentials $\xi_i \sim \text{Exp}(x_i)$, $i \in [n]$, and arrange them in increasing order as $\xi_{v(1)} < \xi_{v(2)} < \cdots < \xi_{v(n)}$. Then $(v(1), \ldots, v(n))$ would be a size-biased permutation of $[n]$ using the weight sequence $\mathbf{x}$. In [7, Section 3.1] Aldous used this to construct $\mathscr{G}(\mathbf{x}, q)$ simultaneously with a breadth-first exploration of the graph such that the vertices in the exploration appeared in a size-biased random order. We describe this construction next. Let $\xi_{i,j} \sim \text{Exp}(qx_j)$, $i \neq j \in [n]$, be independent random variables. The exploration process initializes by selecting a vertex $v(1)$ with $\mathbb{P}(v(1) = k) \propto x_k$ for $k \in [n]$. The neighbors (sometimes referred to as children) of $v(1)$ are the vertices $\{i : \xi_{v(1),i} \leq x_{v(1)}\}$. Write $c(1)$ for the number of children of $v(1)$, labeled as $v(2), v(3), \ldots, v(c(1)+1)$ in increasing order of the $\xi_{v(1),v(i)}$ values. Now move to $v(2)$ and obtain the children of $v(2)$ as the vertices $i \notin \{v(1), \ldots, v(c(1)+1)\}$ such that $\xi_{v(2),i} \leq x_{v(2)}$. Label them as $v(c(1)+2), \ldots, v(c(1)+c(2)+1)$ in increasing order of their $\xi_{v(2),i}$ values. Proceed recursively until the component of $v(1)$ has been explored. Then select a new vertex amongst the unexplored vertices with probability proportional to their weights and proceed until all vertices have been explored.

5.2. **Partition into connected components.** We mainly follow [16]. Recall that $(\mathscr{C}_i, i \geq 1)$ denoted the ranked components of $\mathscr{G}(\mathbf{x}, q)$, and let $\mathscr{V}_i := V(\mathscr{C}_i)$. Thus, $(\mathscr{V}_i, i \geq 1)$ is a random partition of the vertex set $[n]$. For $\mathscr{V} \subseteq [n]$ write $\mathbb{G}_{\mathscr{V}}^{con}$ for the space of simple connected graphs with vertex set $\mathscr{V}$. Given $a > 0$ and a pmf $\mathbf{p} = (p_v, v \in \mathscr{V})$, define the probability distribution $\mathbb{P}_{con}(\cdot; \mathbf{p}, a, \mathscr{V})$ on $\mathbb{G}_{\mathscr{V}}^{con}$ via

$$\mathbb{P}_{con}(G; \mathbf{p}, a, \mathscr{V}) \propto \prod_{(u,v) \in E(G)} (1 - e^{-ap_u p_v}) \prod_{(u,v) \notin E(G)} e^{-ap_u p_v}. \tag{56}$$

For $i \geq 1$, define

$$\mathbf{p}_i = (p_{i,v}, v \in \mathcal{V}_i) := \big(x_v / \sum_{v \in \mathcal{V}_i} x_v, \ v \in \mathcal{V}_i\big) \ \text{and} \ a_i := q \cdot \big(\sum_{v \in \mathcal{V}_i} x_v\big)^2. \tag{57}$$

Proposition 5.1 ([16, Proposition 6.1]). *Let N be the number of components in $\mathcal{G}(\mathbf{x}, q)$. Then for any $G_i \in \mathbb{G}_{\mathcal{V}_i}^{\mathrm{con}}, 1 \leq i \leq N$,*

$$\mathbb{P}\left(\mathcal{C}_i = G_i, \ \forall 1 \leq i \leq N \mid N, (\mathcal{V}_i)_{1 \leq i \leq N}\right) = \prod_{1 \leq i \leq N} \mathbb{P}_{\mathrm{con}}(G_i; \mathbf{p}_i, a_i, \mathcal{V}_i).$$

5.3. **Tilted p-trees and scaling limits of components.** We recall here a result from [16] on scaling limits of graphs with distribution $\mathbb{P}_{\mathrm{con}} = \mathbb{P}_{\mathrm{con}}(\cdot; \mathbf{p}, a, [m])$ under regularity assumptions on $a = a^{(m)} > 0$ and the pmfs $\mathbf{p} = \mathbf{p}^{(m)} = (p_i^{(m)}, 1 \leq i \leq m)$. We suppress dependence on m to simplify notation. Let $\sigma(\mathbf{p}) := (\sum_{i \in [m]} p_i^2)^{1/2}$, $p_{\max} := \max_{i \in [m]} p_i$, and $p_{\min} := \min_{i \in [m]} p_i$. Suppose $\mathcal{G}^{\mathbf{p}} \sim \mathbb{P}_{\mathrm{con}}$. Recall the notation $\tilde{\mathbf{e}}^\theta$ from Section 2.3.

Theorem 5.2 ([16, Theorem 7.3]). *Assume that there exist $\bar{\gamma}, r_0, \eta_0 \in (0, \infty)$ such that*

$$\sigma(\mathbf{p}) \to 0, \ \frac{p_{\max}}{[\sigma(\mathbf{p})]^{3/2+\eta_0}} \to 0, \ \frac{[\sigma(\mathbf{p})]^{r_0}}{p_{\min}} \to 0, \ \text{and} \ a\sigma(\mathbf{p}) \to \bar{\gamma}, \tag{58}$$

as $m \to \infty$. View $\mathcal{G}^{\mathbf{p}}$ as a metric measure space by endowing it with the graph distance and the probability measure that assigns mass p_i to the vertex i. Then

$$\sigma(\mathbf{p}) \cdot \mathcal{G}^{\mathbf{p}} = \mathrm{scl}\big(\sigma(\mathbf{p}), 1\big) \cdot \mathcal{G}^{\mathbf{p}} \xrightarrow{\mathrm{d}} \mathcal{G}(2\tilde{\mathbf{e}}^{\bar{\gamma}}, \bar{\gamma}\tilde{\mathbf{e}}^{\bar{\gamma}}, \mathcal{P}),$$

where the metric measure space $\mathcal{G}(2\tilde{\mathbf{e}}^{\bar{\gamma}}, \bar{\gamma}\tilde{\mathbf{e}}^{\bar{\gamma}}, \mathcal{P})$ is as defined in Section 2.3.

We recall from [16] an ingredient in the proof of this theorem that will be required in this paper. Write $\mathbb{T}_m^{\mathrm{ord}}$ for the space of ordered rooted trees with vertex set $[m]$. Here, "ordered" means that the children of every vertex is arranged from "left" to "right." (In other words, these are rooted labeled trees with a plane embedding.) Let $\mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}$ with root ϕ. For a vertex $v \neq \phi$, let $(\phi = v_0, v_1, \ldots, v_k = v)$ be the path in $\mathbf{t}$ from ϕ to v. Let $R(v; \mathbf{t})$ be the set of all vertices u such that for some $0 \leq i \leq k - 1$, u is a child of v_i and u appears on the right side of v_{i+1} in the plane embedding. Define the collection of *permitted edges* to be

$$\mathcal{P}(\mathbf{t}) := \big\{\{v, u\} : v \in [m] \setminus \phi, u \in R(v; \mathbf{t})\big\}.$$

For $\mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}$ and $v \in [m]$, write $d_v(\mathbf{t})$ for the number of children of v in $\mathbf{t}$. Define a probability measure $\mathbb{P}_{\mathrm{ord}}(\cdot) = \mathbb{P}_{\mathrm{ord}}(\cdot; \mathbf{p})$ on $\mathbb{T}_m^{\mathrm{ord}}$ via

$$\mathbb{P}_{\mathrm{ord}}(\mathbf{t}) = \prod_{v \in [m]} p_v^{d_v(\mathbf{t})} / (d_v(\mathbf{t}))!, \qquad \mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}. \tag{59}$$

A random tree with distribution $\mathbb{P}_{\mathrm{ord}}$ is a $\mathbf{p}$-tree ([8, 54]) together with a plane embedding. Define $L : \mathbb{T}_m^{\mathrm{ord}} \to \mathbb{R}_+$ by

$$L(\mathbf{t}) := \Big(\prod_{\{i,j\} \in E(\mathbf{t})} \Big[\frac{\exp(ap_i p_j) - 1}{ap_i p_j}\Big]\Big) \exp\Big(\sum_{\{i,j\} \in \mathcal{P}(\mathbf{t})} ap_i p_j\Big). \tag{60}$$

Define the *tilted* $\mathbf{p}$-tree distribution $\tilde{\mathbb{P}}_{\mathrm{ord}}(\cdot) = \tilde{\mathbb{P}}_{\mathrm{ord}}(\cdot; \mathbf{p}, a)$ via

$$\frac{d\tilde{\mathbb{P}}_{\mathrm{ord}}}{d\mathbb{P}_{\mathrm{ord}}}(\mathbf{t}) = \frac{L(\mathbf{t})}{\mathbb{E}_{\mathrm{ord}}[L]}, \qquad \mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}, \tag{61}$$

where $\mathbb{E}_{\mathrm{ord}}$ is the expectation operator with respect to $\mathbb{P}_{\mathrm{ord}}$.

Proposition 5.3 ([16, Proposition 7.4])**.** *Fix a pmf* $\mathbf{p}$ *on* $[m]$ *and* $a > 0$. *Then* $\mathcal{G}^{\mathbf{p}} \sim \mathbb{P}_{\mathrm{con}}$ *can be constructed as follows: Generate* $\mathcal{T}^{\mathbf{p}} \sim \tilde{\mathbb{P}}_{\mathrm{ord}}$. *Conditional on* $\mathcal{T}^{\mathbf{p}}$, *add each permitted edge* $\{u, v\} \in \mathcal{P}(\mathcal{T}^{\mathbf{p}})$ *independently with probability* $1 - \exp(-ap_u p_v)$, *and then forget about the root of* $\mathcal{T}^{\mathbf{p}}$ *and the plane embedding of the tree* $\mathcal{T}^{\mathbf{p}}$.

5.4. **Connected random graphs with blobs.** For $m \geq 1$, consider $a = a^{(m)} > 0$, and pmfs $\mathbf{p} = \mathbf{p}^{(m)} = (p_i, i \in [m])$. Suppose we are given the following:

(a) **Blob level superstructure:** Let $\mathcal{G}^{\mathbf{p}} \sim \mathbb{P}_{\mathrm{con}}(\,\cdot\,; \mathbf{p}, a, [m])$.

(b) **Blobs:** Let $\mathbf{M} = \mathbf{M}^{(m)} = \{(M_i, d_i, \mu_i) : i \in [m]\}$ be a family of compact metric measure spaces where μ_i, $i \in [m]$, are probability measures. Recall u_i from (16).

(c) **Junction points:** $\mathbf{X} = (X_{i,j},\ i, j \in [m])$ be a family of independent random variables (that is also independent of $\mathcal{G}^{\mathbf{p}}$) such that for each $i, j \in [m]$, $X_{i,j} \in M_i$ has law μ_i.

Recall the operation Γ from Section 3.2. Define

$$\bar{\mathcal{G}}^{\mathbf{p}} := \Gamma(\mathcal{G}^{\mathbf{p}}, \mathbf{p}, \mathbf{M}, \mathbf{X}), \quad A_m := \sum_{i \in [m]} p_i u_i, \quad \text{and} \quad \mathrm{diam}_{\max} := \max_{i \in [m]} \mathrm{diam}(M_i). \tag{62}$$

Theorem 5.4. *Assume that* (58) *holds and that for some* $\eta_0 \in (0, \infty)$,

$$\lim_{m \to \infty} [\sigma(\mathbf{p})]^{1/2 - \eta_0} \mathrm{diam}_{\max} / (A_m + 1) = 0. \tag{63}$$

Then $\frac{\sigma(\mathbf{p})}{A_m + 1} \cdot \bar{\mathcal{G}}^{\mathbf{p}} \xrightarrow{\mathrm{d}} \mathcal{G}(2\tilde{\mathbf{e}}^{\bar{\gamma}}, \bar{\gamma}\tilde{\mathbf{e}}^{\bar{\gamma}}, \mathcal{P})$, *as* $m \to \infty$.

Proof: We construct $\bar{\mathcal{G}}^{\mathbf{p}}$ in the following way: Consider a coupling $(\mathcal{G}^{\mathbf{p}}, \mathcal{T}^{\mathbf{p}}) \in \mathbb{G}_{[m]}^{\mathrm{con}} \times \mathbb{T}_m^{\mathrm{ord}}$ as in Proposition 5.3 with $\mathcal{G}^{\mathbf{p}} \sim \mathbb{P}_{\mathrm{con}}$ and $\mathcal{T}^{\mathbf{p}} \sim \tilde{\mathbb{P}}_{\mathrm{ord}}$. Generate the random junction points $\mathbf{X}$ independent of $(\mathcal{G}^{\mathbf{p}}, \mathcal{T}^{\mathbf{p}})$. Set $\bar{\mathcal{G}}^{\mathbf{p}} := \Gamma(\mathcal{G}^{\mathbf{p}}, \mathbf{p}, \mathbf{M}, \mathbf{X})$. Further, let $\bar{\mathcal{T}}^{\mathbf{p}} := \Gamma(\mathcal{T}^{\mathbf{p}}, \mathbf{p}, \mathbf{M}, \mathbf{X})$. By Theorem 5.2 we have $\sigma(\mathbf{p}) \cdot \mathcal{G}^{\mathbf{p}} \xrightarrow{\mathrm{d}} \mathcal{G}(2\tilde{\mathbf{e}}^{\bar{\gamma}}, \bar{\gamma}\tilde{\mathbf{e}}^{\bar{\gamma}}, \mathcal{P})$. Thus we only need to prove

$$d_{\mathrm{GHP}}\left(\sigma(\mathbf{p}) \cdot \mathcal{G}^{\mathbf{p}}, \frac{\sigma(\mathbf{p})}{A_m + 1} \cdot \bar{\mathcal{G}}^{\mathbf{p}}\right) \xrightarrow{\mathrm{d}} 0, \text{ as } m \to \infty. \tag{64}$$

It would be convenient to work with slight variants of the original spaces $\mathcal{G}^{\mathbf{p}}$ and $\bar{\mathcal{G}}^{\mathbf{p}}$. Write $\mathcal{G}_*^{\mathbf{p}}$ (resp. $\bar{\mathcal{G}}_*^{\mathbf{p}}$) for the metric measure space obtained from $\mathcal{T}^{\mathbf{p}}$ (resp. $\bar{\mathcal{T}}^{\mathbf{p}}$) by identifying i and j (resp. $X_{i,j}$ and $X_{j,i}$) for each edge $\{i, j\} \in E(\mathcal{G}^{\mathbf{p}}) \setminus E(\mathcal{T}^{\mathbf{p}})$, instead of placing an edge of length one between them. Recall that $\mathrm{spls}(\mathcal{G}^{\mathbf{p}})$ denotes the number of surplus edges in $\mathcal{G}^{\mathbf{p}}$. Then one can check that

$$d_{\mathrm{GHP}}(\mathcal{G}_*^{\mathbf{p}}, \mathcal{G}^{\mathbf{p}}) \leq \mathrm{spls}(\mathcal{G}^{\mathbf{p}}), \quad \text{and} \quad d_{\mathrm{GHP}}(\bar{\mathcal{G}}_*^{\mathbf{p}}, \bar{\mathcal{G}}^{\mathbf{p}}) \leq \mathrm{spls}(\mathcal{G}^{\mathbf{p}}). \tag{65}$$

View $\mathcal{T}^{\mathbf{p}}$ as a metric measure space by endowing it with the tree distance and the probability measure that assigns mass p_i to the vertex i. Let $\mathcal{R}_m \subseteq \mathcal{T}^{\mathbf{p}} \times \bar{\mathcal{T}}^{\mathbf{p}}$ be given by $\mathcal{R}_m := \{(i, x) : i \in [m], x \in M_i\}$. Let ν_m be the measure on $\mathcal{T}^{\mathbf{p}} \times \bar{\mathcal{T}}^{\mathbf{p}}$ given by $\nu_m(\{i\} \times A) := p_i \mu_i(A \cap M_i)$ for $i \in [m]$, $A \subseteq \bar{\mathcal{T}}^{\mathbf{p}}$. Write

$$\mathrm{dis}(\mathcal{R}_m) = \mathrm{dis}\left(\mathcal{R}_m; \sigma(\mathbf{p})\mathcal{T}^{\mathbf{p}}, \frac{\sigma(\mathbf{p})}{A_m + 1}\bar{\mathcal{T}}^{\mathbf{p}}\right) \quad \text{and} \quad \mathrm{dsc}(\nu_m) = \mathrm{dsc}\left(\nu_m; \sigma(\mathbf{p})\mathcal{T}^{\mathbf{p}}, \frac{\sigma(\mathbf{p})}{A_m + 1}\bar{\mathcal{T}}^{\mathbf{p}}\right).$$

By [5, Lemma 4.2], we have

$$d_{\mathrm{GHP}}\left(\sigma(\mathbf{p})\mathcal{G}_*^{\mathbf{p}}, \frac{\sigma(\mathbf{p})}{A_m + 1}\bar{\mathcal{G}}_*^{\mathbf{p}}\right) \leq (\mathrm{spls}(\mathcal{G}^{\mathbf{p}}) + 1) \max\left\{\frac{\mathrm{dis}(\mathcal{R}_m)}{2}, \mathrm{dsc}(\nu_m), \nu_m(\mathcal{R}_m^c)\right\}. \tag{66}$$

It is easy to check that for all m, $\mathrm{dsc}(\nu_m) = 0$ and $\nu_m(\mathcal{R}_m^c) = 0$. Therefore by (66) and (65), to obtain (64) we only need to show that

$$(\mathrm{spls}(\mathcal{G}^{\mathbf{p}}), m \geq 1) \text{ is tight, and } \mathrm{dis}(\mathcal{R}_m) \xrightarrow{\mathrm{d}} 0 \text{ as } m \to \infty. \tag{67}$$

We will prove (67) in Lemma 5.5 and Lemma 5.6 below, which will complete the proof of Theorem 5.4. ∎

Lemma 5.5. *The sequence* $\big(\mathrm{spls}(\mathcal{G}^{\mathbf{p}}); m \geq 1\big)$ *is tight.*

Proof: By [16, Corollary 7.13], there exists a constant $K_1 > 0$ such that for m large,

$$\mathbb{E}_{\mathrm{ord}}[L^2] \leq K_1. \tag{68}$$

Since $\mathcal{G}^{\mathbf{p}}$ is obtained by adding each permitted edge $\{u, v\}$ of $\mathcal{T}^{\mathbf{p}}$ independently with probability $1 - \exp(-a p_u p_v)$, $\mathbb{E}[\mathrm{spls}(\mathcal{G}^{\mathbf{p}})] \leq \mathbb{E}\big[\sum_{\{i,j\} \in \mathscr{P}(\mathcal{T}^{\mathbf{p}})} a p_i p_j\big]$. Using (61) and (68) together with the relations $L(\mathbf{t}) \geq 1$ and $\sum_{\{i,j\} \in \mathscr{P}(\mathbf{t})} a p_i p_j \leq L(\mathbf{t})$, we have, for all large m, $\mathbb{E}[\mathrm{spls}(\mathcal{G}^{\mathbf{p}})] \leq \mathbb{E}[L(\mathcal{T}^{\mathbf{p}})] = \mathbb{E}_{\mathrm{ord}}[L^2]/\mathbb{E}_{\mathrm{ord}}[L] \leq K_1$. Thus $\sup_{m \geq 1} \mathbb{E}[\mathrm{spls}(\mathcal{G}^{\mathbf{p}})] < \infty$, which implies the desired tightness. ∎

Lemma 5.6. *As* $m \to \infty$, $\mathrm{dis}(\mathcal{R}_m) \xrightarrow{\mathrm{d}} 0$.

Proof: Let $L(\cdot)$ be as in (60). Suppose $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space where $\mathcal{T}^{\mathbf{p}}$ and $\bar{\mathcal{T}}^{\mathbf{p}}$ are defined. Let $\mathbb{P}'$ be the probability measure on Ω given by

$$\frac{d\mathbb{P}}{d\mathbb{P}'}(\omega) = \frac{L(\mathcal{T}^{\mathbf{p}}(\omega))}{\mathbb{E}_{\mathrm{ord}}[L]} \text{ for } \omega \in \Omega. \tag{69}$$

By (61), $\mathcal{T}^{\mathbf{p}} \sim \mathbb{P}_{\mathrm{ord}}$ under $\mathbb{P}'$. Suppose we show that for any $\varepsilon > 0$,

$$\mathbb{P}'(\mathrm{dis}(\mathcal{R}_m) > \varepsilon) \to 0 \text{ as } m \to \infty. \tag{70}$$

Then $\big(\mathbb{P}(\mathrm{dis}(\mathcal{R}_m) > \varepsilon)\big)^2 \leq \mathbb{P}'(\mathrm{dis}(\mathcal{R}_m) > \varepsilon) \cdot \mathbb{E}_{\mathrm{ord}}[L^2] \to 0$ as $m \to \infty$, where the first step uses (69), the fact that $L(\mathbf{t}) \geq 1$, and the Cauchy-Schwarz inequality, and the second step uses (68) and (70). So we just have to prove (70).

We introduce some additional random variables: Let J and $Y_1, \ldots, Y_m$ be respectively $[m]$ and $M_1, \ldots, M_m$ valued independent random variables that are also independent of $\mathcal{T}^{\mathbf{p}}$ such that $J \sim \mathbf{p}$, and $Y_i \sim \mu_i$ for $i \in [m]$. We extend the probability space to incorporate these random variables and still write $\mathbb{P}'$ for the underlying probability measure.

For $x \in \bar{\mathcal{T}}^{\mathbf{p}}$, write $i(x)$ for the $i \in [m]$ with $x \in M_i$. Write $d_{\mathcal{T}}$ and $d_{\bar{\mathcal{T}}}$ for the metrics on $\mathcal{T}^{\mathbf{p}}$ and $\bar{\mathcal{T}}^{\mathbf{p}}$ respectively. Then

$$\mathrm{dis}(\mathcal{R}_m) = \sup_{x, y \in \bar{\mathcal{T}}^{\mathbf{p}}} \left\{ \left| \sigma(\mathbf{p}) d_{\mathcal{T}}(i(x), i(y)) - \frac{\sigma(\mathbf{p})}{A_m + 1} d_{\bar{\mathcal{T}}}(x, y) \right| \right\}. \tag{71}$$

We first show for any fixed $i_0 \in [m]$ and $x_0 \in M_{i_0}$,

$$\mathrm{dis}(\mathcal{R}_m) \leq 4 \sup_{y \in \bar{\mathcal{T}}^{\mathbf{p}}} \left\{ \left| \sigma(\mathbf{p}) d_{\mathcal{T}}(i_0, i(y)) - \frac{\sigma(\mathbf{p})}{A_m + 1} d_{\bar{\mathcal{T}}}(x_0, y) \right| \right\} + \frac{2\sigma(\mathbf{p})}{A_m + 1} \mathrm{diam}_{\max}. \tag{72}$$

Indeed, for any two points $x, y \in \bar{\mathcal{T}}^{\mathbf{p}}$, there are two unique paths between i_0 and $i(x)$ and between i_0 and $i(y)$ in $\mathcal{T}^{\mathbf{p}}$. Write $(i_0, \ldots, i_k)$ for the longest common path shared by these

two paths. Let $i_* = i_k$ and $x_* = X_{i_k, i_{k-1}}$. Since $\mathscr{T}^{\mathbf{p}}$ is a tree,

$$d_{\mathscr{T}}(i(x), i(y)) = d_{\mathscr{T}}(i_0, i(x)) + d_{\mathscr{T}}(i_0, i(y)) - 2d_{\mathscr{T}}(i_0, i_*). \tag{73}$$

By a similar observation for $\bar{\mathscr{T}}^{\mathbf{p}}$, we have

$$d_{\bar{\mathscr{T}}}(x, y) \le d_{\bar{\mathscr{T}}}(x_0, x) + d_{\bar{\mathscr{T}}}(x_0, y) - 2d_{\bar{\mathscr{T}}}(x_0, x_*) \le d_{\bar{\mathscr{T}}}(x, y) + 2\text{diam}_{\max}. \tag{74}$$

Hence, (72) follows using (73) and (74) in (71).

Next, replace (i_0, x_0) by (I, Y_I), where $I \in [m]$ is the root of $\mathscr{T}^{\mathbf{p}}$, and replace every $y \in M_j$ in (72) with $Y_j \in M_j$, which incurs an error of at most $4\sigma(\mathbf{p})\text{diam}_{\max}/(A_m + 1)$. Thus,

$$\text{dis}(\mathscr{R}_m) \le 4 \sup_{j \in [m]} \left\{ \left| \sigma(\mathbf{p}) d_{\mathscr{T}}(I, j) - \frac{\sigma(\mathbf{p})}{A_m + 1} d_{\bar{\mathscr{T}}}(Y_I, Y_j) \right| \right\} + \frac{6\sigma(\mathbf{p})\text{diam}_{\max}}{A_m + 1}.$$

Using (63) and $\sigma(\mathbf{p}) \to 0$ we can find m_0 such that $6\sigma(\mathbf{p})\text{diam}_{\max}/(A_m + 1) < \varepsilon/5$ for $m > m_0$. Then for $m > m_0$,

$$\mathbb{P}'(\text{dis}(\mathscr{R}_m) > \varepsilon) \le \mathbb{P}'\left(\sup_{j \in [m]} \left| \sigma(\mathbf{p}) d_{\mathscr{T}}(I, j) - \frac{\sigma(\mathbf{p})}{A_m + 1} d_{\bar{\mathscr{T}}}(Y_I, Y_j) \right| > \frac{\varepsilon}{5} \right)$$

$$\le \frac{1}{p_{\min}} \sum_{j \in [m]} p_j \mathbb{P}'\left(\left| \sigma(\mathbf{p}) d_{\mathscr{T}}(I, j) - \frac{\sigma(\mathbf{p})}{A_m + 1} d_{\bar{\mathscr{T}}}(Y_I, Y_j) \right| > \frac{\varepsilon}{5} \right)$$

$$= \frac{1}{p_{\min}} \mathbb{P}'\left(\left| \sigma(\mathbf{p}) d_{\mathscr{T}}(I, J) - \frac{\sigma(\mathbf{p})}{A_m + 1} d_{\bar{\mathscr{T}}}(Y_I, Y_J) \right| > \frac{\varepsilon}{5} \right). \tag{75}$$

Write $R^* := d_{\mathscr{T}}(I, J) + 1$, and let $(I_0 = I, I_1, ..., I_{R^*-1} = J)$ be the path between I and J in $\mathscr{T}^{\mathbf{p}}$. Define $\xi_0^* := d_I(Y_I, X_{I,I_1})$, $\xi_{R^*-1}^* := d_J(X_{J,I_{R^*-2}}, Y_J)$ and for $1 \le i \le R^*-2$, $\xi_i^* := d_{I_i}(X_{I_i, I_{i-1}}, X_{I_i, I_{i+1}})$. In this notation $d_{\bar{\mathscr{T}}}(Y_I, Y_J) = \sum_{i=0}^{R^*-1} \xi_i^* + (R^* - 1)$ and $d_{\mathscr{T}}(I, J) = R^* - 1$. Thus,

$$\sigma(\mathbf{p}) d_{\mathscr{T}}(I, J) - \frac{\sigma(\mathbf{p})}{A_m + 1} d_{\bar{\mathscr{T}}}(Y_I, Y_J) = \frac{\sigma(\mathbf{p})}{A_m + 1} \left(\left[\sum_{i=0}^{R^*-1} (A_m - \xi_i^*) \right] - A_m \right). \tag{76}$$

The summation above involves the path in a $\mathbf{p}$-tree connecting the root and a random vertex selected according to the distribution $\mathbf{p}$. This admits the following alternate construction: Let $\mathbf{J} := (J_i, i \ge 0)$ be a sequence of iid $[m]$-valued random variables with law $\mathbf{p}$. For each $j \in [m]$, let $\boldsymbol{\xi}^{(j)} := (\xi_i^{(j)}, i \ge 0)$ be an iid sequence with $\xi_1^{(j)} \stackrel{d}{=} d_j(X_{j,1}, X_{j,2})$ such that $\boldsymbol{\xi}^{(1)}, ..., \boldsymbol{\xi}^{(m)}$ and $\mathbf{J}$ are jointly independent. Let $(\Omega'', \mathscr{F}'', \mathbb{P}'')$ be the probability space on which the random variables $\mathbf{J}$ and $(\boldsymbol{\xi}^{(j)}; j \in [m])$ are defined. Let R denote the first repeat time of the sequence $\mathbf{J}$, i.e., $R := \inf\{k \ge 1 : J_k = J_i \text{ for some } i < k\}$. By [27, Corollary 3],

$$(I_0, ..., I_{R^*-1}; R^*)_{\mathbb{P}'} \stackrel{d}{=} (J_0, ..., J_{R-1}; R)_{\mathbb{P}''}.$$

