Dynamic random graphs with vertex removal

Josep Díaz∗1, Lyuben Lichev2,3, and Bas Lodewijks2,3

1Universitat Politècnica de Catalunya, Barcelona, Spain
2Univ. Jean Monnet, Saint-Etienne, France
3Institut Camille Jordan, Lyon, France

July 12, 2022

Abstract

We introduce and analyse a Dynamic Random Graph with Vertex Removal (DRGVR) defined as follows. At every step, with probability \( p > 1/2 \) a new vertex is introduced, and with probability \( 1 - p \) a vertex, chosen uniformly at random among the present ones (if any), is removed from the graph together with all edges adjacent to it. In the former case, the new vertex connects by an edge to every other vertex with probability proportional to the number of vertices already present.

We prove that the DRGVR converges to a local limit and determine this limit. Moreover, we analyse its component structure and distinguish a subcritical and a supercritical regime with respect to the existence of a giant component. As a byproduct of this analysis, we obtain upper and lower bounds for the critical parameter. Furthermore, we provide precise expression of the maximum degree (as well as in- and out-degree for a natural orientation of the DRGVR). Several concentration and stability results complete the study.

Keywords: random graph, complex network, local limit, component structure, maximum degree, Poisson process

MSC Class: 05C80

1 Introduction

Since the appearance of the first random graph models in the late 50’s [13, 15], the study of random graphs has attracted great interest both from theoretical and applied point of view.

In the last twenty years, apart from the study of classical random graph theory, much effort has been made to provide and analyse accurate models for large-scale real-world networks like the internet or different P2P social networks. One key characteristic of such networks is that for most of them, the degree of the vertices follows an inverse power law, that is, the proportion of vertices with degree \( d \) is approximately \( cd^{-α} \) for some constants \( c, α > 0 \). As a consequence, the fact that the classical random graphs as \( G(n,m) \) and \( G(n,p) \) typically have Poisson degree distribution makes them unsuitable as models for real-world networks. An important contribution was made by Albert and Barabási [2] who proposed a scale-free dynamic network model with a preferential attachment rule. In their model, a new vertex is introduced at every step. At its arrival, a vertex chooses to attach to existing vertices with probability proportional to the degree of the existing vertices. Since Albert and Barabási’s breakthrough, the analysis of his scale-free model and its variations has been object of intensive study, see for example [24, 27].

Although the power-law behaviour is an important characteristic of real-world networks, it is not the only one. One natural observation is that besides the creation of vertices in a dynamic network, sometimes

∗Supported by grant MOTION, PID2020-112581GB-C21 from MCIN/ AEI /10.13039/501100011033
vertices could also disappear: for instance, in a social network a user could delete their account, or in a web network a server could break down due to failure. However, most dynamic network models did not consider the possibility of vertex removal. Among the few exceptions are [6, 7], which both consider settings based on Albert and Barabási’s scale-free model. In the model of Chung and Lu [6], each of vertex creation, vertex deletion, edge creation and edge deletion may happen at any given step with probabilities that sum up to 1. When a vertex is created, it connects to an existing vertex by a single edge according to a preferential attachment rule. The authors provide bounds for the diameter, the graph distance between two typical vertices, the connected components and the spectrum of the adjacency matrix. In the similar model of Cooper, Frieze and Vera [7] every vertex is connected to \( m \geq 1 \) neighbours at its arrival. The authors analyse the degree sequence of the obtained dynamic graph and prove its scale-freeness.

The notion of a dynamic network on a set of aging individuals has also been considered. Instead of deleting a vertex at a given step, which forbids later vertices to attach to it, several works [3, 10, 14] considered models in which new vertices connect to any fixed vertex with smaller probability as time goes by. The main motivation of such models comes from real-world networks like the citation network where empirical observations show that papers usually get less popular with time. Contrary to [6, 7] and similarly to the model we introduce below, the authors of [14] consider a setting where every new vertex connects to an edge independently with probability \( \frac{\beta}{\sqrt{n}} \), for which Shepp [26] showed the existence of a sharp threshold at \( \beta = 1/4 \) for the appearance of a giant component, and Dereich and Mörters later recovered the result in a more general framework, see Proposition 1.3 in [8]. Similar results were obtained in [11, 12] and for the closely related CHKNS model.

**Definition 1.1** (Dynamic Random Graph with Vertex Removal (DRGVR)). Fix constants \( \beta > 0, \varepsilon \in (0, 1/2) \) and a sequence \((\xi_n)_{n \geq 0}\) of i.i.d. random variables with Bernoulli distribution with parameter \( p := 1/2 + \varepsilon \). We define a sequence of graphs \((G_n)_{n \geq 0} = ((V_n, E_n))_{n \geq 0}\) by initialising \( G_0 \) to be the empty graph, and for every \( n \geq 1 \), we construct \( G_n \) from \( G_{n-1} \) as follows. If \( \xi_n = 1 \), set \( V_n = V_{n-1} \cup \{n\} \) and conditionally on \( V_{n-1} \), add each of the edges \((\{i, n\})_{i \in V_{n-1}}\) to \( E_{n-1} \) independently with probability \( \min\left\{\frac{\beta}{|V_{n-1}|}, 1\right\} \) to construct \( E_n \). If \( \xi_n = 0 \), select a vertex uniformly at random from \( V_{n-1} \) (if any) and remove it together with all edges incident to it.

Note that the graph \( G_n \) will sometimes be seen as a directed graph where the edges are directed from the vertex arriving later to the vertex arriving earlier. In this case, the edge set will be denoted by \( E_n \), and the directed graph itself by \( \vec{G}_n \).

Finally, we remark that the well-studied Dubin’s model [12, 18, 26] is a particular case of the DRGVR for \( \varepsilon = 1/2 \) (or equivalently \( p = 1 \)). In this particular setting Shepp [26] showed the existence of a sharp threshold at \( \beta = 1/4 \) for the appearance of a giant component, and Dereich and Mörters later recovered the result in a more general framework, see Proposition 1.3 in [8]. Similar results were obtained in [11, 12] and for the closely related CHKNS model.

**Notation** Throughout the paper we set \( \mathbb{N} := \{1, 2, \ldots\} \) denote the set of natural numbers, \( \mathbb{N}_0 := \mathbb{N} \cup \{0\} \), and \([t] := \{i \in \mathbb{N} : i \leq t\}\) for every \( t \geq 1 \). For \( x \in \mathbb{R} \), we let \([x] := \inf\{n \in \mathbb{Z} : n \geq x\}\) and \([x] := \sup\{n \in \mathbb{Z} : n \leq x\}\). For positive real sequences \((a_n, b_n)_{n \in \mathbb{N}}\), we say that \( a_n = o(b_n) \) if \( \lim_{n \to \infty} a_n/b_n = 0 \), \( a_n = o(a_n) \) if \( \lim_{n \to \infty} a_n = 0 \), \( a_n \sim b_n \) if \( \lim_{n \to \infty} a_n/b_n = 1 \), and \( a_n = \Omega(b_n) \) if \( b_n = \Omega(a_n) \). Note that we allow the constant \( C \) to depend on the parameters of the problem that are fixed, that is, do not depend on the choice of a sufficiently large \( n \). For random variables \((X_n)_{n \in \mathbb{N}}\) and \( X \), we let \( X_n \xrightarrow{d} X \), \( X_n \xrightarrow{p} X \), and \( X_n \xrightarrow{a.s.} X \) denote convergence in distribution, in probability and almost sure convergence of \((X_n)_{n \in \mathbb{N}}\) to \( X \), respectively. Furthermore, we denote by \( L(X) \) the distribution of \( X \).

For a graph \( G \), the size of \( G \), often denoted \(|G|\), is the number of vertices in \( G \). For two vertices \( u, v \) in \( G \), we let \( u \to v \) denote the fact that \( u \) is connected to \( v \) by a (directed) edge. Moreover, for two finite rooted graphs \((G, o_G)\) and \((H, o_H)\), we denote by \((G, o_G) \cong (H, o_H)\) the fact that the two rooted graphs are isomorphic.
Finally, for an event $\mathcal{E}$ in a given probability space, we denote by $\mathcal{E}^c$ the complement of $\mathcal{E}$.

1.1 Results

In this section we present our results related to the properties of the DRGVR model. The first theorem in this section is the main result of the paper. To introduce it, we first need to explain the notion of local convergence of a sequence of graphs. Fix a sequence of rooted graphs $(H_n, o_n)_{n \geq 0}$, where $o_n$ is a vertex of $H_n$ chosen uniformly at random. Also, for a graph $H$, a vertex $v$ in $H$ and an integer $r \geq 0$, we denote by $B_r(H, v)$ the graph induced by all vertices at distance at most $r$ from $v$ in $H$ (often referred to as the ball with radius $r$ around $v$ in $H$). For ease of writing we often assume that $B_r(H, v)$ is a graph rooted at $v$.

Given a probability measure $\mu$ on the set of finite rooted graphs $\mathcal{G}$, we say that the sequence $(H_n, o_n)_{n \geq 1}$ converges locally to $\mu$ if for every integer $r \geq 0$ and every finite rooted graph $(H, o_H) \in \mathcal{G}$ we have that

$$\mathbb{P}(B_r(H_n, o_n) \cong (H, o_H)) \xrightarrow{n \to \infty} \mu((G, o_G) \in \mathcal{G} : B_r(G, o_G) \cong (H, o_H))).$$

Often the limiting measure $\mu$ is identified with a (possibly random) rooted graph. For a complete account and applications on the notion of local convergence, see for example [25].

In this paper we provide a stronger result for the local weak convergence in terms of the total variation distance. For two probability distributions $\mu_1$ and $\mu_2$ defined on a common probability space $\Omega$, the total variation distance between $\mu_1$ and $\mu_2$ is defined as

$$d_{TV}(\mu_1, \mu_2) := \sup_{A \subseteq \Omega} |\mu_1(A) - \mu_2(A)| = \inf \{ \mathbb{P}(X \neq Y) : (X, Y) \text{ is a coupling of } (\mu_1, \mu_2) \}.$$

By taking $\Omega = \mathcal{G}$, we may consider the total variation between the distributions of $B_r(G_n, k_0)$ and $B_r(T, 0)$, where $(T, 0)$ is some rooted random graph and $k_0$ is a vertex chosen uniformly at random from $G_n$. In this paper, the definition of the total variation distance in terms of couplings is particularly useful.

The first and main result of this paper deals with the local weak convergence of the DRGVR model. To this end, we define the following multi-type branching process.

Definition 1.2 (Binomial birth-death tree). Fix $\beta > 0$, $\varepsilon \in (0, 1/2]$ and set $p := 1/2 + \varepsilon$. We define a multi-type branching process $\mathcal{T}$ with type space $(0, 1]$ as follows. The root 0 of $\mathcal{T}$ has type $a_0 \sim \text{Beta}(\frac{\beta \varepsilon}{2}, 1)$. Then, any vertex $v$ in $\mathcal{T}$ with type $a_v$ produces an offspring independently of all other vertices according to a Poisson point process on $(0, 1]$. Conditionally on $a_v$, the density of this Poisson process is given by

$$\lambda_v^+(dx) := \frac{\beta p}{2\varepsilon a_v} x^{(1-p)/2\varepsilon} dx \text{ for } x \in (0, a_v), \quad \text{and} \quad \lambda_v^-(dx) := \frac{\beta p}{2\varepsilon} x^{(1-p)/(2\varepsilon) - 1} dx \text{ for } x \in (a_v, 1].$$

The types of the vertices in the offspring are identified by their position in the Poisson point process.

Theorem 1.3. Fix $\beta > 0$ and $\varepsilon \in (0, 1/2]$, and consider the DRGVR model and Binomial birth-death tree given in Definitions 1.1 and 1.2, respectively. Let $k_0$ be a vertex selected uniformly at random from $V_n$. Then, for any $r \in \mathbb{N}$,

$$d_{TV}(\mathcal{L}(B_r(G_n, k_0)), \mathcal{L}(B_r(T, 0))) \leq (\log \log n)^{-(1-p)/(2\varepsilon) - 1/r}.$$

In particular, $(G_n, k_0)_{n \geq 0}$ converges locally to the random rooted tree $(T, 0)$.

The following corollary is obtained by applying Theorem 1.3 for $r = 1$ and integrating the density of the Poisson process, associated to the offspring of the origin, over the interval $(0, 1]$.

Corollary 1.4. Let $D_n^0$ denote the degree of the vertex $k_0$, and conditionally on $a_0$ let

$$X_0 \sim \text{Poi} \left( \frac{\beta p}{1 - p} \left( 1 - \frac{2p - 1}{p} a_0^{(1-p)/(2\varepsilon)} \right) \right).$$

Then, $d_{TV}(\mathcal{L}(D_n^0), \mathcal{L}(X_0)) \leq (\log \log n)^{-p/(2\varepsilon)}$. 
Recall that in the case $p = 1$, the DRGVR recovers Dubin’s model. Though a tailored proof for the case $p = 1$ might result in an improved rate of convergence, the existence of a local weak limit for Dubin’s model by Theorem 1.3 is novel in itself. Furthermore, we directly obtain the limiting degree distribution as well as bounds on the convergence rate of the empirical degree distribution as a result. Though Corollary 1.4 is a result already obtained by Bollobás, Janson and Riordan [5] for a “closely related” inhomogeneous random graph model (see Section 6 for more details), the convergence in total variation distance strengthens this result.

We now present several results that discuss more global properties. First, we discuss the conditions for the emergence of a giant component and the relation with the local weak limit. Define

$$\gamma = \gamma(\beta) := \mathbb{P}(|T| = \infty)$$

(2)
to be the survival probability of the branching process $T$.

**Theorem 1.5.** Fix $\beta > 0$ and $\varepsilon \in (0, 1/2]$, and consider the DRGVR model as in Definition 1.1. Moreover, consider the survival probability $\gamma$ of its local weak limit given in (2). Let $C_1 = C_1(G_n)$ and $C_2 = C_2(G_n)$ denote the first and the second largest component in $G_n$. Then, there exists $\beta_c = \beta_c(p)$ such that:

- if $\beta < \beta_c$, then $|C_1|/n$ converges in probability to 0 as $n \to \infty$;
- if $\beta > \beta_c$, then $|C_1|/n$ converges in probability to $2\varepsilon \gamma \in (0, 1]$, and moreover $|C_2| = O(\log n)$ with high probability.

**Remark 1.6.** (i) It may seem counter-intuitive at first that the size of the largest component is proportional to $2\varepsilon \gamma$. However, heuristically, the proportion of vertices in the giant component, i.e. $|C_1|/|V_n|$, converges to $\gamma$ in probability, and it can be checked that $|V_n|/n$ converges to $2\varepsilon$ almost surely (see Lemma 2.5). Combining the two observations yields the result.

(ii) It follows directly from the definition of the DRGVR model that bond percolation on $G_n$ with retention probability $q$ yields a graph with a giant component only when $\beta > \beta_c$ and $q > \beta_c/\beta$.

Theorem 1.5 raises the natural question of how $\beta_c$ behaves as a function of $p$. The following proposition partially answers this question by providing a lower and upper bound for $\beta_c$ in terms of $p$.

**Proposition 1.7.** Consider the same set-up as Theorem 1.5, the threshold function $\beta_c = \beta_c(p)$ is a non-increasing continuous function over the interval $(1/2, 1]$. Moreover, for every $p \in (1/2, 1]$,

$$\max \left\{ \sqrt{\frac{1 - p}{p}}, \frac{1}{4} \right\} \leq \beta_c(p) \leq \inf_{t \in (-1/2, \infty)} \left( \frac{(1 + 2t)(2t^2 + 7t + 4 + 1/p)}{(1 + t)^2(t + 1/p)(2t + 2/p - 1)} \right)^{-1/2} \leq \sqrt{\frac{2 - p}{p(1 + 4p)}}.$$

**Remark 1.8.** While the tighter upper bound for $\beta_c(p)$, as provided by the infimum in Proposition 1.7 is strictly smaller for all $p \in (1/2, 1]$, we still provide a weaker but also simpler upper bound in terms of $p$. This bound is realised by taking $t = 0$ in the infimum. Note that $\beta_c(1) = 1/4$ and $\lim_{p \downarrow 1/2} \beta_c(p) = 1$, as can be observed in Figure 1.

Once the giant component is established, one may wonder what else can be said about its structure. We are unable to provide a conclusive answer to this question. However, we can say a lot more for the graph that is induced by the vertices born “not too early”. For example, deleting the vertices in the set $V_{\lambda n} \cap V_n$ from $G_n$ does not lead to a “big” change in the size of the giant when $\lambda$ is a “small constant”.

**Theorem 1.9.** Consider the setup of Theorem 1.5 with $\beta > \beta_c(p)$ and recall the constant $\gamma = \gamma(p)$ from (2). Then, for every $\delta \in (0, \gamma)$ there exists $\lambda > 0$ such that the graph $G_n^\lambda$, induced from $G_n$ by the vertices in $V_n \setminus V_{\lambda n}$, satisfies w.h.p. each of the following statements:

1. $(2\varepsilon \gamma - \delta)n \leq |C_1(G_n^\lambda)| \leq (2\varepsilon \gamma + \delta)n,$
Bounds for $\beta_c$

\[
\begin{align*}
\text{Bounds for } \beta_c \\
p & \quad 0.5, 0.6, 0.7, 0.8, 0.9, 1.0 \\
& \quad 0.2, 0.4, 0.6, 0.8, 1.0 \\
\end{align*}
\]

Figure 1: A plot of the lower bound and the two upper bounds for $\beta_c$ from Proposition 1.7. The sharper upper bound given by $f(p) := \inf_{t \in (-1/2, \infty)} \left( \frac{(1+2t)(2t^2+7t+4+1/p)}{(1+t)^2(t+1/p)(2t+2/p-1)} \right)^{-1/2}$ is approximated numerically.

2. $|C_2(G_n^\beta)| = O(\log n)$,

3. there is a constant $\zeta = \zeta(\beta, p) > 0$ such that two vertices of $G_n^\beta$, chosen uniformly at random, have a graph distance in the interval $[(1 - \delta)\zeta \log n, (1 + \delta)\zeta \log n]$ (in $G_n^\beta$).

The graph $G_n$ being constructed recursively, its edges may naturally be equipped with an orientation from the vertex arriving at a later step to the vertex arriving earlier. Therefore, by setting $V_n = \{v_1, \ldots, v_{|V_n|}\}$ with $v_i < v_j$ if $i < j$, for every $j \in [|V_n|]$ we denote by $d_n^+(v_j)$ (respectively $d_n^-(v_j)$) the number of neighbors of $v_j$ in the set $\{v_1, \ldots, v_{j-1}\}$ (respectively in the set $\{v_{j+1}, \ldots, v_{|V_n|}\}$). Also, we denote by $d_n^s(v_j) := d_n^+(v_j) + d_n^-(v_j)$ the total degree of $v_j$ in $G_n$. We remark that in general we identify the vertices in the process in two different ways: as elements of $V_n = \{v_1, \ldots, v_{|V_n|}\}$, and by their labels which indicate the step at which they were born. The latter notation is more general as it is applied to all vertices in the process and not only the ones that survive until step $n$.

The following theorem precisely characterises the maximum degree, the maximum in-degree and the maximum out-degree, as well as the the labels of vertices that attain these maximal degrees.

**Theorem 1.10.** Consider the DRGVR model as in Definition 1.1 with $\beta > 0$ and $\varepsilon \in (0, 1/2)$. Then, for $\Box \in \{s, +\}$,

\[
\frac{(\log \log n)^2}{\log n} \max_{v \in V_n} d_n^\Box(v) - \log \log n - \log \log \log n \xrightarrow{p} 1 + \log \left( \frac{\beta p}{1-p} \right),
\]

and

\[
\frac{(\log \log n)^2}{\log n} \max_{v \in V_n} d_n^- (v) - \log \log n - \log \log \log n \xrightarrow{p} 1 + \log \beta.
\]

Finally, let $I^\Box_n := \{u \in V_n \text{ and } d_n^\Box(u) = \max_{v \in V_n} d_n^\Box(v)\}$ be the set of vertices that attain the maximum (in-/out-)degree. Then, for $\Box \in \{s, +\}$ and for any fixed $\delta \in (0, 1)$,

\[
\frac{\log \min I_n^\Box}{\log n} \xrightarrow{p} 1, \quad \text{and} \quad P(\max I_n^\Box \leq \delta n) = 1 - o(1), \quad \text{and} \quad \frac{\min I_n^-}{n} \xrightarrow{p} 1.
\]
Note that Theorem 1.10 does not include the case $p = 1$ (or, equivalently, $\varepsilon = 1/2$) in which no vertices are removed. Indeed, as can be observed from the results by Lodewijks [22, 23] and Banerjee and Bhamidi [1] (in the case of deterministic out-degree equal to 1), the behaviour of large-degree vertices and their labels in the case $p = 1$ is rather different. Most notably, the vertices with maximal (in-)degree have labels of the order $n^{(1-1/(2\log 2)) + o(1)}$ instead of $n^{1-o(1)}$.

The last result in this section concerns a large family of Lipschitz-type parameters. Fix a function $f$ from the set of finite directed graphs to the real numbers. We say that $f$ is $L$-Lipschitz if for any two directed graphs $H_1$ and $H_2$ that differ in only one edge (that is, $|E(H_1) \setminus E(H_2)| + |E(H_2) \setminus E(H_1)| \leq 1$) we have $|f(H_1) - f(H_2)| \leq L$. We remark that the definition does not make any reference to the vertex sets of $H_1$ and $H_2$, in particular, the family of functions we are interested in is insensitive to isolated vertices.

**Theorem 1.11.** Fix $\beta > 0$ and $\varepsilon \in (0, 1/2]$, and consider the DRVGR model as in Definition 1.1. Fix an integer $L \in \mathbb{N}$ and an $L$-Lipschitz function $f$ defined on the set of directed graphs. Then, for any $t \geq 0$,

(i) $\mathbb{P}\left( \left| f(G_n^\beta) - \mathbb{E}[f(G_n^\beta) \mid |V_n|, (d_n^\beta(v_i))_{i=1}^{|V_n|}] \right| \geq t \mid |V_n|, (d_n^\beta(v_i))_{i=1}^{|V_n|} \right) \leq 2 \exp\left(-\frac{t^2}{8|E_n|L^2}\right)$;

(ii) $\mathbb{P}\left( \left| f(G_n) - \mathbb{E}[f(G_n) \mid (V_i)_{i=1}^{|V_n|}, |E_n|] \right| \geq t \mid (V_i)_{i=1}^{|V_n|}, |E_n| \right) \leq 2 \exp\left(-\frac{t^2}{8|E_n|L^2}\right)$.

Note that despite the fact that both statements show conditional concentration of $f(G_n^\beta)$, the two are quite different in nature. In the first case, the conditioning is done over the out-degrees of all vertices in $G_n^\beta$ and is independent from the vertex process $(V_i)_{i \geq 0}$ conditionally on $V_n$. In the second case, the conditioning is done over the entire vertex process $(V_i)_{i \geq 0}$ but the only structural information about $G_n$ is its number of edges.

**Main ideas of the proofs.** The proofs of the main results use a variety of techniques. First, Theorem 1.3 is obtained by providing a coupling between the Breadth-First Search (BFS) exploration of the neighbourhood of a vertex $k_0$ selected uniformly at random in $G_n$, and the recursive construction of the multi-type branching process $T$ with root 0, as defined in Definition 1.2. This coupling is mainly inspired by techniques used for proving the local weak limit of affine preferential attachment models (see [4, 21]), and ensures that the $r$-neighbourhood of $k_0$ in the BFS is isomorphic to the subtree of $T$, induced by the first $r$ generations, as required for the local weak convergence. The aim is to show that this coupling is successful with high probability. Whilst such an approach is common for random graph models with a “tree-like” structure, and for evolving random graph models in particular (see e.g. [4, 21]), the main technical contribution is to prove that, despite the randomness introduced by the vertex removal, the direct neighbours of vertices with “sufficiently large” labels in $G_n$ can be coupled w.h.p. with a discretised Poisson point process which encodes the offspring of vertices in the multi-type branching process $T$.

The proof of Theorem 1.5 combines a comparison of the DRGVR model $G_n$ with a particular inhomogeneous Erdős-Rényi graph model on the vertex set $V_n$ (of $G_n$) and results by Bollobás, Janson and Riordan [5]. More precisely, we use necessary and sufficient conditions for the existence of a giant component of a wide family of inhomogeneous random graph models from [5], and the comparison between the DRGVR model and an instance of this family allows us to transfer their results to our setting. Proposition 1.7 is a byproduct of this proof and is shown by analysing a related linear operator. Theorem 1.9 combines the fact that the restriction of the inhomogeneous Erdős-Rényi graph on $V_n \setminus V_{n^\lambda}$ satisfies that every two edges appear in $G_n^{\lambda}$ with probabilities which are a constant factor away from each other (this is not the case for $G_n$ itself) with results from [5].

Theorem 1.10 combines the comparison between the DRGVR model and an inhomogeneous Erdős-Rényi graph model with precise bounds on the tail distribution of the vertex-degrees for the latter model. These
bounds are used to show that, when the expected number of vertices with degree at least \( k_n \) either tends to zero (respectively to infinity), then the largest degree is at most (respectively at least) \( k_n \) with high probability. With the growth rate of the maximum degree at hand, the tail distribution bounds can be used to argue that vertices with “too small” or “too large” label have degrees that are substantially smaller than the maximum degree in \( G_n \).

Finally, the proof of Theorem 1.11 is based on constructing martingales with bounded differences and the use of the classical Azuma inequality.

**Organisation of the paper.** We provide some preliminary results in Section 2 that are to be used in the proofs of the results presented in Section 1. Sections 3, 4 and 5 are dedicated to proving our main result (Theorem 1.3) regarding the local weak convergence of the DRGVR model. In Section 6, we compare our model with a particular instance of the inhomogeneous Erdős-Rényi graph, and use established results for inhomogeneous random graphs to prove Theorem 1.5, Proposition 1.7 and Theorem 1.9. Section 7 is then devoted to proving the size of the maximum degree and the labels of the vertices that attain the maximum degree (Theorem 1.10). Finally, in Section 8 we prove Theorem 1.11.

2 Preliminaries

2.1 General probabilistic preliminaries

Let \( \Gamma \) be a probability distribution on \( (0,1] \times \{0,1\} \) defined as

\[
\Gamma(x,y) := px^{(1-p)/(2\epsilon)} \, dx \, \delta_1(y) + \left(1 - px^{(1-p)/(2\epsilon)}\right) \, dx \, \delta_0(y),
\]

where \( \delta_0 \) and \( \delta_1 \) are Dirac measures. As we shall see, the distribution of \( \Gamma \) is tightly connected to the distribution of the birth times of the vertices in \( V_n \). For now, we state and prove a key property of this distribution.

**Lemma 2.1.** Let \( (X,Y) \) be sampled from \( \Gamma \). Then, for every \( x \in (0,1] \),

\[
P(X \leq x \mid Y = 1) = \frac{x}{p/(2\epsilon)}.
\]

In particular, conditionally on \( Y = 1 \), \( X \sim \text{Beta}(\frac{p}{2\epsilon}, 1) \). Equivalently, let \( X \sim \text{Unif}(0,1) \) and, conditionally on \( X, Y \sim \text{Ber}(pX^{(1-p)/(2\epsilon)}) \). Then, conditionally on \( Y = 1 \), \( X \sim \text{Beta}(\frac{p}{2\epsilon}, 1) \).

**Proof.** Recall that \( p = \frac{1}{2} + \epsilon \). Thus, both results readily follow from the fact that

\[
P(X \leq x \mid Y = 1) = \int_0^x \frac{ps^{(1-p)/(2\epsilon)}}{\int_0^1 ps^{(1-p)/(2\epsilon)} \, ds} \, ds = \frac{x}{p/(2\epsilon)},
\]

as desired. \( \square \)

The next lemma provides a useful coupling of the Bernoulli and the Poisson distributions.

**Lemma 2.2 ([20], page 5, (1.11)).** Fix \( \lambda \in (0,1) \), and let \( X \sim \text{Poi}(\lambda) \) and \( Y \) be a Bernoulli random variable with success probability \( \lambda \). There exists a coupling \( (\hat{X}, \hat{Y}) \) of \( (X,Y) \) such that \( \hat{X} \geq 1_{\{Y \leq 1\}} Y \) almost surely and

\[
P(\hat{X} \neq \hat{Y}) \leq 2\lambda^2.
\]

**Lemma 2.3.** Fix \( \lambda_1, \lambda_2 > 0 \). There exists a coupling \( (X_1, X_2) \) of the Poisson distributions with means \( \lambda_1 \) and \( \lambda_2 \) so that

\[
P(X_1 \neq X_2) \leq |\lambda_1 - \lambda_2|.
\]
Lemma 2.4. (i) Denote \( \phi : x \in [-1, \infty) \mapsto (1+x) \log(1+x) - x \). Given a binomial random variable \( X \),
\[
\mathbb{P}(X - \mathbb{E}[X] \geq t) \leq \exp \left( -\mathbb{E}[X] \cdot \phi \left( \frac{t}{\mathbb{E}[X]} \right) \right) \leq \exp \left( -\frac{t^2}{2(\mathbb{E}[X] + t/3)} \right) \quad \text{for every } t \geq 0,
\]
\[
\mathbb{P}(X - \mathbb{E}[X] \leq -t) \leq \exp \left( -\mathbb{E}[X] \cdot \phi \left( -\frac{t}{\mathbb{E}[X]} \right) \right) \leq \exp \left( -\frac{t^2}{2(\mathbb{E}[X] + t/3)} \right) \quad \text{for every } t \in [0, \mathbb{E}[X]].
\]

(ii) Fix a sequence of positive real numbers \( (c_i)_{i=1}^n \) and a random variable \( X \). Suppose that \( (X_n)_{n \geq 0} \) is a martingale such that \( X_n = X \), \( X_0 = \mathbb{E}[X] \) and for every \( i \in [n] \), \( |X_i - X_{i-1}| \leq c_i \). Then, for every \( t \geq 0 \),
\[
\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right).
\]

2.2 Preliminaries for the DRGVR

An alternative viewpoint on the vertex process \((V_n)_{n \geq 0}\). In this section we introduce a formal framework that will help us to keep track of the vertex process \((V_n)_{n \geq 0}\). To begin with, we define recursively a set of marks \( (\nu_{i,n})_{1 \leq i \leq n} \) as follows:

- if \( n = 1 \), set \( \nu_{1,1} = 1 \) if \( \xi_1 = 1 \) and \( \nu_{1,1} = 0 \) otherwise,
- if \( n \geq 2 \) and \( \xi_n = 1 \), set \( \nu_{i,n} = \nu_{i,n-1} \) for all \( i \in [n-1] \) and \( \nu_{n,n} = 1 \),
- if \( n \geq 2 \) and \( \xi_n = 0 \), select a uniformly random element \( i_n \) from \( V_{n-1} \) and set \( \nu_{i,n} = \nu_{i,n} = 0 \) and \( \nu_{i,n} = \nu_{i,n-1} \) for all \( i \in [n-1] \setminus \{i_n\} \).

Note that by definition \( V_n = \{i \in [n], \nu_{i,n} = 1\} \), so the set of all marks contains the entire information for the vertex process \((V_n)_{n \geq 1}\) (and is, in a sense, equivalent to it). From this point, we say that the vertices in \( V_n \) are the ones that are alive after \( n \) steps (of the vertex process), or equivalently survive after \( n \) steps. One advantage of the marks is that they allow to describe the distribution of the alive vertices in a clear way. More precisely, as we shall see in Lemma 4.2, the empirical distribution
\[
\Gamma^{(n)} := \frac{1}{n} \sum_{i=1}^n \delta_{(i/n, \nu_{i,n})}
\]
over the set \( (0,1] \times \{0,1\} \), where \( \delta_{(x,j)} \) denotes a Dirac mass at the point \((x,j) \in (0,1] \times \{0,1\} \), converges in distribution to the distribution \( \Gamma \) defined in (6).
Further preliminary results. We define
\[ Q_n := \{|V_n| - 2\varepsilon n| \leq n^{2/3} \}. \]

**Lemma 2.5.** Fix \( j \in [n] \) and \( S \subseteq [n] \) such that \( j = \omega(1) \) and \( |S| = o(j^{2/3}) \). Then, there exists a constant \( C = C(\varepsilon) > 0 \) such that
\[ \mathbb{P}(Q^c_{j} \mid \forall i \in S : \nu_{i,i} = 1) \leq e^{-Cj^{1/3}}. \]

**Proof.** The proof consists of two parts. For the lower bound, note that the process \((|V_i|)_{i \geq 0}\) dominates a random walk \((S_i)_{i \geq 0}\) on \( \mathbb{Z} \) that makes a step +1 with probability \( p \) and −1 with probability \( 1 - p \). On the other hand, the variable \( \hat{S}_i := (S_i + i)/2 \) has Binomial distribution with parameters \( i \) and \( p \). Using Chernoff’s inequality for the binomial random variable \( \hat{S}_j \), we conclude that
\[ \mathbb{P} \left( |V_j| - 2\varepsilon j \leq j^{2/3} \mid \forall i \in S : \nu_{i,i} = 1 \right) \leq \mathbb{P} \left( |V_j| - 2\varepsilon j \leq j^{2/3} \right) \]
\[ \leq \mathbb{P} \left( \hat{S}_j - (1/2 + \varepsilon)j \leq j^{2/3}/2 \right) = e^{-\Omega(j^{1/3})}. \]

For the upper bound, for all \( i \geq 0 \) we define \( Z_i \) as the number of steps \( r \in [i] \) at which \( S_r = \min_{t \in [r]} S_t \); that is, the last step of the random walk is a minimum of the walk up to now. We now show that conditionally on the event \( \{\forall i \in S : \nu_{i,i} = 1\} \), \( |V_j| \) is dominated by the process \( \hat{M}_j := \sum_{i=1}^{j} Z_j - |S| \). On the one hand, \((|V_i|)_{i \geq 0}\) and \((S_i + Z_i)_{i \geq 0}\) have the same distribution (which is the one of a discrete biased random walk conditioned to stay non-negative). On the other hand, \( |V_j| \) conditioned on births at the final \( |S| \) steps stochastically dominates \( |V_j| \) conditioned on \( \{\forall i \in S : \nu_{i,i} = 1\} \) (i.e. births at steps \( i \) in \( S \)) for any other choice of \( S \subseteq [j] \). Indeed, note that by introducing \( |S| \) vertices at the last \( |S| \) steps we ensure that none of these vertices is removed until step \( j \).

At the same time, the distribution of \( \max_{i \geq 0} Z_i \) is dominated by a geometric random variable since at every step \( i \geq 0 \), the event \( \{\forall t > i, S_t - S_i > 0\} \) has strictly positive probability depending only on \( \varepsilon \), see e.g. Example 2.2 in [19]. Therefore, together with Chernoff’s inequality, we get
\[ \mathbb{P} \left( |V_j| - 2\varepsilon j \geq j^{2/3} \mid \forall i \in S, \nu_{i,i} = 1 \right) \leq \mathbb{P} \left( S_j - |S| + Z_j - |S| + |S| - 2\varepsilon j \geq j^{2/3} \right) \]
\[ \leq \mathbb{P} \left( S_j - |S| - 2\varepsilon j \geq j^{2/3}/3 \right) + \mathbb{P} \left( Z_j - |S| \geq j^{2/3}/3 \right) \]
\[ \leq e^{-\Omega(j^{1/3})} + e^{-\Omega(j^{2/3})} = e^{-\Omega(j^{1/3})}, \]
which finishes the proof. \( \square \)

Now, for \( i \in [n] \) and a set \( S \subseteq [i] \), define the event \( W_i(S) = \{ S \subseteq V_i \} \), that is, at each of the steps in \( S \), a vertex was born, and each of these vertices survives until step \( i \). Similarly to Lemma 2.5, the following lemma allows us to control the number of vertices alive after \( n \) steps conditionally on the introduction of vertices at certain steps.

**Lemma 2.6.** Fix \( j = \omega((\log n)^3) \) such that \( j < n \), \( S \subseteq [n] \) such that \( |S| = o(j^{2/3}) \) and \( R \subseteq [j] \) such that \( R \cap S = \emptyset \) and \( |R| = o \left( \left( \log n \right)^2 \wedge \frac{j^{1/3}}{\log n} \right) \). Then,
\[ \mathbb{P} \left( W_n(R) \mid W_n(S) \cap W_j(R) \right) = \left( \frac{j}{n} \right)^{|R|(1-p)/(2\varepsilon)} \left( 1 + \mathcal{O} \left( \frac{|R| \log n}{j^{1/3}} \right) \right). \]

The lemma implies that, in particular, as long as the set \( S \) of vertices, conditioned to survive after \( n \) steps, is not “too large”, the probability that a vertex is born at step \( j \) and survives after step \( n \) remains the same as in the unconditional setting up to lower order terms.
Proof of Lemma 2.6. First, note that

$$\Pr( W_n(R) \mid W_n(S) \cap W_j(R)) = E \left[ \prod_{\ell \in [j+1,n] \setminus S} \left( 1 - \frac{|R(1-p)}{|V_{\ell-1} \setminus S|} \right) \right]. \tag{7}$$

Indeed, at each step $\ell$ between $j+1$ and $n$ apart from the steps in $S$ (at which we know that no deletion takes place), a vertex in $R$ is deleted with probability $\frac{|R(1-p)}{|V_{\ell-1} \setminus S|}$. Define the event

$$C_n := \bigcap_{\ell \in [j+1,n] \setminus S} \mathcal{Q}_\ell.$$ 

Using that $|V_{\ell-1} \setminus S| \leq |V_{\ell-1}|$ and $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$, we obtain

$$E \left[ \prod_{\ell \in [j+1,n] \setminus S} \left( 1 - \frac{|R(1-p)}{|V_{\ell-1} \setminus S|} \right) \right] \leq E \left[ \mathbb{1}_{C_n} \prod_{\ell \in [j+1,n] \setminus S} \left( 1 - \frac{|R(1-p)}{|V_{\ell-1} \setminus S|} \right) \right] + \Pr(C_n^c)$$

$$\leq \prod_{\ell \in [j+1,n] \setminus S} \exp \left( -\frac{|R(1-p)}{2\varepsilon \ell + \ell^2/3} \right) + \Pr(C_n^c). \tag{8}$$

Moreover, Lemma 2.5 implies that

$$\Pr(C_n^c) \leq \sum_{\ell=j+1}^n \exp \left( -\Omega(\ell^{1/3}) \right) \leq n \exp \left( -\Omega(j^{1/3}) \right) = o\left( n^{-|R(1-p)/(2\varepsilon)} \right), \tag{9}$$

where the third inequality follows from a union bound and Lemma 2.5.

Let us focus on the first term in (8). Transforming the above product into a sum yields

$$\exp \left( -\frac{|R(1-p)}{2\varepsilon} \sum_{\ell \in [j+1,n] \setminus S} \frac{1}{\ell} \frac{1}{1 + (2\varepsilon \ell^{1/3})^{-1}} \right) \leq \exp \left( -\frac{|R(1-p)}{2\varepsilon} \left( 1 - \frac{1}{2\varepsilon j^{1/3}} \right) \sum_{\ell \in [j+1,n] \setminus S} \frac{1}{\ell} \right), \tag{10}$$

where we used that $\frac{1}{1 + (2\varepsilon \ell^{1/3})^{-1}} \geq 1 - \frac{1}{2\varepsilon j^{1/3}}$ for all $\ell \geq j+1$.

Let us set $S_{>j} := \{s \in S : s > j\}$ and write $S_{>j} := \{s_1, \ldots, s_{|S_{>j}|}\}$ such that $s_0 := j < s_1 < \ldots < s_{|S_{>j}|} < s_{|S_{>j}|+1} := n+1$. We can write the sum on the right-hand side of (10) as

$$\sum_{m=1}^{|S_{>j}|+1} \sum_{\ell = s_{m-1}+1}^{s_m} \frac{1}{\ell} \geq \sum_{m=1}^{|S_{>j}|+1} \int_{s_{m-1}+1}^{s_m} \frac{1}{x} \, dx = \sum_{m=1}^{|S_{>j}|+1} \log \left( \frac{s_m}{s_{m-1} + 1} \right) = \log \left( \frac{n+1}{j+1} \right) + \sum_{m=1}^{|S_{>j}|} \log \left( 1 - \frac{1}{s_m + 1} \right).$$

Writing $\log(n+1/j) \geq \log(n/j) + \log(1 - 1/j+1)$ and using that $s_m \geq j$ for all $m \in [|S_{>j}|]$, we thus have

$$\sum_{\ell \in [j+1,n] \setminus S} \frac{1}{\ell} \geq \log \left( \frac{n}{j} \right) + (|S_{>j}| + 1) \log \left( 1 - \frac{1}{j+1} \right) \geq \log \left( \frac{n}{j} \right) - O \left( \frac{|S_{>j}| + 1}{j} \right).$$

Using this bound in (10) then yields the upper bound

$$\exp \left( -\frac{|R(1-p)}{2\varepsilon} \left( 1 - \frac{1}{2\varepsilon j^{1/3}} \right) \left( \log \left( \frac{n}{j} \right) - O \left( \frac{|S_{>j}| + 1}{j} \right) \right) \right)$$

$$= \left( \frac{j}{n} \right)^{|R(1-p)/(2\varepsilon)} \left( 1 + O \left( \frac{|R| \log n}{j^{1/3}} + \frac{|R||S_{>j}| + 1}{j} \right) \right), \tag{11}$$

$$= \left( \frac{j}{n} \right)^{|R(1-p)/(2\varepsilon)} \left( 1 + O \left( \frac{|R| \log n}{j^{1/3}} \right) \right).$$
The last step follows from the fact that \(|S_j| / j \leq |S| / j = o(j^{-1/3})\), which concludes the proof of the upper bound since \(P(C_n)\) is of smaller order than (11).

For a lower bound, we again use (7) to get

\[
P(W_n(R) | W_n(S) \cap W_j(R)) \geq E \left[ 1_{C_n} \prod_{\ell=j+1}^{n} \left( 1 - \frac{|R|(1-p)}{\max\{|V_{\ell-1}| - |S|, 1\}} \right) \right],
\]

which holds since including more terms smaller than one in the product yields a lower bound. On \(C_n\), it then follows that the right hand side of (12) is at least

\[
(1 - P(C_n^c)) \prod_{\ell=j+1}^{n} \left( 1 - \frac{|R|(1-p)}{\max\{2\varepsilon(\ell - 1) - (\ell - 1)^{2/3} - |S|, 1\}} \right).
\]

By the constraints on \(|S|\), the denominator is larger than 1 for any value of \(\ell \geq j + 1\) and for any sufficiently large \(n\), so that the maximum with one can be omitted. Moreover,

\[
2\varepsilon(\ell - 1) - (\ell - 1)^{2/3} - |S| = 2\varepsilon\ell - (1 + o(1))\ell^{2/3}.
\]

Using this together with the inequality \(1 - x \geq e^{-x - x^2}\) for all sufficiently small \(x\) yields the lower bound

\[
\exp\left( - \sum_{\ell=j+1}^{n} \frac{|R|(1-p)}{2\varepsilon\ell - (1 + o(1))\ell^{2/3}} + \left( \frac{|R|(1-p)}{2\varepsilon\ell - (1 + o(1))\ell^{2/3}} \right)^2 \right) (1 - P(C_n^c)).
\]

Since for every large enough \(n\) we have \(2\varepsilon\ell - (1 + o(1))\ell^{2/3} \geq \varepsilon\ell\), the exponential term can be bounded from below by

\[
\exp\left( - \frac{|R|(1-p)}{2\varepsilon} \sum_{\ell=j+1}^{n} \frac{1}{\ell} \frac{1}{1 - (1 + o(1))(2\varepsilon\ell^{1/3})^{-1}} - C_{p,\varepsilon} |R|^2 \sum_{\ell=j+1}^{n} \ell^{-2} \right) (1 - P(C_n^c))
\geq \exp\left( - \frac{|R|(1-p)}{2\varepsilon} \log \left( \frac{n}{j} \right) \left( 1 + \frac{1}{(\varepsilon j)^{1/3}} \right) - C_{p,\varepsilon} |R|^2 j^{-1/3} \right) (1 - P(C_n^c)),
\]

where the final step holds for all sufficiently large \(n\) and a sufficiently large constant \(C_{p,\varepsilon} > 0\). By the upper bound on the probability in (9), we finally arrive at the lower bound

\[
\left( \frac{j}{n} \right)^{|R|(1-p)/(2\varepsilon)} \left( 1 - \mathcal{O}\left( \frac{|R|\log n}{j^{1/3}} \right) \right) \left( 1 - \mathcal{O}(e^{-(j-1)^{1/3}/3(j-1)^{2/3}}) \right)
= \left( \frac{j}{n} \right)^{|R|(1-p)/(2\varepsilon)} \left( 1 - \mathcal{O}\left( \frac{|R|\log n}{j^{1/3}} \right) \right),
\]

which, combined with the upper bound in (11), yields the desired result. \(\square\)

**Lemma 2.7.** Fix \(i = i(n) = \omega((\log n)^2)\) and set \(\{v_1, \ldots, v_{|V_n|}\} = V_n\), where the vertices are listed in increasing order of their labels.