Owing to the independence structure, we further have

$$(I_0, ..., I_{R^*-1}; R^*; \xi_0^*, ..., \xi_{R^*-1}^*)_{\mathbb{P}'} \stackrel{d}{=} (J_0, ..., J_{R-1}; R; \xi_0^{(J_0)}, ..., \xi_{R-1}^{(J_{R-1})})_{\mathbb{P}''}.$$

For $i \ge 0$ write $\Delta_i := \xi_i^{(J_i)} - A_m$. Using (76) and the fact that $\sigma(\mathbf{p}) \to 0$, for large m,

$$\mathbb{P}'\left(\left| \sigma(\mathbf{p}) d_{\mathscr{T}}(I, J) - \frac{\sigma(\mathbf{p})}{A_m + 1} d_{\bar{\mathscr{T}}}(Y_I, Y_J) \right| > \frac{\varepsilon}{5} \right) \le \mathbb{P}''\left(\frac{\sigma(\mathbf{p})}{A_m + 1} \left| \sum_{i=0}^{R-1} \Delta_i \right| > \frac{\varepsilon}{6} \right). \tag{77}$$

Note that $\big(\xi_i^{(J_i)}; i \geq 0\big)$ is a collection of iid random variables with mean $\sum_{k \in [m]} p_k \mathbb{E}[\xi_0^{(k)}] = A_m$. Thus $(\sum_{i=0}^{k} \Delta_i, \ k \geq 0)$ is a martingale with respect to its natural filtration. Then, for any $t > 0$,

$$\mathbb{P}'' \Big(\frac{\sigma(\mathbf{p})}{A_m + 1} \Big| \sum_{i=0}^{R-1} \Delta_i \Big| > \frac{\varepsilon}{6} \Big) \leq \mathbb{P}''(R \geq t) + \mathbb{P}'' \Big(\sup_{0 \leq k \leq t-1} \Big| \sum_{i=0}^{k} \Delta_i \Big| > \frac{\varepsilon(A_m + 1)}{6\sigma(\mathbf{p})} \Big). \tag{78}$$

For the first term, by [16, Lemma 10.7] (also see [27, Equations (26), (29)]),

$$\mathbb{P}''(R \geq t) \leq 2 \exp\big(-t^2 \sigma^2(\mathbf{p})/24\big) \text{ for } t \in (0, 1/p_{\max}). \tag{79}$$

The second term on the right side of (78) can be bounded by using Markov's inequality and the Burkholder-Davis-Gundy (BDG) inequality [25]. For fixed $r \geq 1$, we have

$$\mathbb{P}'' \Big(\sup_{0 \leq k \leq t-1} \Big| \sum_{i=0}^{k} \Delta_i \Big| > \frac{\varepsilon(A_m + 1)}{6\sigma(\mathbf{p})} \Big) \leq \Big(\frac{6\sigma(\mathbf{p})}{\varepsilon(A_m + 1)} \Big)^{2r} \mathbb{E}'' \Big[\sup_{0 \leq k \leq t-1} \Big| \sum_{i=0}^{k} \Delta_i \Big|^{2r} \Big]$$

$$\leq \Big(\frac{6\sigma(\mathbf{p})}{\varepsilon(A_m + 1)} \Big)^{2r} K_2(r) \mathbb{E}'' \Big[\Big(\sum_{i=0}^{t-1} \Delta_i^2 \Big)^r \Big] \leq \Big(\frac{6\sigma(\mathbf{p})}{\varepsilon(A_m + 1)} \Big)^{2r} K_2(r) \cdot t^r \mathrm{diam}_{\max}^{2r}, \tag{80}$$

where $\mathbb{E}''$ denotes expectation with respect to $\mathbb{P}''$, the constant $K_2(r) > 0$ comes from the BDG inequality, and the last step uses $|\Delta_i| \leq \mathrm{diam}_{\max}$. Define $\alpha_m := -24 r \log \sigma(\mathbf{p})$ and $t_m := \alpha_m / \sigma(\mathbf{p})$. Combining (78), (80), (79) with $t = t_m$, we have for sufficiently large m,

$$\mathbb{P}'' \Big(\frac{\sigma(\mathbf{p})}{A_m + 1} \Big| \sum_{i=0}^{R-1} \Delta_i \Big| > \frac{\varepsilon}{6} \Big) \leq 2 \exp \Big(-\frac{t_m^2 \sigma^2(\mathbf{p})}{24} \Big) + K_3(r, \varepsilon) \cdot \frac{\alpha_m^r \sigma^r(\mathbf{p}) \mathrm{diam}_{\max}^{2r}}{(A_m + 1)^{2r}}$$

$$\leq 2[\sigma(\mathbf{p})]^r + K_3(r, \varepsilon) \alpha_m^r [\sigma(\mathbf{p})]^{2r\eta_0}, \tag{81}$$

where $K_3(r, \varepsilon) = 6^{2r} K_2(r)/\varepsilon^{2r}$, and the last line uses (63). (It is easy to check that $t_m p_{\max} \to 0$ as $m \to \infty$. Thus, we can apply (79) for m large, and (81) is valid for large m.) By (81), (77), and (75), taking $r = r_* := \max\{r_0, \lceil (r_0 + 1)/(2\eta_0) \rceil\}$,

$$\mathbb{P}'(\mathrm{dis}(\mathscr{R}_m) > \varepsilon) \leq \Big(2[\sigma(\mathbf{p})]^{r_0} + K_3(r_*, \varepsilon) \alpha_m^{r_*} [\sigma(\mathbf{p})]^{r_0 + 1} \Big) / p_{\min}.$$

By (58), the above expression goes to zero as $m \to \infty$. This completes the proof of (70). ∎

5.5. **Proof of Theorem 3.4.**

The previous section deals with asymptotics for a single connected component. This section leverages the above results to study maximal components of $\bar{\mathscr{G}}(\mathbf{x}, q, \mathbf{M})$ and prove Theorem 3.4. Recall that $\mathbf{x}$ is the weight sequence, $\mathbf{M}$ is the collection of blobs, $\mathbf{X}$ is the collection of random junction points, and $u_1, \ldots, u_n$ are the average distances. Also recall that $\sigma_r = \sum_{i \in [n]} x_i^r, \ r \geq 1$, and $\tau = \sum_{i \in [n]} x_i^2 u_i$. We start with the following auxiliary result.

Proposition 5.7. *Under Conditions 2.1 and 3.3, for each $i \geq 1$,*

$$\frac{\sum_{v \in \mathscr{C}_i} x_v^2}{\sum_{v \in \mathscr{C}_i} x_v} \cdot \frac{\sigma_2}{\sigma_3} \xrightarrow{\mathrm{d}} 1, \quad \text{and} \quad \frac{\sum_{v \in \mathscr{C}_i} x_v u_v}{\sum_{v \in \mathscr{C}_i} x_v} \cdot \frac{\sigma_2}{\tau} \xrightarrow{\mathrm{d}} 1, \quad \text{as } n \to \infty.$$

We will make use of the following lemma in the proof.

Lemma 5.8 ([16], Lemma 8.2). *For $n \geq 1$, fix $\boldsymbol{x} := (x_i > 0, \; i \in [n])$ and $\boldsymbol{a} := (a_i \geq 0, \; i \in [n])$ such that $c_n := \sum_{i \in [n]} x_i a_i / \sum_{i \in [n]} x_i > 0$. Let $(v(i), i \in [n])$ be a size-biased random ordering of $[n]$ using the weights $\boldsymbol{x}$. Let $x_{\max} := \max_{i \in [n]} x_i$ and $a_{\max} := \max_{i \in [n]} a_i$. Let $\ell = \ell(n) \in [n]$ be such that*

$$\ell x_{\max}/\sigma_1 \to 0, \quad \text{and} \quad a_{\max}/(\ell c_n) \to 0 \quad \text{as} \quad n \to \infty. \tag{82}$$

Then

$$\sup_{k \leq \ell} \left| \frac{\sum_{i=1}^{k} a_{v(i)}}{\ell c_n} - \frac{k}{\ell} \right| \xrightarrow{\mathrm{d}} 0 \quad \text{as} \quad n \to \infty.$$

We are now ready to prove Proposition 5.7.

Proof of Proposition 5.7: We only work with $i = 1$ to keep the notation simple. Recall the breadth-first construction of $\mathscr{G}(\mathbf{x}, q)$ from Section 5.1, where vertices are explored in a size-biased order $(v(i), 1 \leq i \leq n)$ using the weight sequence $\mathbf{x}$. The following properties of this exploration were derived in [7]: (a) There exist random variables $m_L, m_R \in [n]$ such that $V(\mathscr{C}_1) = \{v(i) : m_L + 1 \leq i \leq m_R\}$. (b) Under Condition 2.1, $\sum_{i=1}^{m_R} x_{v(i)}$ is tight. (c) By Theorem 2.2, as $n \to \infty$, $\sum_{i=m_L+1}^{m_R} x_{v(i)} \xrightarrow{\mathrm{d}} \gamma_1(\lambda)$. In this notation, the claim in Proposition 5.7, for $i = 1$, is equivalent to

$$\frac{\sum_{i=m_L+1}^{m_R} x_{v(i)}^2}{\sum_{i=m_L+1}^{m_R} x_{v(i)}} \cdot \frac{\sigma_2}{\sigma_3} \xrightarrow{\mathrm{d}} 1, \quad \text{and} \quad \frac{\sum_{i=m_L+1}^{m_R} x_{v(i)} u_{v(i)}}{\sum_{i=m_L+1}^{m_R} x_{v(i)}} \cdot \frac{\sigma_2}{\tau} \xrightarrow{\mathrm{d}} 1. \tag{83}$$

Thus, it is enough to prove (83). We first show that

$$\frac{\sum_{i=m_L+1}^{m_R} x_{v(i)}}{m_R - m_L} \cdot \frac{\sigma_1}{\sigma_2} \xrightarrow{\mathrm{d}} 1, \quad \text{as } n \to \infty. \tag{84}$$

Fix $\eta > 0$. Since $\left(\sum_{i=1}^{m_R} x_{v(i)}; n \geq 1 \right)$ is a tight sequence, there exists $T > 0$ such that for all $n \geq 1$, $\mathbb{P} \left(\sum_{i=1}^{m_R} x_{v(i)} \geq T \right) < \eta$. Let $m_0 := \sigma_1/\sigma_2$ and apply Lemma 5.8 with $\ell = 2Tm_0$ and $a_i = x_i$ (thus $c_n = \sigma_2/\sigma_1$). Condition (82) is equivalent to $x_{\max}/\sigma_2 \to 0$, which follows from Assumption 2.1. By Lemma 5.8, there exists $N_\eta > 0$ such that for all $n > N_\eta$,

$$\mathbb{P} \left(\sup_{k \leq 2Tm_0} \left| \sum_{i=1}^{k} x_{v(i)} - \frac{k}{m_0} \right| > \eta \right) < \eta. \tag{85}$$

On the event $\left\{ \sup_{k \leq 2Tm_0} \left| \sum_{i=1}^{k} x_{v(i)} - \frac{k}{m_0} \right| \leq \eta \right\} \cap \left\{ \sum_{i=1}^{m_R} x_{v(i)} \leq T \right\}$, we must have $m_L < m_R < 2Tm_0$ (assuming $\eta < T$), and hence $\left| \sum_{i=m_L+1}^{m_R} x_{v(i)} - (m_R - m_L)/m_0 \right| < 2\eta$. Since η can be arbitrarily small, using (85) we have

$$\left| \sum_{i=m_L+1}^{m_R} x_{v(i)} - \frac{m_R - m_L}{m_0} \right| \xrightarrow{\mathrm{d}} 0 \text{ as } n \to \infty. \tag{86}$$

By property (c) of the exploration and (86), $(m_R - m_L)/m_0 \xrightarrow{\mathrm{d}} \gamma_1(\lambda)$, which coupled with (86) yields (84).

Now, under Conditions 2.1 and 3.3,

$$\frac{\sigma_2 x_{\max}^2}{\sigma_3} \to 0, \quad \text{and} \quad \frac{\sigma_2 x_{\max} \mathrm{diam}_{\max}}{\tau} \to 0, \quad \text{as } n \to \infty. \tag{87}$$

Thus, repeating the above argument with respectively $x_{v(i)}^2$ and $x_{v(i)} u_{v(i)}$ in place of $x_{v(i)}$ and using Lemma 5.8, we get

$$\frac{\sum_{i=m_L+1}^{m_R} x_{v(i)}^2}{m_R - m_L} \cdot \frac{\sigma_1}{\sigma_3} \xrightarrow{\mathrm{d}} 1, \quad \text{and} \quad \frac{\sum_{i=m_L+1}^{m_R} x_{v(i)} u_{v(i)}}{m_R - m_L} \cdot \frac{\sigma_1}{\tau} \xrightarrow{\mathrm{d}} 1. \tag{88}$$

Combining (84) and (88) completes the proof of (83), and thus of Proposition 5.7. ∎

Proof of Theorem 3.4: Since we work with the product GHP topology, it is enough to prove the assertion for the maximal component $\bar{\mathscr{C}}_1$, i.e.,

$$\mathrm{scl}\Big(\frac{\sigma_2^2}{\sigma_2 + \tau}, 1\Big)\bar{\mathscr{C}}_1 \xrightarrow{\mathrm{d}} \mathscr{G}(2\tilde{\mathbf{e}}_{\gamma_1}, \tilde{\mathbf{e}}_{\gamma_1}, \mathscr{P}_1), \quad \text{as } n \to \infty. \tag{89}$$

In (89) and the proof below, to simplify notation, we suppress dependence on λ and write γ_1 for $\gamma_1(\lambda)$. By Theorem 2.2 and Proposition 5.7, without loss of generality, we may consider a probability space on which the following convergences hold almost surely:

$$\mathrm{mass}(\mathscr{C}_1) \xrightarrow{\mathrm{a.s.}} \gamma_1, \quad \frac{\sum_{v \in \mathscr{C}_1} x_v^2}{\mathrm{mass}(\mathscr{C}_1)} \cdot \frac{\sigma_2}{\sigma_3} \xrightarrow{\mathrm{a.s.}} 1, \quad \text{and} \quad \frac{\sum_{v \in \mathscr{C}_1} x_v u_v}{\mathrm{mass}(\mathscr{C}_1)} \cdot \frac{\sigma_2}{\tau} \xrightarrow{\mathrm{a.s.}} 1. \tag{90}$$

Consider the following construction of $\bar{\mathscr{C}}_1$: Define $\mathbf{p} := \big(x_v/\mathrm{mass}(\mathscr{C}_1), v \in \mathscr{C}_1\big)$ and $a := q \cdot [\mathrm{mass}(\mathscr{C}_1)]^2$. Conditioned on $\mathcal{V}_1 := V(\mathscr{C}_1)$, let $\mathscr{C}_1'$ be a $\mathbb{G}_{\mathcal{V}_1}^{\mathrm{con}}$-valued random variable with distribution $\mathbb{P}_{\mathrm{con}}(\cdot; \mathbf{p}, a, \mathcal{V}_1)$. By Proposition 5.1,

$$\bar{\mathscr{C}}_1 \stackrel{\mathrm{d}}{=} \mathrm{scl}(1, \mathrm{mass}(\mathscr{C}_1))\Gamma(\mathscr{C}_1', \mathbf{p}, \mathbf{M}, \mathbf{X}). \tag{91}$$

In order to apply Theorem 5.4, we will verify the statements in (58) and (63) with $\bar{\gamma} = \gamma_1^{3/2}$. Note that $p_{\max} \le x_{\max}/\mathrm{mass}(\mathscr{C}_1)$, $p_{\min} \ge x_{\min}/\mathrm{mass}(\mathscr{C}_1)$, and by (90),

$$\sigma(\mathbf{p}) = \frac{\sqrt{\sum_{v \in \mathscr{C}_1} x_v^2}}{\mathrm{mass}(\mathscr{C}_1)} \sim \frac{\sigma_2}{\gamma_1^{1/2}}, \quad \text{and} \quad A_m = \frac{\sum_{v \in \mathscr{C}_1} x_v u_v}{\mathrm{mass}(\mathscr{C}_1)} \sim \frac{\tau}{\sigma_2}, \tag{92}$$

where the simplification uses the relation $\sigma_3 \sim \sigma_2^3$ (by Condition 2.1). In addition, using $q \sim 1/\sigma_2$ from Condition 2.1, $a \cdot \sigma(\mathbf{p}) \sim q\gamma_1^2 \cdot \frac{\sigma_2}{\gamma_1^{1/2}} \sim \gamma_1^{3/2}$. The remaining conditions in (58) and (63) follow directly from Conditions 2.1 and 3.3, and the details are omitted.

Applying Theorem 5.4 and using (91), we see that

$$\mathrm{scl}\Big(\frac{\sigma_2^2}{(\sigma_2 + \tau)\gamma_1^{1/2}}, \frac{1}{\gamma_1}\Big)\bar{\mathscr{C}}_1 \xrightarrow{\mathrm{d}} \mathscr{G}(2\tilde{\mathbf{e}}^{\gamma_1^{3/2}}, \gamma_1^{3/2}\tilde{\mathbf{e}}^{\gamma_1^{3/2}}, \mathscr{P}_1).$$

Note that for any excursions h and g and Poisson point process $\mathscr{P}$, for $\alpha, \beta > 0$, we have $\mathrm{scl}(\alpha, \beta)\mathscr{G}(h, g, \mathscr{P}) \stackrel{d}{=} \mathscr{G}\big(\alpha h(\cdot/\beta), \frac{1}{\beta}g(\cdot/\beta), \mathscr{P}\big)$. Thus,

$$\mathrm{scl}(\gamma_1^{1/2}, \gamma_1)\,\mathrm{scl}\Big(\frac{\sigma_2^2}{(\sigma_2 + \tau)\gamma_1^{1/2}}, \frac{1}{\gamma_1}\Big)\bar{\mathscr{C}}_1 \xrightarrow{\mathrm{d}} \mathscr{G}\Big(2\gamma_1^{1/2}\tilde{\mathbf{e}}^{\gamma_1^{3/2}}(\cdot/\gamma_1), \gamma_1^{1/2}\tilde{\mathbf{e}}^{\gamma_1^{3/2}}(\cdot/\gamma_1), \mathscr{P}_1\Big).$$

By [16, Display (4.9)], we have $\gamma_1^{1/2}\tilde{\mathbf{e}}^{\gamma_1^{3/2}}(\cdot/\gamma_1) \stackrel{d}{=} \tilde{\mathbf{e}}_{\gamma_1}(\cdot)$. Combining this observation with the last display completes the proof of (89). ∎

6. Proofs: Associated random graph models

In this section, we will prove the results stated in Section 4. We start in Section 6.1 with the proof of Lemma 4.4. In Section 6.2 we state a result (Proposition 6.3) that claims that the conclusions in Theorems 4.5 and 4.6 hold under a stronger assumption. Here, we also introduce the notation to be used in the sequel. From Section 6.3 through Section 6.9, we prove various estimates working towards the proof of Proposition 6.3. Section 6.10 collects all these results to complete the proof of Proposition 6.3 via an application of Theorem 3.4. We then use Proposition 6.3 in Section 6.11 to prove Theorems 4.5 and 4.6. Finally, the proofs of Theorems 4.9, 4.13, and 4.15 are given in Sections 6.12 and 6.13.

6.1. **Proof of Lemma 4.4.** We start with an elementary result.

Lemma 6.1. *Suppose $K \in L^3([0,1]^2)$. Then $\|T_K\|_{p,q} \le \left(\int_{[0,1]^2} |K|^3\right)^{1/3}$ for any $p, q \in [3/2, 3]$.*

Proof: For any $f \in L^{3/2}([0,1])$,

$$|T_K f(x)|^3 \le \left(\int_0^1 |K(x,y)f(y)|\,dy\right)^3 \le \left(\int_0^1 |K(x,y)|^3\,dy\right)\left(\int_0^1 |f|^{3/2}\right)^2,$$

where the last step uses Hölder's inequality. Hence, $\|T_K f\|_3 \le \left(\int_{[0,1]^2} |K|^3\right)^{1/3}\|f\|_{3/2}$, and the claim follows for $p = 3/2$ and $q = 3$. The proof is completed upon noting that $\|T_K\|_{p,q} \le \|T_K\|_{3/2,3}$ for any $p, q \in [3/2, 3]$. ∎

Proof of Lemma 4.4: Let us first show that T_W is strongly band irreducible, i.e., there do not exist Borel $A_1, A_2 \subseteq [0,1]$ with $\mu_\circ(A_1) > 0$ and $\mu_\circ(A_2) > 0$ such that $W(x,y) = 0$ for $\mu_\square$ a.e. $(x,y) \in A_1 \times A_2$. Assume on the contrary that such A_1 and A_2 exist. Then

$$n^{1/3}\langle \mathbb{1}_{A_2}, T_{W_n}\mathbb{1}_{A_1}\rangle = n^{1/3}\langle \mathbb{1}_{A_2}, T_{W_n}\mathbb{1}_{A_1}\rangle - n^{1/3}\langle \mathbb{1}_{A_2}, T_W\mathbb{1}_{A_1}\rangle \to \langle \mathbb{1}_{A_2}, T_H\mathbb{1}_{A_1}\rangle \qquad (93)$$

as $n \to \infty$, where the last step follows from (20). By Condition 4.3, $\mu_\square\big(\{(x,y) : W_n(x,y) \le n^{-\delta_0}\}\big) = o(1)$. Hence,

$$n^{1/3}\langle \mathbb{1}_{A_2}, T_{W_n}\mathbb{1}_{A_1}\rangle \ge n^{1/3} \cdot n^{-\delta_0} \cdot \mu_\circ(A_1)\mu_\circ(A_2) \cdot \big(1 - o(1)\big),$$

which contradicts (93) as $\delta_0 < 1/3$.

Now, $W \in L^3([0,1]^2) \subseteq L^2([0,1]^2)$ implies that T_W is a compact operator on $L^2([0,1])$. Since $W \ge 0$ and T_W is band irreducible (i.e., there does not exist a Borel $A \subseteq [0,1]$ with $\mu_\circ(A) \in (0,1)$ such that $W(x,y) = 0$ for $\mu_\square$ a.e. $(x,y) \in A \times A^c$), by [61, Theorem 43.8], the spectral radius of T_W (which equals $+1$ under Condition 4.1) is an eigenvalue of T_W. Further, the corresponding eigenspace is one dimensional, and is spanned by an eigenfunction ψ satisfying $\psi(x) > 0$ for $\mu_\circ$ a.e. x. We can further assume that ψ is normalized so that $\|\psi\|_2 = 1$. Finally, $\|\psi\|_3 = \|T_W\psi\|_3 \le \|T_W\|_{2,3} < \infty$, where the last step follows from Lemma 6.1. This proves Lemma 4.4 (a) and (b).

The claim in Lemma 4.4 (c) follows from [61, Theorem 44.10]. In our rather simple setting, this result follows from an easy argument, which we include here. Note that $T_W^2 = T_K$, where $K(x,y) = \int_{z=0}^1 W(x,z)W(z,y)dz$. If T_W^2 is not strongly band irreducible, then we could find Borel $A_1, A_2 \subseteq [0,1]$ with positive measures such that

$$0 = \int_{x \in A_1}\int_{y \in A_2} K(x,y)dxdy = \int_{z=0}^1 \left(\int_{x \in A_1} W(x,z)dx\right)\left(\int_{y \in A_2} W(y,z)dy\right)dz.$$

Hence, we can find a Borel $A_3 \subseteq [0,1]$ such that $W(x,z) = 0$ for $\mu_\square$ a.e. $(x,z) \in A_1 \times A_3$, and $W(y,z) = 0$ for $\mu_\square$ a.e. $(y,z) \in A_1 \times A_3^c$. This contradicts the strong band irreducibility property of T_W, since at least one of the sets A_3 and A_3^c has positive measure. Thus, $T_W^2 = T_K$ is a compact, strongly band irreducible operator on $L^2([0,1])$ with nonnegative kernel K. Since $\|T_W^2\|_{2,2} = 1$ and $T_W^2 \psi = \psi$, by [61, Theorem 43.8], the eigenspace of T_W^2 associated with the eigenvalue 1 is spanned by ψ. If there exists $\phi \in L^2([0,1])$ with $\|\phi\|_2 = 1$ and $T_W \phi = -\phi$, then $T_W^2 \phi = \phi$, and consequently, either $\phi = \psi$ or $\phi = -\psi$ which leads to a contradiction. Hence, -1 is not an eigenvalue of T_W, and this completes the proof of Lemma 4.4 (c). ∎

6.2. **An intermediate result and related notation.** As a first step, we will prove Theorem 4.5 when the following stronger assumption replaces Condition 4.3.

Condition 6.2. *Let θ_0 be as in Condition 4.2. Then there exist $\delta_0 \in (0, 1/3 - 2\theta_0)$ and $\varpi_0 \in [0, 2\delta_0 + 1/3)$ such that the following hold: For each $n \geq 1$, there exists $B_n \subseteq [n]$ satisfying*

$$|B_n| = o(n^{\varpi_0}), \quad \beta_{ij}^{(n)} = \beta_{ji}^{(n)} = 0 \quad \text{if } (i,j) \in B_n \times [n], \tag{94}$$

$$\beta_{ii}^{(n)} = 0 \quad \text{if } i \in [n] \setminus B_n, \quad \text{and} \quad \beta_{ij}^{(n)} > n^{-\delta_0} \quad \text{if } i,j \in [n] \setminus B_n \text{ with } i \neq j. \tag{95}$$

Proposition 6.3. *Under Conditions 4.1, 4.2, and 6.2, the convergences in (26) and (27) hold with the same constants as in (25).*

The next few sections are devoted to the proof of Proposition 6.3. For the reader's convenience, we collect some notation here that will be used frequently in the proof. We define

$$\kappa_{ij}^{(n)} := \begin{cases} e^{-n}, & \text{if } i = j \in [n] \setminus B_n, \\ \beta_{ij}^{(n)} - n^{-\delta_0}, & \text{if } i,j \in [n] \setminus B_n \text{ with } i \neq j, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\kappa_n(x,y) := \begin{cases} \kappa_{ij}^{(n)}, & \text{if } (x,y) \in Q_{ij}^{(n)}, \ 1 \leq i,j \leq n, \\ 0, & \text{if } (x,y) \in [0,1]^2 \setminus \bigcup_{i,j \in [n]} Q_{ij}^{(n)}. \end{cases} \tag{96}$$

Let $\mathcal{G}_n^\bullet$ denote the random graph on the vertex set $[n]$ where the edges $\{i,j\}$ appear independently with respective probabilities $1 \wedge (\kappa_{ij}^{(n)}/n)$, $1 \leq i < j \leq n$. For a finite graph G and $k \geq 1$, define

$$s_k(G) := \sum_{\mathcal{C} \subseteq G} \frac{|\mathcal{C}|^k}{|G|} \quad \text{and} \quad \mathcal{D}(G) := \frac{1}{|G|} \sum_{\mathcal{C} \subseteq G} \sum_{i \in V(\mathcal{C})} \sum_{j \in V(\mathcal{C})} d_{\mathcal{C}}(i,j),$$

where $\sum_{\mathcal{C} \subseteq G}$ denotes the sum over all connected components of G. The function $s_k(G)$ is called the k-th susceptibility function. We will write

$$D_{\max}^\bullet := \text{diam}(\mathcal{G}_n^\bullet), \quad \mathcal{D}^\bullet := \mathcal{D}(\mathcal{G}_n^\bullet), \quad \text{and} \quad s_k^\bullet := s_k(\mathcal{G}_n^\bullet), \ k \geq 1. \tag{97}$$

We write $\mathcal{C}_i^\bullet$ for the i-th largest component in $\mathcal{G}_n^\bullet$, and let $\mathcal{C}^\bullet(j)$ denote the component in $\mathcal{G}_n^\bullet$ containing the vertex $j \in [n]$. For $\ell \in \mathbb{Z}_{>0}$ and $i,j \in [n]$ with $i \neq j$, we write $i \overset{\ell}{\longleftrightarrow} j$ for the event that there is a self-avoiding path with ℓ edges in $\mathcal{G}_n^\bullet$ whose endpoints are i and

j. We will write $i \leftrightarrow j$ for the event $i \in \mathscr{C}^\bullet(j)$ and $i \leftrightsquigarrow j$ for the event that there is an edge between i and j in $\mathscr{G}_n^\bullet$. (Thus, $\{i \leftrightsquigarrow j\} = \{i \xleftrightarrow{1} j\}$.)