(i) \(P(v_1 \text{ is well defined and } v_1 \leq i) \leq \frac{i^{p/(2\varepsilon)}}{n^{1-p/(2\varepsilon)}}\);

(ii) Fix \(\ell = \lceil 4\varepsilon(\log n)^{p/\varepsilon} \rceil\). Then, there is a sequence \((\delta_i)_{i=1}^{\infty}\) satisfying \(\delta_n = o(1)\) such that w.h.p. for all \(j \in [\ell, |V_n|]\) one has

\[
v_j \in \left[ (1 - \delta_n)n^{(1-p)/p} \left( \frac{j}{2\varepsilon} \right)^{(2\varepsilon)/p}, (1 + \delta_n)n^{(1-p)/p} \left( \frac{j}{2\varepsilon} \right)^{(2\varepsilon)/p} \right).
\]
Proof. First, we conduct a first moment computation used in the proof of both Parts (i) and (ii). For every \( i \in \mathbb{N} \), set \( M_{i,n} := |V_i \cap V_n| \) and recall the event \( Q_i := \{||V_i| - 2\varepsilon i| \leq i^2/3\} \). Together with Lemma 2.6 applied with \( i \) instead of \( j \), \( S = \emptyset \) and \( R \) being a single vertex in \( V_i \), this implies that

\[
\mathbb{E}[M_{i,n}] = \sum_{j=1}^{i} \mathbb{P}(\nu_{j,n} = 1) = \sum_{j=1}^{i} \mathbb{P}(\nu_{j,i} = 1) \mathbb{P}(\nu_{j,n} = 1 | \nu_{j,i} = 1) = \mathbb{P}(\nu_{i,n} = 1 | \nu_{i,i} = 1) \mathbb{E}[V_i]
\]

where the third equality uses that the probability of \( \nu_{j,n} = 1 \) is independent of \( j \).

Now, we prove Part (ii) by a second moment argument. From now on, we assume that \( i = \omega(n^{1-p}/p) \) so that \( \mathbb{E}M_{i,n} = \omega(1) \). Similar reasoning as above implies

\[
\sum_{j_1,j_2 \in [i]; j_1 \neq j_2} \mathbb{P}(\nu_{j_1,n} = \nu_{j_2,n} = 1) = \sum_{j_1,j_2 \in [i]; j_1 \neq j_2} \mathbb{P}(\nu_{j_1,i} = \nu_{j_2,i} = 1) \mathbb{P}(\nu_{j_1,n} = \nu_{j_2,n} = 1 | \nu_{j_1,i} = \nu_{j_2,i} = 1).
\]

Using Lemma 2.6 applied with \( i \) instead of \( j \), \( S = \emptyset \) and \( R \) being a pair of vertices in \( V_i \) for the first term, for all \( j_1 \neq j_2 \) we get

\[
\mathbb{P}(\nu_{j_1,n} = \nu_{j_2,n} = 1 | \nu_{j_1,i} = \nu_{j_2,i} = 1) = \left(1 + O\left(\frac{\log n}{i^{1/3}}\right)\right) \frac{i^{1-p}/\varepsilon}{n^{1-p}/\varepsilon}.
\]

Moreover, by Lemma 2.5 applied with \( Q_i \) and \( S = \emptyset \),

\[
\sum_{j_1,j_2 \in [i]; j_1 \neq j_2} \mathbb{P}(\nu_{j_1,i} = \nu_{j_2,i} = 1) = \mathbb{E}[|V_i|^2] - \mathbb{E}[|V_i|] = \mathbb{E}[|V_i|^2 | Q_i] - \mathbb{E}[|V_i| | Q_i] + O\left(i^2 \mathbb{P}(Q_i)\right)
\]

\[
= \left(1 + O\left(\frac{1}{i^{1/3}}\right)\right) 4i^2 \varepsilon^2.
\]

We conclude that

\[
\mathbb{E}[M_{i,n}^2] - \mathbb{E}[M_{i,n}]^2 \leq \mathbb{E}[M_{i,n}] + \sum_{j_1,j_2 \in [i]; j_1 \neq j_2} \mathbb{P}(\nu_{j_1,n} = \nu_{j_2,n} = 1) - \left(1 + O\left(\frac{\log n}{i^{1/3}}\right)\right) \frac{2i^{1-p}/\varepsilon}{n^{1-p}/\varepsilon}
\]

\[
= \mathbb{E}[M_{i,n}] \left(1 + O\left(\frac{\log n}{i^{1/3}}\right)\right) \frac{2i^{1-p}/\varepsilon}{n^{1-p}/\varepsilon} + O\left(\frac{i^{p/\varepsilon - 1/3} \log n}{n^{1-p}/\varepsilon}\right)
\]

\[
= \mathbb{E}[M_{i,n}] \left(1 + O\left(\frac{\log n}{i^{1/3}}\right)\right) \frac{n^{(1-p)/(2\varepsilon)} - \varepsilon^2}{n^{1-p}/\varepsilon}
\]

where we used (13) in the last equality. Define \( \psi(i) := \frac{\log n}{i^{1/3}} + \frac{i^{1-p/(2\varepsilon)}}{np/(2\varepsilon)} \); then, Chebychev’s inequality implies

\[
\mathbb{P}(|M_{i,n} - \mathbb{E}[M_{i,n}]| \geq \psi(i)^{-1/3} \mathbb{E}[M_{i,n}]) \leq \frac{\mathbb{E}(|M_{i,n} - \mathbb{E}[M_{i,n}]|^2)}{(\psi(i)^{-1/3} \mathbb{E}[M_{i,n}])^2} = O\left(\psi(i)^{5/3}\right).
\]

Now, set \( \alpha = \alpha(n) = ((n^{1-p}/p) (\log n)^2)^{-1} n^{1/|\log n|^2} \), and for all integer \( k \) between 0 and \( |\log n|^2 \), set \( i_k = \lceil \alpha^{-2} n^{1-p}/p (\log n)^2 \rceil \). By (14) we have that for all \( k \) as above, \( |M_{i_k,n} - \mathbb{E}[M_{i_k,n}]| \geq \psi(i_k)^{-1/3} \mathbb{E}[M_{i_k,n}] \) holds with probability \( O\left(\psi(i_k)^{5/3}\right) \). Moreover, since \( i_k \) is increasing in \( k \) (for \( n \) sufficiently large), \( \phi(i) \) decreasing in \( i \) and \( 5p > 9\varepsilon \),

\[
\sum_{k=1}^{\lfloor \log n \rfloor} \psi(i_k)^{5/3} \leq \lfloor \log n \rfloor^2 \psi(i_0)^{5/3} = O\left((\log n)^{2-5p/(3\varepsilon)}\right) = o(1).
\]
Thus, a union bound over all \([\log n]^2 + 1\) values of \(k\) shows that the event

\[ \mathcal{G} := \{ \text{for all positive integers } k \leq \lfloor (\log n)^2 \rfloor, M_{ik,n} = (1 + o(1))\mathbb{E}[M_{ik,n}] \} \]

holds w.h.p.

Finally, note that \((M_{i,n})_{i=1}^n\) is an increasing sequence of random variables since for all \(i < j\) one has \(V_i \cap V_n \subseteq V_j \cap V_n\). Thus, conditionally on the event \(\mathcal{G}\), for all positive integers \(k \leq \lfloor (\log n)^2 \rfloor\) and \(i \in [i_{k-1}, i_k]\),

\[ 1 - o(1) = (1 - o(1))\frac{\mathbb{E}[M_{ik-1,n}]}{\mathbb{E}[M_{ik,n}]} \leq M_{ik-1,n} \leq M_{ik,n} \leq (1 + o(1))\frac{\mathbb{E}[M_{ik-1,n}]}{\mathbb{E}[M_{ik,n}]} = (1 + o(1)), \]

where the first and the last equalities follow from the fact that, for every positive integer \(k \leq \lfloor (\log n)^2 \rfloor\),

\[ \mathbb{E}[M_{ik-1,n}] = (1 + o(1))(i_{k-1}/i_k)^p/(2\epsilon) = (1 + o(1)). \]

Hence, w.h.p. \(M_{i,n} = (1 + o(1))\mathbb{E}[M_{i,n}]\) for all \(i \in [i_0, n]\). This allows us to conclude that, since

\[ v_j = \min\{ t \in [n] : |V_t \cap V_n| = j \}, \]

for all vertices but the first \((1 + o(1))\mathbb{E}M_{i_0,n} = (1 + o(1))2\epsilon(\log n)^p/\epsilon \leq \ell\) many, we have that w.h.p. the value of \(v_j\) is given, up to a \(1 + o(1)\) factor, by the solution of the equation

\[ \frac{2\epsilon v_j^{p/(2\epsilon)}}{n^{(1-p)/(2\epsilon)}} = j, \]

which yields \(v_j = n^{(1-p)/p}(j/2\epsilon)^{(2\epsilon)/p}\) and completes the proof of the lemma.

\[ \square \]

3 Theorem 1.3: setting up the framework

To start this section, we introduce some important notations. For any non-negative integer \(r\), any graph \(G\) and any vertex \(v\) in \(G\), we denote by \(\partial B_r(G, v)\) the set of vertices at graph distance exactly \(r\) from \(v\) in \(G\), that is, \(\partial B_r(G, v) = B_r(G, v) \setminus B_{r-1}(G, v)\). Note that \(v\) will often be the root in some rooted graph. Also, some graphs in this paper will be naturally defined as rooted, in which case we often omit the root from the notation. Recall that for a random variable \(X\), we denote by \(\mathcal{L}(X)\) the distribution of \(X\).

The exploration process. Recall the definition of the DRGVR model \(G_n\) and the Binomial birth-death tree \(T\) in Definitions 1.1 and 1.2, respectively. To prove the local weak convergence of \(G_n\) to \(T\), we couple the neighbourhood of the root 0 in \(T\) and the neighbourhood of a vertex \(k_0\) in \(G_n\) chosen uniformly at random. More precisely, we couple the breadth-first search (BFS) exploration of the neighbourhood of \(k_0\) in \(G_n\) with the iterative construction of \(T\), starting from the root vertex 0. In this section we provide some notation for this BFS exploration.

We start by introducing the Ulam-Harris tree, which we use to unify the notation that underpins the construction of the multi-type branching process \(T\) and the BFS exploration of the neighbourhood of \(k_0\) in \(G_n\). The Ulam-Harris tree is an infinite rooted tree constructed as follows: its root vertex is denoted by 0 and for all \(r \geq 0\), the children of any vertex \(u := (0, u_1, \ldots, u_r)\) (where \(u_1, \ldots, u_r \in \mathbb{N}\)) are given by \(((u, i))_{i \in \mathbb{N}}.\) In this way, the vertex \(u := (0, u_1, \ldots, u_r)\) is the \(u_r\)th child of the \(u_{r-1}\)th child of ... of the \(u_1\)th child of the root 0 in the breadth-first order.

Furthermore, we introduce an ordering \(\prec_{UH}\) on the vertices in the Ulam-Harris tree. For two vertices \(u := (0, u_1, \ldots, u_r)\) and \(v := (0, v_1, \ldots, v_k)\) (with \(r, k \in \mathbb{N}_0\) and \((u_i)_{i=1}^k \in \mathbb{N}^k, (v_i)_{i=1}^k \in \mathbb{N}^k\)), we write \(u \prec_{UH} v\) when \(u\) is smaller than \(v\) in BFS order. That is, when either \(r < k\) or \(r = k\) and \(u_j < v_j\), where \(j := \min\{ i \in [r] : u_i \neq v_i \}\). For example \((0, 1, 2, 3) \prec_{UH} (0, 1, 2, 3, 4)\) and \((0, 2, 5, 4) \prec_{UH} (0, 2, 6, 3)\). In a similar manner, we define the ordering \(\leq_{UH}\).

From now on, we see the tree \(T\) as a random sub-tree of the Ulam-Harris tree. More precisely, if a vertex \(u := (0, u_1, \ldots, u_r) \in T\) has \(\tau_u\) children, these are labeled \((u, i)_{i=1}^{\tau_u}\) in increasing order of their types.
Moreover, we slightly abuse notation and use the Ulam-Harris formalism for the BFS exploration of the neighbourhood of $k_0$ in $G_n$. This provides more structure when keeping track of the BFS exploration and the parallel iterative construction of $\mathcal{T}$. For a vertex $v = (0, v_1, \ldots, v_r)$ for some $r \geq 0$ and $v_1, \ldots, v_r \in \mathbb{N}$ in the sub-tree of the Ulam-Harris tree obtained from the BFS exploration of $k_0$ in $G_n$, we let $\kappa_v \in [n]$ denote its label in $G_n$. In what follows, we identify the vertex $v$ in the BFS exploration of $G_n$ by its label $\kappa_v$. To illustrate: 0 has label $k_0$, and this is the root of the BFS exploration. The neighbours of 0 are $((0,i))_{i=1}^{\theta_0}$, where $\theta_0$ denotes the number of neighbours of $k_0$ in $G_n$, and their labels are $(k_{(0,i)})_{i=1}^{\theta_0}$. This is continued throughout the BFS exploration. For convenience of notation we skip the parentheses and write $k_{0,v_1,\ldots,v_r}$ for the label of the vertex $(0, v_1, \ldots, v_r)$.

In the BFS exploration, we consider nodes to be active, probed or neutral and denote by $(A_t, \mathcal{P}_t, N_t)_{t \in \mathbb{N}_0}$ the sets of active, probed and neutral vertices after $t$ steps of the exploration, respectively. We initialise the process by setting

$$(A_0, \mathcal{P}_0, N_0) = ((k_0), \emptyset, V_n \backslash \{k_0\}).$$

We also let $k[t]$ denote the active vertex with smallest BFS order in $A_{t-1}$. That is, $k[1] = k_0$ and for $t \geq 1$, if $\kappa_v \in A_{t-1}$ and $v <_{\text{UF}} u$ for all $u \in A_{t-1} \setminus \{v\}$, then $k[t] = \kappa_v$. Let $\mathcal{D}_t := \{u \in N_{t-1} : \{u, k[t]\} \in E(G_n)\}$ denote the set of neutral vertices attached to $k[t]$. Then, we update

$$(A_{t+1}, \mathcal{P}_{t+1}, N_{t+1}) = ((A_t \backslash \{k[t]\}) \cup \mathcal{D}_t, \mathcal{P}_t \cup \{k[t]\}, N_t \setminus \mathcal{D}_t),$$

and if $A_t = \emptyset$, we set $(A_{t+1}, \mathcal{P}_{t+1}, N_{t+1}) = (A_t, \mathcal{P}_t, N_t)$.

Due to the definition of the DRGVR model it is convenient to split the vertices in the BFS exploration into two parts. Indeed, for a vertex with label $\kappa_v$, the probability to connect to a vertex with a smaller label is different than to be connected to a vertex with a larger label. As a result, for each vertex $\kappa_v$ we consider its neighbours to the left (i.e. vertices with a smaller label than $\kappa_v$) and to the right (i.e. vertices with a larger label than $\kappa_v$). We refer to these as the $L$-neighbours and $R$-neighbours of $\kappa_v$, respectively. Furthermore, we let $\theta^L_v$ and $\theta^R_v$ denote the number of $L$-neighbours and $R$-neighbours of $\kappa_v$, respectively, and set $\theta_v := \theta^L_v + \theta^R_v$ to be the number of neighbours of $v$ in the BFS exploration.

In an equivalent manner, we can define the $L$-children and $R$-children of a vertex $v$ in the multi-type branching process $\mathcal{T}$ as in Definition 1.2. If we let $\tau_v$ denote the number of children of $v$, then the $L$-children (respectively $R$-children) of $v$ are the children $(v,i)_{i=1}^{\tau_v}$ such that $a_{v,i} < a_v$ (respectively $a_{v,i} > a_v$). Let us write $\tau^L_v$ and $\tau^R_v$ for the number of $L$-children and $R$-children of $v$, respectively. As we order the children in increasing order of their types, it follows that $(v,i)_{i=1}^{\tau^L_v}$ are $v$'s $L$-children and $(v,i)_{i=\tau^L_v+1}^{\tau^L_v+\tau^R_v}$ are $v$'s $R$-children.

4 Coupling the 1-neighbourhoods of $k_0$ in $G_n$ and 0 in $\mathcal{T}$

In this section we provide a coupling between the neighbourhoods of $k_0$ in the BFS exploration of $G_n$ and the children of the root 0 of $\mathcal{T}$. The coupling should be such that:

(i) the number of $L$- and $R$-neighbours of $k_0$ in $G_n$ are equal to the number of $L$- and $R$-children of 0,

(ii) the rescaled labels of the neighbours of $k_0$ and the types of the children of 0 are approximately the same.

Whilst (ii) is not directly necessary for the 1-neighbourhoods of $k_0$ and 0 to be isomorphic, it aids in providing a coupling such that the 2-neighbourhoods are isomorphic.

In this section, we prove that the following events occur with high probability:

$$\begin{align*}
H_{1,0} &:= \{a_0 \geq 1 / \log \log n\}, \\
H_{1,1} &:= \{\min_{i \in \tau^L_v} a_{0,i} > 1/(\log \log n)^2\}, \\
H_{1,2} &:= \{(B_1(G_n, k_0), k_0) \cong (B_1(\mathcal{T},0),0)\} \cap \{\forall v \in V(B_1(\mathcal{T},0)), |a_v - \frac{k_0}{n}| \leq \frac{2}{n}\}, \\
H_{1,3} &:= \{\tau^L_0 + \tau^R_0 < \log \log n\}.
\end{align*}$$

15
In words, the events $H_{1,0}$ and $H_{1,1}$ control the types of the root 0 and the $L$-children of the root 0 in $T$, and ensure that these types are not “too small”. Similarly, the event $H_{1,3}$ ensures that the number of children of the root 0 in $T$ is not “too large”. The event $H_{1,2}$ then states that the 1-neighbourhoods of $k_0$ in $G_n$ and 0 in $T$ are isomorphic and that the rescaled labels of the neighbours of $k_0$ in $G_n$ are very close to the types of their corresponding counterparts in $T$. As mentioned, the events $H_{1,0}, H_{1,1}, H_{1,3}$ and the part of the event $H_{1,2}$ that states that the labels $k_v$ are “sufficiently close” to the types $a_v$ are not strictly necessary for the local convergence of the 1-neighbourhoods to hold. However, they are important to control error bounds, which are key to extend the coupling to the $r$-neighbourhood later on (for any $r \geq 2$). More details related to this follow in the upcoming sections.

We now state the main result of this section.

**Lemma 4.1.** Fix $\beta > 0$ and $\varepsilon \in (0, 1/2]$, and consider the DRGVR model and the Binomial birth-death tree model as in Definitions 1.1 and 1.2, respectively. Recall the events $(H_{1,i})_{i=0}^4$ from (15). There exists a constant $C = C(\beta, \varepsilon) > 0$ such that, for all sufficiently large $n$,

$$
\mathbb{P}\left(\left(\bigcap_{i=0}^3 H_{1,i}\right)^c\right) \leq \frac{C}{(\log \log n)^{p/(2\varepsilon)}}.
$$

Note that we can bound the probability in the statement of Lemma 4.1 from above by

$$
\mathbb{P}\left(\left(\bigcap_{i=0}^3 H_{1,i}\right)^c\right) \leq \mathbb{P}(H_{1,0}^c) + \mathbb{P}(H_{1,1}^c \cap H_{1,0}) + \mathbb{P}(H_{1,3}^c) + \mathbb{P}(H_{1,2}^c \cap H_{1,0}).
$$

(16)

The first term on the right-hand side is readily bounded: since $a_0 \sim \text{Beta}(\frac{p}{2\varepsilon}, 1)$, it directly follows that

$$
\mathbb{P}(H_{1,0}^c) = (\log \log n)^{-p/(2\varepsilon)}.
$$

(17)

Let us write $V_0^L$ for the Poisson point process on $(0, a_0)$ with density $\lambda_0^-(dx)$ as in (1). Using the construction of the offspring of the root as in Definition 1.1, we can write

$$
H_{1,1} = \{V_0^L \cap (0, 1/(\log \log n)^2) = \emptyset\}.
$$

As a result, conditionally on $a_0$,

$$
\mathbb{P}(H_{1,1}^c \cap H_{1,0}) = \mathbb{P}(V_0^L \cap (0, 1/(\log \log n)^2) \neq \emptyset) \mid H_{1,0}) \mathbb{P}(H_{1,0})
$$

(18)

Conditionally on $a_0$ and the event $H_{1,0}$, and using that $1 - e^{-x} \leq x$ for all $x \in \mathbb{R}$, we deduce that (18) is bounded from above by

$$
1 - \exp\left(-\frac{\beta}{a_0}(\log \log n)^{-p/\varepsilon}\right) \leq 1 - \exp\left(-\beta(\log \log n)^{1-p/\varepsilon}\right) \leq \beta(\log \log n)^{-1/(2\varepsilon)},
$$

so that taking the expected value with respect to $a_0$ yields

$$
\mathbb{P}(H_{1,1}^c \cap H_{1,0}) = O\left((\log \log n)^{-1/(2\varepsilon)}\right).
$$

(19)

Furthermore, conditionally on $a_0$, $\tau_0^L + \tau_0^R$ is Poisson distributed with rate

$$
\int_0^{a_0} \lambda_0^-(dx) + \int_{a_0}^1 \lambda^+(dx) = \frac{\beta p}{2\varepsilon} \int_0^{a_0} \frac{x^{(1-p)/(2\varepsilon)}}{a_0} dx + \int_{a_0}^1 x^{(1-p)/(2\varepsilon)-1} dx
$$

$$
= \beta a_0^{(1-p)/(2\varepsilon)} + \frac{\beta p}{1-p} \left(1 - a_0^{(1-p)/(2\varepsilon)}\right) \leq \frac{\beta p}{1-p}.
$$

(20)

As a result, $\tau_0^L + \tau_0^R$ is stochastically dominated by $\tau' \sim \text{Poi}(\frac{\beta p}{1-p})$. It follows that for any $t > 0$,

$$
\mathbb{P}(H_{1,3}^c) \leq \mathbb{P}(\tau' \geq \log \log n) \leq \frac{\mathbb{E}[e^{\tau'}]}{e^{t \log \log n}} = \frac{e^{(\beta p/(1-p))(\varepsilon' - 1)}}{(\log n)^t} = o((\log \log n)^{-p/(2\varepsilon)}),
$$

(21)

where the second inequality follows from a Chernoff bound. The main goal in this section is thus to construct a coupling such that the probability of $H_{1,2}^c \cap H_{1,0}$ is $O((\log \log n)^{-p/(2\varepsilon)})$. 