It follows from (19) that $\max_{i,j \in [n]} \beta_{ij}^{(n)} = O(n^{2/3})$. In fact, we will prove later (see (106)) that $\max_{i,j \in [n]} \beta_{ij}^{(n)} = o(n^{2/3})$. Consequently,

$$\max_{i,j \in [n]} \kappa_{ij}^{(n)} \le e^{-n} + \max_{i,j \in [n]} \beta_{ij}^{(n)} = o(n^{2/3}). \tag{98}$$

Hence, we will simply write $(\kappa_{ij}^{(n)}/n)$ and $(\beta_{ij}^{(n)}/n)$ for the edge probabilities in the sequel (i.e., we will omit the '$\wedge 1$'). Further, from now on, we will drop the superscript '(n)' and write β_{ij}, κ_{ij}, and Q_{ij}. We let

$$T_n = T_{\kappa_n} \quad \text{and} \quad \|T_n\| = \|T_{\kappa_n}\|_{2,2}.$$

We make note of a very useful bound here: for any $p, q \in [3/2, 3]$ and for all n,

$$\|T_n\|_{p,q} \le \left(\int_{[0,1]^2} \kappa_n^3 \right)^{1/3} \le \left(e^{-3n} + \int_{[0,1]^2} W_n^3 \right)^{1/3} \le C, \tag{99}$$

where the first step uses Lemma 6.1, and the last step follows from (19).

Note that if Condition 6.2 (or Condition 4.3) is satisfied for some choice of $\{\delta_0, \varpi_0, (B_n, n \ge 1)\}$ with $\delta_0 \in (0, 1/3 - 2\theta_0)$, $\varpi_0 \in [0, 2\delta_0 + 1/3)$, and $B_n \subseteq [n]$, $n \ge 1$, then the condition is also satisfied for the choice $\{\delta_1, \varpi_0, (B_n, n \ge 1)\}$ for each $\delta_1 \in (\delta_0, 1/3 - 2\theta_0)$. Since $\theta_0 < 1/12$ by assumption and consequently $1/3 - 2\theta_0 > 1/6$, we will, from now on, assume that

$$\delta_0 \in (1/6, 1/3 - 2\theta_0). \tag{100}$$

There is an obvious correspondence between the non-zero eigenvalues of the (compact and self-adjoint) operator T_n (resp. their associated eigenfunctions) and the non-zero eigenvalues of the symmetric matrix $((\kappa_{ij}))_{n \times n}$ (resp. their associated eigenvectors). Under Condition 6.2, this matrix has a submatrix of dimension $(n - |B_n|) \times (n - |B_n|)$ with positive entries and all its other entries are zeros. Thus, by the Perron-Frobenius theorem, $\|T_n\| = \|T_n\|_{2,2}$ is a simple eigenvalue of T_n, and the corresponding eigenspace is spanned by a nonnegative eigenfunction. Further, we can write

$$\kappa_n(x,y) = \|T_n\| \psi_n(x)\psi_n(y) + \sum_{j=2}^n \lambda_{n;j} \psi_{n;j}(x)\psi_{n;j}(y), \quad x, y \in [0,1], \tag{101}$$

where (a) $\lambda_{n;j}$, $2 \le j \le n$, are eigenvalues of T_n (some of which may be zeros) satisfying $|\lambda_{n;n}| \le \cdots \le |\lambda_{n;2}| < \|T_n\|$, (b) $\mathfrak{F} := \{\psi_n, \psi_{n;2}, \ldots, \psi_{n;n}\}$ is an orthonormal collection of functions in $L^2([0,1])$ such that each $f \in \mathfrak{F}$ is constant on $((i-1)/n, i/n)$ for $i \in [n]$ and $f(i/n) = 0$ for $i = 0, 1, \ldots, n$, and (c) $\psi_n(x) \ge 0$ for $x \in [0,1]$ with $\mu_\circ(\{x : \psi_n(x) = 0\}) = |B_n|/n$.

Our aim is to use Theorem 3.4 to prove Proposition 6.3. In the present setting, the functionals appearing in Conditions 2.1 and 3.3 will be rescaled versions of $s_2^\bullet$, $s_3^\bullet$, $\mathscr{D}^\bullet$, $D_{\max}^\bullet$, and $|\mathscr{C}_1^\bullet|$, and consequently, we need to derive asymptotics for these random variables. Our strategy is to prove the relevant asymptotics for a closely related branching process and then transfer the results to the above random variables by approximation. We now define this branching process and introduce some associated functions.

Consider the multitype branching process with type space $[n]$ where every vertex of type $i \in [n]$ has Bernoulli(κ_{ij}/n) many type j children for each $j \in [n]$. For $i \in [n]$ and $\ell \ge 0$,

let $\mathfrak{G}_\ell^\triangle(i) = \mathfrak{G}_{n;\ell}^\triangle(i)$ denote the collection of vertices in generation ℓ of such a branching process started from a single vertex of type i. (Thus, $\mathfrak{G}_0^\triangle(i)$ consists of a single vertex of type i.) Let $\mathfrak{X}_n^\triangle(i) := \sum_{\ell \geq 0} |\mathfrak{G}_\ell^\triangle(i)|$ be the total progeny, and let $\mathfrak{H}_n^\triangle(i) := \max\{\ell \geq 0 : \mathfrak{G}_\ell^\triangle(i) \neq \emptyset\}$ denote the height. Further, define

$$\mathfrak{Z}_n^\triangle(i) := \sum_{\ell \geq 1} \ell \cdot |\mathfrak{G}_\ell^\triangle(i)|, \text{ and } g_{n;k}^\triangle(i) := \mathbb{E}[\mathfrak{X}_n^\triangle(i)^k], \quad k \geq 1. \tag{102}$$

We will simply write $g_n^\triangle(i)$ instead of $g_{n;1}^\triangle(i)$. For $x \in [0,1]$, we define

$$g_n(x) := \begin{cases} g_n^\triangle(\lceil nx \rceil), & \text{if } x \in [0,1] \setminus \{0, 1/n, 2/n, \ldots, 1\}, \\ 0, & \text{if } x \in \{0, 1/n, 2/n, \ldots, 1\}. \end{cases} \tag{103}$$

Similarly, for $x \in [0,1]$, $|\mathfrak{G}_\ell(x)|, \mathfrak{X}_n(x), \mathfrak{H}_n(x), \mathfrak{Z}_n(x), g_{n;k}(x)$ etc. are defined in an analogous manner. Note that the functions $g_{n;k}$ can be viewed as elements of $L^p([0,1])$ for any $p \in [1,\infty) \cup \{\infty\}$. Further, if ν is uniformly distributed on $[n]$, then $\mathbb{E}[g_{n;k}^\triangle(\nu)] = \int_0^1 g_{n;k}(x)dx$.

We define $\mathbf{1}_\circ : [0,1] \to \mathbb{R}$ by $\mathbf{1}_\circ(x) := \mathbb{1}_A(x)$ where $A = [0,1] \setminus \{0, 1/n, 2/n, \ldots, 1\}$. We let $\mathbf{1}_\square : [0,1]^2 \to \mathbb{R}$ be the function $\mathbf{1}_\square(x,y) = \mathbf{1}_\circ(x)\mathbf{1}_\circ(y)$.

6.3. **Asymptotics for the eigenfunction** ψ_n.

Throughout this section, we work under Conditions 4.1, 4.2, and 6.2.

Lemma 6.4. *We have, as $n \to \infty$,*

$$\|\psi_n - \psi\|_3 = o(1), \quad and \tag{104}$$

$$\|\psi_n - \psi\|_2 = O(n^{-\delta_0}). \tag{105}$$

The following result will be needed in the proof.

Lemma 6.5. *We have, $\|T_{W_n} - T_W\|_{3,3} \to 0$ as $n \to \infty$.*

Proof: We will first show that

$$\max_{i,j \in [n]} \beta_{ij} = o(n^{2/3}) \quad \text{as} \quad n \to \infty. \tag{106}$$

If (106) fails then there exist $\varepsilon_0 > 0$, $n_1 < n_2 < \cdots$, and $i_k, j_k \in [n_k]$ such that $\beta_{i_k j_k} \geq \varepsilon_0 n_k^{2/3}$ for all $k \geq 1$. For $k \geq 1$, let $f_k := \sqrt{n_k} \cdot \mathbb{1}_{J_k}$, where $J_k = ((j_k - 1)/n_k, j_k/n_k)$. Then

$$|(T_H f_k)(x)| = \sqrt{n_k} \cdot \left| \int_{y \in J_k} H(x,y)dy \right| \leq \left(\int_{y \in J_k} H^2(x,y)dy \right)^{1/2},$$

and consequently,

$$\|T_H f_k\|_2 \leq \left(\int_{[0,1] \times J_k} H^2(x,y)dxdy \right)^{1/2} = o(1) \quad \text{as} \quad k \to \infty. \tag{107}$$

Let $I_k = ((i_k - 1)/n_k, i_k/n_k)$. Then

$$\int_{x \in I_k} (T_W f_k(x))^2 dx = n_k \int_{x \in I_k} \left(\int_{y \in J_k} W(x,y)dy \right)^2 dx$$

$$\leq \int_{I_k \times J_k} W^2 \leq \left(\int_{Q_{i_k j_k}} W^3 \right)^{2/3} \cdot n_k^{-2/3} = o(n_k^{-2/3}), \tag{108}$$

as $k \to \infty$. Hence,

$$
\|T_{W_{n_k}} f_k - T_W f_k\|_2 \geq \Big[\int_{x \in I_k} \Big(T_{W_{n_k}} f_k(x) - T_W f_k(x)\Big)^2 dx \Big]^{1/2}
$$

$$
\geq \Big[\int_{x \in I_k} \Big(T_{W_{n_k}} f_k(x)\Big)^2 dx \Big]^{1/2} - \Big[\int_{x \in I_k} \Big(T_W f_k(x)\Big)^2 dx \Big]^{1/2}
$$

$$
= n_k^{-1} \beta_{i_k j_k} - \Big[\int_{x \in I_k} \Big(T_W f_k(x)\Big)^2 dx \Big]^{1/2} \geq \varepsilon_0 n_k^{-1/3}(1 - o(1)), \qquad (109)
$$

where the second step follows from the Minkowski inequality and the last step uses (108). Combining (109) with (107), we see that

$$
\liminf_{k \to \infty} \big\| n_k^{1/3} \big(T_{W_{n_k}} f_k - T_W f_k\big) - T_H f_k \big\|_2 \geq \varepsilon_0,
$$

which contradicts (20). This proves (106).

Fix $\varepsilon > 0$ and define $W_n^{(\leq)} := W \mathbb{1}_{\{W \leq \varepsilon n^{2/3}\}}$ and $W_n^{(>)} := W - W_n^{(\leq)}$. Then

$$
\|T_{W_n^{(>)}}\|_{3,3} \leq \Big(\int_{W > \varepsilon n^{2/3}} W^3 \Big)^{1/3}, \quad \text{and} \qquad (110)
$$

$$
\|T_{W_n^{(>)}}\|_{2,2} \leq \Big(\int_{W > \varepsilon n^{2/3}} W^2 \Big)^{1/2} \leq \frac{1}{\varepsilon^{1/2} n^{1/3}} \Big(\int_{W > \varepsilon n^{2/3}} W^3 \Big)^{1/2}. \qquad (111)
$$

Let $f \in L^3([0,1])$ with $\|f\|_3 = 1$. Then for all large n and $x \in [0,1]$,

$$
\big|T_{W_n} f(x) - T_{W_n^{(\leq)}} f(x)\big| \leq \int_0^1 \big|W_n(x,y) - W_n^{(\leq)}(x,y)\big| \cdot \big|f(y)\big| dy \leq 2\varepsilon n^{2/3} \|f\|_1 \leq 2\varepsilon n^{2/3}, \quad (112)
$$

where the second step uses (106), and the last step uses the fact $\|f\|_1 \leq \|f\|_3 = 1$. Hence, for large n,

$$
\|T_{W_n} f - T_{W_n^{(\leq)}} f\|_3^3 \leq 2\varepsilon n^{2/3} \|T_{W_n} f - T_{W_n^{(\leq)}} f\|_2^2 \leq 2\varepsilon n^{2/3} \|T_{W_n} - T_{W_n^{(\leq)}}\|_{2,2}^2 \cdot \|f\|_2^2
$$

$$
\leq 2\varepsilon n^{2/3} \|T_{W_n} - T_{W_n^{(\leq)}}\|_{2,2}^2 \leq 4\varepsilon n^{2/3} \Big(\|T_{W_n} - T_W\|_{2,2}^2 + \|T_{W_n^{(>)}}\|_{2,2}^2 \Big)
$$

$$
\leq 4\varepsilon \|T_H\|_{2,2}^2 + \varepsilon + 4 \int_{W > \varepsilon n^{2/3}} W^3, \qquad (113)
$$

where the first step uses (112), the third step uses the fact $\|f\|_2 \leq \|f\|_3 = 1$, and the last step follows from (20) and (111). Since $\|T_{W_n} f - T_W f\|_3 \leq \|T_{W_n} f - T_{W_n^{(\leq)}} f\|_3 + \|T_{W_n^{(>)}}\|_{3,3}$, it follows from (110) and (113) that for all large n,

$$
\|T_{W_n} - T_W\|_{3,3} \leq \Big(4\varepsilon \|T_H\|_{2,2}^2 + \varepsilon + 4 \int_{W > \varepsilon n^{2/3}} W^3 \Big)^{1/3} + \Big(\int_{W > \varepsilon n^{2/3}} W^3 \Big)^{1/3},
$$

which yields the desired claim.   ∎

Proof of Lemma 6.4: Note that for $p = 2, 3$,

$$
\|T_n - T_{W_n}\|_{p,p} \leq \Big[\int_{[0,1]^2} \big(\kappa_n - W_n\big)^p \Big]^{1/p} \lesssim n^{-\delta_0}, \qquad (114)
$$

where the last step uses the definition of κ_n from (96). Hence,

$$
\|T_n - T_W\|_{3,3} = o(1) \quad \text{and} \qquad (115)
$$

$$
\big|1 - \|T_n\|\big| \leq \|T_n - T_W\|_{2,2} = O(n^{-\delta_0}) \qquad (116)
$$

as $n \to \infty$, where (115) uses (114) and Lemma 6.5, and (116) follows from (114) and (20). Now,

$$\|T_n\| \cdot \|\psi_n\|_3 = \|T_n \psi_n\|_3 \le \|T_n\|_{2,3} \le C,$$

where the last step uses (99). It thus follows from (116) that $(\psi_n, n \ge 1)$ is a bounded sequence in $L^3([0,1])$. Consider the sequence $(T_W \psi_n, n \ge 1)$. Since T_W is a compact operator on $L^3([0,1])$, for any subsequence $(T_W \psi_{n_k}, k \ge 1)$, there exists a further subsequence $(T_W \psi_{n_{k_\ell}}, \ell \ge 1)$ and $\phi \in L^3([0,1])$ such that

$$\|T_W \psi_{n_{k_\ell}} - \phi\|_3 \to 0 \quad \text{as} \quad \ell \to \infty. \tag{117}$$

Writing

$$T_W \psi_n = \psi_n + (T_W - T_n)\psi_n + \big(\|T_n\| - 1\big)\psi_n, \tag{118}$$

and using (115), (116) in conjunction with (117) yields

$$\|\psi_{n_{k_\ell}} - \phi\|_3 \to 0 \quad \text{as} \quad \ell \to \infty. \tag{119}$$

This in turn implies that $\|T_W \psi_{n_{k_\ell}} - T_W \phi\|_3 \to 0$, which combined with (117) yields

$$T_W \phi = \phi. \tag{120}$$

Now, (119) shows that $\phi(x) \ge 0$ for $\mu_\circ$ a.e. x and $\|\phi\|_2 = 1$. Hence, using (120) and Lemma 4.4, we conclude that $\phi = \psi$. Thus, for any subsequence we can extract a further subsequence along which $\|\psi_n - \psi\|_3 \to 0$. This proves (104).

Since T_W is a compact and self-adjoint operator on $L^2([0,1])$, we can write $T_W = \langle \psi, \cdot \rangle \psi + \sum_{j \ge 2} \lambda_j \langle \phi_j, \cdot \rangle \phi_j$, where λ_j, $j \ge 2$, are the non-zero eigenvalues (if any) of T_W other than 1 counted according to their multiplicity, and $\{\psi\} \cup \{\phi_j : j \ge 2\}$ is an orthonormal collection of elements of $L^2([0,1])$. Let $\mathscr{P} : L^2([0,1]) \to \text{kernel}(T_W)$ denote the orthogonal projection operator onto the kernel of T_W. Then

$$\psi_n = \langle \psi, \psi_n \rangle \psi + \sum_{j \ge 2} \langle \phi_j, \psi_n \rangle \phi_j + \mathscr{P}\psi_n, \text{ and } T_W \psi_n = \langle \psi, \psi_n \rangle \psi + \sum_{j \ge 2} \lambda_j \langle \phi_j, \psi_n \rangle \phi_j. \tag{121}$$

Hence,

$$Cn^{-2\delta_0} \ge \big\|(T_W - T_n)\psi_n + \big(\|T_n\| - 1\big)\psi_n\big\|_2^2 = \sum_{j \ge 2} \big(1 - \lambda_j\big)^2 \langle \phi_j, \psi_n \rangle^2 + \|\mathscr{P}\psi_n\|_2^2$$

$$\ge \Delta_W^2 \Big(\sum_{j \ge 2} \langle \phi_j, \psi_n \rangle^2 + \|\mathscr{P}\psi_n\|_2^2 \Big) = \Delta_W^2 \big(1 - \langle \psi, \psi_n \rangle^2\big), \tag{122}$$

where the first step uses (116), the second step uses (121), and the third step uses Lemma 4.4 (c). We know from (104) that $\|\psi_n - \psi\|_2 = o(1)$, which combined with (122) yields the lower bound $\langle \psi, \psi_n \rangle \ge 1 - Cn^{-2\delta_0}$. Consequently,

$$\|\psi - \psi_n\|_2^2 = \|\psi\|_2^2 + \|\psi_n\|_2^2 - 2\langle \psi, \psi_n \rangle = 2 - 2\langle \psi, \psi_n \rangle = O(n^{-2\delta_0}),$$

as claimed in (105). ∎

6.4. **Asymptotics for $\|T_n\|$ and $\|g_n\|_1$.** Throughout this section, we work under Conditions 4.1, 4.2, and 6.2. Recall the constant ζ from (25) and the function $g_n = g_{n;1}$ from (102). Recall also from (100) that we are working with $\delta_0 > 1/6$. The aim of this section is to prove the following result:

Lemma 6.6. *We have*

$$\lim_{n\to\infty} \frac{\left(\int_0^1 g_n(x)\,dx\right) - n^{\delta_0}}{n^{2\delta_0 - 1/3}} = \zeta.$$

We will make use of the following result:

Lemma 6.7. *Suppose A_1, A_2, and A_3 are compact self-adjoint operators on $L^2([0,1])$. Further, assume that $\|A_1\|_{2,2}$ is an eigenvalue of A_1 with multiplicity one, and $-\|A_1\|_{2,2}$ is not an eigenvalue of A_1. Then*

$$\lim_{(x,y)\to(0,0)} \frac{1}{y}\cdot\left(\|A_1 + xA_2 + yA_3\|_{2,2} - \|A_1 + xA_2\|_{2,2}\right) = \left\langle \phi_{A_1}, A_3\phi_{A_1}\right\rangle, \tag{123}$$

where ϕ_{A_1} is an eigenfunction of A_1 corresponding to the eigenvalue $\|A_1\|_{2,2}$ with $\|\phi_{A_1}\|_2 = 1$.

Proof: We claim that there exist an open $U \subseteq \mathbb{R}^2$ containing the origin, and infinitely differentiable functions $\Lambda : U \to \mathbb{R}$ and $\Phi : U \to L^2([0,1])$ (the latter differentiability to be interpreted in the Fréchet sense) such that $\Lambda(0,0) = \|A_1\|_{2,2}$, $\Phi(0,0) = \phi_{A_1}$, and $\|\Phi(x,y)\|_2 = 1$ for $(x,y) \in U$. Further, for each $(x,y) \in U$, $\Lambda(x,y)$ is an eigenvalue of $(A_1 + xA_2 + yA_3)$, and $\Phi(x,y)$ is an associated eigenfunction.

Let us first complete the proof assuming the above claim. Choose $\varepsilon > 0$ small so that $(-\varepsilon,\varepsilon)^2 \subseteq U$. For $(x,y) \in (-\varepsilon,\varepsilon)^2$,

$$\left\|\left(A_1 + xA_2 + yA_3\right) - \Lambda(x,y)\left\langle\Phi(x,y),\cdot\right\rangle\Phi(x,y)\right\|_{2,2}$$
$$\leq \varepsilon\|A_2\|_{2,2} + \varepsilon\|A_3\|_{2,2} + \left\|A_1 - \|A_1\|_{2,2}\left\langle\phi_{A_1},\cdot\right\rangle\phi_{A_1}\right\|_{2,2}$$
$$+ \left\|\|A_1\|_{2,2}\left\langle\phi_{A_1},\cdot\right\rangle\phi_{A_1} - \Lambda(x,y)\left\langle\Phi(x,y),\cdot\right\rangle\Phi(x,y)\right\|_{2,2}. \tag{124}$$

By our assumption, $\left\|A_1 - \|A_1\|_{2,2}\left\langle\phi_{A_1},\cdot\right\rangle\phi_{A_1}\right\|_{2,2} < \|A_1\|_{2,2} = \Lambda(0,0)$, and consequently, the right side of (124) is strictly smaller that $\Lambda(x,y)$ (by choosing a sufficiently small ε). Hence, for each $(x,y) \in (-\varepsilon,\varepsilon)^2$, $\Lambda(x,y) = \|A_1 + xA_2 + yA_3\|_{2,2}$ and further, it is an eigenvalue of $(A_1 + xA_2 + yA_3)$ of multiplicity one and the corresponding (one-dimensional) eigenspace is spanned by $\Phi(x,y)$. Thus, the left side of (123) is simply $\frac{\partial}{\partial y}\Lambda(0,0)$, so we just need to show that

$$\frac{\partial}{\partial y}\Lambda(0,0) = \left\langle\phi_{A_1}, A_3\phi_{A_1}\right\rangle. \tag{125}$$

To this end, for $y \in (-\varepsilon,\varepsilon)$, write $\Phi_2(y) = \Phi(0,y)$, $D_y\Phi_2$ for the Fréchet derivative of Φ_2, and identify the linear operator $D_y\Phi_2(y) : \mathbb{R} \to L^2([0,1])$ with $\left[D_y\Phi_2(y)\right](1) \in L^2([0,1])$. Then differentiating both sides of the relation $\langle\Phi_2(y),\Phi_2(y)\rangle = 1$, $y \in (-\varepsilon,\varepsilon)$, we get

$$\left\langle\Phi_2(y), D_y\Phi_2(y)\right\rangle = 0 \quad\text{for all}\quad y \in (-\varepsilon,\varepsilon). \tag{126}$$

Hence,

$$\frac{1}{y}\Big(\big\langle\Phi_2(y),\,A_1\Phi_2(y)\big\rangle-\Lambda(0,0)\Big)\xrightarrow{y\to 0}2\Lambda(0,0)\cdot\big\langle\Phi_2(0),\,D_y\Phi_2(0)\big\rangle=0, \qquad (127)$$

where the last equality follows from (126). Now, $\big\langle\Phi_2(y),\,(A_1+yA_3)\Phi_2(y)\big\rangle=\Lambda(0,y)$. Consequently,

$$\begin{aligned}
y\cdot\big\langle\Phi_2(y),\,A_3\Phi_2(y)\big\rangle&=\Lambda(0,y)-\big\langle\Phi_2(y),\,A_1\Phi_2(y)\big\rangle\\
&=\Big(\Lambda(0,y)-\Lambda(0,0)\Big)+\Big(\Lambda(0,0)-\big\langle\Phi_2(y),\,A_1\Phi_2(y)\big\rangle\Big).
\end{aligned}$$

Now (125) follows upon dividing by y and letting $y\to 0$ in the last equation and using (127).

It remains to prove the existence of the functions Λ and Φ. This is well-known in the setting of one parameter perturbations of matrices, and a proof using the implicit function theorem can be found in the note [49]. A similar approach can be taken in the present setting; we only outline the steps. Consider the function $F:\mathbb{R}\times\mathbb{R}\times L^2([0,1])\times\mathbb{R}\to L^2([0,1]\times\mathbb{R}$ given by

$$\big(x,y,\phi,\lambda\big)\mapsto\Big(\big(A_1+xA_2+yA_3\big)\phi-\lambda\phi,\ \|\phi\|_2^2\Big).$$

Then $F\big(0,0,\phi_{A_1},\|A_1\|_{2,2}\big)=(\mathbf{0},1)$, where $\mathbf{0}$ is the element of $L^2([0,1])$ that is identically zero. Further, F is infinitely differentiable. A direct calculation shows that the Fréchet derivative $D_{\phi,\lambda}F$, at the point $\big(0,0,\phi_{A_1},\|A_1\|_{2,2}\big)$, is the following linear transformation on $L^2([0,1])\times\mathbb{R}$:

$$L^2([0,1])\times\mathbb{R}\ni(\phi,\lambda)\mapsto\big(A_1\phi-\|A_1\|_{2,2}\phi-\lambda\phi_{A_1},\,2\langle\phi,\phi_{A_1}\rangle\big).$$

It can be easily shown that this is a one-to-one and onto linear transformation (and consequently its inverse is a bounded linear operator). Thus, the implicit function theorem in Banach spaces [50, Chapter I, Theorem 5.9] guarantees the existence of an open $U\ni(0,0)$, and infinitely differentiable functions $\Lambda:U\to\mathbb{R}$ and $\Phi:U\to L^2([0,1])$ having the desired properties. This completes the proof. ∎

Proof of Lemma 6.6: Recall the notation $\mathbf{1}_\circ$ and $\mathbf{1}_\square$ introduced at the end of Section 6.2. Let

$$\kappa_n^-:=W+n^{-1/3}\cdot H-n^{-\delta_0}\cdot\mathbf{1}_\square\quad\text{and}\quad\kappa_n^=:=W_n-n^{-\delta_0}\cdot\mathbf{1}_\square.$$

Then

$$\begin{aligned}
\big|\|T_n\|-\|T_{\kappa_n^-}\|_{2,2}\big|^2&\le 2\|T_n-T_{\kappa_n^=}\|_{2,2}^2+2\|T_{\kappa_n^=}-T_{\kappa_n^-}\|_{2,2}^2\\
&\le 2\int_{[0,1]^2}\big(\kappa_n-\kappa_n^=\big)^2+o(n^{-2/3})\le\frac{4}{n^{2\delta_0}}\cdot\frac{|B_n|+1}{n}+4e^{-2n}+o(n^{-2/3})\\
&=o(n^{-2\delta_0-1+\varpi_0})+o(n^{-2/3})=o(n^{-2/3}), \qquad (128)
\end{aligned}$$

where the second step uses (20) and the last two steps use Condition 6.2. Now, as seen in the proof of Lemma 6.7, the function $x\mapsto\|T_W-xT_{\mathbf{1}_\square}\|_{2,2}$ is infinitely differentiable in a neighborhood of the origin. Thus, using Lemma 6.7 and the fact $\|T_W\|_{2,2}=1$, we get

$$1-\|T_W-n^{-\delta_0}T_{\mathbf{1}_\square}\|_{2,2}=n^{-\delta_0}\langle\psi,\,T_{\mathbf{1}_\square}\psi\rangle+O(n^{-2\delta_0})=n^{-\delta_0}\Big(\int_0^1\psi\Big)^2+O(n^{-2\delta_0}), \qquad (129)$$

Further,

$$n^{\delta_0}\big(1 - \|T_n\|\big) = n^{\delta_0}\big(1 - \|T_{\kappa_n^-}\|_{2,2}\big) + o(n^{\delta_0 - 1/3})$$

$$= n^{\delta_0}\big(1 - \|T_W - n^{-\delta_0}T_{\mathbf{1}_\square}\|_{2,2}\big) + O(n^{\delta_0 - 1/3}) \to \Big(\int_0^1 \psi\Big)^2 \tag{130}$$

as $n \to \infty$, where the first step uses (128), the second step uses the deifinition of κ_n^-, and the last step uses (129) and the fact $\delta_0 < 1/3$. Next, recall (101), and note that

$$|\lambda_{n;2}| = \big\|T_n - \|T_n\|\langle \psi_n, \cdot \rangle \psi_n \big\|_{2,2} \xrightarrow{n \to \infty} \big\|T_W - \langle \psi, \cdot \rangle \psi\big\|_{2,2} \leq 1 - \Delta_W,$$

where the penultimate step uses (105) and (116), and the last step follows from Lemma 4.4 (c). In view of the last display and (130), we can assume without loss of generality that

$$1/2 \leq \|T_n\| < 1 \quad \text{and} \quad |\lambda_{n;2}| \leq 1 - (\Delta_W/2) \quad \text{for all} \quad n. \tag{131}$$