15
4.1 Coupling the roots of \( G_n \) and \( T \)

We start by coupling the root vertex 0 of \( T \) and the uniform vertex \( k_0 \) from \( V_n \), which in this model is non-trivial. Note that in the proofs of the local weak convergence of preferential attachment models to the Pólya point tree, which are similar in nature, the vertex \( k_0 \) is uniform among \([n]\) and the root of the tree has a type \( a_0 \sim \text{Unif}(0,1) \). Hence, a simple coupling via \( k_0 := \lfloor a_0 n \rfloor \) provides that \( k_0/n \) and \( a_0 \) are sufficiently close. In our case, since \( V_n \) is a random subset of \([n]\), this coupling requires more work.

Let us define
\[
\mathcal{R}_0 := \{a_0 \leq \frac{k_0}{n} \leq a_0 + \frac{1}{n}\}.
\]
We show the following result.

**Lemma 4.2.** Consider the DRGVR model and the Binomial birth-death tree model as in Definitions 1.1 and 1.2, respectively. Let \( k_0 \) be a uniform vertex from \( V_n \). Then, there exists a constant \( c = c(\beta, \varepsilon) > 0 \) and a coupling of \( k_0 \) and the root vertex 0 of \( T \) such that
\[
\mathbb{P}(\mathcal{R}_0^c) \leq \frac{c}{\log n}.
\]

**Proof.** We sample an i.i.d. sequence \((X_j^0, Y_j^0)_{j \in \mathbb{N}}\) with distribution \( \Gamma \) (defined in (6)). Let \( \kappa_0 = \inf\{j \in \mathbb{N} : Y_j^0 = 1\} \) and set \( a_0 := X_\kappa^0 \) as the type of 0. By Lemma 2.1, \( a_0 \) has the desired distribution. We define \( k_0^\Psi := \lfloor nX_j^0 \rfloor \) as a uniform vertex in \([n]\), and write its mark as \( \nu_j^0 := \nu_{k_0^\Psi, n} \) for brevity. We also introduce \( \psi_0 := \inf\{j \in \mathbb{N} : \nu_j^0 = 1\} \) and set \( k_0 := k_0^{\psi_0} \). Note that since the mark of \( k_0 \) equals one, \( k_0 \) is uniform among all vertices with mark one, so \( k_0 \) also has the desired distribution. Our aim is to couple \( \kappa_0 \) and \( \psi_0 \) so that \( \kappa_0 = \psi_0 \) with probability at least \( 1 - c/\log n \). Indeed, conditionally on the equality \( \kappa_0 = \psi_0 \), \( a_0 \leq \frac{k_0}{n} \leq a_0 + \frac{1}{n} \) holds a.s.

We couple \( \nu_j^0 \) with \( Y_j^0 \sim \text{Ber}(p(X_j^0)^{(1-p)/(2\varepsilon)}) \) using a standard Bernoulli coupling, which conditionally on \( X_j^0 \) fails with probability
\[
|\mathbb{P}(\nu_j^0 = 1 | X_j^0) - \mathbb{P}(Y_j^0 = 1 | X_j^0)|. \tag{23}
\]

We say the coupling is successful at step \( j \) when \( \nu_j^0 = Y_j^0 \), and that the entire coupling is successful when \( \kappa_0 = \psi_0 \), i.e. \( \nu_j^0 = Y_j^0 \) for all \( j \in [\kappa_0] \). Hence,
\[
\mathbb{P} \text{ (Coupling fails)} = \mathbb{P}(\kappa_0 \neq \psi_0) = \mathbb{E} \left[ \mathbb{P} \left( \bigcup_{j=1}^{\kappa_0} \{ \text{Coupling fails at step } j \} \mid \kappa_0 \right) \right].
\]

Since we have coupled \( \nu_j^0 \) and \( Y_j^0 \) at every step with a probability of failure as in (23), it follows that independently of the value of \( Y_j^0 \) (which is known for each \( j \in [\kappa_0] \) when we condition on \( \kappa_0 \)), the probability of the event \( \{\nu_j^0 \neq Y_j^0\} \) can be bounded from above by (23). Using a union bound and observing that the expression in (23) is in fact independent of \( j \), we thus obtain the upper bound
\[
\mathbb{E} \left[ \kappa_0 \mathbb{E} \left[ |\mathbb{P}(\nu_j^1 = 1 | X_j^0) - \mathbb{P}(Y_j^1 = 1 | X_j^0)| \right] \right],
\]
where the inner expected value is with respect to \( X_j^1 \) and the outer expected value with respect to \( \kappa_0 \). Conditionally on \( X_j^1 \), it follows from Lemma 2.6 and Lemma 2.1 that
\[
\mathbb{P}(\nu_j^1 = 1 | X_j^0) = p\left(\frac{\lfloor nX_j^1 \rfloor}{n}\right)^{(1-p)/(2\varepsilon)} \left( 1 + \mathcal{O}\left(\log n \frac{1}{\lfloor nX_j^1 \rfloor^{1/3}}\right)\right) = \mathbb{P}(Y_j^1 = 1 | X_j^0) \left( 1 + \mathcal{O}\left(\log n \frac{1}{\lfloor nX_j^1 \rfloor^{1/3}}\right)\right).
\]

On the event \( \{X_j^1 \geq \frac{1}{\log n}\} \), this yields
\[
\mathbb{P}(Y_j^1 = 1 | X_j^0) + \mathcal{O}\left(\frac{(\log n)^{4/3}}{n^{1/3}}\right).
\]
Hence, for some positive constants $C, c > 0$,
\[
\mathbb{P} (\text{Coupling fails}) \leq \mathbb{P} (\text{Coupling fails} \mid X_0^1 \geq \frac{1}{\log n}) + \mathbb{P} (X_0^1 < \frac{1}{\log n}) \leq C \left( \frac{\log n}{n^{1/3}} \right)^4 \mathbb{E} [\kappa_0] + \frac{1}{\log n} \leq \frac{c}{\log n},
\]
where we used that $\mathbb{E} [\kappa_0] < \infty$ since $\kappa_0$ has geometric distribution with parameter $2\varepsilon$. This concludes the proof.

Having coupled the roots, we can start coupling their direct neighbours. Recall that $\theta_0^L$ and $\theta_0^R$ denote the number of $L$-neighbours and $R$-neighbours of $k_0$, respectively. The aim is to ensure that $\theta_0^L = \tau_0^L$, $\theta_0^R = \tau_0^R$, and that for any neighbour $((0, i))_{i \in \{\tau_0^L, \tau_0^R\}}$, it holds that $\kappa_{0,i} \approx a_{0,i}$. Observe that the latter is not required for the local weak convergence of the $1$-neighbourhood to hold but is needed to ensure that the $2$-neighbourhoods of the root $0$ in $T$ and $k_0$ in $G_n$ can be coupled.

4.2 The $1$-neighbourhood of the root

The children of the root $0$ in $T$ are distributed according to a Poisson point process, whilst the direct neighbours of the root $k_0$ in $G_n$ are characterised by a Bernoulli point process (that is, a sequence of Bernoulli random variables). To provide a coupling between these characterisations, we split the neighbours of $k_0$ into two groups: the $L$-neighbours are those $\kappa_{0,i}$, and that for any neighbour $((0, i))_{i \in \{\tau_0^L, \tau_0^R\}}$, it holds that $\kappa_{0,i} \approx a_{0,i}$. Observe that the latter is not required for the local weak convergence of the $1$-neighbourhood to hold but is needed to ensure that the $2$-neighbourhoods of the root $0$ in $T$ and $k_0$ in $G_n$ can be coupled.

4.2.1 $L$-children of the root $0$ and $L$-neighbours of $k_0$ in $G_n$

The $L$-neighbours of the root $k_0$ of $G_n$ can be encoded by a Bernoulli point process $I_{k_0}^L = (I_{k_0,j})_{j=1}^{k_0-1}$, where $I_{k_0,j}$ equals one if $k_0$ connects to $j$ in $G_n$. Similarly, the $L$-children of the root $0$ of $T$ can be encoded by a Poisson point process $V_{k_0}^L$. We couple the two processes by discretising $V_{k_0}^L$ as follows: for all $j \in [k_0 - 1]$, define $V_{k_0,j} \sim \text{Poi}(\lambda_j)$ where
\[
\lambda_j := \int_{k_0 - 2}^{k_0} \frac{\beta p}{2 \varepsilon a_0} x^{(1-p)/(2\varepsilon)} \, dx \quad \text{if} \quad j \in [k_0 - 2], \quad \text{and} \quad \lambda_{k_0-1} := \int_{k_0-2}^{a_0} \frac{\beta p}{2 \varepsilon a_0} x^{(1-p)/(2\varepsilon)} \, dx. \quad (24)
\]
Then, we write $V_{k_0}^L = (V_{k_0,j})_{j=1}^{k_0-1}$.

The following lemma states that $I_{k_0}^L$ and $V_{k_0}^L$ can be coupled successfully on the event $\mathcal{H}_{1,0} \cap \mathcal{R}_0$.

**Lemma 4.3.** Consider the Bernoulli point process $I_{k_0}^L = (I_{k_0,j})_{j=1}^{k_0-1}$ and the discretised Poisson point process $V_{k_0}^L := (V_{k_0,j})_{j=1}^{k_0-1}$ with means defined in (24). Recall the events $\mathcal{H}_{1,0}$ and $\mathcal{R}_0$ from (15) and (22), respectively, and fix $\mathcal{E}_0 := \mathcal{R}_0 \cap \mathcal{H}_{1,0}$ and $\alpha := \frac{1-p}{p^2}$. Then, there exists a coupling of $I_{k_0}^L$ and $V_{k_0}^L$ such that
\[
\mathbb{P} \left( \{ I_{k_0}^L \neq V_{k_0}^L \} \cap \mathcal{E}_0 \right) = O \left( \frac{\log n}{n^{\alpha/6}} \right).
\]

**Remark 4.4.** When the coupling of the Bernoulli process and the (discretised) Poisson point process is successful, we immediately obtain that $\tau_0^L = \theta_0^L$, that is, the number of $L$-neighbours of the root $0$ in $G_n$ and the number of $L$-children of the root $0$ in $T$ are the same. Moreover, for each $L$-child $(0, i)$ in $T$ (with $i \in [\tau_0^L]$) we have $|a_{0,i} - \kappa_{0,i}| \leq \frac{1}{n}$.}

**Proof of Lemma 4.3.** The coupling consists of several intermediate steps, which we outline first.

1. We couple the Bernoulli process $I_{k_0}^L$ with another Bernoulli process $\overline{I}_{k_0}^L := (\overline{I}_{k_0,j})_{j=1}^{k_0-1}$ of independent variables where $\overline{I}_{k_0,j} = 0$ for all $j \leq n^{\alpha/2}$, and for all $j \in (n^{\alpha/2}, k_0 - 1)$,
\[
\mathbb{P} \left( \overline{I}_{k_0,j} = 1 \mid a_0, k_0 \right) = \lambda_j := \frac{\beta p}{2 \varepsilon k_0} \left( \frac{j}{n} \right)^{(1-p)/(2\varepsilon)}.
\]
2. We couple $\tilde{I}_{k_0}^L$ with a sequence $\tilde{V}_{k_0}^L := (\tilde{V}_{k_0,j})_{j=1}^{k_0-1}$ of independent variables where $\tilde{V}_{k_0,j} = 0$ for all $j \leq n^{\alpha/2}$, and for all $j \in (n^{\alpha/2}, k_0 - 1]$, $\tilde{V}_{k_0,j}$ is a Poisson random variable with mean $\lambda_j$.

3. We couple $\tilde{V}_{k_0}^L$ with $V_{k_0}^L$.

We now provide details for each of the steps. Throughout the proof we use that on the event $E_0$ we have that $k_0 \geq n/\log \log n$.

**Step 1.** Recall $v_1$ from Lemma 2.7. It follows that

$$
P \left( \forall j \leq n^{\alpha/2} : I_{k_0 \rightarrow j} = 0 \mid a_0, k_0 \right) = 1 - P \left( v_1 \leq n^{\alpha/2} \right) E \left[ \min \left\{ \frac{\beta}{|V_{k_0-1}|}, 1 \right\} \mid a_0, k_0 \right].$$

(25)

Using that by Lemma 2.7 (i) one has $P \left( v_1 \leq n^{\alpha/2} \right) = O(n^{-\alpha p/(4\varepsilon)})$, and by Lemma 2.5

$$
E \left[ \min \left\{ \frac{\beta}{|V_{k_0-1}|}, 1 \right\} \mid a_0, k_0 \right] \leq E \left[ 1_{Q_{k_0-1}} \min \left\{ \frac{\beta}{|V_{k_0-1}|}, 1 \right\} \mid a_0, k_0 \right] + P \left( Q_{k_0-1} \right) = (1 + o(1)) \frac{\beta}{2\varepsilon k_0},
$$

we get that

$$
P \left( \forall j \leq n^{\alpha/2} : I_{k_0 \rightarrow j} = 0 \mid a_0, k_0 \right) = 1 - O(n^{-\alpha p/(4\varepsilon)k_0^{-1}}).
$$

It thus remains to show that $I_{k_0 \rightarrow j} = \tilde{I}_{k_0 \rightarrow j}$ for all $j \in (n^{\alpha/2}, k_0 - 1]$ with high probability. By Lemmas 2.5, 2.6 and on the event $E_0$ we have that for all $j > n^{\alpha/2}$,

$$
P \left( I_{k_0 \rightarrow j} = 1 \mid a_0, k_0 \right) = P \left( v_{j,n} = 1 \mid a_0, k_0 \right) E \left[ \min \left\{ \frac{\beta}{|V_{k_0-1}|}, 1 \right\} \mid a_0, k_0 \right]
$$

$$
= p \left( \frac{j}{n} \right) (1-p)^{(1-p)/(2\varepsilon)} \left( 1 + O \left( \frac{\log n}{n^{\alpha/6}} \right) \right) \frac{\beta}{2\varepsilon k_0} \left( 1 + O \left( k_0^{-1/3} \right) \right)
$$

$$
= \frac{\beta p}{2\varepsilon k_0} \left( \frac{j}{n} \right) (1-p)^{(1-p)/(2\varepsilon)} + O \left( \frac{\log n}{n^{\alpha/6}k_0} + \frac{1}{k_0^{3/4}} \left( \frac{j}{n} \right)^{(1-p)/(2\varepsilon)} \right)
$$

$$
= \frac{\beta p}{2\varepsilon k_0} \left( \frac{j}{n} \right) (1-p)^{(1-p)/(2\varepsilon)} + O \left( \frac{\log n}{n^{\alpha/6}k_0} \left( \frac{j}{n} \right)^{(1-p)/(2\varepsilon)} \right),
$$

where the last equality follows from the fact that $k_0 \geq n^{\alpha/2}$. As a result, using a standard Bernoulli coupling yields

$$
P \left( I_{k_0}^L \neq \tilde{I}_{k_0}^L \mid a_0, k_0 \right) \leq O(n^{-\alpha p/(4\varepsilon)k_0^{-1}}) + \sum_{j=[n^{\alpha/2}]}^{k_0-1} P \left( I_{k_0 \rightarrow j} \neq \tilde{I}_{k_0 \rightarrow j} \mid a_0, k_0 \right)
$$

$$
\leq O \left( \frac{\log n}{n^{\alpha/6}k_0} \right) + \sum_{j=[n^{\alpha/2}]}^{k_0-1} O \left( \left( \frac{j}{n} \right) (1-p)^{(1-p)/(2\varepsilon)} \right)
$$

$$
= O \left( n^{-\alpha p/(4\varepsilon)k_0^{-1}} \right) + O \left( \frac{\log n}{n^{\alpha/6}} \right) = O \left( \frac{\log n}{n^{\alpha/6}} \right),
$$

where for the last equality we used that $k_0 \geq n^{\alpha/6} \log n$. This finally yields that

$$
P \left( \left\{ I_{k_0}^L \neq \tilde{I}_{k_0}^L \right\} \cap E_0 \right) = E \left[ 1_{E_0} P \left( I_{k_0}^L \neq \tilde{I}_{k_0}^L \mid a_0, k_0 \right) \right] = O \left( \frac{\log n}{n^{\alpha/6}} \right).
$$

**Step 2.** Using Lemma 2.2 directly yields

$$
1_{E_0} P \left( I_{k_0}^L \neq \tilde{V}_{k_0}^L \mid a_0, k_0 \right) \leq 1_{E_0} \sum_{j=1}^{k_0-1} 2 \left( \frac{\beta p}{2\varepsilon k_0} \left( \frac{j}{n} \right) (1-p)^{(1-p)/(2\varepsilon)} \right)^2 \leq 1_{E_0} \frac{\beta^2 p^2}{2\varepsilon^2 k_0^2} \sum_{j=1}^{k_0-1} \left( \frac{j}{k_0} \right) (1-p)^{(1-p)/\varepsilon} = O \left( \frac{\log n}{n} \right).
$$
Step 3. Via a standard Poisson coupling (see Lemma 2.3),

\[
\mathbb{P} \left( \tilde{V}_{k_0}^L \neq V_{k_0}^L \mid a_0, k_0 \right) \leq \sum_{j=0}^{[n^{\alpha/2}]} \mathbb{P} \left( V_{k_0} \to j \geq 1 \mid a_0, k_0 \right) + \sum_{j=[n^{\alpha/2}]}^{k_0-1} \mathbb{P} \left( \tilde{V}_{k_0} \neq V_{k_0} \to j \mid a_0, k_0 \right)
\]

\[
\leq \int_0^{[n^{\alpha/2}] / n} \frac{\beta p \cdot x^{(1-p)/(2c)}}{2 \varepsilon a_0} \, dx + \sum_{j=[n^{\alpha/2}]}^{k_0-1} |\lambda_j - \tilde{\lambda}_j|.
\]

(26)

Recall that for \( j \in (n^{\alpha/2}, k_0 - 2] \),

\[
\lambda_j = \frac{\beta}{a_0} \left( \left( \frac{j}{n} \right)^{p/(2c)} - \left( \frac{j-1}{n} \right)^{p/(2c)} \right) = \frac{\beta p}{2 \varepsilon a_0} \left( \frac{j}{n} \right)^{p/(2c)} \left( \frac{1}{j} + O \left( \frac{1}{j^2} \right) \right) = \frac{\beta p}{2 \varepsilon a_0 n} \left( \frac{j}{n} \right)^{(1-p)/(2c)} \left( 1 + O \left( \frac{1}{j} \right) \right).
\]

It follows that on the event \( \mathcal{E}_0 \) and for all \( \varepsilon \in (0, 1/2] \) and \( j \in (n^{\alpha/2}, k_0 - 1] \),

\[
|\lambda_j - \tilde{\lambda}_j| \leq \frac{\beta p}{2 \varepsilon} \left( \left( \frac{j}{n} \right)^{(1-p)/(2c)} \right) \left( \frac{1}{k_0} - \frac{1}{a_0 n} \right) + \frac{\beta p}{2 \varepsilon a_0 j n} \left( \left( \frac{j}{n} \right)^{(1-p)/(2c)} \right)
\]

\[
= O \left( \frac{1}{a_0 n^2} \left( \left( \frac{j}{n} \right)^{(1-p)/(2c)} \right) \right) + O \left( \frac{1}{a_0 j n} \left( \left( \frac{j}{n} \right)^{(1-p)/(2c)} \right) \right)
\]

(27)

When \( j = k_0 - 1 \),

\[
\lambda_{k_0-1} = \frac{\beta}{a_0} \left( \left( \frac{k_0 - 2}{n} \right)^{p/(2c)} - \left( \frac{k_0 - 1}{n} \right)^{p/(2c)} \right)
\]

\[
= \frac{\beta}{a_0} \left( \frac{k_0 - 1}{n} \right)^{(1-p)/(2c)} \left( \frac{k_0 - 1}{a_0 n} \right) \left( \left( \frac{k_0 - 1}{k_0 - 1} \right)^{p/(2c)} - \left( \frac{k_0 - 2}{k_0 - 1} \right)^{p/(2c)} \right).
\]

(28)

and on the event \( \mathcal{E}_0 \),

\[
k_0 - 1 \left( \frac{a_0 n}{k_0 - 1} \right)^{p/(2c)} - \left( \frac{k_0 - 2}{k_0 - 1} \right)^{p/(2c)} = \frac{p}{2 \varepsilon k_0} \left( 1 + O \left( \frac{1}{k_0} \right) \right).
\]

As a result, there is a constant \( C > 0 \),

\[
\mathbb{1}_{\mathcal{E}_0} \sum_{j=[n^{\alpha/2}]}^{k_0-1} |\lambda_j - \tilde{\lambda}_j| \leq C \frac{(\log \log n)^2}{n^{1+(1-p)/(2c)}} \sum_{j=1}^{k_0-1} j^{(1-p)/(2c)-1} \leq \int_0^{k_0} x^{(1-p)/(2c)-1} \, dx \leq \frac{2 C \varepsilon (\log \log n)^2}{1-p} \frac{n}{n}.
\]

Using this together with (26) and observing that \( \left( \frac{n^{\alpha/2}}{n} \right)^{p/(2c)} \leq n^{(a-1)p/(2c)} = n^{-1} \) yields

\[
\mathbb{P} \left( \{ \tilde{V}_{k_0}^L \neq V_{k_0}^L \} \cap \mathcal{E}_0 \right) = \mathbb{E} \left[ \mathbb{1}_{\mathcal{E}_0} \mathbb{P} \left( \tilde{V}_{k_0}^L \neq V_{k_0}^L \mid a_0, k_0 \right) \right] = O \left( \frac{(\log \log n)^2}{n} \right).
\]
Combining all three steps then yields
\[
\mathbb{P}\left( \{I_{k_0}^L \neq I_{k_0}^{|E_0}\} \right) 
\leq \mathbb{P}\left( \{I_{k_0}^L \neq \bar{I}_{k_0}^L \} \right) + \mathbb{P}\left( \{\bar{I}_{k_0}^L \neq V_{k_0}^L \} \right) 
+ \mathbb{P}\left( \{V_{k_0}^L \neq V_{k_0} \} \right) 
= O\left( \log n \over n^{3/6} \right),
\]
which concludes the proof.

4.2.2 R-children of the root 0 and R-neighbours of k_0 in G_n

The R-neighbours of the root k_0 of G_n can be encoded by a Bernoulli point process \( I_{k_0}^R = (I_{j \to k_0})_{j=k_0+1}^n \), where \( I_{j \to k_0} \) equals one if \( j \) connects to \( k_0 \) in \( G_n \). Similarly, the R-children of the root 0 of \( T \) can be encoded by a Poisson point process \( V_{k_0}^R \).