Recall the definition of $|\mathfrak{G}_\ell(\cdot)|$ from around (103). It is easy to see that for any $\ell \geq 1$,

$$\mathbb{E}\big(|\mathfrak{G}_\ell(x)|\big) = T_n^\ell \mathbf{1}_\circ(x), \ x \in [0,1], \quad \text{and} \quad T_n^\ell = \|T_n\|^\ell \langle \psi_n, \cdot \rangle \psi_n + \sum_{j=2}^n \lambda_{n;j}^\ell \langle \psi_{n;j}, \cdot \rangle \psi_{n;j}. \tag{132}$$

It thus follows that

$$g_n = \frac{1}{1 - \|T_n\|}\langle \psi_n, \mathbf{1}_\circ \rangle \psi_n + \sum_{j=2}^n \frac{1}{1 - \lambda_{n;j}}\langle \psi_{n;j}, \mathbf{1}_\circ \rangle \psi_{n;j}, \tag{133}$$

and consequently

$$\int_0^1 g_n(x)\,dx = \frac{1}{1 - \|T_n\|}\langle \psi_n, \mathbf{1}_\circ \rangle^2 + \sum_{j=2}^n \frac{1}{1 - \lambda_{n;j}}\langle \psi_{n;j}, \mathbf{1}_\circ \rangle^2 = \frac{1}{1 - \|T_n\|}\Big(\int_0^1 \psi_n\Big)^2 + O(1),$$

where we have used (131) and the fact that $\sum_{j=2}^n \langle \psi_{n;j}, \mathbf{1}_\circ \rangle^2 \leq 1$. Hence,

$$\frac{\big(\int_0^1 g_n\big) - n^{\delta_0}}{n^{2\delta_0 - 1/3}} = \frac{\big(\int_0^1 \psi_n\big)^2 - n^{\delta_0}\big(1 - \|T_n\|\big)}{(1 - \|T_n\|) \cdot n^{2\delta_0 - 1/3}} + o(1). \tag{134}$$

It follows from (105) that

$$\int_0^1 \psi_n = \|\psi_n\|_1 = \|\psi\|_1 + O(n^{-\delta_0}). \tag{135}$$

Hence,

$$\frac{\big(\int_0^1 \psi_n\big)^2 - n^{\delta_0}\big(1 - \|T_n\|\big)}{(1 - \|T_n\|) \cdot n^{2\delta_0 - 1/3}} \sim \frac{\big(\int_0^1 \psi_n\big)^2 - n^{\delta_0}\big(1 - \|T_n\|\big)}{\big(\int_0^1 \psi\big)^2 \cdot n^{\delta_0 - 1/3}} = \frac{\big(\int_0^1 \psi\big)^2 - n^{\delta_0}\big(1 - \|T_n\|\big)}{\big(\int_0^1 \psi\big)^2 \cdot n^{\delta_0 - 1/3}} + o(1)$$

$$= \frac{n^{\delta_0}\big(1 - \|T_W - n^{-\delta_0}T_{\mathbf{1}_\square}\|_{2,2}\big) - n^{\delta_0}\big(1 - \|T_n\|\big)}{\big(\int_0^1 \psi\big)^2 \cdot n^{\delta_0 - 1/3}} + o(1)$$

$$= \frac{n^{\delta_0}\Big(\|T_{\kappa_n^-}\|_{2,2} - \|T_W - n^{-\delta_0}T_{\mathbf{1}_\square}\|_{2,2}\Big)}{\big(\int_0^1 \psi\big)^2 \cdot n^{\delta_0 - 1/3}} + o(1) = \frac{\langle \psi, T_H \psi \rangle}{\big(\int_0^1 \psi\big)^2} + o(1), \tag{136}$$

where the first step uses (130), the second step uses (135), the third step uses (129), the fourth step uses (128), and the last step uses Lemma 6.7; in addition, we have repeatedly used the fact that $\delta_0 > 1/6$. Combining (134) and (136) completes the proof. ∎

6.5. **Norm estimates for the functions $g_{n;k}$.** Throughout this section, we work under Conditions 4.1, 4.2, and 6.2. Recall the functions $g_{n;k}$ from (102). Recall also that we are working under the assumptions of (131) and (100). The following lemma is the main result of this section:

Lemma 6.8. *The following hold:*

(a) $\|g_n\|_3 = O(n^{\delta_0})$;

(b) $\|g_{n;2}\|_3 = O(n^{3\delta_0})$ *and further,*

$$\frac{1}{n^{3\delta_0}} \cdot \int_0^1 g_{n;2} \rightarrow \frac{\int_0^1 \psi^3}{(\int_0^1 \psi)^3} \quad as \quad n \rightarrow \infty; \tag{137}$$

(c) $\|g_{n;k}\|_{3/2} = O\big(n^{(2k-1)\delta_0}\big)$ *for $k = 3, 4, 5$.*

We will need the next two lemmas in the proof. Recall the notation used in (101), and define

$$\widetilde{h}_n(x, y) := \sum_{j=2}^n \Big(\frac{\lambda_{n;j}}{1 - \lambda_{n;j}}\Big) \psi_{n;j}(x) \psi_{n;j}(y), \quad x, y \in [0, 1]. \tag{138}$$

Lemma 6.9. *We have,* $\sup_{n \geq 1} \int_{[0,1]^2} |\widetilde{h}_n|^3 < \infty$.

Lemma 6.10. (a) *For any $p \in [3/2, 3]$, $(I - T_n)$ is an invertible operator on $L^p([0, 1])$, where I denotes the identity operator. Further, there exists $C > 0$ such that $\|(I - T_n)^{-1}\|_{p,p} \leq Cn^{\delta_0}$ for all n.*

(b) *We have,* $\|T_n\|_{3,\infty} = o(n^{\theta_0})$.

Proof of Lemma 6.9: Define

$$h_n(x, y) := \sum_{j=2}^n \lambda_{n;j} \psi_{n;j}(x) \psi_{n;j}(y) = \kappa_n(x, y) - \|T_n\| \psi_n(x) \psi_n(y), \quad x, y \in [0, 1]. \tag{139}$$

Then note that

$$\int_{y=0}^1 h_n^2(x, y) dy \lesssim \int_{y=0}^1 \kappa_n^2(x, y) dy + \psi_n^2(x) \leq \int_{y=0}^1 W_n^2(x, y) dy + \psi_n^2(x), \tag{140}$$

and

$$\int_{y=0}^1 \widetilde{h}_n^2(x, y) dy = \sum_{j=2}^n \Big(\frac{\lambda_{n;j}}{1 - \lambda_{n;j}}\Big)^2 \psi_{n;j}^2(x) \lesssim \sum_{j=2}^n \lambda_{n;j}^2 \psi_{n;j}^2(x)$$

$$= \int_{y=0}^1 h_n^2(x, y) dy \leq \int_{y=0}^1 W_n^2(x, y) dy + \psi_n^2(x), \tag{141}$$

where the second step uses (131), and the last step uses (140). Now,

$$\widetilde{h}_n(x, y) = h_n(x, y) + \int_{z=0}^1 h_n(x, z) \widetilde{h}_n(z, y) dz = h_n(x, y) + \int_{z=0}^1 \widetilde{h}_n(x, z) h_n(z, y) dz, \tag{142}$$

and consequently,

$$|\widetilde{h}_n(x, y)| \leq e^{-n} + W_n(x, y) + \psi_n(x) \psi_n(y) + \Big(\int_{z=0}^1 h_n^2(x, z) dz\Big)^{1/2} \cdot \Big(\int_{z=0}^1 \widetilde{h}_n^2(z, y) dz\Big)^{1/2}. \tag{143}$$

We thus have

$$\int_{[0,1]^2} |\widetilde{h}_n|^3 \lesssim e^{-3n} + \int_{[0,1]^2} W_n^3 + \left(\int_0^1 \psi_n^3\right)^2$$
$$+ \left[\int_{x=0}^1 \left(\psi_n^2(x) + \int_{z=0}^1 W_n^2(x,z)dz\right)^{3/2} dx\right] \cdot \left[\int_{y=0}^1 \left(\psi_n^2(y) + \int_{z=0}^1 W_n^2(y,z)dz\right)^{3/2} dy\right]$$
$$\lesssim e^{-3n} + \int_{[0,1]^2} W_n^3 + \left(\int_0^1 \psi_n^3\right)^2 + \left[\int_0^1 \psi_n^3 + \int_{[0,1]^2} W_n^3\right]^2, \tag{144}$$

where the first step uses (143), (140), and (141). By (104), $(\|\psi_n\|_3, \, n \geq 1)$ is a bounded sequence, and hence the desired result follows from (144) and (19). ∎

Proof of Lemma 6.10 (a): Fix $p \in [3/2, 3]$, and regard all operators as acting on $L^p([0,1])$ in this proof. It follows from (139) and (142) that

$$T_n = T_{h_n} + \|T_n\|\langle\psi_n, \cdot\rangle\psi_n \quad \text{and} \quad T_{\widetilde{h}_n} = T_{h_n} + T_{h_n}T_{\widetilde{h}_n} = T_{h_n} + T_{\widetilde{h}_n}T_{h_n}.$$

From this, it can be seen easily that $(I - T_n)$ is an invertible operator on $L^p([0,1])$, and

$$(I - T_n)^{-1} = I + T_{\widetilde{h}_n} + \frac{\|T_n\|}{1 - \|T_n\|}\langle\psi_n, \cdot\rangle\psi_n. \tag{145}$$

Let $p' := p/(p-1)$. Since $p, p' \in [3/2, 3]$, it follows from (104) that

$$\|\psi_n\|_p = O(1) \quad \text{and} \quad \|\psi_n\|_{p'} = O(1). \tag{146}$$

It thus follows from (145) that

$$\|(I - T_n)^{-1}\|_{p,p} \leq 1 + \left(\int_{[0,1]^2} |\widetilde{h}_n|^3\right)^{1/3} + Cn^{\delta_0}\|\psi_n\|_p \cdot \|\psi_n\|_{p'} \leq C'n^{\delta_0},$$

where the first step uses Lemma 6.1 and (130), and the second step uses Lemma 6.9 and (146). This yields the claim in Lemma 6.10 (a). ∎

Proof of Lemma 6.10 (b): For any $f \in L^3([0,1])$ and $x \in [0,1]$,

$$|T_n f(x)| \leq \int_0^1 \kappa_n(x,y)|f(y)|dy \leq \|f\|_3 \cdot \max_{x \in [0,1]} \left(\int_0^1 W_n^{3/2}(x,y)dy\right)^{2/3}.$$

Now the desired result follows from Condition 4.2. ∎

Proof of Lemma 6.8 (a): Let ε_{ij}, $i,j \in [n]$, and $\mathfrak{X}'_n(j)$, $j \in [n]$, be independent random variables such that $\varepsilon_{ij} \sim \text{Bernoulli}(\kappa_{ij}/n)$ and $\mathfrak{X}'_n(j) \stackrel{\text{d}}{=} \mathfrak{X}_n^{\triangle}(j)$. Then for each $i \in [n]$,

$$\mathfrak{X}_n^{\triangle}(i) - 1 \stackrel{\text{d}}{=} \sum_{j \in [n]} \varepsilon_{ij} \cdot \mathfrak{X}'_n(j). \tag{147}$$

Taking expectation on both sides yields $g_n(x) = \mathbf{1}_\circ(x) + T_n g_n(x)$, $x \in [0,1]$. Hence,

$$g_n = (I - T_n)^{-1}\mathbf{1}_\circ, \tag{148}$$

and consequently, $\|g_n\|_3 \leq \|(I - T_n)^{-1}\|_{3,3} \cdot \|\mathbf{1}_\circ\|_3 \lesssim n^{\delta_0}$, where we have used Lemma 6.10 (a). This completes the proof. ∎

Proof of Lemma 6.8 (b): Squaring both sides of (147) and taking expectation yields

$$g_{n;2}^{\triangle}(i) - 2g_n^{\triangle}(i) + 1 = \sum_{j \in [n]} \frac{\kappa_{ij}}{n} g_{n;2}^{\triangle}(j) + \left(\sum_{j \in [n]} \frac{\kappa_{ij}}{n} g_n^{\triangle}(j)\right)^2 - \sum_{j \in [n]} \frac{\kappa_{ij}^2}{n^2}\left(g_n^{\triangle}(j)\right)^2,$$

for $i \in [n]$. Hence, for $x \in [0,1]$,

$$g_{n;2}(x) = T_n g_{n;2}(x) + \big(T_n g_n(x)\big)^2 + 2g_n(x) - \mathbf{1}_\circ(x) - \frac{1}{n}\int_{y=0}^1 \kappa_n^2(x,y) g_n^2(y)\,dy \qquad (149)$$

$$\leq T_n g_{n;2}(x) + \big(T_n g_n(x)\big)^2 + 2g_n(x). \qquad (150)$$

Since $\|T_n\| < 1$, $(I - T_n)^{-1}$ viewed as an operator on $L^2([0,1])$ admits the expansion

$$(I - T_n)^{-1} = I + T_n + T_n^2 + \dots.$$

Thus, if $f_1 \leq f_2$ are two functions in $L^2([0,1])$, then $T_n^\ell f_1 \leq T_n^\ell f_2$ for each $\ell \geq 1$ and consequently, $(I - T_n)^{-1} f_1 \leq (I - T_n)^{-1} f_2$. Hence, (150) implies that

$$g_{n;2} \leq 2(I - T_n)^{-1} g_n + (I - T_n)^{-1}\big[(T_n g_n)^2\big], \qquad (151)$$

which in turn yields

$$\|g_{n;2}\|_{3/2} \lesssim \|(I - T_n)^{-1}\|_{3/2,3/2} \cdot \|g_n\|_{3/2} + \|(I - T_n)^{-1}\|_{3/2,3/2} \cdot \Big(\int_0^1 (T_n g_n)^3\Big)^{2/3}$$

$$\lesssim n^{2\delta_0} + n^{\delta_0} \cdot \|T_n g_n\|_3^2 \lesssim n^{2\delta_0} + n^{\delta_0} \cdot \|T_n\|_{3,3}^2 \|g_n\|_3^2 \lesssim n^{3\delta_0}, \qquad (152)$$

where the second step uses Lemma 6.10 (a) and Lemma 6.8 (a), and the last step follows from (99) and Lemma 6.8 (a). Going back to (150), we see that

$$\|g_{n;2}\|_3 \lesssim \|T_n\|_{3/2,3} \cdot \|g_{n;2}\|_{3/2} + \|g_n\|_3 + \Big(\int_0^1 (T_n g_n)^6\Big)^{1/3}$$

$$\lesssim n^{3\delta_0} + n^{\delta_0} + \|T_n g_n\|_\infty \cdot \|T_n g_n\|_3$$

$$\lesssim n^{3\delta_0} + n^{\theta_0} \|g_n\|_3 \cdot \|T_n\|_{3,3} \cdot \|g_n\|_3 \lesssim n^{3\delta_0} + n^{2\delta_0 + 1/12} = O(n^{3\delta_0}),$$

where the second step uses (99), (152), Lemma 6.8 (a), the third step uses Lemma 6.10 (b), the fourth step uses (99), Lemma 6.8 (a), and the relation $\theta_0 < 1/12$ from Condition 4.2, and the last step uses the fact that $\delta_0 > 1/6$ from (100). This proves the first claim made in Lemma 6.8 (b).

We now turn to the proof of (137). To this end, we examine the terms appearing on the right side of (149). It follows from (133) and (138) that

$$T_n g_n = \frac{\|T_n\|}{1 - \|T_n\|}\big\langle \psi_n, \mathbf{1}_\circ\big\rangle \psi_n + T_{\widetilde{h}_n} \mathbf{1}_\circ. \qquad (153)$$

Hence,

$$(T_n g_n)^2 = \frac{\|T_n\|^2}{(1 - \|T_n\|)^2}\Big(\int_0^1 \psi_n\Big)^2 \cdot \psi_n^2 + a_n, \qquad (154)$$

where

$$\|a_n\|_{3/2} \lesssim n^{\delta_0}\|\psi_n\|_3 \cdot \big\|T_{\widetilde{h}_n}\mathbf{1}_\circ\big\|_3 + \big\|\big(T_{\widetilde{h}_n}\mathbf{1}_\circ\big)^2\big\|_{3/2} \lesssim n^{\delta_0}; \qquad (155)$$

here, we have used (130), (104), Lemma 6.1, and Lemma 6.9. Now, using (145), we see that

$$(I - T_n)^{-1}\big[\psi_n^2\big] = \frac{\|T_n\|}{1 - \|T_n\|}\Big(\int_0^1 \psi_n^3\Big) \cdot \psi_n + b_n, \qquad (156)$$

where

$$\|b_n\|_{3/2} \leq \big\|I + T_{\widetilde{h}_n}\big\|_{3/2,3/2} \cdot \big\|\psi_n^2\big\|_{3/2} \leq C. \qquad (157)$$

Using Lemma 6.10 (a) and (155), we get

$$\left\| (I - T_n)^{-1} a_n \right\|_{3/2} \lesssim n^{2\delta_0}. \tag{158}$$

Combining (154), (156), (157), and (158) with (104) and (130) yields

$$n^{-3\delta_0} \int_0^1 (I - T_n)^{-1} \left[(T_n g_n)^2 \right] \to \left(\int_0^1 \psi^3 \right) \left(\int_0^1 \psi \right)^{-3} \quad \text{as} \quad n \to \infty. \tag{159}$$

Next,

$$\left| \int_0^1 (I - T_n)^{-1} [2g_n - 1] \right| \le \left\| (I - T_n)^{-1} \right\|_{3/2,3/2} \cdot \|2g_n - 1\|_{3/2} \lesssim n^{2\delta_0}, \tag{160}$$

where we have used Lemma 6.8 (a) and Lemma 6.10 (a). Finally,

$$\left\| \frac{1}{n} \int_{y=0}^1 \kappa_n^2(\cdot, y) g_n^2(y) dy \right\|_{3/2} \lesssim \frac{1}{n^{1/3}} \| T_n g_n^2 \|_{3/2} \lesssim \frac{1}{n^{1/3}} \| g_n^2 \|_{3/2} \lesssim n^{2\delta_0 - 1/3}, \tag{161}$$

where the first step uses (98), the second step uses (99), and the last step uses Lemma 6.8 (a). Combining (149) with (159), (160), and (161) yields the convergence in (137). ∎

Proof of Lemma 6.8 (c)**:** The results of Lemma 6.8 (a) and (b) give the asymptotic behavior for $\|g_{n;k}\|_3$, $k = 1, 2$. Now the asymptotics for $\|g_{n;k}\|_{3/2}$, $k = 3, 4, 5$, can be obtained inductively. We will only give the details for $k = 5$, i.e., we will assume that

$$\|g_{n;k}\|_3 = O\big(n^{(2k-1)\delta_0}\big), \ k = 1, 2, \quad \text{and} \quad \|g_{n;k}\|_{3/2} = O\big(n^{(2k-1)\delta_0}\big), \ k = 3, 4, \tag{162}$$

and from this we will deduce the desired result for $\|g_{n;5}\|_{3/2}$.

Raising both sides of (147) to the power 5 and taking expectation, and then using arguments similar to the one leading to (151), we get

$$g_{n;5} \lesssim (I - T_n)^{-1} \Big[(T_n g_{n;3})(T_n g_{n;2}) + (T_n g_n)^5 + (T_n g_n)^3 (T_n g_{n;2})$$

$$+ (T_n g_{n;4})(T_n g_n) + (T_n g_{n;2})^2 (T_n g_n) + (T_n g_n)^2 (T_n g_{n;3}) + g_{n;4} \Big]. \tag{163}$$

We will estimate the contribution coming from the first three terms on the right side of (163); the other terms can be handled similarly. We start with

$$\left\| (I - T_n)^{-1} \left[(T_n g_{n;3})(T_n g_{n;2}) \right] \right\|_{3/2} \lesssim n^{\delta_0} \left\| (T_n g_{n;3})(T_n g_{n;2}) \right\|_{3/2} \le n^{\delta_0} \| T_n g_{n;3} \|_3 \cdot \| T_n g_{n;2} \|_3$$

$$\le n^{\delta_0} \| T_n \|_{3/2,3}^2 \cdot \| g_{n;3} \|_{3/2} \cdot \| g_{n;2} \|_{3/2} \le C n^{9\delta_0},$$

where the first step uses Lemma 6.10 (a), and the last step uses (162) and (99). Next,

$$\left\| (I - T_n)^{-1} \left[(T_n g_n)^5 \right] \right\|_{3/2} \lesssim n^{\delta_0} \left(\int_0^1 (T_n g_n)^{15/2} \right)^{2/3} \le n^{\delta_0} \left(\| T_n g_n \|_\infty^{9/2} \int_0^1 (T_n g_n)^3 \right)^{2/3}$$

$$\lesssim n^{\delta_0} \left(n^{\theta_0} \| g_n \|_3 \right)^3 \| T_n g_n \|_3^2 \lesssim n^{6\delta_0 + 3\theta_0} = o(n^{9\delta_0}),$$

where the third step uses Lemma 6.10 (b), and the last step uses the fact $\theta_0 < 1/12 < 1/6 < \delta_0$. Similarly,

$$\left\| (I - T_n)^{-1} \left[(T_n g_n)^3 (T_n g_{n;2}) \right] \right\|_{3/2} \lesssim n^{\delta_0} \left\| (T_n g_n)^3 (T_n g_{n;2}) \right\|_{3/2} \le n^{\delta_0} \left\| (T_n g_n)^3 \right\|_3 \cdot \left\| T_n g_{n;2} \right\|_3$$

$$\le n^{\delta_0} \left\| T_n g_n \right\|_\infty^2 \cdot \left\| T_n g_n \right\|_3 \cdot \left\| T_n g_{n;2} \right\|_3 \lesssim n^{7\delta_0 + 2\theta_0} = o(n^{9\delta_0}).$$

As mentioned before, the other terms on the right side of (163) can be bounded in a similar fashion. The claimed result follows upon combining all these bounds. ∎

6.6. **The second susceptibility** $s_2^\bullet$. Throughout this section, we work under Conditions 4.1, 4.2, and 6.2. Recall the constant ζ from (25) and the susceptibility $s_2^\bullet$ from (97). Recall also that we are working under the assumptions of (131) and (100). The aim of this section is to prove the following result:

Lemma 6.11. *We have,* $n^{1/3 - 2\delta_0}\big(s_2^\bullet - n^{\delta_0}\big) \xrightarrow{\mathrm{d}} \zeta$, *as* $n \to \infty$.

Lemma 6.11 follows upon combining Lemma 6.6 with the next lemma and using the fact $\delta_0 < 1/3$.

Lemma 6.12. *The following hold:*
(a) $0 \le (\int_0^1 g_n) - \mathbb{E}[s_2^\bullet] \lesssim n^{4\delta_0 - 1}$, *and*
(b) $\mathrm{Var}(s_2^\bullet) \lesssim n^{5\delta_0 - 1}$.

The rest of this section is devoted to the proof of Lemma 6.12. For $x, y \in [0,1]$ and $\ell \in \mathbb{Z}_{>0}$, define

$$\rho_{n;\ell}(x,y) := \|T_n\|^\ell \psi_n(x)\psi_n(y) + \sum_{j=2}^{n} \lambda_{n;j}^\ell \psi_{n;j}(x)\psi_{n;j}(y), \quad \text{and} \tag{164}$$

$$\rho_n(x,y) := \frac{\|T_n\|}{1 - \|T_n\|}\psi_n(x)\psi_n(y) + \widetilde{h}_n(x,y), \tag{165}$$

where $\widetilde{h}_n$ is as given in (138). Recall the notation $\leftrightarrow$ and $\xrightarrow{\ell}$ introduced below (97). We will need the following simple bounds on the connection probabilities in $\mathscr{G}_n^\bullet$.

Lemma 6.13. *For* $1 \le i \ne j \le n$,

$$\mathbb{P}\big(i \xrightarrow{\ell} j\big) \le \frac{1}{n}\rho_{n;\ell}(x,y) \quad and \quad \mathbb{P}\big(i \leftrightarrow j\big) \le \frac{1}{n}\rho_n(x,y)$$

for any $x \in \big((i-1)/n, \, i/n\big)$ *and* $y \in \big((j-1)/n, \, j/n\big)$.

Proof: Clearly,

$$\mathbb{P}\big(i \xrightarrow{\ell} j\big) \le \sum_{k_1, \dots, k_{\ell-1}} \frac{1}{n^\ell} \cdot \kappa_{ik_1} \cdot \kappa_{k_1 k_2} \dots \kappa_{k_{\ell-1} j}$$

$$= \frac{1}{n}\int_{[0,1]^{\ell-1}} \kappa_n(x,z_1)\kappa_n(z_1,z_2)\dots\kappa_n(z_{\ell-1},y)\, dz_1 dz_2 \dots dz_{\ell-1} = \frac{1}{n}\rho_{n;\ell}(x,y).$$

The bound on $\mathbb{P}\big(i \leftrightarrow j\big)$ follows if we sum over all $\ell \ge 1$. ∎

Proof of Lemma 6.12 (a): Consider $i \in [n]$, and explore its component in $\mathscr{G}_n^\bullet$, namely $\mathscr{C}^\bullet(i)$, in a breadth-first fashion to build its breadth-first spanning tree (BFS tree). The exploration starts at i which is designated as the root of the BFS tree. In each step of this exploration, one looks for edges between the current vertex being explored, say j, and all vertices that have not been discovered up to that point. The vertices found are added as children of j in the BFS tree and they are ordered from left to right using their labels. The vertices in each generation of the BFS tree are explored from left to right, and the exploration process moves to the next generation when all vertices in a generation have been explored. There is an obvious way of coupling this BFS tree with a multitype branching process started from a vertex of type i as described above (102): For each j in the BFS tree
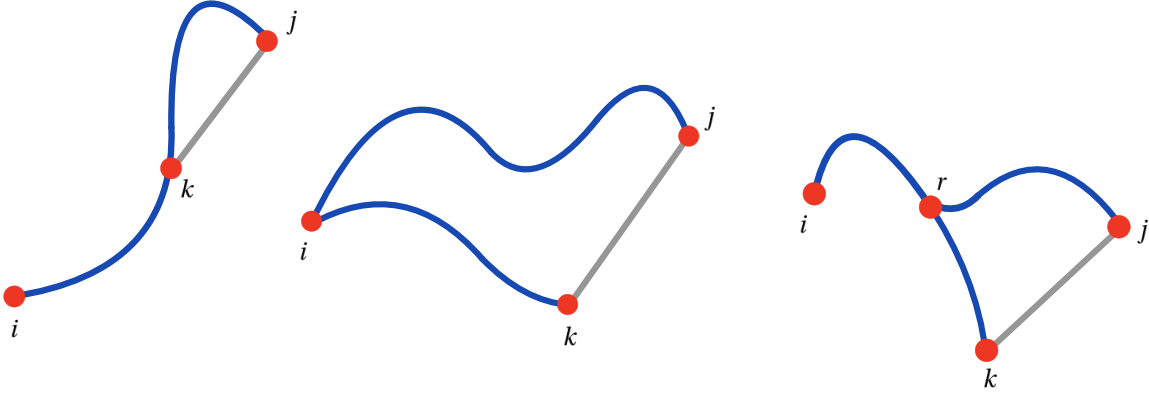
FIGURE 1. The blue curves represent paths of positive length in $\mathscr{G}_n^\bullet$. The gray lines indicate that j gets an extra child of type k with probability κ_{jk}/n.

and each k discovered before the exploration of the vertex j (this includes the vertices j and i), we add an extra child of type k to j with probability κ_{jk}/n, and we do this independently for each such pair (j, k) of vertices. Then we run independent branching processes starting from each new vertex added to the BFS tree. This will yield a multitype branching process tree that contains the BFS tree as a subtree. In particular,

$$\text{diam}\big(\mathscr{C}^\bullet(i)\big) \le 2\mathfrak{H}_n^\triangle(i) \quad \text{and} \quad |\mathscr{C}^\bullet(i)| \le \mathfrak{X}_n^\triangle(i) \quad \text{in this coupling.} \tag{166}$$

Let v be uniformly distributed on $[n]$. Note that $\mathbb{E}[s_2^\bullet] = \mathbb{E}\big(|\mathscr{C}^\bullet(v)|\big)$ and $\mathbb{E}\big[\mathfrak{X}_n^\triangle(v)\big] = \int_0^1 g_n$. It then follows from the above coupling that

$$0 \le \int_0^1 g_n - \mathbb{E}[s_2^\bullet] \le \frac{1}{n}\left(\sum_{1 \le i \ne j \le n} \mathbb{P}\big(i \leftrightarrow j\big) \cdot \left(\frac{\kappa_{ij}}{n} g_n^\triangle(i) + \frac{\kappa_{jj}}{n} g_n^\triangle(j)\right)\right.$$

$$\left. + \sum_{i,j,k \text{ (p.d.)}} \mathbb{P}\big(i \leftrightarrow j, \, i \leftrightarrow k, \, k \text{ before } j\big)\frac{\kappa_{jk}}{n} g_n^\triangle(k)\right), \tag{167}$$

where 'i, j, k (p.d.)' stands for the sum over all $i, j, k \in [n]$ that are pairwise distinct, and 'k before j' represents the event that k is discovered before j is explored in the breadth-first exploration.

We first bound the second term appearing on the right side of (167). Note that

$$\{i \leftrightarrow j, \, i \leftrightarrow k, \, k \text{ before } j\} \subseteq \left(\bigcup_{r \ne i,j,k} \{i \leftrightarrow r\} \circ \{r \leftrightarrow k\} \circ \{r \leftrightarrow j\}\right)$$

$$\bigcup \Big(\{i \leftrightarrow k\} \circ \{k \leftrightarrow j\}\Big) \bigcup \Big(\{i \leftrightarrow k\} \circ \{i \leftrightarrow j\}\Big),$$

where '$\circ$' represents disjoint occurrence; these events are depicted in Figure 1. By the BK inequality,

$$\mathbb{P}\big(i \leftrightarrow j, \, i \leftrightarrow k, \, k \text{ before } j\big) \le \mathbb{P}\big(i \leftrightarrow k\big)\mathbb{P}\big(k \leftrightarrow j\big) + \mathbb{P}\big(i \leftrightarrow k\big)\mathbb{P}\big(i \leftrightarrow j\big)$$

$$+ \sum_{r \ne i,j,k} \mathbb{P}\big(i \leftrightarrow r\big)\mathbb{P}\big(r \leftrightarrow j\big)\mathbb{P}\big(r \leftrightarrow k\big).$$

Combining this bound with Lemma 6.13, we see that

$$\sum_{i,j,k \text{ (p.d.)}} \mathbb{P}\Big(i \leftrightarrow j, \; i \leftrightarrow k, \; k \text{ before } j\Big)\frac{\kappa_{jk}}{n} g_n^{\triangle}(k) \le F_1 + F_2 + F_3\,, \tag{168}$$

where

$$F_1 := \int_{[0,1]^3} \rho_n(x,z)\rho_n(y,z)\kappa_n(y,z)g_n(z)\,dxdydz\,,$$

$$F_2 := \int_{[0,1]^3} \rho_n(x,z)\rho_n(x,y)\kappa_n(y,z)g_n(z)\,dxdydz\,, \text{ and}$$

$$F_3 := \int_{[0,1]^4} \rho_n(x,w)\rho_n(w,y)\rho_n(w,z)\kappa_n(y,z)g_n(z)\,dxdydzdw\,.$$

In order to bound F_1, recall from (148) that $g_n = \mathbf{1}_\circ + T_n g_n$. Hence,

$$\|g_n\|_\infty \le 1 + \|T_n\|_{3,\infty}\cdot\|g_n\|_3 \lesssim n^{\theta_0+\delta_0}\,, \tag{169}$$

where the last step uses Lemma 6.8 (a) and Lemma 6.10 (b). Further, from the expression for ρ_n given in (165), we see that

$$\int_{[0,1]^3} \rho_n^3 \lesssim n^{3\delta_0}\Big(\int_0^1 \psi_n^3\Big)^2 + \int_{[0,1]^3} |\widetilde{h}_n|^3 \lesssim n^{3\delta_0}\,, \tag{170}$$

where the first step uses (130), and the second step uses Lemma 6.9 and (104). Hence,

$$F_1 \lesssim n^{\theta_0+\delta_0}\int_{[0,1]^3}\rho_n(x,z)\rho_n(y,z)\kappa_n(y,z)\,dxdydz \lesssim n^{\delta_0+\theta_0}\cdot n^{2\delta_0} = o(n^{4\delta_0})\,, \tag{171}$$

where the first step uses (169), the second step uses Hölder's inequality in conjuction with (170) and the fact $\int_{[0,1]^2}\kappa_n^3 = O(1)$, and the last step uses the relation $\theta_0 < 1/12 < 1/6 < \delta_0$.

Turning to F_3, recall the function h_n from (139) and note that

$$\int_{[0,1]^2} |h_n|^3 \lesssim \int_{[0,1]^2}\kappa_n^3 + \Big(\int_0^1 \psi_n^3\Big)^2 = O(1)\,. \tag{172}$$

Now,

$$\widetilde{\rho}_n(z,w) := \int_{y=0}^1 \kappa_n(z,y)\rho_n(w,y)\,dy = \frac{\|T_n\|^2}{1-\|T_n\|}\psi_n(w)\psi_n(z) + \sum_{j=2}^n \frac{\lambda_{n;j}^2}{1-\lambda_{n;j}}\psi_{n;j}(w)\psi_{n;j}(z)$$

$$= \frac{\|T_n\|^2}{1-\|T_n\|}\psi_n(w)\psi_n(z) + \widetilde{h}_n(w,z) - h_n(w,z)\,,$$

and consequently,

$$\int_{[0,1]^2}\widetilde{\rho}_n^3 \lesssim n^{3\delta_0}\Big(\int_0^1\psi_n^3\Big)^2 + \int_{[0,1]^3}|\widetilde{h}_n|^3 + \int_{[0,1]^3}|h_n|^3 \lesssim n^{3\delta_0}\,. \tag{173}$$

Hence,

$$F_3 = \int_{[0,1]^3}\big(\rho_n(x,w)g_n(z)\big)\cdot\widetilde{\rho}_n(w,z)\cdot\rho_n(w,z)\,dxdzdw$$

$$\le \Big(\int_{[0,1]^2}\rho_n^3\cdot\int_0^1 g_n^3\Big)^{1/3}\cdot\Big(\int_{[0,1]^2}\widetilde{\rho}_n^3\Big)^{1/3}\cdot\Big(\int_{[0,1]^2}\rho_n^3\Big)^{1/3} \lesssim n^{4\delta_0}\,, \tag{174}$$

where we have used (170), (173), and Lemma 6.8 (a). Similarly,

$$F_2 = \int_{[0,1]^2}\rho_n(x,z)\cdot\widetilde{\rho}_n(x,z)\cdot g_n(z)\,dxdz \lesssim n^{3\delta_0}\,. \tag{175}$$

Combining (168), (171), (174), and (175) yields

$$\sum_{i,j,k \text{ (p.d.)}} \mathbb{P}\left(i \leftrightarrow j, \ i \leftrightarrow k, \ k \text{ before } j\right) \frac{\kappa_{jk}}{n} g_n^{\triangle}(k) \lesssim n^{4\delta_0}. \tag{176}$$

Using similar arguments and the fact that $\kappa_{jj} = e^{-n}$, we can show that

$$\sum_{1 \leq i \neq j \leq n} \mathbb{P}\left(i \leftrightarrow j\right) \cdot \left(\frac{\kappa_{ij}}{n} g_n^{\triangle}(i) + \frac{\kappa_{jj}}{n} g_n^{\triangle}(j)\right) \lesssim n^{2\delta_0},$$

which combined with (176) and (167) completes the proof. ∎

Proof of Lemma 6.12 (b)**:** Let u and v be uniformly distributed on $[n]$ such that u, v, and $\mathscr{G}_n^{\bullet}$ are independent. Note that for any $k \geq 1$,

$$s_{k+1}^{\bullet} = \mathbb{E}\left(|\mathscr{C}^{\bullet}(v)|^k \,\big|\, \mathscr{G}_n^{\bullet}\right) \quad \text{and} \quad \mathbb{E}\left[\left(s_{k+1}^{\bullet}\right)^2\right] = \mathbb{E}\left(|\mathscr{C}^{\bullet}(v)|^k \cdot |\mathscr{C}^{\bullet}(u)|^k\right). \tag{177}$$

For any nonempty $A \subseteq [n]$, write $\mathscr{G}^{A-}$ for the graph induced by $\mathscr{G}_n^{\bullet}$ on the vertex set $[n] \setminus A$. Then

$$\mathbb{E}\left(|\mathscr{C}^{\bullet}(u)|^k \,\big|\, V\left(\mathscr{C}^{\bullet}(v)\right) = A\right) = \frac{|\mathscr{C}^{\bullet}(v)|^{k+1}}{n} + \frac{1}{n}\mathbb{E}\left(\sum_{\substack{\mathscr{C} \text{ compo-}\\ \text{nent of } \mathscr{G}^{A-}}} |\mathscr{C}|^{k+1}\right) \leq \frac{1}{n}|\mathscr{C}^{\bullet}(v)|^{k+1} + \mathbb{E}(s_{k+1}^{\bullet}).$$

Multiplying both sides by $|\mathscr{C}^{\bullet}(v)|^k$, taking expectation, and using (177), we get $\mathbb{E}\left[\left(s_{k+1}^{\bullet}\right)^2\right] \leq n^{-1}\mathbb{E}\left(|\mathscr{C}^{\bullet}(v)|^{2k+1}\right) + \left(\mathbb{E}s_{k+1}^{\bullet}\right)^2$. This leads to the following useful bound:

$$\text{Var}\left(s_{k+1}^{\bullet}\right) \leq n^{-1}\mathbb{E}\left(|\mathscr{C}^{\bullet}(v)|^{2k+1}\right) \leq n^{-1}\int_0^1 g_{n;2k+1} \leq n^{-1} \cdot \|g_{n;2k+1}\|_{3/2}, \tag{178}$$

where the second step follows from (166). Now an application of Lemma 6.8 (c) shows that $\text{Var}\left(s_2^{\bullet}\right) \leq n^{-1} \cdot \|g_{n;3}\|_{3/2} \lesssim n^{5\delta_0-1}$, which completes the proof. ∎

6.7. **Tail bounds for $|\mathscr{C}_1^{\bullet}|$ and $D_{\max}^{\bullet}$.** Throughout this section, we work under Conditions 4.1, 4.2, and 6.2. We will further assume that (131) and (100) hold. The following lemma is the main result of this section:

Lemma 6.14. *For each $r \geq 1$, there exist $C_1, C_2 > 0$ such that for all n,*

$$\mathbb{P}\left(D_{\max}^{\bullet} \geq C_1 n^{\delta_0} \log n\right) \leq C_2 n^{-r} \quad \text{and} \tag{179}$$

$$\mathbb{P}\left(|\mathscr{C}_1^{\bullet}| \geq C_1 n^{2\delta_0+\theta_0} \log n\right) \leq C_2 n^{-r}. \tag{180}$$

Proof: It follows from (166) that for any $\ell \geq 1$,

$$\mathbb{P}\left(D_{\max}^{\bullet} \geq 2\ell\right) \leq \sum_{i \in [n]} \mathbb{P}\left(\mathfrak{H}_n^{\triangle}(i) \geq \ell\right) = n\mathbb{P}\left(\mathfrak{H}_n^{\triangle}(v) \geq \ell\right), \tag{181}$$

where v is uniformly distributed on $[n]$. Now,

$$\mathbb{P}\left(\mathfrak{H}_n^{\triangle}(v) \geq \ell\right) = \mathbb{P}\left(|\mathfrak{G}_{\ell}^{\triangle}(v)| \geq 1\right) \leq \mathbb{E}\left(|\mathfrak{G}_{\ell}^{\triangle}(v)|\right) = \int_0^1 T_n^{\ell} \mathbf{1}_{\circ} \leq \|T_n^{\ell} \mathbf{1}_{\circ}\|_2 \leq \|T_n\|^{\ell}, \tag{182}$$

where the third step uses (132). By (130) and (131),

$$\|T_n\| \leq 1 - C_{183} n^{-\delta_0} \quad \text{for all} \quad n, \tag{183}$$

which combined (182) and (181) yields the claim in (179).

We now turn to the proof of (180). Note that for any $\ell \geq 2$ and any n,

$$\|T_n^\ell \mathbf{1}_\circ\|_\infty \leq C n^{\theta_0} \|T_n^{\ell-1}\mathbf{1}_\circ\|_3 \leq C' n^{\theta_0}\|T_n^{\ell-2}\mathbf{1}_\circ\|_2 \leq C' n^{\theta_0}\|T_n\|^{\ell-2}, \tag{184}$$

where the first step uses Lemma 6.10 (b), and the second step uses (99). Define

$$\Phi_{n,k} := \mathbf{1}_\circ + \sum_{\ell=1}^{k}\left(1+\mathfrak{b}_n\right)^\ell T_n^\ell \mathbf{1}_\circ \;\; \text{and} \;\; \Phi_n := \mathbf{1}_\circ + \sum_{\ell=1}^{\infty}\left(1+\mathfrak{b}_n\right)^\ell T_n^\ell \mathbf{1}_\circ, \;\; \text{where} \;\; \mathfrak{b}_n = \frac{C_{183}}{2n^{\delta_0}}.$$

It follows from (183) and (184) that

$$\|\Phi_n\|_\infty \leq 1 + \left(1+\frac{C_{183}}{2n^{\delta_0}}\right)\|T_n\mathbf{1}_\circ\|_\infty + C'n^{\theta_0}\sum_{\ell=2}^{\infty}\left(1+\frac{C_{183}}{2n^{\delta_0}}\right)^\ell\left(1-\frac{C_{183}}{n^{\delta_0}}\right)^{\ell-2} \leq C_{185}n^{\theta_0+\delta_0}. \tag{185}$$

Let

$$\varepsilon_n := C_{186}n^{-\theta_0-2\delta_0}, \;\; \text{where} \;\; C_{186} = \frac{C_{183}}{2eC_{185}}. \tag{186}$$

Using (185), we can choose $n_0 \geq 1$ be such that $\varepsilon_n\|\Phi_n\|_\infty \leq 1$ for all $n \geq n_0$. Now we claim that for any $n \geq n_0, r \in [n]$, and $k \geq 1$,

$$\mathbb{E}\exp\left(\varepsilon_n\sum_{\ell=0}^{k}|\mathfrak{G}_\ell^\triangle(r)|\right) \leq \exp\left(\varepsilon_n\Phi_{n,k}\left(\tfrac{r}{n}-\tfrac{1}{2n}\right)\right). \tag{187}$$

From this claim it follows, by letting $k \to \infty$, that for any $n \geq n_0$ and $r \in [n]$,

$$\mathbb{E}\exp\left(\varepsilon_n\mathfrak{X}_n^\triangle(r)\right) \leq \exp\left(\varepsilon_n\Phi_n\left(\tfrac{r}{n}-\tfrac{1}{2n}\right)\right).$$

This in turn implies that for any $C > 0, n \geq n_0$, and $r \in [n]$,

$$\mathbb{P}\left(\mathfrak{X}_n^\triangle(r) \geq C(\log n)/\varepsilon_n\right) \leq \exp(-C\log n)\cdot\exp\left(\varepsilon_n\|\Phi_n\|_\infty\right) \leq \exp\left(-C\log n + 1\right), \tag{188}$$

which combined with (166) and a union bound yields the claim in (180).

It remains to prove (187). Fix $k \geq 1, n \geq n_0$, and $r \in [n]$. Let us make note of the following simple inequality: $e^x - 1 \leq x(1+xe)$ for all $x \in [0,1]$. Hence, for $x \in [0,1]$ and $m \in \mathbb{Z}_{\geq 0}$,

$$\exp\left(\varepsilon_n\Phi_{n,m}(x)\right) - 1 \leq \varepsilon_n\Phi_{n,m}(x)\cdot\left(1+\varepsilon_n\Phi_{n,m}(x)e\right)$$
$$\leq \varepsilon_n\Phi_{n,m}(x)\cdot\left(1+\varepsilon_n\|\Phi_n\|_\infty e\right) \leq \varepsilon_n\Phi_{n,m}(x)\cdot(1+\mathfrak{b}_n), \tag{189}$$

where the last step uses (185) and (186). Consider a multitype branching process starting from a vertex of type $r \in [n]$ as described above (102), and for any vertex w in this branching process tree, let $i_w \in [n]$ denote its type, and let $x_w := n^{-1}(i_w - 1/2)$; we will similarly write $x_j := n^{-1}(j - 1/2)$ for $j \in [n]$. We will prove inductively that for $t = k-1, k-2, \ldots, 1, 0$, on the event $\{\mathfrak{G}_t^\triangle(r) \neq \varnothing\}$,

$$\mathbb{E}\left(\exp\left(\varepsilon_n\sum_{m=t}^{k}|\mathfrak{G}_m^\triangle(r)|\right)\Big|\mathfrak{G}_t^\triangle(r)\right) \leq \prod_{w\in\mathfrak{G}_t^\triangle(r)}\exp\left(\varepsilon_n\Phi_{n,k-t}(x_w)\right); \tag{190}$$

the inequality in (187) corresponds to $t = 0$. Starting with $t = k-1$, we see that on the event $\{\mathfrak{G}^{\triangle}_{k-1}(r) \neq \emptyset\}$,

$$\mathbb{E}\Big(\exp\big(\varepsilon_n|\mathfrak{G}^{\triangle}_k(r)|\big)\,\Big|\,\mathfrak{G}^{\triangle}_{k-1}(r)\Big) = \prod_{w \in \mathfrak{G}^{\triangle}_{k-1}(r)} \prod_{j \in [n]} \Big(1 - \frac{\kappa_{i_w j}}{n} + \frac{\kappa_{i_w j}}{n} e^{\varepsilon_n}\Big)$$

$$\leq \prod_{w \in \mathfrak{G}^{\triangle}_{k-1}(r)} \prod_{j \in [n]} \exp\Big(\kappa_{i_w j}\big(e^{\varepsilon_n} - 1\big)/n\Big) \leq \prod_{w \in \mathfrak{G}^{\triangle}_{k-1}(r)} \prod_{j \in [n]} \exp\Big(\varepsilon_n(1 + \mathfrak{b}_n)\kappa_{i_w j}/n\Big)$$

$$= \prod_{w \in \mathfrak{G}^{\triangle}_{k-1}(r)} \exp\Big(\varepsilon_n(1 + \mathfrak{b}_n)\,T_n \mathbf{1}_{\circ}(x_w)\Big),$$

where the third step uses (189) with $m = 0$. This implies (190) for $t = k - 1$. Suppose (190) holds for $t = k-1, k-2, \ldots, \ell+1$ for some $\ell \geq 0$. Then on the event $\{\mathfrak{G}^{\triangle}_\ell(r) \neq \emptyset\}$,

$$\mathbb{E}\Big(\exp\Big(\varepsilon_n \sum_{m=\ell+1}^{k} |\mathfrak{G}^{\triangle}_m(r)|\Big)\,\Big|\,\mathfrak{G}^{\triangle}_\ell(r)\Big) \leq \mathbb{E}\Big(\prod_{\widetilde{w} \in \mathfrak{G}^{\triangle}_{\ell+1}(r)} \exp\big(\varepsilon_n \Phi_{n,k-\ell-1}(x_{\widetilde{w}})\big)\,\Big|\,\mathfrak{G}^{\triangle}_\ell(r)\Big)$$

$$= \prod_{w \in \mathfrak{G}^{\triangle}_\ell(r)} \prod_{j \in [n]} \Big(1 - \frac{\kappa_{i_w j}}{n} + \frac{\kappa_{i_w j}}{n} \exp\big(\varepsilon_n \Phi_{n,k-\ell-1}(x_j)\big)\Big)$$

$$\leq \prod_{w \in \mathfrak{G}^{\triangle}_\ell(r)} \prod_{j \in [n]} \exp\Big(\frac{\kappa_{i_w j}}{n}\Big[\exp\big(\varepsilon_n \Phi_{n,k-\ell-1}(x_j)\big) - 1\Big]\Big)$$

$$\leq \prod_{w \in \mathfrak{G}^{\triangle}_\ell(r)} \prod_{j \in [n]} \exp\Big(\frac{\kappa_{i_w j}}{n}\Big[\varepsilon_n \Phi_{n,k-\ell-1}(x_j) \cdot (1 + \mathfrak{b}_n)\Big]\Big)$$

$$= \prod_{w \in \mathfrak{G}^{\triangle}_\ell(r)} \exp\Big(\varepsilon_n(1 + \mathfrak{b}_n)\,T_n \Phi_{n,k-\ell-1}(x_w)\Big) = \prod_{w \in \mathfrak{G}^{\triangle}_\ell(r)} \exp\Big(\varepsilon_n\big(\Phi_{n,k-\ell}(x_w) - 1\big)\Big),$$

where the fourth step uses (189). This completes the proof. ∎

6.8. The third susceptibility $s_3^{\bullet}$.

Throughout this section, we work under Conditions 4.1, 4.2, and 6.2. We will further assume that (131) and (100) hold. Our goal in this section is to prove the following lemma:

Lemma 6.15. *We have, $n^{-3\delta_0} \cdot s_3^{\bullet} \xrightarrow{\text{d}} \int_0^1 \psi^3 \cdot \big(\int_0^1 \psi\big)^{-3}$ as $n \to \infty$.*

Note that

$$\text{Var}(s_3^{\bullet}) \leq n^{-1} \|g_{n;5}\|_{3/2} \lesssim n^{9\delta_0 - 1} = o(n^{6\delta_0}),$$

where the first step follows from (178), the second step uses Lemma 6.8 (c), and the last step uses the relation $\delta_0 < 1/3$. Thus, the claim in Lemma 6.15 follows from the next result and (137).

Lemma 6.16. *We have, $0 \leq \int_0^1 g_{n;2} - \mathbb{E}[s_3^{\bullet}] = o(n^{3\delta_0})$.*

Proof: Let v be uniformly distributed over $[n]$. Then $\mathbb{E}[s_3^{\bullet}] = \mathbb{E}\big(|\mathscr{C}^{\bullet}(v)|^2\big)$. Suppose $\mathscr{C}^{\bullet}(v)$ is coupled with a multitype branching process started from a vertex of type v as described above (166). In this coupling, $\mathfrak{X}^{\triangle}_n(v) \geq |\mathscr{C}^{\bullet}(v)|$ and consequently,

$$\int_0^1 g_{n;2} - \mathbb{E}[s_3^{\bullet}] = \mathbb{E}\big(\mathfrak{X}^{\triangle}_n(v)^2\big) - \mathbb{E}\big(|\mathscr{C}^{\bullet}(v)|^2\big) = \mathbb{E}\big(\mathfrak{X}^{\triangle}_n(v)^2 - |\mathscr{C}^{\bullet}(v)|^2\big) \geq 0.$$

Further, for any $m, n \geq 1$,

$$
\begin{aligned}
\mathbb{E}\big(\mathfrak{X}_n^{\triangle}(v)^2 - |\mathscr{C}^{\bullet}(v)|^2\big) &= \mathbb{E}\big[\big(\mathfrak{X}_n^{\triangle}(v) + |\mathscr{C}^{\bullet}(v)|\big) \cdot \big(\mathfrak{X}_n^{\triangle}(v) - |\mathscr{C}^{\bullet}(v)|\big)\big] \\
&\leq 2 \cdot \mathbb{E}\big[\mathfrak{X}_n^{\triangle}(v) \cdot \big(\mathfrak{X}_n^{\triangle}(v) - |\mathscr{C}^{\bullet}(v)|\big)\big] \\
&\leq 2m \cdot \mathbb{E}\big(\mathfrak{X}_n^{\triangle}(v) - |\mathscr{C}^{\bullet}(v)|\big) + 2 \cdot \mathbb{E}\big(\mathfrak{X}_n^{\triangle}(v)^2 \mathbb{1}_{\{\mathfrak{X}_n^{\triangle}(v) \geq m\}}\big) \\
&\leq 2m \cdot \Big(\int_0^1 g_n - \mathbb{E}[s_2^{\bullet}]\Big) + 2\Big(\int_0^1 g_{n;4}\Big)^{1/2} \mathbb{P}\big(\mathfrak{X}_n^{\triangle}(v) \geq m\big)^{1/2} \\
&\leq C\Big(mn^{4\delta_0 - 1} + n^{7\delta_0/2} \cdot \mathbb{P}\big(\mathfrak{X}_n^{\triangle}(v) \geq m\big)^{1/2}\Big),
\end{aligned}
$$

where the last step uses Lemma 6.12 (a) and Lemma 6.8 (c). Using (188), we conclude that

$$
0 \leq \int_0^1 g_{n;2} - \mathbb{E}[s_3^{\bullet}] \lesssim n^{\theta_0 + 6\delta_0 - 1} \cdot \log n = o(n^{3\delta_0}),
$$

since $3\delta_0 < 1 - \theta_0$. This completes the proof. ∎

6.9. **Asymptotics for $\mathscr{D}^{\bullet}$.** Throughout this section, we work under Conditions 4.1, 4.2, and 6.2. Further, we assume that (131) and (100) hold. The following lemma is the main result of this section.

Lemma 6.17. *We have,* $n^{-2\delta_0} \cdot \mathscr{D}^{\bullet} \xrightarrow{\mathrm{d}} \big(\int_0^1 \psi\big)^{-2}$ *as* $n \to \infty$.

This result follows immediately from the next lemma and the fact that $\delta_0 < 1/3$.

Lemma 6.18. *The following hold:*

(a) $n^{-2\delta_0} \cdot \mathbb{E}\big[\int_0^1 \mathfrak{Z}_n\big] \to \big(\int_0^1 \psi\big)^{-2}$ *as* $n \to \infty$;

(b) $0 \leq \mathbb{E}\big[\int_0^1 \mathfrak{Z}_n\big] - \mathbb{E}\big(\mathscr{D}^{\bullet}\big) \lesssim n^{5\delta_0 - 1} \cdot \log n$; *and*

(c) $\mathrm{Var}\big(\mathscr{D}^{\bullet}\big) \lesssim n^{7\delta_0 - 1}(\log n)^2$.