We couple the two processes by discretising \( V_{k_0}^R \) as follows: for all \( j \in \{k_0 + 1, \ldots, n\} \), define \( V_{k_0 \to j} \sim \text{Poi}(\lambda_j) \) where
\[
\lambda_{k_0+1} := \int_{a_0}^{(k_0+1)/n} \frac{\beta \lambda}{2e} x^{(1-p)/(2\varepsilon) - 1} \, dx \quad \text{and} \quad \lambda_j := \int_{(j-1)/n}^{j/n} \frac{\beta \lambda}{2e} x^{(1-p)/(2\varepsilon) - 1} \, dx \quad \text{if} \quad j \in \{k_0 + 2, \ldots, n\}. \quad (29)
\]

Then, we write \( V_{k_0}^R = (V_{j \to k_0})_{j=k_0+1}^n \). Moreover, recall the event \( R_0 \) from (22) and note that \( k_{n+1}/n \geq a_0 \) on this event (so \( \lambda_{k_0+1} \) is non-negative). The following lemma shows that \( I_{k_0}^R \) and \( V_{k_0}^R \) can be coupled successfully on the event \( \mathcal{H}_{1,0} \cap \mathcal{R}_0 \).

**Lemma 4.5.** Consider the Bernoulli point process \( I_{k_0}^R := (I_{j \to k_0})_{j=k_0+1}^n \) and the discretised Poisson process \( V_{k_0}^R := (V_{j \to k_0})_{j=k_0+1}^n \) with means given in (29). Recall the events \( \mathcal{H}_{1,0} \) and \( \mathcal{R}_0 \) from (15) and (22), respectively, and set \( E_0 := \mathcal{H}_{1,0} \cap \mathcal{R}_0 \). Then, there exists a coupling of \( I_{k_0}^R \) and \( V_{k_0}^R \) such that
\[
\mathbb{P}\left( \{I_{k_0}^R \neq V_{k_0}^R \} \cap E_0 \right) \leq \frac{(\log n)^2}{n^{1/3}}.
\]

**Remark 4.6.** When the coupling of the Bernoulli process and the (discretised) Poisson point process is successful, we immediately obtain that \( \theta_0^R = \tau_0^R \), that is, the number of R-neighbours of the root \( k_0 \) in \( G_n \) and the number of R-children of the root 0 in \( T \) are the same. Moreover, for each R-child \( 0, i \) in \( T \) (with \( i \in [r_0^L + 1, \tau_0^L + \tau_0^R] \)) we have \( |a_{0,i} - k_{n+1}/n| \leq \frac{1}{n} \).

**Proof of Lemma 4.5.** As in the proof of Lemma 4.3, the coupling follows several intermediate steps, which we outline first.

1. We couple the Bernoulli process \( I_{k_0}^R \) with another Bernoulli process \( \bar{I}_{k_0}^R := (\bar{I}_{j \to k_0})_{j=k_0+1}^n \), where the variables \( (\bar{I}_{j \to k_0})_{j=k_0+1}^n \) are independent and have success probabilities
\[
\mu_j := \frac{\beta \lambda}{2e} \frac{1}{n} \left( \frac{j}{n} \right)^{(1-p)/(2\varepsilon) - 1}.
\] \( (30) \)

2. We couple \( \bar{I}_{k_0}^R \) with a sequence of independent Poisson random variables \( \bar{V}_{k_0}^R := (\bar{V}_{j \to k_0})_{j=k_0+1}^n \) with means given in (30).

3. We couple \( \bar{V}_{k_0}^R \) with \( V_{k_0}^R \).

We now provide the details for each of the coupling steps. Throughout the proof of the steps, we work on the event \( E_0 \).
Step 1. We couple $I_{k_0}^R$ to $\bar{I}_{k_0}^R$. Since both are Bernoulli processes, a standard coupling yields

$$
\mathbb{P}\left(I_{k_0}^R \neq \bar{I}_{k_0}^R \mid a_0, k_0 \right) \leq \sum_{j=k_0+1}^{n} \mathbb{P}(I_j = 1 \mid a_0, k_0) - \mathbb{P}(\bar{I}_j = 1 \mid a_0, k_0). \tag{31}
$$

Recall that $I_j = 1$ when at step $j$ a vertex is born, it connects to $k_0$ and survives after step $n$. Hence,

$$
\mathbb{P}(I_j = 1 \mid k_0) = \mathbb{P}(\{ j \text{ connects to } k_0 \text{ in } G_j \cap \{ \nu_j = 1 \} \mid k_0)
= \mathbb{E}\left[ \min\left\{ \frac{\beta}{|V_j - 1|}, 1 \right\} \right] \mathbb{P}(\nu_j = 1 \mid k_0).
$$

Furthermore, by Lemma 2.5 (applied with $S = \{k_0\}$), Lemma 2.6 and on the event $E_0$,

$$
\mathbb{1}_{E_0} \mathbb{P}(\bar{I}_j = 1 \mid a_0, k_0) = \mathbb{1}_{E_0} \mathbb{P}(\{ j \text{ connects to } k_0 \text{ in } G_j \cap \{ \nu_j = 1 \} \mid k_0)
= \mathbb{1}_{E_0} \mathbb{P}(\nu_j = 1 \mid k_0).
$$

As a result, using (30) and (31) yields that

$$
\mathbb{P}\left(I_{k_0}^R \neq \bar{I}_{k_0}^R \cap E_0 \right) = \mathbb{E}\left[ \mathbb{1}_{E_0} \mathbb{P}(I_{k_0}^R \neq \bar{I}_{k_0}^R \mid a_0, k_0) \right]
= O\left(\sum_{j=1}^{n} \log n \left( \frac{j}{n} \right)^{1-p}/(2\varepsilon) \right) = O\left(\log n \frac{1}{n^{1/3}} \sum_{j=1}^{n} \left( \frac{j}{n} \right)^{-4/3} \right) = O\left(\frac{\log n^2}{n^{1/3}} \right).
$$

Step 2. For the coupling of $\bar{I}_{k_0}^R$ and $\bar{V}_{k_0}^R$, we use Lemma 2.2. As a result,

$$
\mathbb{1}_{E_0} \mathbb{P}(\bar{I}_{k_0}^R \neq \bar{V}_{k_0}^R \mid a_0, k_0) \leq \sum_{j=[n/\log \log n]}^{n} 2^{\left( \frac{\beta p j}{2\varepsilon n} \left( \frac{j}{n} \right)^{1-p}/(2\varepsilon) \right)^2} \leq \frac{(\beta p)^2}{2^{2}\varepsilon^2 n^2} \sum_{j=[n/\log \log n]}^{n} \left( \frac{j}{n} \right)^{1-p}/\varepsilon.
$$

Using that

$$
\frac{1}{n} \sum_{j=[n/\log \log n]}^{n} \left( \frac{j}{n} \right)^{1-p}/\varepsilon \leq \frac{1}{n} \sum_{j=1}^{n} \left( \frac{j}{n} \right)^{1-p}/\varepsilon = O(1),
$$

we deduce that

$$
\mathbb{P}\left(I_{k_0}^R \neq \bar{V}_{k_0}^R \cap E_0 \right) = \mathbb{E}\left[ \mathbb{1}_{E_0} \mathbb{P}(I_{k_0}^R \neq \bar{V}_{k_0}^R \mid a_0, k_0) \right] = O(n^{-1}).
$$

Step 3. Recall $(\lambda_j)_{j=k_0+1}^{n}$ from (29) and $(\mu_j)_{j=k_0+1}^{n}$ from (30). Via a standard coupling of Poisson random variables (see Lemma 2.3),

$$
\mathbb{P}\left(V_{k_0}^R \neq \bar{V}_{k_0}^R \mid a_0, k_0 \right) \leq \sum_{j=k_0+1}^{n} |\lambda_j - \mu_j|. \tag{33}
$$

It remains to bound $|\lambda_j - \mu_j|$, which follows by computing the integral in the expression of $\lambda_j$. For $j \in \{k_0 + 2, \ldots, n\}$,

$$
\lambda_j = \int_{(j-1)/n}^{j/n} \frac{\beta p x^{1-p}/(2\varepsilon) - 1}{2\varepsilon} dx = \frac{\beta p}{1-p} \left( \frac{j}{n} \right)^{1-p}/(2\varepsilon) \left( 1 - \left( 1 - \frac{1}{j} \right)^{1-p}/(2\varepsilon) \right)
= \frac{\beta p}{1-p} \left( \frac{j}{n} \right)^{1-p}/(2\varepsilon) \left( \frac{1-p}{2\varepsilon j} - O(j^{-2}) \right)
= \mu_j + O\left( j^{-2} \left( \frac{j}{n} \right)^{1-p}/(2\varepsilon) \right). \tag{34}
$$
Hence, 
\[
\sum_{j=k_0+2}^{n} |\lambda_j - \mu_j| = \mathcal{O}\left(\frac{1}{n}\right) \tag{35}
\]
almost surely. For \( j = k_0 + 1 \), we find in a similar way that 
\[
\lambda_{k_0+1} = \int_{a_0}^{(k_0+1)/n} \frac{\beta p}{2\varepsilon} x^{(1-p)/(2\varepsilon)-1} \, dx = \frac{\beta p}{1-p} \left( \frac{k_0 + 1}{n} \right)^{(1-p)/(2\varepsilon)} \left( 1 - \left( \frac{a_0 n}{k_0 + 1} \right)^{(1-p)/(2\varepsilon)} \right). \tag{36}
\]
Having that \( n/\log \log n \leq a_0 n \leq k_0 \leq a_0 n - 1 \) on the event \( \mathcal{E}_0 \) implies that 
\[
\left( \frac{a_0 n}{k_0 + 1} \right)^{(1-p)/(2\varepsilon)} = \left( 1 - \frac{k_0 + 1 - a_0 n}{k_0 + 1} \right)^{(1-p)/(2\varepsilon)} = 1 - \mathcal{O}\left(\frac{\log \log n}{n}\right).
\]
Combining the above inequality with (36), and then with (33) and (35) yields 
\[
\mathbb{P}\left( \{ \tilde{V}_{k_0}^R \neq V_{k_0}^R \} \cap \mathcal{E}_0 \right) = \mathbb{E} \left[ \mathbf{1}_{\mathcal{E}_0} \mathbb{P}\left( \tilde{V}_{k_0}^R \neq V_{k_0}^R \mid a_0, k_0 \right) \right] = \mathcal{O}\left(\frac{\log \log n}{n}\right).
\]
Combining steps 1 through 3 via a union bound, we arrive at 
\[
\mathbb{P}\left( \{ I_{k_0}^R \neq V_{k_0}^R \} \cap \mathcal{E}_0 \right) \leq \frac{(\log n)^2}{n^{1/3}},
\]
which concludes the proof. \( \square \)

### 4.3 Summary

We are now ready to prove Lemma 4.1.

**Proof of Lemma 4.1.** Part of the proof has already been demonstrated in (16),(17), (19) and (21). Furthermore, we observe that 
\[
\mathcal{H}_{1,2} \supseteq \mathcal{R}_0 \cap \{ I_{k_0}^L = V_{k_0}^L \} \cap \{ I_{k_0}^R = V_{k_0}^R \},
\]
where we recall the events \( \mathcal{R}_0, \{ I_{k_0}^L = V_{k_0}^L \} \) and \( \{ I_{k_0}^R = V_{k_0}^R \} \) from (22) and Lemmas 4.3 and 4.5, respectively. These events ensure that the immediate neighbourhoods of \( k_0 \) in \( G_n \) and \( 0 \) in \( T \) are isomorphic, since the number of \( L \)- and \( R \)-neighbours are the same. Thus, from Lemmas 4.2, 4.3 and 4.5 we obtain that 
\[
\mathbb{P}\left( \mathcal{H}_{1,2}^c \cap \mathcal{H}_{1,0} \right) \leq \mathbb{P}\left( \mathcal{R}_0^c \right) + \mathbb{P}\left( \{ I_{k_0}^L \neq V_{k_0}^L \} \cap \mathcal{R}_0 \cap \mathcal{H}_{1,0} \right) + \mathbb{P}\left( \{ I_{k_0}^R \neq V_{k_0}^R \} \cap \mathcal{R}_0 \cap \mathcal{H}_{1,0} \right) = \mathcal{O}\left(\frac{1}{\log n}\right).
\]
As a result, combining this with the steps from (16), (17), (19) and (21) we finally obtain 
\[
\mathbb{P}\left( \left( \bigcap_{i=0}^{3} \mathcal{H}_{1,i} \right)^c \right) \leq \mathbb{P}\left( \mathcal{H}_{1,0}^c \right) + \mathbb{P}\left( \mathcal{H}_{1,1}^c \right) + \mathbb{P}\left( \mathcal{H}_{1,2}^c \cap \mathcal{H}_{1,0} \right) + \mathbb{P}\left( \mathcal{H}_{1,3}^c \right) = \mathcal{O}\left(\frac{1}{(\log \log n)^p/(2\varepsilon)}\right),
\]
which finishes the proof. \( \square \)

### 5 Continuing the coupling

We continue the construction of the coupling described in the previous sections by providing a coupling of the \( r \)-neighbourhoods of \( k_0 \) in \( G_n \) and \( 0 \) in \( T \) that is successful with high probability for any \( r \geq 2 \). Since the number of vertices in \( G_n \) is “large” w.h.p, exploring a finite neighbourhood of \( G_n \) changes the distribution of the remaining graph “very little”. As a result, the coupling of the \( r \)-neighbourhoods follows a similar
approach as the coupling of the 1-neighbourhood in the previous section. Modifications are necessary, though, which make the proofs somewhat more technical and involved.

As in (20), fix \( r \in \mathbb{N} \) and for all \( q \in \{r - 1\} \), define

\[
\mathcal{H}_{q,1} := \{ v \in V(\partial B_{q-1}(T, 0)) : \min_{i \in [r]} a_{v,i} \geq (\log \log n)^{-q} \},
\]

\[
\mathcal{H}_{q,2} := \{ (B_q(G_n, k_0), k_0) \cong (B_q(T, 0), 0) \} \cap \{ v \in V(\partial B_q(T, 0)) : |a_v - k_v/q| \leq \frac{q+1}{n} \},
\]

\[
\mathcal{H}_{q,3} := \{ v \in V(B_q(T, 0)) : \tau_v^L + \tau_v^R < (\log \log n)^{1/r} \}.
\]

Note that for \( r = 1 \) and \( q = 0 \), these events coincide with the last three events in (15). When \( r > 1 \), only minor modifications in (21) are required to show that, when \( q = 0 \), \( \mathcal{H}_{1,3} \) holds with probability at least \( 1 - O((\log \log n)^{-p/(2e)}) \). We now state the main result of this section.

**Proposition 5.1.** Fix \( \varepsilon \in (0, \frac{1}{2}] \), \( r \in \mathbb{N}\setminus\{1\} \) and \( q \in \{r - 1\} \) and set \( p := \frac{1}{2} + \varepsilon \). Recall the events in (37), and assume that for all \( n \) sufficiently large there exists a coupling of \((G_n, k_0)\) and \((T, 0)\) and a constant \( C_q > 0 \) such that

\[
\mathbb{P}\left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{q,i} \right)^c \right) \leq C_q (\log \log n)^{(q-1)/r-p/(2e)}.
\]

Then, there exists a coupling of \((G_n, k_0)\) and \((T, 0)\) and a constant \( C_{q+1} > 0 \) such that

\[
\mathbb{P}\left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{q+1,i} \right)^c \right) \leq C_{q+1} (\log \log n)^{q/r-p/(2e)}.
\]

Proposition 5.1 combined with Lemma 4.1 has the following result as an immediate corollary.

**Corollary 5.2.** Fix \( \varepsilon \in (0, \frac{1}{2}] \) and set \( p := \frac{1}{2} + \varepsilon \). For every fixed \( r \in \mathbb{N} \) there exists a coupling of \((G_n, k_0)\) and \((T, 0)\) such that

\[
\mathbb{P}\left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{r,i} \right)^c \right) = O((\log \log n)^{-(1-p)/(2e)-1/r}).
\]

To prove Proposition 5.1, we bound

\[
\mathbb{P}\left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{q+1,i} \right)^c \right) \leq \mathbb{P}\left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{q+1,i} \right)^c \cap \bigcap_{i=1}^{3} \mathcal{H}_{q,i} \right) + \mathbb{P}\left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{q,i} \right)^c \right)
\]

\[
\leq \mathbb{P}\left( \mathcal{H}_{q+1,3}^c \cap \mathcal{H}_{q,3} \right) + \mathbb{P}\left( \mathcal{H}_{q+1,1}^c \cap \mathcal{H}_{q+1,3} \cap \mathcal{H}_{q,1} \right) + \mathbb{P}\left( \mathcal{H}_{q+1,2}^c \cap \mathcal{H}_{q+1,1} \cap \mathcal{H}_{q+1,3} \cap \mathcal{H}_{q,3} \right) + \mathbb{P}\left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{q,i} \right)^c \right).
\]

First, let \( (P_i)_{i \in \mathbb{N}} \) be i.i.d. \( \text{Poi}(\frac{\varepsilon}{1-\varepsilon}) \) random variables. Recall that a vertex \( v \) in \( T \) has \( \tau_v^L + \tau_v^R \) many children. As in (20), it is readily checked that for any type \( a_v \in (0, 1] \), \( \tau_v^L + \tau_v^R \) is stochastically dominated by \( P_1 \). Since for every \( q \in \{r - 1\} \), conditionally on \( \mathcal{H}_{q,3} \), we have that \( |\partial B_q(T, 0)| \leq (\log \log n)^{q/r} \), a union bound and a similar approach as in (21) with \( t = 1 \) yield

\[
\mathbb{P}\left( \mathcal{H}_{q+1,3}^c \cap \mathcal{H}_{q,3} \right) \leq \max_{i \in ([\log \log n])^{q/r} \cap \mathbb{N}} P_i \geq (\log \log n)^{1/r}
\]

\[
\leq O((\log \log n)^{q/r - (\log \log n)^{1/r}}) = o((\log \log n)^{-s}),
\]

where the final equality holds for any \( s > 0 \).
Similarly, for every \( q \in [r-1] \), by using a similar approach as in (18) and a union bound over \( |\partial B_q(T,0)| \leq (\log \log n)^{\frac{q}{r}} \) vertices, we obtain
\[
\mathbb{P}\left( H^c_{q+1,1} \cap H_{q+1,3} \cap H_{q,1} \right) \leq \beta (\log \log n)^{\frac{q+q-(q+1)p/(2e)}{r}} \leq \beta (\log \log n)^{\frac{q-p/(2e)}{r}},
\]
(40)
since \( \frac{p}{2e} \geq 1 \). By the induction hypothesis, we know that the last term in (38) is at most \( C_q (\log \log n)^{(q-1)/r-p/(2e)} \) for some constant \( C_q > 0 \), so only the third term on the right-hand side of (38) is left to analyse.

To extend the coupling from the \( q \)-neighbourhoods of \( k_0 \) and 0 to their respective \( (q+1) \)-neighbourhoods, we can assume that a coupling of \( B_q(G_n, k_0) \) and \( B_q(T,0) \) exists such that \( \bigcap_{i=1}^{q+3} H_{q,i} \) holds. Then, for each \( v \in V(\partial B_q(T,0)) \), we want to couple its children in \( T \) to the neighbours of \( k_v \in V(\partial B_q(G_n, k_0)) \) in \( G_n \) so that the number of children and neighbours, respectively, are equal, and their types and rescaled labels are sufficiently close.

Recall that \( k[t] \) is the smallest vertex in \( A_{t-1} \) with respect to the BFS order, and let, for ease of writing, \( k_v(t) := k[t] \). Suppose that \( k_v(t) \in V(\partial B_q(G_n, k_0)) \). We then define the event
\[
K_{t,q} := \{ \theta^L_{v(t)} = \tau^L_{v(t)} , \theta^R_{v(t)} = \tau^R_{v(t)} \} \cap \{ \forall i \in [\tau_{v(t)}] : a_{v(t),i} - \frac{k_{\tau_{v(t)+1}},i}{n} \leq \frac{q+2}{n}, k_{v(t),i} \notin A_{t-1} \},
\]
and the collection of random variables
\[
\tilde{C}_{t(q)} := (k_v, a_v)_{k_v \in A_{t-1} \cup \mathcal{P}_{t-1}}, ((\theta^L_{w(t)}, \tau^L_{w(t)}), (\theta^R_{w(t)}, \tau^R_{w(t)}))_{w \in \mathcal{P}_{t-1}}.
\]
Note that the event \( \{ \forall i \in [\tau_{v(t)}] : k_{\tau_{v(t)+1},i} \notin A_{t-1} \} \) is included in (41) since the formation of a cycle in the BFS exploration in \( G_n \) would immediately invalidate the coupling.

Furthermore, for every \( q \in [r-1] \), define \( \rho[q] \) as
\[
\rho[q] = \left\{ \begin{array}{ll}
|V(B_q-1(G_n, k_0))| + 1, & \text{if } V(B_q(G_n, k_0)) \neq \emptyset, \\
\infty, & \text{otherwise}.
\end{array} \right.
\]
(42)
When \( \rho[q] < \infty \), it denotes the time step of the BFS exploration at which the smallest vertex \( k[\rho[q]] \) (with respect to the BFS order) in \( V(\partial B_q(G_n, k_0)) \) is explored. Also observe that, for all \( q \in [r-1] \),
\[
(A_{\rho[q]-1}, P_{\rho[q]-1}, N_{\rho[q]-1}) = (V(\partial B_q(G_n, k_0)), V(B_q-1(G_n, k_0)), V_n \setminus V(B_q(G_n, k_0))).
\]
We observe that
\[
H_{q+1,2} = H_{q,2} \cap \bigcap_{t=\rho[q]}^{\rho[q+1]-1} K_{t,q}.
\]
(43)
As a result, to provide an upper bound for the probability
\[
\mathbb{P}\left( H^c_{q+1,2} \cap H_{q+1,3} \cap H_{q,1} \cap \bigcap_{t=\rho[q]}^{\rho[q+1]-1} K_{t,q} \right)
\]
we first bound
\[
\mathbb{P}\left( K^c_{q,q} \cap H_{q+1,1} \cap H_{q+1,3} \cap \bigcap_{t=1}^{3} H_{q,i} \right)
\]
for every \( t \in \{\rho[q], \ldots, \rho[q+1] - 1\} \). As for a fixed \( q \) and any \( t \in \{\rho[q], \ldots, \rho[q+1] - 1\} \) these bounds are rather similar, we focus on the case \( t = \rho[q] \). That is, we consider the vertex \( k_{\rho[q]} := k[\rho[q]] \) and couple the neighbours of \( k_{\rho[q]} \) in \( G_n \) to the children of \( k_{\rho[q]} \) in \( T \). For ease of writing, let us set
\[
K_{\rho[q]} := K_{\rho[q],q} = \{ \theta^L_{v[\rho[q]]} = \tau^L_{v[\rho[q]]}, \theta^R_{v[\rho[q]]} = \tau^R_{v[\rho[q]]} \} \cap \{ \forall i \in [\tau_{v[\rho[q]]}] : a_{v[\rho[q]],i} - \frac{k_{\tau_{v[\rho[q]]}},i}{n} \leq \frac{q+2}{n}, k_{v[\rho[q]],i} \notin A_{\rho[q]-1} \},
\]
(44)
To do this, we show the following lemma, which is similar in nature to Lemma 4.3.

We only provide the details for the coupling of the neighbourhoods of \( v[q] \) in \( B_{q+1}(T,0) \) and of \( k_{v[q]} \) in \( B_{q+1}(G_n,k_0) \) and briefly touch upon the coupling for the remaining vertices in \( V(\partial B_q(G_n,k_0)) \), as the coupling for these other vertices follows in a similar way.

To couple the neighbours of \( k_{v[q]} \) in \( G_n \) and \( v[q] \) in \( T \), we assume that the random variables in \( \mathcal{C}_q \) have been coupled such that \( \mathcal{E}_q := \cap_{i=1}^3 \mathcal{H}_{q,i} \) holds. In the upcoming sections we prove the following result.

**Lemma 5.3.** Fix \( \varepsilon \in (0, \frac{1}{2}] \), \( r \in \mathbb{N} \setminus \{1\} \), \( q \in [r-1] \), and set \( \rho := \frac{1}{2} + \varepsilon \) and \( \alpha := \frac{1-\rho}{p} \). Recall \( (\mathcal{H}_{q,i})_{i\in[3]} \) and \( \mathcal{K}_{\rho[q]} \) from (37) and (44), respectively. There exists a coupling of \((G_n,k_0) \) and \((T,0) \) such that

\[
\Pr(\mathcal{K}_{\rho[q]} \cap \mathcal{E}_q) = O\left(\frac{\log n}{n^{\alpha/6}}\right).
\]

### 5.1 Coupling the \( L \)-children of \( v[q] \) in \( T \) with the \( L \)-neighbours of \( k_{v[q]} \) in \( G_n \)

To prove Lemma 5.3, we first couple the \( L \)-neighbours of \( k_{v[q]} \) and the \( L \)-children of \( v[q] \) in this section. We define

\[
\mathcal{L}_{\rho[q]} := \{\theta_{v[q]}^L = \tau_{v[q]}^L\} \cap \{\forall i \in [\tau_{v[q]}^L] : |a_{v[q],i} - \frac{k_{v[q],i}}{n}| \leq \frac{\rho+2}{n} \text{ and } k_{v[q],i} \in \mathcal{N}_{\rho[q]} - 1\}.
\]

(45)

to be the event that the coupling of the \( L \)-neighbours of \( k_{v[q]} \) and the \( L \)-children of \( v[q] \) is successful. In words, the first two parts of \( \mathcal{L}_{\rho[q]} \) imply that the number of \( L \)-neighbours of \( k_{v[q]} \) and the \( L \)-children of \( v[q] \) are the same, and that their respective rescaled labels and types are sufficiently close. The last part ensures that the exploration of the \( L \)-neighbours of \( k_{v[q]} \) does not lead to cycles.

We now define an encoding of the \( L \)-neighbours of \( k_{v[q]} \) in \( G_n \) and the \( L \)-children of \( v[q] \) in \( T \) that allows us to construct a coupling such that \( \mathcal{L}_{\rho[q]} \) holds with high probability. Unlike in the proof of Lemma 4.3, we cannot ensure that \( \frac{k_{v[q],-1}}{n} \leq a_{v[q]} \) holds on the event \( \mathcal{E}_q = \cap_{i=1}^3 \mathcal{H}_{q,i} \). Instead, we let \( N_{v[q]} := \lceil a_{v[q]} n \rceil \) and \( k^*_q := \max\{N_{v[q]}, k_{v[q] - 1}\} \) and use \( k^*_q \) to construct the necessary processes that we couple further on in this section. If \( N_{v[q]} \geq k_{v[q]} - 1 \), we define conditionally on \( a_{v[q]} \),

\[
\lambda^q_{N_{v[q]} := \int_{\frac{a_{v[q]} + 1}{n}} a_{v[q]} + \frac{\rho p}{2} x^{(1-p)/2} \, dx, \text{ and for all } j \in [N_{v[q]} - 1], \lambda^q_j := \int_{\frac{j-1}{n}}^{\frac{j}{n}} \frac{\rho p}{2} x^{(1-p)/2} \, dx.}
\]

If instead \( N_{v[q]} \leq k_{v[q]} - 2 \), we additionally set \( \lambda^q_j := 0 \) for all integers \( j \in [N_{v[q]} + 1, k^*_q] \). This defines \( \lambda^q_j \) for all \( j \in [0, k^*_q] \), and allows us to define the discretised Poisson point process \( \mathbf{V}^L_{k_{v[q]} := (V^L_{k_{v[q]} \to j})_{j=0}^{k^*_q}} \), where the \( V^L_{k_{v[q]} \to j} \) are independent Poisson random variables with means \( \lambda^q_{k_{v[q]} \to j} \) (where a Poisson variable with mean zero is equal to zero almost surely). Note that \( \mathbf{V}^L_{k_{v[q]} \to j} \) is a discretisation of the Poisson point process on \((0, a_{v[q]}) \) with density \( \lambda^q_{v[q]} (dx) \) that determines the \( L \)-children of \( v[q] \) (see Definition 1.2).

Similarly, we define the Bernoulli point process \( \mathbf{I}^L_{k_{v[q]} := (I^L_{k_{v[q]} \to j})_{j=1}^{k^*_q}} \) as follows. For all \( j \in [k_{v[q]} - 1] \), \( I^L_{k_{v[q]} \to j} \) equals one if \( k_{v[q]} \) connects to \( j \) by an edge in \( G_n \) and \( j \in \mathcal{N}_{\rho[q]} - 1 \cup \mathcal{A}_{\rho[q]} - 1 \), and equals zero otherwise. For \( j \in [k_{v[q]}, k^*_q] \) (that is, when \( N_{v[q]} > k_{v[q]} - 1 \), we additionally set \( I^L_{k_{v[q]} \to j} := 0 \).

Our aim is to couple the Bernoulli point process \( \mathbf{I}^L_{k_{v[q]} \to j} \) and the discretised Poisson point process \( \mathbf{V}^L_{k_{v[q]} \to j} \).

To do this, we show the following lemma, which is similar in nature to Lemma 4.3.

**Lemma 5.4.** Fix \( \alpha = \frac{1-p}{p} \) and recall that \( \mathcal{E}_q = \cap_{i=1}^3 \mathcal{H}_{q,i} \), where the events \( (\mathcal{H}_{q,i})_{i\in[3]} \) were defined in (37). Consider the processes \( \mathbf{I}^L_{k_{v[q]} \to j} \) and \( \mathbf{V}^L_{k_{v[q]} \to j} \). There exists a coupling such that

\[
\Pr(\{\mathbf{I}^L_{k_{v[q]} \to j} = \mathbf{V}^L_{k_{v[q]} \to j} \cap \{\forall j \in \mathcal{A}_{\rho[q]} - 1 \cup \mathcal{N}_{\rho[q]} : I^L_{k_{v[q]} \to j} = 0\} \cap \mathcal{E}_q \}) = O\left(\frac{\log n}{n^{\alpha/6}}\right).
\]
Remark 5.5. When the coupling is successful, in the sense that the event
\[ \{ I_{\hat{V}_{k[q]}}^L = V_{k[q]}^L \} \cap \{ \forall j \in A_{\rho[q]-1} \cap [k_q^*]: I_{\hat{V}_{k[q]}}^L \to j = 0 \} \]
holds, it directly follows that the event $\mathcal{L}_{\rho[q]}$ holds as well. As a result,
\[ \mathbb{P}\left( \mathcal{L}_{\rho[q]}^c \cap \mathcal{E}_q \right) = O\left( \frac{\log n}{n^{\alpha/6}} \right). \]

Proof of Lemma 5.4. Throughout the proof, we assume for simplicity that $k_q^* = N_{v[q]} = k_{v[q]} - 1$. The other cases can be dealt with in a similar way with only minor modifications. We write
\begin{equation}
\begin{aligned}
\mathbb{P}\left( \left( \{ I_{\hat{V}_{k[q]}}^L = V_{k[q]}^L \} \cap \{ \forall j \in A_{\rho[q]-1} \cap [k_q^*]: I_{\hat{V}_{k[q]}}^L \to j = 0 \} \right)^c \cap \mathcal{E}_q \right)
&= \mathbb{P}\left( \{ I_{\hat{V}_{k[q]}}^L \neq V_{k[q]}^L \} \cap \mathcal{E}_q \right) + \mathbb{P}\left( \{ \exists j \in A_{\rho[q]-1} \cap [k_q^*]: I_{\hat{V}_{k[q]}}^L \to j = 1 \} \cap \{ I_{\hat{V}_{k[q]}}^L = V_{k[q]}^L \} \cap \mathcal{E}_q \right),
\end{aligned}
\end{equation}

We divide the coupling into several steps, similar to the proof of Lemma 4.3.

1. We couple $I_{\hat{V}_{k[q]}}^L$ with a Bernoulli process $I_{\hat{V}_{k[q]}}^L = (\hat{I}_{k[q]})_{j=1}^{k_{v[q]} - 1}$, where the variables $(\hat{I}_{k[q]})_{j=1}^{k_{v[q]} - 1}$ are independent, $\hat{I}_{k[q] \to j} = 0$ for all $j \leq n^{\alpha/2}$ and otherwise
\[ \mathbb{P}\left( \hat{I}_{k[q] \to j} = 1 \mid \mathcal{C}_q \right) = \lambda_{j}^{\rho[q]} := \frac{\beta p}{2 \varepsilon k_{v[q]}} \left( \frac{j}{n} \right)^{(1-p)/(2\varepsilon)} \text{ for all } j \in (n^{\alpha/2}, k_{v[q]}]. \]

2. We couple $\hat{V}_{k[q]}^L$ with a sequence $V_{k[q]}^L = (\hat{V}_{k[q]})_{j=1}^{k_{v[q]} - 1}$ of independent variables where $\hat{V}_{k[q] \to j} = 0$ for all $j \leq n^{\alpha/2}$, and for $j \in (n^{\alpha/2}, k_{v[q]} - 1]$, the $\hat{V}_{k[q] \to j}$ is a Poisson random variables with means $\lambda_{j}^{\rho[q]}$.

3. We couple $\hat{V}_{k[q]}^L$ with $V_{k[q]}^L$.

We now provide details for these steps.

Step 1. Recall $v_1$ from Lemma 2.7. Then, by a computation similar to (25) we conclude that
\[ \mathbb{P}\left( \exists j \leq n^{\alpha/2} : I_{k[q] \to j} = 1 \mid \mathcal{C}_q \right) = O\left( n^{-\alpha p/(4\varepsilon) k_{v[q]}^{-1}} \right). \]

Now, let $j \in (n^{\alpha/2}, k_{v[q]} - 1]$. We observe that all vertices in $A_{\rho[q]-1} \cup P_{\rho[q]-1}$ survive up to step $n$. We use Lemmas 2.5 and 2.6 with $S = A_{\rho[q]-1} \cup P_{\rho[q]-1}$ and $R = \{ j \}$. Since $|S| = O((\log \log n)^q/r) = O(j^{2/3})$ for any $j > n^{\alpha/2}$, we obtain for all $j \not\in A_{\rho[q]-1} \cup P_{\rho[q]-1}$,
\[ 1_{\mathcal{E}_q} \mathbb{P}\left( I_{k[q] \to j} = 1 \mid \mathcal{C}_q \right) = 1_{\mathcal{E}_q} \beta p \left( \frac{j}{n} \right)^{(1-p)/(2\varepsilon)} \left( 1 + O\left( \frac{\log n}{n^{\alpha/6}} \right) \right) \left( 1 + O\left( k_{v[q]}^{-1/3} \right) \right)
= 1_{\mathcal{E}_q} \beta p \left( \frac{j}{n} \right)^{(1-p)/(2\varepsilon)} + O\left( 1_{\mathcal{E}_q} \left( \frac{\log n}{n^{\alpha/6}} k_{v[q]} + \frac{1}{k_{v[q]}^{1/3}} \right) \right)
= 1_{\mathcal{E}_q} \beta p \left( \frac{j}{n} \right)^{(1-p)/(2\varepsilon)} + O\left( 1_{\mathcal{E}_q} \left( \frac{\log n}{n^{\alpha/6}} k_{v[q]} \right) \right), \]

where for the last equality we used that $k_{v[q]} \geq n^{\alpha/3}$. Note also that for $j \in P_{\rho[q]-1}$, we know that $I_{k[q] \to j} = 0$ by definition. Hence, the coupling is successful for these $j$ only when $\hat{I}_{k[q] \to j} = 0$. Combining
the above observations with a standard Bernoulli coupling yields

\[ 1_{\mathcal{E}_q} \mathbb{P}(I_{k_{v[q]}^L} \neq \bar{I}_{k_{v[q]}^L} \mid C_q) \leq 1_{\mathcal{E}_q} \mathbb{P}\left( \bigcup_{j \leq \frac{n^\alpha}{2}} \{I_{k_{v[q]}^L} \to j = 1\} \mid C_q \right) + 1_{\mathcal{E}_q} \sum_{j = \lfloor \frac{n^\alpha}{2} \rfloor}^{k_{v[q]} - 1} \mathbb{P}(I_{k_{v[q]}^L} \to j \neq \bar{I}_{k_{v[q]}^L} \mid C_q) \]

\[ + 1_{\mathcal{E}_q} \sum_{j = \lfloor \frac{n^\alpha}{2} \rfloor}^{k_{v[q]} - 1} \mathbb{P}(\bar{I}_{k_{v[q]}^L} \to j = 1 \mid C_q) \]

\[ \leq \mathcal{O}(1_{\mathcal{E}_q} n^{-\alpha p/(4\varepsilon)} k_{v[q]}^{-1}) + 1_{\mathcal{E}_q} \sum_{j = \lfloor \frac{n^\alpha}{2} \rfloor}^{k_{v[q]} - 1} \mathbb{P}(I_{k_{v[q]}^L} = 1 \mid C_q) - \mathbb{P}(\bar{I}_{k_{v[q]}^L} \to j = 1 \mid C_q) \]

\[ + 1_{\mathcal{E}_q} \sum_{j = \lfloor \frac{n^\alpha}{2} \rfloor}^{k_{v[q]} - 1} \frac{\beta_p}{2 \varepsilon k_{v[q]}} \left( \frac{j}{n} \right) (1-p)/(2\varepsilon) \]

\[ \leq 1_{\mathcal{E}_q} \left( \mathcal{O}(n^{-\alpha p/(4\varepsilon)} k_{v[q]}^{-1}) + \sum_{j = \lfloor \frac{n^\alpha}{2} \rfloor}^{k_{v[q]} - 1} \mathcal{O}\left( \frac{\log n}{n^{\alpha/6} k_{v[q]}} \right) + \mathcal{O}\left( \frac{|\mathcal{P}_{\rho[q]} - 1|}{k_{v[q]}} \right) \right) \]

\[ \leq 1_{\mathcal{E}_q} \left( \mathcal{O}(n^{-\alpha p/(4\varepsilon)} k_{v[q]}^{-1}) + \mathcal{O}\left( \frac{\log n}{n^{\alpha/6}} \right) + \mathcal{O}\left( \frac{|\mathcal{P}_{\rho[q]} - 1|}{k_{v[q]}} \right) \right). \]

Finally, using the bounds on $k_{v[q]}$ and $|\mathcal{P}_{\rho[q]} - 1|$ ensured by the event $\mathcal{E}_q$ yields

\[ \mathbb{P}\left( \{I_{k_{v[q]}^L} \neq \bar{I}_{k_{v[q]}^L}\} \cap \mathcal{E}_0 \right) = \mathbb{E}\left[ 1_{\mathcal{E}_q} \mathbb{P}(I_{k_{v[q]}^L} \neq \bar{I}_{k_{v[q]}^L} \mid C_q) \right] = \mathcal{O}\left( \frac{\log n}{n^{\alpha/6}} \right). \]

**Step 2.** By the Poisson-Bernoulli coupling given in Lemma 2.2, we obtain

\[ \mathbb{P}\left( \{I_{k_{v[q]}^L} \neq \bar{I}_{k_{v[q]}^L}\} \cap \mathcal{E}_q \right) = \mathbb{E}\left[ 1_{\mathcal{E}_q} \mathbb{P}(I_{k_{v[q]}^L} \neq \bar{I}_{k_{v[q]}^L} \mid C_q) \right] \leq \mathbb{E}\left[ 1_{\mathcal{E}_q} \sum_{j = \lfloor \frac{n^\alpha}{2} \rfloor}^{k_{v[q]} - 1} \left( \frac{\beta_p}{2 \varepsilon k_{v[q]}} \left( \frac{j}{n} \right) (1-p)/(2\varepsilon) \right)^2 \right]. \]

As $k_{v[q]} \leq n$ almost surely,

\[ \sum_{j = \lfloor \frac{n^\alpha}{2} \rfloor}^{k_{v[q]} - 1} \left( \frac{\beta_p}{2 \varepsilon k_{v[q]}} \left( \frac{j}{n} \right) (1-p)/(2\varepsilon) \right)^2 = \frac{1}{k_{v[q]} \sum_{j = 1}^{k_{v[q]}} \left( \frac{j}{k_{v[q]}} \right) (1-p)/\varepsilon} \]

\[ \text{(47)} \]

Using that the term in the $\mathcal{O}$ in (47) is at most of constant order, we get

\[ \mathbb{P}\left( \{I_{k_{v[q]}^L} \neq \bar{I}_{k_{v[q]}^L}\} \cap \mathcal{E}_q \right) \leq \mathbb{E}\left[ 1_{\mathcal{E}_q} \mathcal{O}(k_{v[q]}^{-1}) \right] = \mathcal{O}\left( \frac{(\log \log n)^{q+1}}{n} \right). \]

**Step 3.** We first consider the case $j \leq \frac{n^{\alpha/2}}{2}$. As in (26),

\[ \mathbb{P}\left( \forall j \leq \frac{n^{\alpha/2}}{2} : V_{k_{v[q]}^L} \to j = 0 \right) \geq 1 - \int_{0}^{\left\lfloor \frac{n^{\alpha/2}}{2} \right\rfloor / n} \frac{\beta p}{2 \varepsilon a_{v[q]}^q} x^{(1-p)/(2\varepsilon)} \, dx \geq 1 - \frac{2\beta}{a_{v[q]}} \left( \frac{n^{\alpha/2}}{2} \right)^{p/(2\varepsilon)}, \]

\[ \text{(48)} \]

where for the first inequality we used our assumption that $N_{v[q]} = k_{v[q]} - 1$. Moreover, with a similar approach as in (27), for all $j \in (\frac{n^{\alpha/2}}{2}, k_{v[q]} - \frac{2}{2\varepsilon a_{v[q]}^q})$ we have

\[ |\lambda_{j}^{\rho[q]} - \bar{\lambda}_{j}^{\rho[q]}| \leq \frac{\beta p}{2 \varepsilon} \left( \frac{j}{n} \right) (1-p)/(2\varepsilon) \]

\[ + 1 \cdot \frac{1}{a_{v[q]}^q n} \left| \frac{1}{a_{v[q]}^q n} \right| + \frac{\beta p}{2 \varepsilon a_{v[q]}^q j n} \left( \frac{j}{n} \right) (1-p)/(2\varepsilon). \]
Hence, on the event \( E_q \), for \( j \in (n^{\alpha/2}, k_v[q] - 2] \),
\[
|\lambda_j^\rho[q] - \bar{\lambda}_j^\rho[q]| = O\left( \frac{(\log \log n)^{(q+1)}}{n^{1+(1-p)/(2\varepsilon) - 1}} \right).
\] (49)

For \( j = k_v[q] - 1 \), with a similar approach as in (28) and on the event \( E_q \),
\[
|\lambda_j^\rho[q] - \bar{\lambda}_j^\rho[q]| = O\left( \frac{(k_v[q] - 1 - \varepsilon.a_v[q])}{n^{1+(1-p)/(2\varepsilon) - 1}} \right).
\] (50)

Combining the estimates in (48), (49) and (50) then yields
\[
\mathbb{P}\left( \{ V_{k_v[q]}^L \neq \bar{V}_{k_v[q]}^L \} \cap E_q \right) = \mathbb{E}\left[ 1_{E_q} \mathbb{P}\left( \{ V_{k_v[q]}^L \neq \bar{V}_{k_v[q]}^L \} \mid C_q \right) \right]
\]
\[
= O\left( \frac{(\log \log n)^{q+1} + n^{-(1-\alpha/2)p/(2\varepsilon)}) + \frac{(\log \log n)^{q+1} + n^{-(1-p)/(2\varepsilon)}}{\sum_{j=1}^{n} j^{-(1-p)/(2\varepsilon) - 1}} \right)
\]
\[
= O\left( \frac{(\log \log n)^{2(q+1)}}{n} \right).
\] Combining all three steps via a union, we finally arrive at
\[
\mathbb{P}\left( \{ I_{k_v[q]}^L \neq V_{k_v[q]}^L \} \cap E_q \right) = O\left( \frac{\log n}{n^{\alpha/6}} \right),
\] (51)
so it remains to bound the second term on the right-hand side of (46). Via a union bound and Markov’s inequality we obtain
\[
\mathbb{P}\left( \{ \exists j \in A_{\rho[q]-1} \cap [k_v[q]^*] : I_{j-\tau_{k_v[q]}^L}^L = 1 \} \cap \{ I_{k_v[q]}^L = V_{k_v[q]}^L \} \mid E_q \right)
\]
\[
\leq \mathbb{E}\left[ 1_{E_q} \mathbb{P}\left( \{ \exists j \in A_{\rho[q]-1} \cap [k_v[q]^*] : V_{k_v[q]}^L \leq 1 \mid C_q \right) \right]
\]
\[
\leq \mathbb{E}\left[ 1_{E_q} |A_{\rho[q]-1}| \max_{j \in [k_v[q]-1]} \lambda_j^\rho[q] \right].
\] (52)

We then observe that for all \( j \in [k_v[q]^*] \),
\[
\lambda_j^\rho[q] \leq \frac{\beta p}{2 \varepsilon a_v[q] n} \left( \frac{j + 1}{n} \right)^{(1-p)/(2\varepsilon)} \leq \frac{\beta p}{2 \varepsilon a_v[q] n} \left( \frac{k_v[q]}{n} \right)^{(1-p)/(2\varepsilon)} \leq \frac{\beta p}{\varepsilon n} \frac{(1-p)/(2\varepsilon) - 1}{a_v[q]}
\] where the final inequality holds on the event \( E_q \) for all sufficiently large \( n \). Regardless of the value of \( \varepsilon \in (0, 1/2) \), we can bound \( O_v[q]^{(1-p)/(2\varepsilon) - 1} \) from above by \( (\log \log n)^{q+1} \) on the event \( E_q \). Using this upper bound in (52) and bounding \( |A_{\rho[q]-1}| = O\left( (\log \log n)^{q/r} \right) \) on \( E_q \), this finally leads to
\[
\mathbb{P}\left( \{ \exists j \in A_{\rho[q]-1} \cap [k_v[q], n] : I_{j-\tau_{k_v[q]}^L}^L = 1 \} \cap \{ I_{k_v[q]}^L = V_{k_v[q]}^L \} \mid E_q \right) = O\left( \frac{(\log \log n)^{q/r}}{n} \right).
\]

Combined with (51), this yields the desired result. \( \square \)

5.2 Coupling the \( R \)-children of \( v[q] \) in \( \mathcal{T} \) with the \( R \)-neighbours of \( k_v[q] \) in \( G_n \)

In this section we couple the \( R \)-neighbours of \( k_v[q] \) and the \( R \)-children of \( v[q] \). We define
\[
\mathcal{R}_v[q] := \{ \theta_v[q] = T_v[q] \} \cap \{ \forall i \in [T_v[q]] : a_v[q], i = \frac{k_v[q] + 1}{n} \leq \frac{q + 1}{n} \text{ and } k_v[q], i \in \mathcal{N}_{v[q]-1} \}.
\] (53)
to be the event that the coupling of the R-neighbours of \( k_{v[q]} \) and the R-children of \( v[q] \) is successful. In words, the first two parts of \( \mathcal{R}_{\rho[q]} \) imply that the number of R-neighbours of \( k_{v[q]} \) and the R-children of \( v[q] \) are the same, and that their respective rescaled labels and types are sufficiently close. The last part ensures that the exploration of the L-neighbours of \( k_{v[q]} \) of does not lead to cycles.