Proof of Lemma 6.18 (a): Note that

$$
\mathbb{E}\Big[\int_0^1 \mathfrak{Z}_n\Big] = \int_0^1 \sum_{\ell=1}^{\infty} \ell \cdot \mathbb{E}\big(|\mathfrak{G}_{\ell}(x)|\big)\,dx = \frac{1}{(1 - \|T_n\|)^2}\big\langle \psi_n, \mathbf{1}_{\circ}\big\rangle^2 + \sum_{j=2}^{n} \frac{1}{(1 - \lambda_{n;j})^2}\big\langle \psi_{n;j}, \mathbf{1}_{\circ}\big\rangle^2
$$

$$
= \frac{1}{(1 - \|T_n\|)^2}\Big(\int_0^1 \psi_n\Big)^2 + O(1),
$$

where we have used (132), (131), and the fact that $\sum_{j=2}^{n}\big\langle \psi_{n;j}, \mathbf{1}_{\circ}\big\rangle^2 \leq 1$. The desired result now follows from (130). ∎

Proof of Lemma 6.18 (b): For $i \in [n]$, let

$$
\mathscr{D}^{\bullet}(i) := \sum_{j \in \mathscr{C}^{\bullet}(i)} d_{\mathscr{G}_n^{\bullet}}(i, j) = \sum_{\ell=1}^{\infty} \ell \cdot \mathscr{Y}_{\ell}^{\bullet}(i), \tag{191}
$$

where $\mathscr{Y}_{\ell}^{\bullet}(i)$ denotes the number of vertices in $\mathscr{C}^{\bullet}(i)$ at distance ℓ from i. Let v be uniformly distributed over $[n]$. Then note that $\mathbb{E}\big(\mathscr{D}^{\bullet}\big) = \mathbb{E}\big(\mathscr{D}^{\bullet}(v)\big)$ and $\mathbb{E}\big[\int_0^1 \mathfrak{Z}_n\big] = \mathbb{E}[\mathfrak{Z}_n^{\triangle}(v)]$. Assume that $\mathscr{C}^{\bullet}(v)$ is coupled with a multitype branching process started from a vertex of type v as explained above (166), and recall the construction of the BFS tree of $\mathscr{C}^{\bullet}(v)$

rooted at v. Note that for any $j \in \mathscr{C}^\bullet(v)$, the distance between v and j in the BFS tree equals $d_{\mathscr{G}_n^\bullet}(v, j)$. Consequently, in the above coupling,

$$|\mathfrak{G}_\ell^\triangle(v)| \geq \mathscr{Y}_\ell^\bullet(v) \text{ for each } \ell \geq 1 \text{ and } \mathfrak{Z}_n^\triangle(v) \geq \mathscr{D}^\bullet(v). \tag{192}$$

In particular, $\mathbb{E}\big[\int_0^1 \mathfrak{Z}_n\big] = \mathbb{E}\big[\mathfrak{Z}_n^\triangle(v)\big] \geq \mathbb{E}\big(\mathscr{D}^\bullet(v)\big) = \mathbb{E}\big(\mathscr{D}^\bullet\big)$. Further,

$$\mathfrak{Z}_n^\triangle(v) - \mathscr{D}^\bullet(v) = \sum_{\ell \geq 1} \ell\big(|\mathfrak{G}_\ell^\triangle(v)| - \mathscr{Y}_\ell^\bullet(v)\big) \leq \mathfrak{H}_n^\triangle(v)\big(\mathfrak{X}_n^\triangle(v) - |\mathscr{C}^\bullet(v)|\big).$$

Thus, for any $m, n \geq 1$,

$$\mathbb{E}\big[\mathfrak{Z}_n^\triangle(v)\big] - \mathbb{E}\big(\mathscr{D}^\bullet(v)\big) \leq m\Big(\int_0^1 g_n - \mathbb{E}[s_2^\bullet]\Big) + \mathbb{E}\Big(\mathfrak{H}_n^\triangle(v) \cdot \mathfrak{X}_n^\triangle(v) \cdot \mathbb{1}_{\{\mathfrak{H}_n^\triangle(v) \geq m\}}\Big)$$

$$\leq Cmn^{4\delta_0 - 1} + \Big[\mathbb{E}\Big(\mathfrak{H}_n^\triangle(v)^2 \cdot \mathbb{1}_{\{\mathfrak{H}_n^\triangle(v) \geq m\}}\Big)\Big]^{1/2} \cdot \Big(\int_0^1 g_{n;2}\Big)^{1/2},$$

where the last step uses Lemma 6.12 (a). The claimed result now follows from (182), (183), and Lemma 6.8 (b). ∎

Proof of Lemma 6.18 (c)**:** Recall the notation $\mathscr{D}^\bullet(\cdot)$ from (191). Let u and v be uniformly distributed over $[n]$ such that u, v, and $\mathscr{G}_n^\bullet$ are independent. Then note that $\mathbb{E}\big[\mathscr{D}^\bullet(u) \mid \mathscr{G}_n^\bullet\big] = \mathscr{D}^\bullet = \mathbb{E}\big[\mathscr{D}^\bullet(v) \mid \mathscr{G}_n^\bullet\big]$. Hence,

$$\mathbb{E}\big[(\mathscr{D}^\bullet)^2\big] = \mathbb{E}\big[\mathscr{D}^\bullet(u)\mathscr{D}^\bullet(v)\big] = \mathbb{E}\Big(\mathscr{D}^\bullet(v) \cdot \mathbb{E}\Big[\mathscr{D}^\bullet(u) \mid \mathscr{C}^\bullet(v)\Big]\Big). \tag{193}$$

For any nonempty $A \subseteq [n]$, let $\mathscr{G}^{A-}$ be the graph induced by $\mathscr{G}_n^\bullet$ on the vertex set $[n] \setminus A$, as introduced below (177). For $i \in [n] \setminus A$, let $\mathscr{D}^{A-}(i) := \sum_{j \in \mathscr{C}^{A-}(i)} d_{\mathscr{G}^{A-}}(i, j)$ where $\mathscr{C}^{A-}(i)$ denotes the component of $\mathscr{G}^{A-}$ containing i. Similarly, for $i \in [n] \setminus A$, let $\mathfrak{Z}^{A-}(i)$ be the analogue of $\mathfrak{Z}_n^\triangle(i)$ (defined in (102)) for a multitype branching process with type space $[n] \setminus A$ started from a vertex of type i where each vertex of type $j \in [n] \setminus A$ has Bernoulli(κ_{jk}/n) many type k children for each $k \in [n] \setminus A$. Then similar to (192), we have $\mathfrak{Z}^{A-}(i) \geq \mathscr{D}^{A-}(i)$ in the obvious coupling. Thus, on the event $\{V(\mathscr{C}^\bullet(v)) = A\}$,

$$\mathbb{E}\Big[\mathscr{D}^\bullet(u) \mid \mathscr{C}^\bullet(v)\Big] = \frac{1}{n}\sum_{i \in A}\sum_{j \in A} d_{\mathscr{C}^\bullet(v)}(i, j) + \frac{1}{n}\sum_{i \in [n] \setminus A}\mathbb{E}\big(\mathscr{D}^{A-}(i)\big)$$

$$\leq \frac{1}{n}\sum_{i \in A}\sum_{j \in A} d_{\mathscr{C}^\bullet(v)}(i, j) + \frac{1}{n}\sum_{i \in [n] \setminus A}\mathbb{E}\big(\mathfrak{Z}^{A-}(i)\big)$$

$$\leq \frac{1}{n}\sum_{i \in A}\sum_{j \in A} d_{\mathscr{C}^\bullet(v)}(i, j) + \frac{1}{n}\sum_{i \in [n]}\mathbb{E}\big(\mathfrak{Z}_n^\triangle(i)\big) =: F_1 + F_2, \tag{194}$$

where the third step uses the relation $\mathfrak{Z}^{A-}(i) \leq \mathfrak{Z}_n^\triangle(i)$ for each $i \in [n] \setminus A$ in the obvious coupling between the two branching processes. Now,

$$F_1 \leq \frac{1}{n}\sum_{i \in A}\sum_{j \in A}\big(d_{\mathscr{C}^\bullet(v)}(i, v) + d_{\mathscr{C}^\bullet(v)}(j, v)\big) = \frac{2}{n}|\mathscr{C}^\bullet(v)| \cdot \mathscr{D}^\bullet(v), \tag{195}$$

and

$$F_2 = \mathbb{E}\Big[\int_0^1 \mathfrak{Z}_n\Big] \leq \mathbb{E}\big[\mathscr{D}^\bullet(v)\big] + Cn^{5\delta_0 - 1}\log n, \tag{196}$$

where the last step uses Lemma 6.18 (b).

Combining (193), (194), (195), and (196), we see that

$$\mathbb{E}\big[(\mathscr{D}^{\bullet})^2\big] \le \frac{2}{n}\mathbb{E}\big(|\mathscr{C}^{\bullet}(v)| \cdot \mathscr{D}^{\bullet}(v)^2\big) + \big(\mathbb{E}[\mathscr{D}^{\bullet}(v)]\big)^2 + C n^{5\delta_0-1}\log n \cdot \mathbb{E}\big[\mathscr{D}^{\bullet}(v)\big]$$

$$\le \frac{2}{n}\mathbb{E}\big[\mathfrak{X}_n^{\triangle}(v) \cdot \mathfrak{Z}_n^{\triangle}(v)^2\big] + \big(\mathbb{E}[\mathscr{D}^{\bullet}]\big)^2 + C n^{7\delta_0-1}\log n, \tag{197}$$

where the last step uses Lemma 6.18 (a), (b) and the fact that $\mathbb{E}[\mathscr{D}^{\bullet}] = \mathbb{E}[\mathscr{D}^{\bullet}(v)]$. Using the relation $\mathfrak{Z}_n^{\triangle}(v) \le \mathfrak{H}_n^{\triangle}(v) \cdot \mathfrak{X}_n^{\triangle}(v)$, we see that for any $n, m \ge 1$,

$$\mathbb{E}\big[\mathfrak{X}_n^{\triangle}(v) \cdot \big(\mathfrak{Z}_n^{\triangle}(v)\big)^2\big] \le \mathbb{E}\big[\mathfrak{X}_n^{\triangle}(v)^3 \cdot \big(\mathfrak{H}_n^{\triangle}(v)\big)^2\big]$$

$$\le m^2 \mathbb{E}\big[\mathfrak{X}_n^{\triangle}(v)^3\big] + \mathbb{E}\big[\mathfrak{X}_n^{\triangle}(v)^3 \cdot \big(\mathfrak{H}_n^{\triangle}(v)\big)^2 \cdot \mathbb{1}_{\{\mathfrak{H}_n^{\triangle}(v) \ge m\}}\big]$$

$$\le m^2\big(\int_0^1 g_{n;3}\big) + \big(\int_0^1 g_{n;4}\big)^{3/4} \cdot \mathbb{E}\big(\mathfrak{H}_n^{\triangle}(v)^8 \cdot \mathbb{1}_{\{\mathfrak{H}_n^{\triangle}(v) \ge m\}}\big)^{1/4}. \tag{198}$$

Now the claimed bound on $\mathrm{Var}\big(\mathscr{D}^{\bullet}\big)$ follows upon combining (197) with (198), and using Lemma 6.8 (c), (182), and (183).   ∎

6.10. **Proof of Proposition 6.3.** Throughout this section, we will work under Conditions 4.1, 4.2, and 6.2. Further, we assume that (131) and (100) hold. Recall the connection probabilities $p_{ij}^{(n)}$ and $\widetilde{p}_{ij}^{(n)}$ from (4). Then

$$\sum_{\substack{i,j \in [n] \setminus B_n \\ i \ne j}} \frac{\big(p_{ij}^{(n)} - \widetilde{p}_{ij}^{(n)}\big)^2}{p_{ij}^{(n)}} \le \frac{C}{n^3} \sum_{i,j \in [n]} \beta_{ij}^3 = O(1/n), \tag{199}$$

where the last step uses (19). Hence, by [46, Corollary 2.12 (i)], for all n there exists a coupling of $\mathscr{G}_n^{\mathrm{gphn}}$ and $\widetilde{\mathscr{G}}_n^{\mathrm{gphn}}$ such that

$$\mathbb{P}\big(\mathscr{G}_n^{\mathrm{gphn}} \ne \widetilde{\mathscr{G}}_n^{\mathrm{gphn}}\big) \to 0 \quad \text{as} \quad n \to \infty. \tag{200}$$

Thus, it is enough to show that (26) and (27) hold for $\widetilde{\mathscr{G}}_n^{\mathrm{gphn}}$.

Let $\widetilde{\mathscr{G}}_n^{\bullet}$ be the random graph with vertex set $[n]$ where the edges $\{i, j\}$ appear independently with respective probabilities $1 - \exp\big(-\kappa_{ij}/n\big)$, $1 \le i < j \le n$. By an argument similar to the one leading to (200), we can couple $\widetilde{\mathscr{G}}_n^{\bullet}$ with $\mathscr{G}_n^{\bullet}$ for all n such that

$$\mathbb{P}\big(\widetilde{\mathscr{G}}_n^{\bullet} \ne \mathscr{G}_n^{\bullet}\big) \to 0 \quad \text{as} \quad n \to \infty. \tag{201}$$

Let $\widetilde{D}_{\max}^{\bullet} := \mathrm{diam}(\widetilde{\mathscr{G}}_n^{\bullet})$, $\widetilde{\mathscr{D}}^{\bullet} := \mathscr{D}(\widetilde{\mathscr{G}}_n^{\bullet})$, and $\widetilde{s}_k^{\bullet} := s_k(\widetilde{\mathscr{G}}_n^{\bullet})$, $k \ge 1$. Excluding the singleton vertices belonging to B_n from $\widetilde{\mathscr{G}}_n^{\bullet}$, arrange the remaining (say m_n many) components in decreasing order of their sizes with ties broken following some deterministic rule; let $\widetilde{\mathscr{C}}_i^{\bullet}$ denote the i-th largest component, $1 \le i \le m_n$. Our plan is to apply Theorem 3.4 with $\mathbf{x} = (x_1, \ldots, x_{m_n})$, q, and $\mathbf{M} = (M_1, \ldots, M_{m_n})$, where

$$x_i = n^{-2/3}\chi^{1/3}|\widetilde{\mathscr{C}}_i^{\bullet}|, \ q = \chi^{-2/3}n^{1/3-\delta_0}, \ \text{and} \ M_i = \mathrm{scl}\big(1, 1/|\widetilde{\mathscr{C}}_i^{\bullet}|\big)\widetilde{\mathscr{C}}_i^{\bullet}, \ 1 \le i \le m_n. \tag{202}$$

Writing $\sigma_k = \sum_{i=1}^{m_n} x_i^k$, we have

$$\sigma_2 = \frac{\chi^{2/3}}{n^{1/3}}\Big(\widetilde{s}_2^{\bullet} - \frac{|B_n|}{n}\Big), \ \sigma_3 = \frac{\chi}{n}\Big(\widetilde{s}_3^{\bullet} - \frac{|B_n|}{n}\Big), \ x_{\max} = \frac{\chi^{1/3}|\widetilde{\mathscr{C}}_1^{\bullet}|}{n^{2/3}}, \ \text{and} \ x_{\min} \ge \frac{\chi^{1/3}}{n^{2/3}}. \tag{203}$$

Using (201) in conjunction with Lemma 6.11, (180), and Lemma 6.15, we see that Condition 2.1 holds with $\lambda = \zeta\chi^{-2/3}$. Next, we can clearly choose $r_0 > 0$ large so that $\sigma_2^{r_0}/x_{\min} \xrightarrow{d} 0$ as $n \to \infty$. Further, (201) and (180) show that for any $\eta_0 > 0$,

$$x_{\max}/\sigma_2^{3/2+\eta_0} = O_P\Big(n^{-\frac{1}{6}+\frac{\delta_0}{2}+\theta_0+\eta_0\left(\frac{1}{3}-\delta_0\right)} \cdot \log n\Big).$$

Since $\delta_0 < 1/3 - 2\theta_0$, it follows that $x_{\max}/\sigma_2^{3/2+\eta_0} \xrightarrow{d} 0$ as $n \to \infty$ if we choose η_0 sufficiently small. This shows that Condition 3.1 holds. Turning to the verification of Condition 3.3, note that

$$\tau = \sum_{\ell=1}^{m_n} x_\ell^2 u_\ell = n^{-4/3}\chi^{2/3} \sum_{\ell=1}^{m_n} \sum_{i,j\in\widetilde{\mathscr{C}}_\ell^\bullet} d_{\widetilde{\mathscr{C}}_\ell^\bullet}(i,j) = n^{-1/3}\chi^{2/3}\widetilde{\mathscr{D}}^\bullet \sim \alpha\chi^{2/3}n^{2\delta_0-1/3}$$

and $\text{diam}_{\max} = \widetilde{D}_{\max}^\bullet = O_P(n^{\delta_0}\log n)$, where we have used (201), Lemma 6.17, and (179). Using these relations together with the fact that $x_{\max} = O_P(n^{-2/3+2\delta_0+\theta_0}\log n)$ and $\sigma_2 \sim \chi^{2/3}n^{\delta_0-1/3}$, we see that

$$\frac{\sigma_2^{\frac{3}{2}-\eta_0} \cdot \text{diam}_{\max}}{\tau + \sigma_2} = O_P\Big(n^{-\left(\frac{1}{3}-\delta_0\right)\left(\frac{1}{2}-\eta_0\right)} \cdot \log n\Big) \quad\text{and}\quad \frac{\sigma_2 x_{\max}\text{diam}_{\max}}{\tau} = O_P\Big(\frac{(\log n)^2}{n^{\frac{2}{3}-2\delta_0-\theta_0}}\Big).$$

Since $\delta_0 < 1/3 - 2\theta_0$, it follows that Condition 3.3 holds for any choice of $\eta_0 < 1/2$.

Similar to the notation introduced at the beginning of Section 2.4, we let $\mathscr{C}_i$ be the i-th maximal component (with respect to the mass given by $\text{mass}(\mathscr{C}) = \sum_{i\in\mathscr{C}} x_i$ for each component $\mathscr{C}$) in $\mathscr{G}(\mathbf{x}, q)$. Following the notation introduced around (16), let $\bar{\mathscr{C}}_i$ be the component in $\bar{\mathscr{G}}(\mathbf{x}, q, \mathbf{M})$ corresponding to the component $\mathscr{C}_i$ in $\mathscr{G}(\mathbf{x}, q)$. We will write $\bar{\mathscr{C}}_i$ to denote both the graph and the metric measure space (where the metric is given by (15) and coincides with the graph distance, and the measure is as described in (14) and coincides with $n^{-2/3}\chi^{1/3}$ times the counting measure on the vertices). Note that $\bar{\mathscr{C}}_i$ is the i-th maximal component (in terms of the number of vertices) of $\bar{\mathscr{G}}(\mathbf{x}, q, \mathbf{M})$. Since Condition 2.1 and Condition 3.3 are satisfied, the conclusions of Theorem 2.2 and Theorem 3.4 hold for $(\mathscr{C}_i; i \geq 1)$ and $(\bar{\mathscr{C}}_i; i \geq 1)$. Thus,

$$\Big(\big(n^{-2/3}\chi^{1/3}|\bar{\mathscr{C}}_i|, \text{spls}(\mathscr{C}_i)\big); i \geq 1\Big) \xrightarrow{d} \Xi\big(\zeta\chi^{-2/3}\big) \tag{204}$$

as $n \to \infty$ with respect to the product topology and

$$\Big(\text{scl}\Big(\frac{\chi^{2/3}}{\alpha n^{1/3}}, 1\Big)\bar{\mathscr{C}}_i; i \geq 1\Big) \xrightarrow{d} \mathbf{Crit}\big(\zeta\chi^{-2/3}\big) \tag{205}$$

as $n \to \infty$, where we have used the fact that $\sigma_2^2/(\sigma_2 + \tau) \sim \chi^{2/3}/(\alpha n^{1/3})$.

We will now use (204) and (205) to deduce the convergences (26) and (27) for $\widetilde{\mathscr{G}}_n^{\text{gphn}}$. For simplicity, we only work with the maximal component $\mathscr{C}_1(\widetilde{\mathscr{G}}_n^{\text{gphn}})$. Note that $\widetilde{\mathscr{G}}_n^{\text{gphn}}$ can be constructed from $\widetilde{\mathscr{G}}_n^\bullet$ as follows: Consider all pairs $i, j \in [n] \setminus B_n$ with $i \neq j$ such that there is no edge between i and j in $\widetilde{\mathscr{G}}_n^\bullet$, and then place edges between them independently with probability $1 - \exp\big(-1/n^{1+\delta_0}\big)$. On the other hand, we construct $\bar{\mathscr{G}}(\mathbf{x}, q, \mathbf{M})$ by placing one edge between distinct components $\widetilde{\mathscr{C}}_i^\bullet$ and $\widetilde{\mathscr{C}}_j^\bullet$ independently with probability $1 - \exp(-|\widetilde{\mathscr{C}}_i^\bullet| \cdot |\widetilde{\mathscr{C}}_j^\bullet|/n^{1+\delta_0})$, $1 \leq i \neq j \leq m_n$. There is a natural way of coupling $\widetilde{\mathscr{G}}_n^\bullet$, $\mathscr{G}(\mathbf{x}, q)$,

$\bar{\mathcal{G}}(\mathbf{x}, q, \mathbf{M})$, and $\widetilde{\mathcal{G}}_n^{\mathrm{gphn}}$ such that $V\big(\mathscr{C}_1(\widetilde{\mathcal{G}}_n^{\mathrm{gphn}})\big) = V(\bar{\mathscr{C}}_1)$, and $E\big(\mathscr{C}_1(\widetilde{\mathcal{G}}_n^{\mathrm{gphn}})\big) \supseteq E(\bar{\mathscr{C}}_1)$. It suffices to show that (a) $E\big(\mathscr{C}_1(\widetilde{\mathcal{G}}_n^{\mathrm{gphn}})\big) = E(\bar{\mathscr{C}}_1)$ whp, and (b) the blobs $\widetilde{\mathscr{C}}_i^{\bullet}$ that become part of $\bar{\mathscr{C}}_1$ are all trees whp. Thus, the following lemma completes the proof.

Lemma 6.19. *The following hold in the above coupling:*
(a) *With high probability, $E\big(\mathscr{C}_1(\widetilde{\mathcal{G}}_n^{\mathrm{gphn}})\big) = E(\bar{\mathscr{C}}_1)$.*
(b) *With high probability, $\widetilde{\mathscr{C}}_i^{\bullet}$ is a tree for all i such that $\widetilde{\mathscr{C}}_i^{\bullet} \subseteq \bar{\mathscr{C}}_1$.*

Proof of Lemma 6.19 (a)**:** Let $\mathscr{F}$ denote the σ-field generated by $\widetilde{\mathcal{G}}_n^{\bullet}$, $\mathcal{G}(\mathbf{x}, q)$, and $\bar{\mathcal{G}}(\mathbf{x}, q, \mathbf{M})$. Write $\sum_{\dagger}$ for the sum over all $k \in [m_n]$ such that $k \in V(\mathscr{C}_1)$, and $\sum_{\dagger\dagger}$ for the sum over all $1 \le i < j \le m_n$ such that $\{i, j\} \in E(\mathscr{C}_1)$. Then conditional on $\mathscr{F}$, $\#E\big(\mathscr{C}_1(\widetilde{\mathcal{G}}_n^{\mathrm{gphn}})\big) - \#E(\bar{\mathscr{C}}_1)$ is stochastically bounded by

$$\sum\nolimits_{\dagger\dagger} (N_{ij} - 1) + \sum\nolimits_{\dagger} N_{kk}, \tag{206}$$

where (i) for each $i < j$ with $\{i, j\} \in E(\mathscr{C}_1)$, N_{ij} is distributed as a Binomial$\big(|\widetilde{\mathscr{C}}_i^{\bullet}| \cdot |\widetilde{\mathscr{C}}_j^{\bullet}|, 1 - \exp(-1/n^{1+\delta_0})\big)$ random variable conditioned to be at least one, (ii) for each $k \in V(\mathscr{C}_1)$, N_{kk} is distributed as a Binomial$\big(|\widetilde{\mathscr{C}}_k^{\bullet}|^2, 1 - \exp(-1/n^{1+\delta_0})\big)$ random variable, and (iii) all the random variables from (i) and (ii) are jointly independent. Then conditional on $\mathscr{F}$, on the event $\{|\widetilde{\mathscr{C}}_1^{\bullet}| \le n^{2\delta_0 + \theta_0}(\log n)^2\}$, the expectation of the expression in (206) is bounded by

$$Cn^{-1-\delta_0}\Big(\sum\nolimits_{\dagger\dagger} |\widetilde{\mathscr{C}}_i^{\bullet}| \cdot |\widetilde{\mathscr{C}}_j^{\bullet}| + \sum\nolimits_{\dagger} |\widetilde{\mathscr{C}}_k^{\bullet}|^2\Big) \le Cn^{-1-\delta_0} \cdot |\widetilde{\mathscr{C}}_1^{\bullet}| \cdot \Big(|\bar{\mathscr{C}}_1| + \mathrm{spls}(\mathscr{C}_1)|\widetilde{\mathscr{C}}_1^{\bullet}| + |\bar{\mathscr{C}}_1|\Big).$$

It follows from (204) that $|\bar{\mathscr{C}}_1| = O_P(n^{2/3})$ and $\mathrm{spls}(\mathscr{C}_1) = O_P(1)$, and consequently, using the fact that $\delta_0 < 1/3 - \theta_0$, we see the right side of the last display is $o_P(1)$. Since $\{|\widetilde{\mathscr{C}}_1^{\bullet}| \le n^{2\delta_0 + \theta_0}(\log n)^2\}$ whp by (201) and (180), the proof is complete. ∎

Proof of Lemma 6.19 (b)**:** Recall the notation $\leftrightarrow$, $\xleftrightarrow{\ell}$, and $\leftrightsquigarrow$ introduced below (97). Note that if $\mathrm{spls}\big(\mathscr{C}^{\bullet}(i)\big) \ge 1$ for some $i \in [n]$, then either there exist $j, k \in [n]$ with i, j, k pairwise distinct such that the event $\{i \leftrightarrow j\} \circ \{j \xleftrightarrow{\ell} k\} \circ \{j \leftrightsquigarrow k\}$ occurs for some $\ell \ge 2$, or there exists $k \ne i$ such that the event $\{i \xleftrightarrow{\ell} k\} \circ \{i \leftrightsquigarrow k\}$ occurs for some $\ell \ge 2$. Let v be uniformly distributed over $[n]$ independent of $\mathcal{G}_n^{\bullet}$. Using the BK inequality and Lemma 6.13, and writing 'p.d.' for 'pairwise distinct,' we get

$$\mathbb{P}\big(\mathrm{spls}\big(\mathscr{C}^{\bullet}(v)\big) \ge 1\big) \le F_1 + F_2, \tag{207}$$

where

$$F_1 = n^{-1} \sum_{\ell \ge 2} \sum_{i,j,k \text{ (p.d.)}} \mathbb{P}\big(i \leftrightarrow j\big) \mathbb{P}\big(j \xleftrightarrow{\ell} k\big) \mathbb{P}\big(j \leftrightsquigarrow k\big)$$

$$\le n^{-1} \int_{[0,1]^3} \rho_n(x, y)\rho_n(y, z)\kappa_n(y, z)\, dx\, dy\, dz, \tag{208}$$

and similarly

$$F_2 \le n^{-1} \int_{[0,1]^2} \rho_n(x, z)\kappa_n(x, z)\, dx\, dz.$$

It follows from (208) that

$$F_1 \le n^{-1}\Big(\int_{[0,1]^2} \rho_n^3\Big)^{1/3} \cdot \Big(\int_{[0,1]^2} \rho_n^3\Big)^{1/3} \cdot \Big(\int_{[0,1]^2} \kappa_n^3\Big)^{1/3} \lesssim n^{2\delta_0 - 1}, \tag{209}$$

where we have used (170) and the fact $\int_{[0,1]^2} \kappa_n^3 = O(1)$. Similarly $F_2 \lesssim n^{\delta_0-1}$, which combined with (209) and (207) yields

$$\mathbb{E}\Big(\frac{1}{n}\sum_i |\mathscr{C}_i^\bullet| \cdot a_i^\bullet\Big) = \mathbb{P}\big(\mathrm{spls}\big(\mathscr{C}^\bullet(v)\big) \geq 1\big) \lesssim n^{2\delta_0-1}, \tag{210}$$

where $\mathscr{C}_i^\bullet$ is i-th largest component of $\mathscr{G}_n^\bullet$ (as defined below (97)) and $a_i^\bullet := \mathbb{1}_{\{\mathscr{C}_i^\bullet \text{ is not a tree}\}}$. For $1 \leq i \leq m_n$, define $\widetilde{a}_i^\bullet := \mathbb{1}_{\{\widetilde{\mathscr{C}}_i^\bullet \text{ is not a tree}\}}$. Then using (201), (210), and the fact that $|B_n| = o(n)$, we get

$$\frac{\sum_{i=1}^{m_n} x_i \widetilde{a}_i^\bullet}{\sum_{i=1}^{m_n} x_i} = \frac{\sum_{i=1}^{m_n} |\widetilde{\mathscr{C}}_i^\bullet| \cdot \widetilde{a}_i^\bullet}{n - |B_n|} = o_P\big(n^{2\delta_0-1}\log n\big). \tag{211}$$

Consider the breadth-first exploration of $\mathscr{G}(\mathbf{x}, q)$ as described in Section 5.1, and let $v(1),\dots,v(m_n)$ denote the order in which the blobs appear in this exploration. Now, with $\mathbf{x}$ as in (202), we have $\sigma_1/\sigma_2 = o_P(n^{2/3-\delta_0}\log n)$. Thus, by the observation made below (85), whp the exploration of the maximal component $\mathscr{C}_1$ is completed in the first $n^{2/3-\delta_0}\log n$ many steps, i.e.,

$$\mathbb{P}\Big(\{i \in [m_n] : i \in V(\mathscr{C}_1)\} \subseteq \{v(i) : 1 \leq i \leq n^{2/3-\delta_0}\log n\}\Big) \to 1 \tag{212}$$

as $n \to \infty$. Since $v(1),\dots,v(m_n)$ is a size-biased permutation of the blobs using the weight sequence $\mathbf{x}$, we have, on the event

$$\Big\{x_{\max} \leq n^{2\delta_0+\theta_0-2/3}(\log n)^2 \text{ and } \big(\sum_{i=1}^{m_n} x_i \widetilde{a}_i^\bullet\big)/\big(\sum_{i=1}^{m_n} x_i\big) \leq n^{2\delta_0-1}\log n\Big\}, \tag{213}$$

$$\mathbb{P}\big(\widetilde{a}_{v(i)}^\bullet = 1 \text{ for some } i \leq n^{2/3-\delta_0}\log n \,\big|\, \widetilde{\mathscr{G}}_n^\bullet\big) \leq \frac{n^{2/3-\delta_0}\log n \cdot \sum_{i=1}^{m_n} x_i \widetilde{a}_i^\bullet}{\sum_{i=1}^{m_n} x_i - x_{\max}\cdot n^{2/3-\delta_0}\log n}$$

$$\leq 2n^{2/3-\delta_0}\log n \cdot \frac{\sum_{i=1}^{m_n} x_i \widetilde{a}_i^\bullet}{\sum_{i=1}^{m_n} x_i} \leq 2n^{\delta_0-1/3}(\log n)^2 = o(1), \tag{214}$$

where the second step uses the relation $\delta_0 < 1/3 - \theta_0$ and holds for all large n. By (201) and (180), $x_{\max} = o_P\big(n^{2\delta_0+\theta_0-2/3}(\log n)^2\big)$ which combined with (211) shows that the event in (213) occurs whp. Thus, combining (212) and (214) completes the proof. ∎

6.11. **Proofs of Theorem 4.5 and Theorem 4.6.** We will first prove the following result:

Lemma 6.20. *Suppose Conditions* 4.1 *and* 4.2 *hold, and in addition, with θ_0 as in Condition* 4.2*, there exist $\delta_0 \in (0, 1/3 - 2\theta_0)$ and $\varpi_0 \in [0, 2\delta_0 + 1/3)$ such that the following hold: For each $n \geq 1$, there exists $B_n \subseteq [n]$ satisfying* (22)*, and further, $\beta_{ij} > n^{-\delta_0}$ whenever $i, j \in [n] \setminus B_n$ with $i \neq j$. Then the convergences in* (26) *and* (27) *hold for $\mathscr{G}_n^{gphn}$ with the same constants as in* (25).