We now define an encoding of the R-neighbours of \( k_{v[q]} \) in \( G_n \) and the R-children of \( v[q] \) in \( T \) that allows us to construct a coupling such that \( \mathcal{R}_{\rho[q]} \) holds with high probability. As in Section 5.1, we cannot ensure that \( k_{v[q]} + 1 \geq a_{v[q]} \) holds on the event \( \mathcal{E}_q = \cap_{i=1}^3 \mathcal{H}_{q,i} \). Instead, we let \( N_{v[q]} := \lceil a_{v[q]} \rceil \) and \( k_q^* := \min\{N_{v[q]}, k_{v[q]} + 1\} \) and use \( k_q^* \) in constructing the necessary processes that we couple further on in this section.

If \( N_{v[q]} \leq k_{v[q]} + 1 \), we define conditionally on \( a_{v[q]} \)

\[
\lambda^{\rho[q]}_{N_{v[q]}} := \int_{a_{v[q]}}^{N_{v[q]}/n} \frac{\beta p_x}{2 \varepsilon} \lambda^{1-p}/(2\varepsilon) - 1 \, dx, \text{ and for all } j \in \{N_{v[q]} + 1, \ldots, n\}, \quad \lambda^j_{N_{v[q]}} := \int_{j-1}^{j} \frac{\beta p_x}{2 \varepsilon} \lambda^{1-p}/(2\varepsilon) - 1 \, dx.
\]

If instead \( N_{v[q]} \geq k_{v[q]} + 2 \), we additionally set \( \lambda^j_{N_{v[q]}} := 0 \) for all integers \( j \in [k_q^*, N_{v[q]} - 1] \). This defines \( \lambda^j_{N_{v[q]}} \) for all \( j \in [k_q^*, n] \) and allows us to define the discretised Poisson point process \( V^R_{k_{v[q]}} := (V^R_{j \to k_{v[q]}}^q)^{n}_{j=k_q^*} \), where the \( V^R_{j \to k_{v[q]}}^q \) are independent Poisson random variables with means \( \lambda^j_{N_{v[q]}} \) (again, a Poisson variable with mean zero is equal to zero almost surely). Note that \( V^R_{k_{v[q]}} \) is a discretisation of the Poisson point process on \( (a_{v[q]}, 1) \) with density \( \lambda^+ (dz) \) that determines the R-children of \( v[q] \) (see Definition 1.2).

Similarly, we define the Bernoulli point process \( I^R_{k_{v[q]}} := (I^R_{j \to k_{v[q]}})^{n}_{j=k_q^*} \) as follows. For all \( j \in [k_q^*, n] \), \( I^R_{j \to k_{v[q]}} \) equals one if \( j \) connects to \( k_{v[q]} \) by an edge in \( G_n \) and \( j \in N_{\rho[q]} - 1 \cup A_{\rho[q]} - 1 \), and equals zero otherwise. If \( N_{v[q]} < k_{v[q]} + 1 \), for \( j \in [k_q^*, k_{v[q]}] \) we additionally set \( I^R_{j \to k_{v[q]}} := 0 \).

Our aim is to couple the Bernoulli point process \( I^R_{k_{v[q]}} \) and the discretised Poisson point process \( V^R_{k_{v[q]}} \). To do this, we show the following lemma, which is similar in nature to Lemma 4.5.

**Lemma 5.6.** Recall the event \( \mathcal{E}_q := \cap_{i=1}^3 \mathcal{H}_{q,i} \) where the events \((\mathcal{H}_{q,i})_{i\in[3]}\) were defined in (37), and consider the processes \( I^R_{k_{v[q]}} \) and \( V^R_{k_{v[q]}} \). There exists a coupling such that

\[
\mathbb{P} \left( \left( \{I^R_{k_{v[q]}} = V^R_{k_{v[q]}} \} \cap \{\forall j \in A_{\rho[q]-1} \cap [k_q^*, n]: V^R_{j \to k_{v[q]}} = 0\} \right) \cap \mathcal{E}_q \right) = O \left( \frac{(\log n)^2}{n^{1/3}} \right).
\]

**Remark 5.7.** When the coupling is successful, in the sense that the event

\[
\{I^R_{k_{v[q]}} = V^R_{k_{v[q]}} \} \cap \{\forall j \in A_{\rho[q]-1} \cap [k_q^*, n]: V^R_{j \to k_{v[q]}} = 0\}
\]

holds, it directly follows that the event \( \mathcal{R}_{\rho[q]} \) holds as well. As a result,

\[
\mathbb{P} \left( \mathcal{R}^c_{\rho[q]} \cap \mathcal{E}_q \right) = O \left( \frac{(\log n)^2}{n^{1/3}} \right).
\]

**Proof of Lemma 5.6.** We write

\[
\mathbb{P} \left( \left( \{I^R_{k_{v[q]}} = V^R_{k_{v[q]}} \} \cap \{\forall j \in A_{\rho[q]-1} \cap [k_q^*, n]: V^R_{j \to k_{v[q]}} = 0\} \right) \cap \mathcal{E}_q \right) = \mathbb{P} \left( \left( \{I^R_{k_{v[q]}} \neq V^R_{k_{v[q]}} \} \cap \mathcal{E}_q \right) + \mathbb{P} \left( \exists j \in A_{\rho[q]-1} \cap [k_q^*, n]: V^R_{j \to k_{v[q]}} = 1 \right) \cap \{I^R_{k_{v[q]}} = V^R_{k_{v[q]}} \} \right) \cap \mathcal{E}_q
\]

and start by bounding the first term on the right-hand side. As in the proof of Lemma 4.5, we use several intermediate steps to couple \( I^R_{k_{v[q]}} \) and \( V^R_{k_{v[q]}} \).

1. We define \( \bar{I}^R_{k_{v[q]}} := \{I^R_{j \to k_{v[q]}}^n \}_{j=k_q^*} \) as a sequence of independent Bernoulli random variables with success probabilities \( \{\mu_j\}_{j=k_q^*} \) defined in (30) and couple \( I^R_{k_{v[q]}} \) with \( \bar{I}^R_{k_{v[q]}} \).
2. We couple $\tilde{I}^R_{v[q]}$ with a sequence of independent Poisson random variables $\tilde{V}^R_{k[q]} := (\tilde{V}_{j \to k[q]}^R)_{j = k[q]}^n$ with means $(\mu_j)_{j = k[q]}^n$.

3. We couple $\tilde{V}^R_{k[q]}$ with $V^R_{k[q]}$.

Throughout the proof, we assume that $N_v[q] \geq k_v[q] + 2$ so that $k_v^* = k_v[q] + 1$. The cases $N_v[q] = k_v[q] + 1$ and $N_v[q] < k_v[q] + 1$ can be dealt with in a similar way with only minor modifications. We now provide details for the three steps.

Step 1. We start by observing that for all $j \in [k_v^*, n] \cap \mathcal{P}_{\rho[q]-1}$, $I^R_{j \to k_v[q]} = 0$ almost surely. We first bound the probability that $\tilde{I}^R_{j \to k_v[q]} \neq I^R_{j \to k_v[q]}$ for any $j \in [k_v^*, n] \cap \mathcal{P}_{\rho[q]-1}$, which occurs when the former is equal to 1. By a union bound,

$$\mathbb{P} \left( \bigcup_{j = k_v^*}^{n} \left\{ \tilde{I}^R_{j \to k_v[q]} = 1 \right\} \cap \mathcal{E}_q \right) = \mathbb{E} \left[ \chi_{\mathcal{E}_q} \mathbb{P} \left( \bigcup_{j = k_v^*}^{n} \left\{ \tilde{I}^R_{j \to k_v[q]} = 1 \right\} \bigg| \mathcal{C}_q \right) \right]$$

$$\leq \mathbb{E} \left[ \chi_{\mathcal{E}_q} \sum_{j = k_v^*}^{n} \frac{\beta p \left( \frac{j}{n} \right) (1-p)/(2e)-1}{2e n} \right]$$

$$\leq \mathbb{E} \left[ \chi_{\mathcal{E}_q} \frac{\beta p \left| \mathcal{P}_{\rho[q]-1} \right|}{2e k_v[q]} \right] = O \left( \frac{(\log \log n)^{q+1+q/r}}{n} \right),$$

where in the last inequality we used that $\frac{1}{n} \left( \frac{j}{n} \right) (1-p)/(2e)-1 \leq \frac{1}{2}$ for all $j \leq n$, and the last equality is due to the bounds on $k_v[q]$ and $\left| \mathcal{P}_{\rho[q]-1} \right|$ on the event $\mathcal{E}_q$. For the remaining indices, we use a standard Bernoulli coupling to obtain

$$\mathbb{P} \left( \bigcup_{j \in \mathcal{P}_{\rho[q]-1}}^{n} \left\{ \tilde{I}^R_{j \to k_v[q]} \neq I^R_{j \to k_v[q]} \right\} \cap \mathcal{E}_q \right) \leq \mathbb{E} \left[ \chi_{\mathcal{E}_q} \sum_{j \in \mathcal{P}_{\rho[q]-1}}^{n} \frac{\beta p \left( \frac{j}{n} \right) (1-p)/(2e)-1}{2e n} \right] - \mathbb{P} \left( I^R_{j \to k_v[q]} = 1 \bigg| \mathcal{C}_q \right) \right] \right)$$

As in (32), we show that the absolute value is small by providing a precise asymptotic value for the probability. Using Lemmas 2.6 and 2.5 with $S = \mathcal{A}_{\rho[q]-1,j} \cup \mathcal{P}_{\rho[q]-1,j}$ and $R = \{j\}$ (note that $j \geq k_v[q] = \omega((\log n)^3)$ and $|S| = o(j^{2/3})$ on the event $\mathcal{E}_q$), we arrive at

$$\mathbb{P} \left( \bigcup_{j \in \mathcal{P}_{\rho[q]-1}}^{n} \left\{ \tilde{I}^R_{j \to k_v[q]} \neq I^R_{j \to k_v[q]} \right\} \cap \mathcal{E}_q \right) \leq \mathbb{E} \left[ \chi_{\mathcal{E}_q} \sum_{j = k_v^*}^{n} \mathcal{O} \left( \log \frac{n}{j^{1/3}} \right) \right] = \mathbb{E} \left[ \chi_{\mathcal{E}_q} \mathcal{O} \left( (k_v^*)^{-1/3} \log n \right) \right] = O \left( \frac{(\log n)^2}{n^{1/3}} \right),$$

Combining this with the bound in (56), we get

$$\mathbb{P} \left( \bigcup_{j = k_v^*}^{n} \left\{ \tilde{I}^R_{j \to k_v[q]} \neq I^R_{j \to k_v[q]} \right\} \cap \mathcal{E}_q \right) \leq \mathbb{E} \left[ \chi_{\mathcal{E}_q} \mathcal{O} \left( \log n \right) \right] = \mathbb{E} \left[ \chi_{\mathcal{E}_q} \mathcal{O} \left( (k_v^*)^{-1/3} \log n \right) \right] = O \left( \frac{(\log n)^2}{n^{1/3}} \right),$$

where the last equality holds by the lower bound on $k_v^*$ ensured by the event $\mathcal{E}_q$. Combining this with (55) finishes the first step.

Step 2. This step is identical to step 2 in the proof of Lemma 4.5 and we only state the main result:

$$\mathbb{P} \left( \tilde{I}^R_{k_v[q]} \neq \tilde{V}^R_{k_v[q]} \right) \cap \mathcal{E}_q) = O \left( \frac{(\log \log n)^{q+1}}{n} \right).$$
Step 3. We couple each $V^R_{j \rightarrow k_{v[q]}}$ and $V^R_{j \rightarrow k_{v[q]}}$ via a standard coupling between Poisson random variables (see Lemma 2.3). Recall that we assume that $N_{v[q]} \geq k_{v[q]} + 2$, so that $\lambda^j_{v[q]} = 0$ for all $j \in [k_{v[q]}^*, N_{v[q]} - 1]$. Also, we define $\mu^j_{v[q]}$ as in (30) but for $j \in [k_{v[q]}^*, n]$ instead. With steps equivalent to the ones in (34), it follows that for all $j \in [N_{v[q]}, n]$,

$$|\lambda^j_{v[q]} - \mu^j_{v[q]}| = O\left(\frac{1}{n^2}\left(\frac{j}{n}\right)^{(1-p)/(2\varepsilon)-2}\right).$$

Moreover, for all $j \in [k^*_{v[q]}, N_{v[q]} - 1]$, $\mu^j_{v[q]} \leq \frac{\beta p}{2\varepsilon j} \leq \frac{\beta p}{2\varepsilon k_{v[q]}^*}$ again. Again, we write

$$P\left(\{\tilde{V}^R_{k_{v[q]}} \neq V^R_{k_{v[q]}}\} \cap \mathcal{E}_q\right) = \mathbb{E}\left[\mathbb{I}_{\mathcal{E}_q} P\left(\tilde{V}^R_{k_{v[q]}} \neq V^R_{k_{v[q]}} | \mathcal{C}_q\right)\right],$$

and the previous observations imply that

$$P\left(\tilde{V}^R_{k_{v[q]}} \neq V^R_{k_{v[q]}} | \mathcal{C}_q\right) \leq \sum_{j=1}^{N_{v[q]}-1} \mu^j_{v[q]} + \sum_{j=N_{v[q]}}^{n} |\lambda^j_{v[q]} - \mu^j_{v[q]}| \leq \sum_{j=k_{v[q]}^*}^{N_{v[q]}-1} \frac{\beta p}{2\varepsilon k_{v[q]}^*} + O\left(\sum_{j=N_{v[q]}}^{n} \frac{1}{n^2}\left(\frac{j}{n}\right)^{(1-p)/(2\varepsilon)-2}\right) \leq \left([a_{v[q]} n] - 1 - k_{v[q]}\right)\frac{\beta p (\log \log n)^q + 1}{2\varepsilon n} + O\left(\frac{1}{n} \int_{N_{v[q]}/n}^{1} x^{(1-p)/(2\varepsilon)-2} \, dx\right).$$

Note that on the event $\mathcal{E}_q$ one has $|a_{v[q]} n - k_{v[q]}| \leq q + 1$, and the integral in the $\mathcal{O}$ satisfies:

- if $\frac{1-p}{2\varepsilon} > 1$, its value is $\mathcal{O}(1)$;
- if $\frac{1-p}{2\varepsilon} = 1$, its value is $\mathcal{O}(\log n)$;
- if $\frac{1-p}{2\varepsilon} < 1$, its value is $\mathcal{O}\left(\left(\frac{N_{v[q]} - 1}{n}\right)^{(1-p)/(2\varepsilon)-1}\right)$. 

Hence, on the event $\mathcal{E}_q$, in all three cases we obtain the upper bound

$$P\left(\{\tilde{V}_{k_{v[q]}} \neq V_{k_{v[q]}}\} \cap \mathcal{E}_q\right) = \mathcal{O}\left(\frac{(\log \log n)^{q+1}}{n}\right).$$

Combining steps 1 through 3 shows that

$$P\left(\{I_{k_{v[q]}} \neq V_{k_{v[q]}}\} \cap \mathcal{E}_q\right) = \mathcal{O}\left(\frac{(\log n)^2}{n^{1/3}}\right).$$

(57)

It remains to bound from above the second term on the right-hand side of (54). On the occurrence of a successful coupling, it follows from Markov’s inequality that

$$P\left(\{\exists j \in A_{\rho[q] - 1} \cap [k_{v[q]} + 1, n] : I_{j \rightarrow k_{v[q]}} \geq 1\} \cap \{I_{k_{v[q]}} = V_{k_{v[q]}}\} \cap \mathcal{E}_q\right)$$

$$\leq \mathbb{E}\left[\mathbb{I}_{\mathcal{E}_q} P\left(\exists j \in A_{\rho[q] - 1} \cap [k_{v[q]} + 1, n] : V_{j \rightarrow k_{v[q]}} \geq 1 | \mathcal{C}_q\right)\right] \leq \mathbb{E}\left[\mathbb{I}_{\mathcal{E}_q} |A_{\rho[q] - 1}| \max_{j \in [k_{v[q]} + 1, n]} |\lambda^j_{v[q]}|\right].$$

Moreover, on the event $\mathcal{E}_q$ we have that for every $j \in [k_{v[q]} + 1, n]$,

$$|A_{\rho[q] - 1}| \lambda^j_{v[q]} = \mathcal{O}\left(\frac{(\log \log n)^q}{j}\right) = \mathcal{O}\left(\frac{(\log \log n)^q}{k_{v[q]}}\right) = \mathcal{O}\left(\frac{(\log \log n)^{q+1}}{n}\right),$$

Combined with (54) and (57), this yields the desired result and concludes the proof.
5.3 Proof of Lemma 5.3

In this section we prove Lemma 5.3 using the coupling of the immediate neighbours of \( k_{v[q]} \) in \( G_n \) and \( v[q] \) in \( T \) constructed in the previous sections.

Proof of Lemma 5.3. Observe that conditionally on \( \mathcal{E}_q = \bigcap_{i=1}^{3} \mathcal{H}_{q,i} \), \( \mathcal{K}_{\rho[q]} \) is implied by \( \mathcal{L}_{\rho[q]} \cap \mathcal{R}_{\rho[q]} \), where we recall \( \mathcal{L}_{\rho[q]}, \mathcal{R}_{\rho[q]} \) from (45) and (53), respectively. It thus follows from Lemmas 5.4 and 5.6 that

\[
\mathbb{P} \left( \mathcal{K}_{\rho[q]}^c \cap \mathcal{E}_q \right) \leq \mathbb{P} \left( \mathcal{L}_{\rho[q]}^c \cap \mathcal{E}_q \right) + \mathbb{P} \left( \mathcal{R}_{\rho[q]}^c \cap \mathcal{E}_q \right) = O \left( \frac{\log n}{n^{\alpha/6}} \right),
\]

as desired. \( \Box \)

5.4 Coupling the remaining vertices in \( B_q(G_n, k_0) \) and \( B_q(T, 0) \)

In the previous sections we covered the exploration of the neighbours of \( k_{v[q]} \), the smallest vertex in \( V(\partial B_q(G_n, k_0)) \) with respect to the BFS ordering, and coupled this process with the construction of the children of \( v[q] \) in \( T \). As discussed prior to this coupling, the same proofs (with only minor modifications) can be used to couple all the vertices in \( V(\partial B_q(G_n, k_0)) \) and in \( V(\partial B_q(T, 0)) \). The main difference is the number of proved and active vertices in the BFS exploration of \( G_n \). However, the bounds used in the proofs of the lemmas in this section still suffice for the coupling of all vertices.

5.5 Summary

With the coupling of the immediate neighbours of the \( L \)- and the \( R \)-neighbours in \( V(\partial B_q(G_n, k_0)) \) and \( V(\partial B_q(T, 0)) \) described in the previous sections, we are now ready to prove Proposition 5.1.

Proof of Proposition 5.1. We start by recalling the bounds in (38), (39) and (40) together with the hypothesis in Proposition 5.1, which yields

\[
\mathbb{P} \left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{q+1,i} \right)^c \right) \leq \mathbb{P} \left( \mathcal{H}_{q+1,1}^c \cap \mathcal{H}_{q+1,2} \cap \mathcal{H}_{q+1,3} \cap \bigcap_{i=1}^{3} \mathcal{H}_{q,i} \right) + 2\beta (\log \log n)^{q/r-p/(2\varepsilon)}
\]

for all sufficiently large \( n \). Recall \( \mathcal{K}_{t,q} \) from (41), \( \mathcal{K}_{\rho[q]} = \mathcal{K}_{\rho[q],q} \) from (44), and that for all \( t \in [\rho[q], \rho[q+1]-1] \) (with \( \rho[q] \) defined in (42)) we have \( k_{v(t)} = k[t] \). That is, \( k_{v(t)} \) is the vertex in \( \mathcal{A}_{t-1} \) with smallest BFS order whose immediate neighbours are to be explored at step \( t \) of the BFS exploration. With \( \mathcal{E}_q := \bigcap_{i=1}^{3} \mathcal{H}_{q,i} \), it follows from (43) that

\[
\mathbb{P} \left( \mathcal{H}_{q+1,1}^c \cap \mathcal{E}_q \right) = \mathbb{P} \left( \left( \bigcap_{t=\rho[q]}^{\rho[q+1]-1} \mathcal{K}_{t,q}^c \cap \mathcal{E}_q \right) \right) = \mathbb{E} \left[ \sum_{t=\rho[q]}^{\rho[q+1]-1} \mathbb{1}_{\mathcal{E}_q} \mathbb{1}_{\mathcal{H}_t} \mathbb{1}_{\mathcal{K}_{t,q}} \mathbb{P} \left( \mathcal{K}_{t,q}^c \big| \overline{C_t} \right) \right].
\]

As discussed after (44) and in Section 5.4, we can bound

\[
\mathbb{1}_{\mathcal{E}_q} \mathbb{1}_{\mathcal{K}_{t,q}} \mathbb{P} \left( \mathcal{K}_{t,q}^c \big| \overline{C_t} \right)
\]

in the same manner as we bound \( \mathbb{1}_{\mathcal{E}_q} \mathbb{P} \left( \mathcal{K}_{t,q}^c \big| \mathcal{C}_t \right) \) in Lemma 5.3 for any \( t \in [\rho[q], \rho[q+1]-1] \). Moreover, conditionally on \( \mathcal{E}_q \cap \bigcap_{s=\rho[q]}^{t-1} \mathcal{K}_{s,q} \) it follows that \( \rho[q+1] - \rho[q] \leq q(\log \log n)^{q/r} \), so we obtain the upper bound

\[
\mathbb{P} \left( \mathcal{H}_{q+1,1}^c \cap \bigcap_{i=1}^{3} \mathcal{H}_{q,i} \right) = O \left( \frac{(\log \log n)^{q/r} \log n}{n^{\alpha/6}} \right).
\]
Using this bound in (58) finally yields
\[ \mathbb{P} \left( \left( \bigcap_{i=1}^{3} \mathcal{H}_{q+1,i} \right)^c \right) = \mathcal{O}(\log \log n)^{q/p}/(2\varepsilon), \]
which concludes the proof.