Proof: Let $\beta_{ii}^\pm := 0$ for all $i \in [n]$, and for $1 \leq i \neq j \leq n$, define

$$\beta_{ij}^- := \begin{cases} \beta_{ij}, & \text{if } i, j \in [n] \setminus B_n, \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad \beta_{ij}^+ := \begin{cases} \beta_{ij}, & \text{if } i, j \in [n] \setminus B_n, \\ \beta_{ij} + n^{-\delta_0}, & \text{otherwise.} \end{cases}$$

Let $\boldsymbol{\beta}_n^{\pm} := \left(\beta_{ij}^{\pm}; i, j \in [n]\right)$. Clearly, $\boldsymbol{\beta}_n^{+}$ and $\boldsymbol{\beta}_n^{-}$ satisfy Conditions 4.2 and 6.2. Define the functions W_n^{-} and W_n^{+} on $[0,1]^2$ similar to (18) using the edge weights $\boldsymbol{\beta}_n^{-}$ and $\boldsymbol{\beta}_n^{+}$ respectively. Then (19) is satisfied by both W_n^{-} and W_n^{+}. Next, note that

$$\left|\left\|T_{H-n^{1/3}(W_n-W)}\right\|_{2,2} - \left\|T_{H-n^{1/3}(W_n^{-}-W)}\right\|_{2,2}\right|^2 \le n^{\frac{2}{3}} \int_{[0,1]^2} \left(W_n - W_n^{-}\right)^2 \le \frac{2}{n^{4/3}} \sum_{i \in B_n} \sum_{j=1}^{n} \beta_{ij}^2 = o(1),$$

where the last step uses the second relation in (22). Thus, (20) is satisfied by the sequence $\left(W_n^{-}; n \ge 1\right)$. We can similarly show that (20) is satisfied by the sequence $\left(W_n^{+}; n \ge 1\right)$ by using the first relation in (22). Hence, both the sequences $\left(\boldsymbol{\beta}_n^{+}\right)$ and $\left(\boldsymbol{\beta}_n^{-}\right)$ satisfy Conditions 4.1, 4.2, and 6.2, and consequently, Proposition 6.3 is applicable in both cases.

Let $\mathscr{G}_n^{\pm}$ denote the random graph on $[n]$ where the edges $\{i, j\}$ appear independently with respective probabilities $1 \wedge (\beta_{ij}^{\pm}/n)$, $1 \le i < j \le n$. Since $\beta_{ij}^{-} \le \beta_{ij} \le \beta_{ij}^{+}$ for all $i, j \in [n]$, there is a natural coupling in which

$$\mathscr{G}_n^{-} \subseteq \mathscr{G}_n^{\mathrm{gphn}} \subseteq \mathscr{G}_n^{+}. \tag{215}$$

Assume that $\mathscr{G}_n^{-}, \mathscr{G}_n^{\mathrm{gphn}}$, and $\mathscr{G}_n^{+}$ are coupled in this way for all n. For $i \ge 1$, let $\mathscr{C}_i(\mathscr{G}_n^{\pm})$ denote the i-th largest component of $\mathscr{G}_n^{\pm}$, and view it as a metric measure space using the counting measure on the vertices. In the coupling of (215), we have

$$|\mathscr{C}_1(\mathscr{G}_n^{-})| \le |\mathscr{C}_1(\mathscr{G}_n^{\mathrm{gphn}})| \le |\mathscr{C}_1(\mathscr{G}_n^{+})| \quad \text{for all} \quad n. \tag{216}$$

Recall the random variables $\gamma_i(\cdot)$ and $\mathrm{Pois}_i(\cdot)$ defined below (11), and write $\lambda = \chi^{-2/3}\zeta$. It follows from Proposition 6.3 that $n^{-2/3}\chi^{1/3}|\mathscr{C}_1(\mathscr{G}_n^{\pm})| \xrightarrow{\mathrm{d}} \gamma_1(\lambda)$, which combined with (216) yields $n^{-2/3}\left(|\mathscr{C}_1(\mathscr{G}_n^{+})| - |\mathscr{C}_1(\mathscr{G}_n^{-})|\right) \xrightarrow{\mathrm{d}} 0$. Since, $\gamma_1(\lambda) > \gamma_2(\lambda)$ a.s., we get $\mathscr{C}_1(\mathscr{G}_n^{-}) \subseteq \mathscr{C}_1(\mathscr{G}_n^{\mathrm{gphn}}) \subseteq \mathscr{C}_1(\mathscr{G}_n^{+})$ whp. This implies that $\mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{-})\right) \le \mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{\mathrm{gphn}})\right) \le \mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{+})\right)$ whp. Now, by Proposition 6.3, $\mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{\pm})\right) \xrightarrow{\mathrm{d}} \mathrm{Pois}_i(\lambda)$, and consequently, $\mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{-})\right) = \mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{\mathrm{gphn}})\right) = \mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{+})\right)$ whp. Proceeding inductively, we can show that for each $i \ge 1$,

$$\mathscr{C}_i(\mathscr{G}_n^{-}) \subseteq \mathscr{C}_i(\mathscr{G}_n^{\mathrm{gphn}}) \subseteq \mathscr{C}_i(\mathscr{G}_n^{+}) \ \text{whp}, \ n^{-2/3}\left(|\mathscr{C}_i(\mathscr{G}_n^{+})| - |\mathscr{C}_i(\mathscr{G}_n^{-})|\right) \xrightarrow{\mathrm{d}} 0, \ \text{and} \tag{217}$$

$$\mathrm{spls}\left(\mathscr{C}_i(\mathscr{G}_n^{-})\right) = \mathrm{spls}\left(\mathscr{C}_i(\mathscr{G}_n^{\mathrm{gphn}})\right) = \mathrm{spls}\left(\mathscr{C}_i(\mathscr{G}_n^{+})\right) \ \text{whp}. \tag{218}$$

This shows that (27) holds.

Turning to the proof of (26), let

$$\mathscr{E}_n := \left\{\mathscr{C}_1(\mathscr{G}_n^{-}) \subseteq \mathscr{C}_1(\mathscr{G}_n^{\mathrm{gphn}}) \subseteq \mathscr{C}_1(\mathscr{G}_n^{+}), \ \mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{-})\right) = \mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{\mathrm{gphn}})\right) = \mathrm{spls}\left(\mathscr{C}_1(\mathscr{G}_n^{+})\right)\right\}.$$

Note that by Proposition 6.3,

$$\mathrm{scl}\left(\frac{\chi^{2/3}}{\alpha n^{1/3}}, \frac{\chi^{1/3}}{n^{2/3}}\right)\mathscr{C}_1(\mathscr{G}_n^{\pm}) \xrightarrow{\mathrm{d}} \mathrm{Crit}_1(\lambda) \quad \text{as} \quad n \to \infty \tag{219}$$

with respect to the GHP topology, which combined with [6, Proposition 6.5] yields

$$n^{-1/3} \cdot \mathbb{1}_{\mathscr{E}_n} \cdot d_{\mathrm{H}}\left(\mathscr{C}_1(\mathscr{G}_n^{-}), \mathscr{C}_1(\mathscr{G}_n^{+})\right) \xrightarrow{\mathrm{d}} 0. \tag{220}$$

On the event $\mathscr{E}_n$, $\mathscr{C}_1(\mathscr{G}_n^{+})$ can be obtained from $\mathscr{C}_1(\mathscr{G}_n^{-})$ by attaching finitely many, say k_n trees $\mathfrak{T}_1, \dots, \mathfrak{T}_{k_n}$; define $k_n = 0$ on $\mathscr{E}_n^c$. Then (220) implies that $\max_{1 \le i \le k_n} \mathrm{diam}(\mathfrak{T}_i) =$

$o_P(n^{1/3})$. Now, on the event $\mathcal{E}_n$, $\mathscr{C}_1(\mathscr{G}_n^{\text{gphn}})$ is the space $\mathscr{C}_1(\mathscr{G}_n^-)$ with some additional trees attached each of which is a subtree of $\mathfrak{T}_i$ for some $i \leq k_n$. We thus have

$$\mathbb{1}_{\mathcal{E}_n} \cdot d_{\mathrm{H}}\big(\mathscr{C}_1(\mathscr{G}_n^-), \mathscr{C}_1(\mathscr{G}_n^{\text{gphn}})\big) = o_P(n^{1/3}) \ \text{ and } \ |\mathscr{C}_1(\mathscr{G}_n^{\text{gphn}})| - |\mathscr{C}_1(\mathscr{G}_n^-)| = o_P(n^{2/3}), \quad (221)$$

where the second relation uses (217). Using (221) and the fact that $\mathbb{P}(\mathcal{E}_n) \to 1$, it is easy to show that

$$d_{\mathrm{GHP}}\Big(\mathrm{scl}\big(n^{-1/3}, n^{-2/3}\big)\mathscr{C}_1(\mathscr{G}_n^-), \ \mathrm{scl}\big(n^{-1/3}, n^{-2/3}\big)\mathscr{C}_1(\mathscr{G}_n^{\text{gphn}})\Big) \xrightarrow{\mathrm{d}} 0.$$

It thus follows that (219) holds if we replace $\mathscr{C}_1(\mathscr{G}_n^\pm)$ by $\mathscr{C}_1(\mathscr{G}_n^{\text{gphn}})$. This argument can be extended in a straighforward way to establish the convergence in (26) with respect to the product topology. ∎

Completing the proofs of Theorems 4.5 and 4.6: Suppose $(\boldsymbol{\beta}_n)$ satisfies Conditions 4.1, 4.2, and 4.3. For each n, let $\beta_{ii}^\odot := 0$ for all $i \in [n]$, and for $1 \leq i \neq j \leq n$, define

$$\beta_{ij}^\odot := \begin{cases} 2n^{-\delta_0}, & \text{if } i, j \in [n] \setminus B_n \ \text{ and } \ \beta_{ij} \leq n^{-\delta_0}, \\ \beta_{ij}, & \text{otherwise.} \end{cases}$$

Define the function $W_n^\odot$ similar to (18) using the edge weights $\beta_{ij}^\odot$, $i, j \in [n]$. Then

$$\big\| T_{n^{1/3}(W_n^\odot - W_n)} \big\|_{2,2}^2 \leq n^{2/3} \int_{[0,1]^2} \big(W_n^\odot - W_n\big)^2 \leq n^{2/3}\big(2/n^{\delta_0}\big)^2 \cdot n^{-2} \cdot o(n^{1+\delta_0}) = o(1),$$

where the second step uses (23). Hence, the assumptions in Lemma 6.20 are satisfied by $\big(\beta_{ij}^\odot; i, j \in [n]\big)$, $n \geq 1$, and consequently, the convergences in (26) and (27) hold for the sequence of random graphs with these edge weights. Since

$$\sum_{i,j \in [n] \setminus B_n, \ i \neq j} \frac{\big(n^{-1}(\beta_{ij} - \beta_{ij}^\odot)\big)^2}{(n^{-1}\beta_{ij}^\odot)} \leq \frac{\big(2/n^{1+\delta_0}\big)^2}{2/n^{1+\delta_0}} \cdot o(n^{1+\delta_0}) = o(1),$$

where in the first step we have usesd(23), an application of [46, Corollary 2.12 (i)] shows that the convergences in (26) and (27) hold for $\mathscr{G}_n^{\text{gphn}}$. Finally, the same result can be transferred to $\widetilde{\mathscr{G}}_n^{\text{gphn}}$ by another application of [46, Corollary 2.12 (i)]. ∎

6.12. **Proofs of Theorem 4.9 and Theorem 4.13.** We will verify that Conditions 4.1, 4.2, and 4.3 hold in the settings of both these theorems.

Proof of Theorem 4.9: Let $U_1, \ldots, U_n$ and $V_1^{(n)}, \ldots, V_n^{(n)}$ be as defined before (41). Let W_n be as in (18) corresponding to the edge weights in (41). Similarly define $F_n(x, y) = W\big(V_i^{(n)}, V_j^{(n)}\big)$ and $G_n(x, y) = H\big(V_i^{(n)}, V_j^{(n)}\big)$ if $(x, y) \in Q_{ij}$ for some $1 \leq i \neq j \leq n$; let $F_n(x, y) = G_n(x, y) = 0$ if $(x, y) \in [0,1]^2 \setminus \big(\bigcup_{1 \leq i \neq j \leq n} Q_{ij}\big)$. Clearly, $W_n = \big(F_n + n^{-1/3}G_n\big)^+$. Further, $\max_{i \in [n]} \big|V_i^{(n)} - i/n\big| = o_P\big(n^{-1/2} \cdot \log n\big)$ whp, and consequently,

$$\mathcal{E}_n := \big\{\text{for all } 1 \leq i \neq j \leq n, \ |V_i^{(n)} - x| + |V_j^{(n)} - y| \leq n^{-1/2} \cdot \log n \text{ for all } (x, y) \in Q_{ij}\big\} \quad (222)$$

occurs whp.

Verification of Condition 4.1: By the main theorem of [42], we have

$$\int_{[0,1]^2} F_n^3 \xrightarrow{\mathrm{d}} \int_{[0,1]^2} W^3 \ \text{ as } \ n \to \infty. \quad (223)$$

Next, writing $\varepsilon_n = 1/n$, we note that

$$\int_{[0,1]^2 \setminus D_n} \left(W - F_n\right)^2 = \int_{[0,1]^2 \setminus (D_n \cup \mathscr{A}_{\varepsilon_n})} \left(W - F_n\right)^2 + \int_{\mathscr{A}_{\varepsilon_n} \setminus D_n} \left(W - F_n\right)^2 =: J_1^{(n)} + J_2^{(n)}, \qquad (224)$$

where $D_n := \bigcup_{i=1}^n Q_{ii}$ as defined below (20). Now, $J_1^{(n)} \le \int_{[0,1]^2 \setminus \mathscr{A}_{\varepsilon_n}} \mathrm{Osc}_{\varepsilon_n;W}^2(x, y)\, dx dy$ on the event $\mathscr{E}_n$. Using Condition 4.8 (a), we get

$$n^{2/3} \cdot J_1^{(n)} \xrightarrow{\ \mathrm{d}\ } 0 \ \ \text{as} \ \ n \to \infty. \qquad (225)$$

To bound $J_2^{(n)}$, note that

$$\int_{\mathscr{A}_{\varepsilon_n}} F_n^2 = \sum_{1 \le i \ne j \le n} W^2\big(V_i^{(n)}, V_j^{(n)}\big) \cdot \mu_\square\big(\mathscr{A}_{\varepsilon_n} \cap Q_{ij}\big)$$

$$= \sum_{1 \le i \ne j \le n} W^2\big(V_i^{(n)}, V_j^{(n)}\big) \cdot \mu_\square\big(\mathscr{A}_{\varepsilon_n} \cap Q_{ij}\big) \cdot \mathbb{1}_{\left\{\big(V_i^{(n)}, V_j^{(n)}\big) \in \widetilde{\mathscr{A}}_{\varepsilon_n}\right\}}$$

$$\le \frac{1}{n^2} \sum_{1 \le i \ne j \le n} W^2\big(V_i^{(n)}, V_j^{(n)}\big) \cdot \mathbb{1}_{\left\{\big(V_i^{(n)}, V_j^{(n)}\big) \in \widetilde{\mathscr{A}}_{\varepsilon_n}\right\}}$$

$$= \frac{1}{n^2} \sum_{1 \le i \ne j \le n} W^2\big(U_i, U_j\big) \cdot \mathbb{1}_{\{(U_i, U_j) \in \widetilde{\mathscr{A}}_{\varepsilon_n}\}},$$

where the second equality holds on the event $\mathscr{E}_n$. Thus, $\mathbb{E}\big(\mathbb{1}_{\mathscr{E}_n} \cdot \int_{\mathscr{A}_{\varepsilon_n}} F_n^2\big) \le \int_{\widetilde{\mathscr{A}}_{\varepsilon_n}} W^2 = o(n^{-2/3})$, where we have again used Condition 4.8 (a). This last observation combined with the fact that $J_2^{(n)} \le 2 \int_{\mathscr{A}_{\varepsilon_n}} (W^2 + F_n^2)$, and (224) and (225) yields

$$n^{2/3} \int_{[0,1]^2 \setminus D_n} \left(W - F_n\right)^2 \xrightarrow{\ \mathrm{d}\ } 0 \ \ \text{as} \ \ n \to \infty. \qquad (226)$$

If $H = \lambda W$ for some $\lambda \in \mathbb{R}$, combining (223) and (226) with the fact that (21) implies (20), we see that the functions W_n, $n \ge 1$, can be constructed on the same probability space where Condition 4.1 is satisfied almost surely.

If H satisfies the condition in Theorem 4.9 (b), then it suffices to show that

$$\int_{[0,1]^2 \setminus D_n} \left(H - G_n\right)^2 \xrightarrow{\ \mathrm{d}\ } 0 \qquad \text{and} \qquad n^{2/3} \int_{[0,1]^2} \left(\left(F_n + n^{-1/3} G_n\right) \wedge 0\right)^2 \xrightarrow{\ \mathrm{d}\ } 0. \qquad (227)$$

To prove the first convergence, fix $\eta > 0$ and let S_η be a closed subset of $[0,1]^2$ such that $\mu_\square(S_\eta^c) \le \eta$ and the restriction of H to S_η is continuous (and hence uniformly continuous). Then for all large n,

$$\int_{[0,1]^2 \setminus D_n} \mathbb{P}\big(\left|H(x, y) - G_n(x, y)\right| \ge \eta\big) dx dy$$

$$\le \mu_\square(S_\eta^c) + \mathbb{P}(\mathscr{E}_n^c) + \sum_{1 \le i \ne j \le n} \int_{Q_{ij} \cap S_\eta} \mathbb{P}\big((V_i^{(n)}, V_j^{(n)}) \notin S_\eta\big) dx dy$$

$$\le \eta + \mathbb{P}(\mathscr{E}_n^c) + \frac{1}{n^2} \sum_{1 \le i \ne j \le n} \mathbb{P}\big((U_i, U_j) \notin S_\eta\big) \le 2\eta + o(1). \qquad (228)$$

Now (228) combined with the facts $\int_{[0,1]^2} |H^3| < \infty$ and

$$\mathbb{E}\Big(\int_{[0,1]^2} |G_n^3|\Big) = n^{-1}(n-1) \int_{[0,1]^2} |H^3| = O(1) \qquad (229)$$

yields $\mathbb{E}\big[\int_{[0,1]^2\setminus D_n}(H-G_n)^2\big]\to 0$, which in particular implies the first convergence in (227).

To prove the second convergence in (227), note that

$$n^{2/3}\int_{[0,1]^2}\Big((F_n+n^{-1/3}G_n)\wedge 0\Big)^2 \le \int_{F_n\le n^{-\delta_0}}G_n^2 + \int_{|G_n|\ge n^{1/3-\delta_0}}G_n^2$$

$$\le \Big(\int_{[0,1]^2}|G_n|^3\Big)^{2/3}\cdot\mu_\square\big(\{(x,y):F_n(x,y)\le n^{-\delta_0}\}\big)^{1/3} + n^{\delta_0-1/3}\int_{[0,1]^2}|G_n|^3. \tag{230}$$

Let $B_n := \big\{i\in[n] : V_i^{(n)}\in\mathscr{B}_{2^{1/\delta_0}\cdot n^{-1}}\big\}$, where $\mathscr{B}_\bullet$ is as in Condition 4.8 (c). Then

$$\mathbb{E}\big(|B_n|\big)=o(n^{\varpi_0}) \quad\text{and}\quad \mathbb{E}\big(\#\{(i,j):i,j\notin B_n \text{ and } W(V_i^{(n)},V_j^{(n)})\le 2n^{-\delta_0}\}\big)=o(n^{1+\delta_0}), \tag{231}$$

where we have used Condition 4.8 (c). Now the second convergence in (227) follows upon combining (230), (229), and (231).

Verification of Condition 4.2: Define

$$f(x):=\int_0^1 W^{3/2}(x,y)\,dy,\ x\in[0,1],\ \text{and}\ W_n^{(\wedge)}(x,y):=W(x,y)\wedge n^{2/3},\ x,y\in[0,1].$$

Let p_1 be as in Condition 4.8 (b), and let $\theta_1=1/p_1$. It follows from Condition 4.8 (b) that for each $\varepsilon>0$,

$$n\cdot\mu_\circ(S_{n;\varepsilon})\to 0 \quad\text{as}\quad n\to\infty \quad\text{where}\quad S_{n;\varepsilon}:=\{x\in[0,1]:f(x)\ge\varepsilon n^{\theta_1}\}. \tag{232}$$

Let $r\in\mathbb{Z}_{>0}$ be such that $r\le p_1<2r$. Then for each $\varepsilon>0$,

$$\mathbb{P}\Big(\max_{x\in[0,1]}\int_0^1 F_n^{3/2}(x,y)\,dy\ge 2\varepsilon n^{\theta_1}\Big)=\mathbb{P}\Big(\sum_{j\in[n]\setminus\{i\}}\frac{1}{n}\big(W_n^{(\wedge)}(U_i,U_j)\big)^{\frac{3}{2}}\ge 2\varepsilon n^{\theta_1}\text{ for some }i\in[n]\Big)+o(1)$$

$$\le n\cdot\mathbb{P}\Big(\sum_{j=2}^n\frac{1}{n}\big(W_n^{(\wedge)}(U_1,U_j)\big)^{3/2}\ge 2\varepsilon n^{\theta_1}\Big)+o(1)$$

$$\le n\cdot\mathbb{P}(U_1\in S_{n;\varepsilon})+n\cdot\mathbb{P}\Big(U_1\notin S_{n;\varepsilon},\ \sum_{j=2}^n\frac{1}{n}\big(W_n^{(\wedge)}(U_1,U_j)\big)^{3/2}\ge 2\varepsilon n^{\theta_1}\Big)+o(1)$$

$$\le o(1)+n\cdot\int_{x\in[0,1]\setminus S_{n;\varepsilon}}\mathbb{P}\Big(\sum_{j=2}^n\frac{1}{n}\Delta_j(x)\ge\varepsilon n^{\theta_1}\Big)\,dx$$

$$\le o(1)+n^{1-2r\cdot\theta_1}\cdot\varepsilon^{-2r}\int_{x\in[0,1]\setminus S_{n;\varepsilon}}\mathbb{E}\Big[\Big(\sum_{j=2}^n\frac{1}{n}\Delta_j(x)\Big)^{2r}\Big]\,dx, \tag{233}$$

where the first step uses the fact that $\{W(U_i,U_j)\le n^{2/3}$ for all $1\le i\ne j\le n\}$ occurs whp, and the fourth step uses the notation $\Delta_j(x):=\big(W_n^{(\wedge)}(x,U_j)\big)^{3/2}-\mathbb{E}\big[\big(W_n^{(\wedge)}(x,U_j)\big)^{3/2}\big]$, $j=2,\dots,n$, together with (232) and the relation $\mathbb{E}\big[\big(W_n^{(\wedge)}(x,U_j)\big)^{3/2}\big]\le f(x)\le\varepsilon n^{\theta_1}$ when $x\in[0,1]\setminus S_{n;\varepsilon}$. We claim that

$$\mathbb{E}\Big[\Big(\sum_{j=2}^n\frac{1}{n}\Delta_j(x)\Big)^{2r}\Big]\le C(r)\cdot\big(1+f(x)^r\big) \tag{234}$$

for a constant $C(r)$ depending only on r. For simplicity, we prove (234) when $r = 5$. Writing Δ_j for $\Delta_j(x)$, we see that

$$\mathbb{E}\Big[\Big(\sum_{j=2}^n \Delta_j\Big)^{10}\Big] \le C \cdot \mathbb{E}\Big[\Big(\sum_{j=2}^n \Delta_j^2\Big)^5\Big] \le C' \cdot \mathbb{E}\big[n\Delta_2^{10} + n^2\Delta_2^8\Delta_3^2 + n^2\Delta_2^6\Delta_3^4 + n^3\Delta_2^4\Delta_3^4\Delta_4^2$$
$$+ n^4\Delta_2^4\Delta_3^2\Delta_4^2\Delta_5^2 + n^5\Delta_2^2\Delta_3^2\Delta_4^2\Delta_5^2\Delta_6^2\big], \qquad (235)$$

where the first step uses the Burkholder-Davis-Gundy inequality [25]. Note that $\mathbb{E}\big(\Delta_2^2\big) \le \mathbb{E}\big[\big(W_n^{(\wedge)}(x, U_2)\big)^3\big] \le n f(x)$, and consequently $\mathbb{E}\big[n^5\Delta_2^2\Delta_3^2\Delta_4^2\Delta_5^2\Delta_6^2\big] \le n^{10}f(x)^5$. Similarly, using the bound $|\Delta_2| \le 2n$, we get $\mathbb{E}\big[n^4\Delta_2^4\Delta_3^2\Delta_4^2\Delta_5^2\big] \le 4n^6\big(\mathbb{E}[\Delta_2^2]\big)^4 \le 4n^{10}f(x)^4$. Treating the other terms on the right side of (235) in a similar manner, we get $\mathbb{E}\big[\big(\sum_{j=2}^n \Delta_j\big)^{10}\big] \le Cn^{10}\big(1 + f(x)^5\big)$. The claim in (234) follows similarly for a general r. Now (233) combined with (234) yields

$$\max_{x \in [0,1]} \int_0^1 F_n^{3/2}(x, y)\,dy = o_P(n^{\theta_1}).$$

This completes the verification of Condition 4.2 when $H = \lambda W$.

If H satisfies the condition in Theorem 4.9 (b), then for any sequence $\omega_n \to \infty$,

$$\mathbb{P}\Big(\max_{x \in [0,1]} \int_0^1 \Big|\frac{G_n(x, y)}{n^{1/3}}\Big|^{3/2}\,dy \ge \omega_n\Big) = \mathbb{P}\Big(\exists i \in [n] \text{ such that } \sum_{j \in [n]\setminus\{i\}} |H(U_i, U_j)|^{3/2} \ge n^{3/2}\omega_n\Big)$$
$$\le n \cdot \mathbb{P}\Big(\sum_{j=2}^n |H(U_1, U_j)|^{3/2} \ge n^{3/2}\omega_n\Big) \le n \cdot \mathbb{P}\Big(\sum_{j=2}^n |H(U_1, U_j)|^3 \ge n^2\omega_n^2\Big) \le \frac{1}{\omega_n^2}\int_{[0,1]^2} |H|^3 = o(1),$$

which completes the verification of Condition 4.2.

Verification of Condition 4.3: Note that with B_n as defined below (230), we have

$$\mathbb{E}\Big[\sum_{i \in B_n} \sum_{j \in [n]\setminus\{i\}} W\big(V_i^{(n)}, V_j^{(n)}\big)^2\Big] = n(n-1)\int_{\mathscr{B}_{2^{1/\delta_0} \cdot n^{-1}}} \int_{[0,1]} W(x, y)^2\,dx\,dy = o(n^{4/3}),$$

where the last step uses the second relation in (43). This combined with (231) completes the verification of Condition 4.3 when $H = \lambda W$.

If H satisfies the condition in Theorem 4.9 (b), then

$$\mathbb{E}\Big[\sum_{i \in B_n} \sum_{j \in [n]\setminus\{i\}} \big(n^{-1/3}H(V_i^{(n)}, V_j^{(n)})\big)^2\Big] = n(n-1)n^{-2/3}\int_{[0,1]\times\mathscr{B}_{2^{1/\delta_0} \cdot n^{-1}}} H^2 = o(n^{4/3}). \quad (236)$$

Further, if β_{ij} is as in (41), then $\beta_{ij} \le n^{-\delta_0}$ implies that either $W(V_i^{(n)}, V_j^{(n)}) \le 2n^{-\delta_0}$ or $H(V_i^{(n)}, V_j^{(n)}) \le -n^{1/3-\delta_0}$. Thus, in view of (231), the following bound completes the verification of Condition 4.3:

$$\mathbb{E}\big[\#\big\{(i, j) : 1 \le i \ne j \le n, \ H(V_i^{(n)}, V_j^{(n)}) \le -n^{1/3-\delta_0}\big\}\big]$$
$$= n(n-1) \cdot \mathbb{P}\big(H^-(U_1, U_2) \ge n^{1/3-\delta_0}\big) \le \frac{n^2}{n^{(1/3-\delta_0)q}}\int_{H^- \ge n^{1/3-\delta_0}} (H^-)^q = o(n^{1+\delta_0}).$$

Finally, the claimed result follows in an identical way when W is bounded away from zero and $H \in L^3([0,1]^2)$; we do not repeat the argument. ∎

Proof of Theorem 4.13: Let W_n be as in (18) corresponding to the edge weights in (42). Similarly define $F_n(x, y) = \fint_{Q_{ij}} W$ and $G_n(x, y) = \fint_{Q_{ij}} H$ if $(x, y) \in Q_{ij}$ for some $1 \le i \ne j \le n$; let $F_n(x, y) = G_n(x, y) = 0$ if $(x, y) \in [0, 1]^2 \setminus \left(\bigcup_{1 \le i \ne j \le n} Q_{ij} \right)$. Clearly, $W_n = \left(F_n + n^{-1/3} G_n \right)^+$. We make note of the following elementary inequality: for any Borel measurable $f : [0, 1]^2 \to [0, \infty)$, $p \ge 1$, and $i, j \in [n]$,

$$\left(\fint_{Q_{ij}} f \right)^p \le \fint_{Q_{ij}} f^p. \tag{237}$$

Verification of Condition 4.1: That (19) holds follows immediately from (237). Further, it follows from (21) and (50) that (20) holds when $H = \lambda W$ for some $\lambda \in \mathbb{R}$. If H is as in Theorem 4.13 (b), to complete the verification of Condition 4.1, it suffices to prove (227) in the present setting.

To prove the first convergence in (227), note that $\lim_{n \to \infty} G_n(x, y) = H(x, y)$ for $\mu_\square$-a.e. (x, y) (see, e.g., [57, Theorem 7.10]). Further, $(H - G_n)^2 \le 2H^2 + 2(\mathcal{M} H)^2$, where $\mathcal{M}$ denotes the Hardy-Littlewood maximal function given by $\mathcal{M} H(x, y) = \sup_{Q \ni (x,y)} \left(\mu_\square(Q) \right)^{-1} \int_Q |H|$, $x, y \in [0, 1]$, with the supremum taken over all squares Q containing (x, y). Now, $\mathcal{M} H \in L^2([0, 1]^2)$ since $H \in L^2([0, 1]^2)$ (see. e.g., [34, Theorem 2.5]), and consequently the first convergence in (227) holds.

To prove the second convergence in (227), using an argument similar to the one in (230), it suffices to show that $\mu_\square\left(\{ (x, y) : F_n(x, y) \le n^{-\delta_0} \} \right) = o(1)$. To this end, note that (52) implies that we can assume, by choosing a slighly larger δ_0 if necessary, that

$$\mu_\square\left(\{ (x, y) \in \left([0, 1] \setminus \mathscr{B}_{1/n} \right)^2 : W(x, y) \le 4n^{-\delta_0} \} \right) = o(n^{\delta_0 - 1}) \quad \text{as} \quad n \to \infty. \tag{238}$$

Now, if $\fint_{Q_{ij}} W \le 2n^{-\delta_0}$, then $\mu_\square\left(\{ (x, y) \in Q_{ij} : W(x, y) \le 4n^{-\delta_0} \} \right) \ge 1/(2n^2)$. Using (238), we see that

$$\#\left\{ (i, j) : i, j \in [n] \setminus B_n, \fint_{Q_{ij}} W \le 2n^{-\delta_0} \right\} = o(n^{1+\delta_0}). \tag{239}$$

Consequently, $\mu_\square\left(\{ (x, y) : F_n(x, y) \le 2n^{-\delta_0} \} \right) \le 2|B_n|/n + o(n^{\delta_0 - 1}) = o(1)$, where we have used (51) and (239). Thus, the second convergence in (227) follows.

Verification of Condition 4.2: Write $\theta_1 = 1/p_1$, and note that

$$\max_{x \in [0,1]} \int_0^1 F_n(x, y)^{3/2} dy \le n \cdot \max_{i \in [n]} \int_{(i-1)/n}^{i/n} \left(\int_0^1 W(x, y)^{3/2} dy \right) dx$$

$$\le n \cdot \left(\max_{i \in [n]} \int_{(i-1)/n}^{i/n} \left(\int_0^1 W(x, y)^{3/2} dy \right)^{p_1} dx \right)^{\theta_1} \cdot (1/n)^{1-\theta_1} = o(n^{\theta_1}),$$

where the first step uses (237), and the last step uses Condition 4.11 (b). This shows that Condition 4.2 holds when $H = \lambda W$. If H is as in Theorem 4.13 (b), then proceeding similarly, we get

$$\max_{x \in [0,1]} \int_0^1 \left| \frac{G_n(x, y)}{n^{1/3}} \right|^{3/2} dy \le \left(\max_{i \in [n]} \int_{(i-1)/n}^{i/n} \int_0^1 |H(x, y)|^3 dy dx \right)^{1/2} = o(1),$$

which completes the verification of Condition 4.2.

Verification of Condition 4.3: If $H = \lambda W$, then combining (237) with (51) ensures that (22) holds, and further, (23) follows from (239). If H is as in Theorem 4.13 (b), then analogous to the argument below (236), it would suffice to show that

$$\#\left\{(i,j) : i, j \in [n], \int_{Q_{ij}} H \le -n^{1/3-\delta_0}\right\} = o(n^{1+\delta_0}). \tag{240}$$

Note that the left side of (240) is bounded by $m_n := \#\left\{(i,j) : i, j \in [n], f_{Q_{ij}}(H^-)^q \ge n^{1-\delta_0}\right\}$. Further, letting $S_n := \bigcup\left\{Q_{ij} : i, j \in [n], f_{Q_{ij}}(H^-)^q \ge n^{1-\delta_0}\right\}$, we have

$$m_n \le n^{1+\delta_0} \int_{S_n} (H^-)^q \le n^{1+\delta_0} \int_{[0,1]^2} (H^-)^q. \tag{241}$$

This in turn implies that $\mu_\square(S_n) = m_n/n^2 = o(1)$, which combined with the first bound in (241) shows that $m_n = o(n^{1+\delta_0})$. This completes the proof of (240), and consequently, the verification of Condition 4.3.

Finally, when W is bounded away from zero and $H \in L^3([0,1]^2)$, the desired result follows in a similar way, so we do not repeat the argument. ∎

6.13. **Proof of Theorem 4.15.** Let $t_{n,\lambda}$ be as in (5). Then the number of vertices in $\mathscr{G}_n^{\mathrm{rgiv}}(\lambda)$ is given by

$$N_n = n t_{n,\lambda} + O_P(\sqrt{n}). \tag{242}$$

Let $\mathscr{G}_{n,\otimes}^{\mathrm{rgiv}}(\lambda)$ be $\mathscr{G}_n^{\mathrm{rgiv}}(\lambda)$ where we ignore multiple edges between any two vertices.

Now, conditional on $N_n = m$, and labeling the vertices as $m, m-1, \ldots, 2, 1$ (i.e., the last vertex to appear is labeled 1), the arrival times of the vertices have the same joint distribution as $t_{n,\lambda}\left(1 - V_i^{(m)}\right)$, $1 \le i \le m$, where $V_1^{(m)} < \ldots < V_m^{(m)}$ are m i.i.d. Uniform$[0,1]$ random variables arranged in increasing order. Further, the edge set of $\mathscr{G}_{n,\otimes}^{\mathrm{rgiv}}(\lambda)$ can be generated by placing edges independently with respective probabilities $1 - \exp\left(-t_{n,\lambda}(V_i^{(m)} \wedge V_j^{(m)})/n\right)$ between the vertices i and j, $1 \le i < j \le m$.

Thus, conditional on N_n, $\mathscr{G}_{n,\otimes}^{\mathrm{rgiv}}(\lambda)$ has the same distribution as $\widetilde{\mathscr{G}}_{N_n}^{\mathrm{gphn}}$ constructed using the edge weights

$$\beta_{ij}^{(N_n)} := \frac{N_n}{n} \cdot t_{n,\lambda} \cdot \left(V_i^{(N_n)} \wedge V_j^{(N_n)}\right) = \left(t_{n,\lambda} + O_P(n^{-1/2})\right) \cdot t_{n,\lambda} \cdot \left(V_i^{(N_n)} \wedge V_j^{(N_n)}\right) \tag{243}$$

$$= \left(\frac{\pi^2}{4} + \frac{\lambda\pi}{n^{1/3}} + O_P\left(\frac{1}{\sqrt{n}}\right)\right) \cdot \left(V_i^{(N_n)} \wedge V_j^{(N_n)}\right) = \left(\frac{\pi^2}{4} + \frac{\lambda\pi}{N_n^{1/3}}\left(\frac{\pi}{2}\right)^{\frac{1}{3}} + O_P\left(\frac{1}{\sqrt{n}}\right)\right) \cdot \left(V_i^{(N_n)} \wedge V_j^{(N_n)}\right),$$

where we have used (5) and (242). Let

$$W(x,y) := \pi^2 (x \wedge y)/4 \quad \text{and} \quad H(x,y) := \lambda\pi^{4/3}(x \wedge y)/2^{1/3}, \quad x, y \in [0,1]. \tag{244}$$

It can be shown by a direct calculation that $\|T_W\|_{2,2} = 1$, and the corresponding nonnegative eigenfuction normalized to have L^2 norm one is given by

$$\psi(x) := \sqrt{2} \cdot \sin\left(\pi x/2\right), \quad x \in [0,1]. \tag{245}$$

By the discussion around (48), W given by (244) satisfies Condition 4.8. As seen in the proof of Theorem 4.9, this implies that the edge weights in (41) corresponding to the pair (W, H) in (244) satisfy Conditions 4.1, 4.2, and 4.3. Consequently, these conditions are

also satisfied by the edge weights in (243) (i.e., the $O_P(n^{-1/2})$ term appearing on the right side of (243) does not make any difference). Hence, an application of Theorem 4.5 yields the claimed result in Theorem 4.15 for $\mathscr{G}_{n,\otimes}^{\mathrm{rgiv}}(\lambda)$. Finally, it is routine to show that for each $i \geq 1$, whp there are no multiple edges in $\mathscr{C}_i\big(\mathscr{G}_n^{\mathrm{rgiv}}(\lambda)\big)$, and consequently, $\mathscr{C}_i\big(\mathscr{G}_n^{\mathrm{rgiv}}(\lambda)\big) = \mathscr{C}_i\big(\mathscr{G}_{n,\otimes}^{\mathrm{rgiv}}(\lambda)\big)$ whp. This completes the proof.

7. DISCUSSION

An important class of edge weights to which Theorems 4.5 and 4.6 do not apply directly are edge weights of the form (30), (41), and (42) when the kernel W does not satisfy Condition 4.3, Condition 4.8 (c), and Condition 4.11 (c) respectively, but is otherwise well-behaved (so that it satisfies the other required conditions). The underlying kernels here are not bounded away from zero and the mass distribution on the set where it assumes small values cannot be handled using the above assumptions. However, as discussed around Conjecture 4.7, Theorems 4.5 and 4.6 may be applicable in such cases if we modify the edge weights by taking maximum with $n^{-\delta}$ for some $\delta \in (1/6, 1/3)$. So, if one can prove a convergence result with the original edge weights similar to (39) that only concerns the component sizes and the surplus edges (which are usually easier to deal with compared to the metric space structure), then that would yield the metric scaling limit as well using a coupling argument. Thus, a general result of this form (dealing only with the component sizes and the surplus edges) applicable to a broad collection of kernels that are not bounded away from zero would significantly extend the class of edge weights for which the convergence in (26) holds.

Consider the setting in Theorem 4.9. As can be seen from its proof, Condition 4.8 (a) was only needed to ensure that (226) holds. Now, the left side of (226) equals

$$n^{2/3}\Big[\int_{[0,1]^2 \setminus D_n} W^2 + \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} W^2(U_i, U_j) - \frac{4}{n^2} \sum_{1 \leq i < j \leq n} W\big(V_i^{(n)}, V_j^{(n)}\big) \cdot \fint_{Q_{ij}} W \Big], \quad (246)$$

where $U_1, \ldots, U_n$ and $V_1^{(n)}, \ldots, V_n^{(n)}$ are as defined before (41). Note that the third term appearing in (246) is a weighted U-statistic. Thus, using the expression in (246), it should be possible to prove that (226) holds for a class of kernels much bigger than the collection of kernels that satisfy Condition 4.8 (a), and consequently, the assumption in Theorem 4.9 can be relaxed. However, we are not aware of a result on U-statistics that can be directly applied here.

Fix $d \in \mathbb{Z}_{\geq 3}$, and consider a random graph process on $[n]$ that starts with the empty graph on $[n]$ and evolves as follows: at each step, an available pair of vertices is selected uniformly and an edge is added between them, where a pair of vertices is designated 'available' if they are non-adjacent and have degrees strictly smaller than d. This dynamic random graph model with degree restrictions was introduced by Ruciński and Wormald [56]. For $t \geq 0$, let $\mathscr{G}_n^{\mathrm{dr}}(t)$ denote the above random graph process at time $\lfloor nt \rfloor$. The evolution rule for $\mathscr{G}_n^{\mathrm{dr}}(\cdot)$ is a variation of that of the dynamic configuration model. However, this process lacks the nice exchangeability properties of the configuration model [43, Chapter 7], which makes it harder to analyze. In particular, it is not clear to us if the near-critical behavior of $\mathscr{G}_n^{\mathrm{dr}}(\cdot)$ can be studied using the exploration process technique or the spanning

tree technique mentioned in Section 1. However, we believe that Theorem 3.4 can be applied to this process to get the metric scaling limit inside the critical window. To do this, one would require the following result: Let $|\mathscr{C}_1^{\mathrm{dr}}(t)|$ denote the size of the largest component of $\mathscr{G}_n^{\mathrm{dr}}(t)$, $t \geq 0$. Then there exists $C > 0$ and $\delta \in (1/6, 1/3)$ such that

$$\lim_{n\to\infty} \mathbb{P}\left(|\mathscr{C}_1^{\mathrm{dr}}(t)| \leq \frac{C\log n}{(t_c - t)^2} \quad \text{and} \quad \mathrm{diam}\big(\mathscr{G}_n^{\mathrm{dr}}(t)\big) \leq \frac{C\log n}{(t_c - t)} \quad \text{for all} \quad t \leq t_c - n^{-\delta}\right) = 1,$$

where t_c denotes the point of phase transition for this process. It seems to us that the rest of the program can be carried out relatively easily.

We close this section with the result on the critical scaling limit of the stochastic block model mentioned in Section 4.2.1. Fix $k \in \mathbb{Z}_{>0}$, and let $\kappa_n : [k] \times [k] \to (0, \infty)$ be symmetric kernels for each $n \geq 1$. Suppose each vertex $i \in [n]$ has a "type" $x_i \in [k]$. For $1 \leq i \neq j \leq n$, define $\beta_{ij}^{(n)} := \kappa_n(x_i, x_j)$, and let $\mathscr{G}_n^{\mathrm{sbm}}$ and $\widetilde{\mathscr{G}}_n^{\mathrm{sbm}}$ respectively be the random graphs $\mathscr{G}_n^{\mathrm{gphn}}$ and $\widetilde{\mathscr{G}}_n^{\mathrm{gphn}}$ constructed using these edge weights. For each $x \in [k]$ and each n, let $\mu_n(x) := n^{-1} \cdot \#\{i \in [n] : x_i = x\}$.

Condition 7.1. (a) **Convergence of the kernels:** *There exists a kernel* $\kappa(\cdot, \cdot) : [k] \times [k] \to (0, \infty)$ *and a matrix* $A = (a_{xy})_{x,y\in[k]}$ *with real-valued (not necessarily positive) entries such that for each* $x, y \in [k]$, $\lim_{n\to\infty} n^{1/3}\big(\kappa_n(x, y) - \kappa(x, y)\big) = a_{xy}$.

(b) **Convergence of the empirical measures:** *There exists a probability measure* μ *on* $[k]$ *and a vector* $\mathbf{b} = (b_1, \ldots, b_k)^t$ *such that* $\min_{x\in[k]} \mu(x) > 0$ *and for each* $x \in [k]$,

$$n^{1/3}\big(\mu_n(x) - \mu(x)\big) \to b_x \quad \text{as} \quad n \to \infty.$$

(c) **Criticality:** *The operator norm of* T_κ *in* $L^2([k], \mu)$ *equals one.*

We will simply write κ for the $k \times k$ matrix with entries $\kappa(i, j)$. Define $\boldsymbol{\mu}$ (resp. $\mathbf{1}$) to be the vector $(\mu(1), \ldots, \mu(k))^t$ (resp. the $k \times k$ column vector where each coordinate is one). Let $D := \mathrm{Diag}(\boldsymbol{\mu})$–the $k \times k$ diagonal matrix with diagonal entries $\mu(1), \ldots, \mu(k)$. Similarly define $B := \mathrm{Diag}(\mathbf{b})$. Define $m_{ij} := \mu(j)\kappa(i, j)$ for $i, j \in [k]$ and let $M = \big(m_{ij}\big)_{i,j\in[K]}$. Note that Condition 7.1 (c) is equivalent to the Perron root of M being equal to one. Let $\mathbf{u}$ and $\mathbf{v}$ respectively be the right and left eigenvectors of M corresponding to the eigenvalue one normalized so that $\mathbf{v}^t\mathbf{u} = 1$ and $\mathbf{u}^t\mathbf{1} = 1$, *i.e.*,

$$M\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^t M = \mathbf{v}^t, \quad \mathbf{u}^t\mathbf{1} = 1, \quad \text{and} \quad \mathbf{v}^t\mathbf{u} = 1. \tag{247}$$

Writing $\mathbf{u} = (u_1, \ldots, u_K)^t$ and $\mathbf{v} = (v_1, \ldots, v_K)^t$, define

$$\alpha := \frac{1}{(\mathbf{v}^t\mathbf{1}) \cdot (\boldsymbol{\mu}^t\mathbf{u})}, \quad \chi := \frac{\sum_{x\in[K]} v_x u_x^2}{(\mathbf{v}^t\mathbf{1}) \cdot (\boldsymbol{\mu}^t\mathbf{u})^2}, \quad \text{and} \quad \zeta := \alpha \cdot \big[\mathbf{v}^t(AD + \kappa B)\mathbf{u}\big]. \tag{248}$$

Theorem 7.2. *Under Condition 7.1, the convergences in* (26) *and* (27) *continue to hold if we replace* G_n^{gphn} *by* $\mathscr{G}_n^{\mathrm{sbm}}$ *or* $\widetilde{\mathscr{G}}_n^{\mathrm{sbm}}$, *where the constants are as in* (248).

ACKNOWLEDGMENTS

References

[1] R. Abraham, J.-F. Delmas, and P. Hoscheit, *A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces*, Electron. J. Probab. **18** (2013), no. 14, 1–21.

[2] L. Addario-Berry, S. Bhamidi, and S. Sen, *A probabilistic approach to the leader problem in random graphs*, Random Structures Algorithms **58** (2021), no. 1, 34–67. MR4180253

[3] L. Addario-Berry, N. Broutin, and C. Goldschmidt, *Critical random graphs: limiting constructions and distributional properties*, Electron. J. Probab **15** (2010), no. 25, 741–775.

[4] L. Addario-Berry, N. Broutin, and C. Goldschmidt, *The continuum limit of critical random graphs*, Probability Theory and Related Fields **152** (2012), no. 3-4, 367–406.

[5] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont, *The scaling limit of the minimum spanning tree of the complete graph*, The Annals of Probability **45** (2017), no. 5, 3075–3144.

[6] L. Addario-Berry and S. Sen, *Geometry of the minimal spanning tree of a random 3-regular graph*, Probab. Theory Related Fields **180** (2021), no. 3-4, 553–620. MR4288328

[7] D. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, The Annals of Probability (1997), 812–854.

[8] D. Aldous, G. Miermont, and J. Pitman, *The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin's local time identity*, Probability Theory and Related Fields **129** (2004), no. 2, 182–218.

[9] D. Aldous and B. Pittel, *On a random graph with immigrating vertices: Emergence of the giant component*, Random Structures and Algorithms **17** (2000), no. 2, 79–102.

[10] S. Bhamidi, N. Broutin, S. Sen, and X. Wang, *Scaling limits and universality II: geometry of maximal components in dynamic random graph models in the critical regime*, In preparation.

[11] S. Bhamidi, S. Dhara, R. v. d. Hofstad, and S. Sen, *Universality for critical heavy-tailed network models: metric structure of maximal components*, Electron. J. Probab. **25** (2020), Paper No. 47, 57. MR4092766

[12] S. Bhamidi, S. Dhara, R. v. d. Hofstad, and S. Sen, *Global lower mass-bound for critical configuration models in the heavy-tailed regime*, Electron. J. Probab. **27** (2022), Paper No. 103, 29. MR4460270

[13] S. Bhamidi, R. v. d. Hofstad, and S. Sen, *The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs*, Probab. Theory Related Fields **170** (2018), no. 1-2, 387–474. MR3748328

[14] S. Bhamidi and S. Sen, *Geometry of the minimal spanning tree in the heavy-tailed regime: new universality classes*, arXiv preprint arXiv:2009.10696 (2020).

[15] S. Bhamidi and S. Sen, *Geometry of the vacant set left by random walk on random graphs, Wright's constants, and critical random graphs with prescribed degrees*, Random Structures Algorithms **56** (2020), no. 3, 676–721. MR4084187

[16] S. Bhamidi, S. Sen, and X. Wang, *Continuum limit of critical inhomogeneous random graphs*, Probability Theory and Related Fields **169** (2017), no. 1-2, 565–641.

[17] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan, *Percolation on dense graph sequences*, Ann. Probab. **38** (2010), no. 1, 150–183. MR2599196

[18] B. Bollobás, S. Janson, and O. Riordan, *The phase transition in inhomogeneous random graphs*, Random Structures & Algorithms **31** (2007), no. 1, 3–122.

[19] B. Bollobás, S. Janson, and O. Riordan, *The cut metric, random graphs, and branching processes*, J. Stat. Phys. **140** (2010), no. 2, 289–335. MR2659281

[20] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, *Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), no. 6, 1801–1851. MR2455626

[21] L. A. Braunstein, S. V. Buldyrev, R. Cohen, S. Havlin, and H. E. Stanley, *Optimal paths in disordered complex networks*, Physical review letters **91** (2003), no. 16, 168701.

[22] L. A Braunstein, Z. Wu, Y. Chen, S. V Buldyrev, T. Kalisky, S. Sreenivasan, R. Cohen, E. Lopez, S. Havlin, and H E. Stanley, *Optimal path and minimal spanning trees in random weighted networks*, International Journal of Bifurcation and Chaos **17** (2007), no. 07, 2215–2255.

[23] N. Broutin, T. Duquesne, and M. Wang, *Limits of multiplicative inhomogeneous random graphs and Lévy trees: limit theorems*, Probab. Theory Related Fields **181** (2021), no. 4, 865–973. MR4344135

[24] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, Vol. 33, American Mathematical Society Providence, 2001.

[25] D. L. Burkholder, B. J. Davis, and R. F. Gundy, *Integral inequalities for convex functions of operators on martingales*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, 1972, pp. 223–240. MR0400380

[26] D. S Callaway, J. E Hopcroft, J. M Kleinberg, M. E. Newman, and S. H Strogatz, *Are randomly grown graphs really random?*, Physical Review E **64** (2001), no. 4, 041902.

[27] M. Camarri and J. Pitman, *Limit distributions and random trees derived from the birthday problem with unequal probabilities*, Electron. J. Probab **5** (2000), no. 2, 1–18.

[28] M. Capobianco and O. Frank, *Graph evolution by stochastic additions of points and lines*, Discrete Math. **46** (1983), no. 2, 133–143. MR710884

[29] Y. Chen, E. López, S. Havlin, and H E. Stanley, *Universal behavior of optimal paths in weighted networks with general disorder*, Physical review letters **96** (2006), no. 6, 068702.

[30] G. Conchon-Kerjan and C. Goldschmidt, *The stable graph: the metric space scaling limit of a critical random graph with iid power-law degrees*, The Annals of Probability **51** (2023), no. 1, 1–69.

[31] S. Dhara, R. v. d. Hofstad, J. S. H. v. Leeuwaarden, and S. Sen, *Critical window for the configuration model: finite third moment degrees*, Electron. J. Probab. **22** (2017), 33 pp.

[32] S. Dhara, R. v. d. Hofstad, and J. S. H. van Leeuwaarden, *Critical percolation on scale-free random graphs: new universality class for the configuration model*, Comm. Math. Phys. **382** (2021), no. 1, 123–171. MR4223472

[33] S. Dhara, R. v. d. Hofstad, J. S. H. van Leeuwaarden, and S. Sen, *Heavy-tailed configuration models at criticality*, Ann. Inst. Henri Poincaré Probab. Stat. **56** (2020), no. 3, 1515–1558. MR4116701

[34] J. Duoandikoetxea, *Fourier analysis*, Graduate Studies in Mathematics, vol. 29, American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe. MR1800316

[35] R. Durrett, *Rigorous result for the CHKNS random graph model*, Discrete random walks (Paris, 2003), 2003, pp. 95–104. MR2042377

[36] L. C. Evans, *Partial differential equations*, Second Edition, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR2597943

[37] S. N. Evans, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR2351587 (2009d:60014)

[38] A. Frieze and R. Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), no. 2, 175–220. MR1723039

[39] C. Goldschmidt, B. Haas, and D. Sénizergues, *Stable graphs: distributions and line-breaking construction*, Ann. H. Lebesgue **5** (2022), 841–904. MR4526241

[40] M. Heydenreich and R. v. d. Hofstad, *Random graph asymptotics on high-dimensional tori*, Comm. Math. Phys. **270** (2007), no. 2, 335–358. MR2276449

[41] M. Heydenreich and R. v. d. Hofstad, *Random graph asymptotics on high-dimensional tori II: volume, diameter and mixing time*, Probab. Theory Related Fields **149** (2011), no. 3-4, 397–415. MR2776620

[42] W. Hoeffding, *The strong law of large numbers for U-statistics.*, North Carolina State University. Dept. of Statistics, 1961.

[43] R. v. d. Hofstad, *Random graphs and complex networks. Vol. 1* (2017), xvi+321. MR3617364

[44] R. v. d. Hofstad and A. Nachmias, *Hypercube percolation*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 3, 725–814. MR3612867

[45] R. v. d. Hofstad and A. Sapozhnikov, *Cycle structure of percolation on high-dimensional tori*, Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 3, 999–1027. MR3224297

[46] S. Janson, *Asymptotic equivalence and contiguity of some random graphs*, Random Structures Algorithms **36** (2010), no. 1, 26–45. MR2591045 (2011h:60021)

[47] S. Janson, *Graphons, cut norm and distance, couplings and rearrangements*, New York Journal of Mathematics. NYJM Monographs, vol. 4, State University of New York, University at Albany, Albany, NY, 2013. MR3043217

[48] A. Joseph, *The component sizes of a critical random graph with given degree sequence*, Ann. Appl. Probab. **24** (2014), no. 6, 2560–2594. MR3262511

[49] J. L Kazdan, *Matrices $A(t)$ depending on a parameter t*, unpublished note (1995).

[50] S. Lang, *Fundamentals of differential geometry*, Graduate Texts in Mathematics, vol. 191, Springer-Verlag, New York, 1999. MR1666820

[51] J.-F. Le Gall, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311. MR2203728 (2007h:60078)

[52] E. Lubetzky and Y. Peled, *Noise sensitivity of critical random graphs*, Israel J. Math. **252** (2022), no. 1, 187–214. MR4526830

[53] G. Miermont and S. Sen, *On breadth-first constructions of scaling limits of random graphs and random unicellular maps*, Random Structures Algorithms **61** (2022), no. 4, 803–843. MR4504844

[54] J. Pitman, *Random mappings, forests, and subsets associated with Abel-Cayley-Hurwitz multinomial expansions*, Sém. Lothar. Combin. **46** (2001/02), Art. B46h, 45. MR1877634

[55] O. Riordan, *The phase transition in the configuration model*, Combinatorics, Probability and Computing **21** (2012), no. 1-2, 265–299.

[56] A. Ruciński and N. C. Wormald, *Random graph processes with degree restrictions*, Combin. Probab. Comput. **1** (1992), no. 2, 169–180. MR1179247

[57] W. Rudin, *Real and complex analysis*, Third, McGraw-Hill Book Co., New York, 1987. MR924157

[58] J. Spencer and N. Wormald, *Birth control for giants*, Combinatorica **27** (2007), no. 5, 587–628.

[59] J. L. Spouge, *Solutions and critical times for the monodisperse coagulation equation when $a_{ij} = A + B(i + j) + Cij$*, J. Phys. A **16** (1983), no. 4, 767–773. MR706194

[60] Z. Wu, L. A Braunstein, S. Havlin, and H E. Stanley, *Transport in weighted networks: partition into superhighways and roads*, Physical review letters **96** (2006), no. 14, 148702.

[61] A. C. Zaanen, *Introduction to operator theory in Riesz spaces*, Springer-Verlag, Berlin, 1997. MR1631533