With Proposition 5.1 and, in particular, Corollary 5.2 at hand, we can prove Theorem 1.3.

Proof of Theorem 1.3. Fix any \( r \in \mathbb{N} \) and use the coupling of \( B_r(G_n, k_0) \) and \( B_r(T, 0) \) from Corollary 5.2. We readily have that
\[ d_{TV}(\mathcal{L}(B_r(G_n, k_0)), \mathcal{L}(B_r(T, 0))) \leq \mathbb{P}(B_r(G_n, k_0) \not\equiv B_r(T, 0)) = \mathcal{O}(\log \log n)^{-(1-p)/(2\varepsilon)-1/r}, \]
and hence \((T, 0)\) is the local weak limit of \((G_n, k_0)\). \( \square \)

6 The DRGVR model as an inhomogeneous random graph

In this section we compare our model to an inhomogeneous version of the Erdős-Rényi random graph and then provide some applications. Such a comparison method was previously introduced and used by Chung and Lu in [6] to compare preferential attachment models with vertex and edge removal to inhomogeneous random graphs. Here, we are able to provide a closer comparison due to the more tractable dynamics of the DRGVR model, which allows us to obtain precise results.

Recall the notations \( V_n, \beta \) and \( \varepsilon \). We define the following inhomogeneous random graph model.

Definition 6.1 (Birth-death inhomogeneous Erdős-Rényi graph). Given a constant \( \Delta > 0 \) and an integer \( n \), the Birth-death inhomogeneous Erdős-Rényi graph, denoted \( G_{BD}(\Delta, n) \), has vertex set \( V_n \) and for all \( i, j \leq |V_n| \) satisfying \( i < j \), the edge between the \( i^{th} \) and the \( j^{th} \) vertex in \( V_n \) is sampled independently of other edges with probability
\[ p_{ij} = p_{ij}(\Delta) = \Delta \beta(2\varepsilon)^{-(1-p)/p} \left( \frac{1}{n} \frac{j}{n} \right)^{-2\varepsilon/p}. \] (59)

Remark that the above definition is done conditionally on \( V_n \). Moreover, on the event \( Q_n = \{|V_n| - \mathbb{E}|V_n|\leq n^{2/3}\} \) (which holds w.h.p. by Lemma 2.5), (59) rewrites
\[ p_{ij} = (1 + o(1)) \frac{\Delta \beta}{|V_n|} \left( \frac{j}{|V_n|} \right)^{-2\varepsilon/p}. \] (60)

The motivation of introducing this Erdős-Rényi-type model is that w.h.p. one may “sandwich” \( G_n \) between \( G_{BD}(1 - \delta_n, n) \) and \( G_{BD}(1 + \delta_n, n) \), where \((\delta_n)_{n \geq 1}\) is a sequence satisfying \( \delta_n = o(1) \), while at the same time the vertices of these graphs “carry no randomness” in the sense that, as opposed to the DRGVR, these do not have random arrival times and the birth-death process serving to define other edges with probability \( \leq |i,j| \) and hence \( \delta_n = 0 \).

Lemma 6.2. There is a sequence \((\delta_n)_{n \geq 1}\) such that \( \delta_n = o(1) \) and w.h.p
\[ G_{BD}(1 - \delta_n, n) \subseteq G_n \subseteq G_{BD}(1 + \delta_n, n). \]

Proof. Denote \( \{v_1, \ldots, v_{|V_n|}\} = V_n \) and recall the integer \( \ell = \ell(n) \) and the sequence \((\delta_i)_{i=1}^{\infty}\) from Part (ii) of Lemma 2.7, and the event \( Q_i = \{|V_i| - 2\varepsilon i \leq i^{2/3}\} \). We prove the lemma for \( \delta_n = \delta_n + \max\{\delta_n^{1/2}, (\log n)^{-1}\} \). Also, for any \( j \in \{|V_n|\} \), set
\[ r_j = \left[ (1 - \delta_n)n^{1-p/p} \left( \frac{j}{2\varepsilon} \right)^{2\varepsilon/p} \right] \quad \text{and} \quad R_j = \left[ (1 + \delta_n)n^{1-p/p} \left( \frac{j}{2\varepsilon} \right)^{2\varepsilon/p} \right], \]
and define the events \( W_1 = \{ v_1 \geq n^{(1-p)/(2p)} \} \), \( W_2 = \bigcap_{j=\ell}^{\lvert V_1 \rvert} \{ v_j \in [r_j, R_j] \} \), \( C = \bigcap_{i=[n(1-p)/(2p)]}^{n} Q_i \), and \( W = W_1 \cap W_2 \cap C \).

Then, \( W_1 \) and \( W_2 \) are both w.h.p. events by Lemma 2.7, and \( C \) as well by a computation similar to (9).

Moreover, note that each of \( W_1, W_2 \) and \( C \) is measurable with respect to the process \((V_i)_{i=1}^n\) (and hence \( W \) as well). Then, conditionally on \( W \), for all \( j \geq \ell + 1 \) and \( i < j \),

\[
    p_{ij}(1 - \hat{\delta}_n) = (1 - \hat{\delta}_n)\beta(2\varepsilon)^{-(1-p)/p} \frac{1}{n} \left( \frac{j}{n} \right)^{-2\varepsilon/p} \leq \frac{\beta}{2\varepsilon R_j + R_j^{2/3}} \leq \frac{\beta}{2\varepsilon r_j - r_j^{2/3}} \leq (1 + \hat{\delta}_n)\beta(2\varepsilon)^{-(1-p)/p} \frac{1}{n} \left( \frac{j}{n} \right)^{-2\varepsilon/p} = p_{ij}(1 + \hat{\delta}_n),
\]

where for the first inequality we used that

\[
    (1 - \hat{\delta}_n)(1 + \delta_n)(1 + R_j^{-1/3}/(2\varepsilon)) \leq (1 + \delta_n - \hat{\delta}_n)(1 + n^{\Omega(1)}) \leq (1 - (\log n)^{-1})(1 + n^{\Omega(1)}) \leq 1,
\]

for the last inequality we used that

\[
    (1 + \hat{\delta}_n)(1 - \delta_n)(1 - r_j^{-1/3}/(2\varepsilon)) \geq (1 + \max\{\delta_n^{1/2}, (\log n)^{-1}\} - \delta_n)(1 - n^{\Omega(1)}) \geq (1 + \max\{\delta_n^{1/2}, (\log n)^{-1}\}/2)(1 - n^{\Omega(1)}) \geq 1,
\]

and moreover the second and the penultimate inequality are justified by the fact that the sequences \((2\varepsilon r_j - r_j^{2/3})_{j \geq 1}\) and \((2\varepsilon R_j + R_j^{2/3})_{j \geq 1}\) are increasing and therefore the event \( C \subseteq W \) ensures that

\[
    2\varepsilon r_j - r_j^{2/3} = \min_{r_j \leq t \leq R_j} \{ 2\varepsilon t - t^{2/3} \} \leq \min_{r_j \leq t \leq R_j} \left| V_t \right| \leq \max_{r_j \leq t \leq R_j} \left| V_t \right| \leq \max_{r_j \leq t \leq R_j} \{ 2\varepsilon t + t^{2/3} \} = 2\varepsilon R_j + R_j^{2/3}.
\]

As usual, for every pair of vertices \( \{ v_i, v_j \} \), we couple the states of the edge \( \{ v_i, v_j \} \) in \( GB(1 - \hat{\delta}_n, n) \), \( G_n \) and \( GB(1 + \hat{\delta}_n, n) \) by sampling a random variable \( U_{i,j} \sim \text{Unif}[0, 1] \) and setting

\[
    \{ v_i, v_j \} \in GB(1 - \hat{\delta}_n, n) \iff U_{i,j} \leq p_{ij}(1 - \hat{\delta}_n); \quad \{ v_i, v_j \} \in G_n \iff U_{i,j} \leq \mathbb{P}(\{ v_i, v_j \} \in E(G_n) | W); \quad \{ v_i, v_j \} \in GB(1 + \hat{\delta}_n, n) \iff U_{i,j} \leq p_{ij}(1 + \hat{\delta}_n).
\]

By (61) this coupling ensures that, under the event \( W \), all edges with an endvertex among \( (v_j)_{j=\ell+1}^{\lvert V_1 \rvert} \) satisfy

\[
    \{ v_i, v_j \} \in GB(1 - \hat{\delta}_n, n) \subseteq \{ v_i, v_j \} \in G_n \subseteq \{ v_i, v_j \} \in GB(1 + \hat{\delta}_n, n).
\]

Of course, there are \( \ell^2 = (\log n)^{\Omega(1)} \) remaining edges. However, note that on the event \( W \) (and in particular, \( W_1 \)) each of them appears with probability \( n^{-\Omega(1)} \). Since \( \ell^2 n^{-\Omega(1)} = o(1) \), a union bound implies that w.h.p. none of these edges appears in any of \( GB(1 - \hat{\delta}_n, n) \), \( G_n \) and \( GB(1 + \hat{\delta}_n, n) \). Hence, on the event \( W \), we constructed a coupling of the three graphs that holds w.h.p. The fact that \( W \) holds w.h.p. itself finishes the proof of the lemma.

\[ \square \]

### 6.1 Applications of Lemma 6.2: proofs of Theorems 1.5 and 1.9 and Proposition 1.7

In this section we combine Lemma 6.2 with results from [5] to derive Theorems 1.5 and 1.9 and Proposition 1.7. We start with a brief summary of the notation and the theorems from [5] that we need in the sequel.
A closer look at the paper of Bollobás, Janson, and Riordan. To begin with, we introduce some notation from [5]. A ground space is a pair \((S, \mu)\) where \(S\) is a separable metric space and \(\mu\) is a Borel probability measure on \(S\). Also, a vertex space \(V\) is a triplet \((S, \mu, (x_n)_{n \geq 1})\), where \((S, \mu)\) is a ground space and \(x_n = (x^n_i)_{i=1}^n\) is a random sequence of \(n\) points of \(S\) such that for every measurable set \(A \subseteq S\) with boundary \(\partial A\) satisfying \(\mu(\partial A) = 0\),
\[
\frac{1}{n} \sum_{i=1}^n 1_{x^n_i \in A} \xrightarrow{P} \mu(A). \tag{62}
\]

A kernel \(\kappa\) on a ground space \((S, \mu)\) is a symmetric non-negative measurable function on \(S \times S\). By a kernel on a vertex space \((S, \mu, (x_n)_{n \geq 1})\) we mean a kernel on \((S, \mu)\). Also, given a vertex space \(V\), a kernel \(\kappa\) on \(V\) and a sequence \(x_n\), we denote by \(G^V(n, \kappa)\) the random graph with vertex set \(x_n\), and where the edge between \(x^n_i\) and \(x^n_j\) appears independently of other edges with probability \(\min\{\kappa(x^n_i, x^n_j)/n, 1\}\).

A kernel \(\kappa\) on a vertex space \(V = (S, \mu, (x_n)_{n \geq 1})\) is called graphical if the following conditions hold simultaneously:

1. \(\kappa\) is continuous almost everywhere on \(S \times S\);
2. \(\kappa\) is integrable over \(S \times S\) with respect to the product measure \(\mu \times \mu\),
3. \[
\frac{1}{n} \mathbb{E}[|E(G^V(n, \kappa))|] \xrightarrow{n \to \infty} \frac{1}{2} \int_{S \times S} \kappa(x, y)\mu(x)\mu(y). \tag{63}
\]

Moreover, for a kernel \(\kappa\) and a sequence of kernels \((\kappa_n)_{n \geq 1}\) on \(V\), we say that \((\kappa_n)_{n \geq 1}\) is graphical on \(V\) with limit \(\kappa\) if for almost every \((x, y) \in S \times S\)
\[
x_n \xrightarrow{n \to \infty} x \quad \text{and} \quad y_n \xrightarrow{n \to \infty} y \quad \text{in} \quad S \implies \kappa_n(x_n, y_n) \xrightarrow{n \to \infty} \kappa(x, y),
\]
\(\kappa\) satisfies conditions (1) and (2) above, and
\[
\frac{1}{n} \mathbb{E}[|E(G^V(n, \kappa_n))|] \xrightarrow{n \to \infty} \frac{1}{2} \int_{S \times S} \kappa(x, y)\mu(x)\mu(y). \tag{64}
\]

For a kernel \(\kappa\) on \((S, \mu)\), define \(T_\kappa\) as the integral operator defined by
\[
(T_\kappa f)(x) := \int_S \kappa(x, y)f(y)\mu(y)
\]
where \(f : S \to \mathbb{R}\) is any measurable function for which the integral is defined (finite or \(+\infty\)) for almost every \(x\), and let also
\[
||T_\kappa|| := \sup\{|||T_\kappa f||_2 : f \geq 0, ||f||_2 \leq 1\} \leq \infty.
\]
Then, a kernel \(\kappa\) is called subcritical if \(||T_\kappa|| < 1\) and supercritical if \(||T_\kappa|| > 1\). Moreover, a kernel is irreducible if the fact that \(\kappa\) is almost everywhere 0 on \(A \times (S \setminus A)\) for some measurable set \(A\) implies that \(\mu(A) = 0\) or \(\mu(S \setminus A) = 0\).

Finally, for a kernel \(\kappa\) on a ground space \((S, \mu)\) and a point \(x \in S\), consider the following multi-type Galton-Watson process \(X_x\). At the beginning, we mark a single point \(x \in S\). Then, at each step and every point \(y \in S\) marked at that step, unmark \(y\) and mark new points according to a Poisson process with density \(\kappa(y, z)\mu(z)\). Observe that the Poisson processes on \(S\) for different points \(y\) are independent, and the point configuration of all marked points at the next step is the union of all Poisson processes, associated to the points \(y \in S\) marked at the current step. Finally, denote by \(\rho(\kappa, x)\) the probability that the process \(X_x\) survives to eternity, and denote
\[
\rho(\kappa) := \int_S \rho(\kappa, x)\mu(x). \tag{65}
\]

One of the main results from [5] characterises the component structure of the graph \(G^V(n, \kappa)\). Denote by \(C_1 = C_1(V, \kappa_n)\) and \(C_2 = C_2(V, \kappa_n)\) the largest and the second-largest component in \(G^V(n, \kappa_n)\).
Theorem 6.3 (see Theorems 3.1 and 3.12 in [5]). Let \((\kappa_n)\) be a graphical sequence of kernels on a vertex space \(V\) with limit \(\kappa\).

1. If \(\|T_\kappa\| \leq 1\), then \(|C_1|/n\) converges in probability to 0. Moreover, if \(\kappa\) is subcritical, that is, \(\|T_\kappa\| < 1\), and \(\sup_{x,y,n} \kappa_n(x,y) < \infty\), then \(|C_1| = \mathcal{O}(\log n)\) w.h.p.

2. If \(\|T_\kappa\| > 1\), \(\kappa\) is irreducible and either \(\inf_{x,y,n} \kappa_n(x,y) > 0\) or \(\sup_{x,y,n} \kappa_n(x,y) < \infty\), then \(|C_1|/n\) converges in probability to \(\rho(\kappa)\), and moreover \(|C_2| = \mathcal{O}(\log n)\) w.h.p.

Another main result of [5] concerns the graph distance between two uniformly chosen vertices in \(G^V(n, \kappa_n)\).

Theorem 6.4 (Parts of Theorem 3.14 in [5]). Let \((\kappa_n)_{n \geq 1}\) be a graphical sequence of kernels on a vertex space \(V\) with limit \(\kappa\), where \(\|T_\kappa\| > 1\). Fix any constant \(\varepsilon \in (0,1)\). Denote by \(d\) the graph distance in \(G^V(n, \kappa_n)\), where \(n\) and \(\kappa_n\) are spared in the notation for convenience. If \(\kappa\) is irreducible, then

\[
\frac{1}{n^2} |\{v,w\} : d(v,w) < \infty \} | \xrightarrow{P} \frac{\rho(\kappa)^2}{2}.
\]

If moreover \(\|T_\kappa\| < \infty\), then

\[
\frac{1}{n^2} \left| \left\{ v,w : d(v,w) < (1 + \varepsilon) \frac{\log n}{\log \|T_\kappa\|} \right\} \right| \xrightarrow{P} \frac{\rho(\kappa)^2}{2},
\]

and if \(\sup_{x,y,n} \kappa_n(x,y) < \infty\), then also

\[
\frac{1}{n^2} \left| \left\{ v,w : d(v,w) < (1 - \varepsilon) \frac{\log n}{\log \|T_\kappa\|} \right\} \right| \xrightarrow{P} 0.
\]

Adaptation to our setting. Our next aim is to show that conditionally on \(V_n\), the random graph \(G_{BD}(1,n)\) is a particular case of the general setting described above. In fact, we work on the event \(Q_n = \{ |V_n - 2\varepsilon n| \leq n^{2/3} \}\), which allows us to use the more convenient expression (60) for \(p_{ij}\). Despite the fact that this choice leads to a certain rescaling (e.g. the sequence \(x_n\) has only \(|V_n|\) terms), no substantial modifications are needed to apply the theorems from [5].

First of all, for any \(i < j \leq |V_n|\), \(p_{ij}\) can be rewritten as

\[
p_{ij} = (1 + o(1)) \frac{\beta}{|V_n|} \left( \frac{\max\{i, j\}}{|V_n|} \right)^{-2\varepsilon/p} = (1 + o(1)) \frac{\beta}{|V_n|} \max\left\{ \frac{i}{|V_n|}, \frac{j}{|V_n|} \right\}^{-2\varepsilon/p}.
\]

In accordance with the above expression, we consider the vertex space \(V : (S, \mu, (x_n)_{n \geq 1})\) where \(S \equiv (0,1)\), \(\mu := \text{Unif}(S)\) and for every \(n \geq 1\), \(x_n = (x_n^i)_{i=1}^{V_n}\) is defined so that for all \(i \in [|V_n|]\), \(x_n^i \sim \text{Unif}((i-1)/|V_n|, i/|V_n|)]\). Let us show that this vertex space is well-defined in the sense of (62). Indeed, for any measurable set \(A \subseteq (0,1]\), the fact that \(\max_{i \in [|V_n|]} \mu(A \cap ((i-1)/|V_n|, i/|V_n|]) \rightarrow 0\) as \(n \rightarrow \infty\) and the weak law of large numbers imply that

\[
\frac{1}{|V_n|} \sum_{i=1}^{V_n} 1_{x_n^i \in A} = \frac{1}{|V_n|} \sum_{i=1}^{V_n} 1_{x_n^i \in A \cap ((i-1)/|V_n|, i/|V_n|]} \xrightarrow{P} \mu(A).
\]

Furthermore, for every \(n \geq 1\), consider a kernel

\[
\kappa_n : (x, y) \in (0,1] \times (0,1] \mapsto (1 + o(1)) \beta \max\left\{ \frac{|x|}{|V_n|}, \frac{|y|}{|V_n|} \right\}^{-2\varepsilon/p}.
\]

It is easy to verify that, up to choosing the error term appropriately, the kernel \(\kappa_n\) is the one that defines the probabilities \((p_{ij})_{i,j \in [|V_n|]}\). From the expression of \(\kappa_n\) it is natural to believe that \((\kappa_n)_{n \geq 1}\) is a graphical sequence on \(V\) with limit

\[
\kappa : (x, y) \in (0,1] \times (0,1] \mapsto \beta \max\{x, y\}^{-2\varepsilon/p}.
\]

The next result shows this formally.
Proof of Theorem 1.5. The sequence \((\kappa_n)_{n \geq 1}\) is graphical on \(\mathcal{V}\) with limit \(\kappa\).

Proof. Firstly, let \(x, y \in (0, 1]\) and let \((x_n)_{n \geq 1}\) and \((y_n)_{n \geq 1}\) be two sequences converging to \(x\) and \(y\), respectively. We have

\[
\left| \frac{|x_n| |V_n|}{|V_n|} - x \right| = \left| \frac{|x_n| |V_n|}{|V_n|} - x_n + (x_n - x) \right| \leq \frac{1}{|V_n|} + \frac{|x_n - x|}{|V_n|} \xrightarrow{n \to \infty} 0,
\]

and the same holds for \((y_n)_{n \geq 1}\) and \(y\). Distinguishing the cases \(x \neq y\) and \(x = y\) and using the continuity of the functions \(t \in (0, 1] \mapsto t^{-2\varepsilon/p}\) and \(\max(\cdot, \cdot)\) shows that \((\kappa_n)_{n \geq 1}\) and \(\kappa\) satisfy (63).

At the same time, clearly \(\kappa\) is continuous on \((0, 1] \times (0, 1]\) and

\[
\frac{1}{2} \int_{(0,1] \times (0,1]} \kappa(x,y) d\mu(x) d\mu(y) = \beta \int_{0 < x < y < 1} y^{-2\varepsilon/p} dy dx = \beta \int_{x=0}^{1} \frac{1 - x^{-2\varepsilon/p}}{1 - 2\varepsilon/p} dx = \beta p < \infty.
\]

On the other hand, conditionally on \(|V_n|\), the expected number of edges in \(G_{BD}(1, n)\) is

\[
\frac{1}{|V_n|} \sum_{j=1}^{n} j \beta \left( \frac{j}{|V_n|} \right)^{-2\varepsilon/p} = \frac{\beta}{|V_n|^2} \sum_{j=1}^{n} (j - 1) \left( \frac{j}{|V_n|} \right)^{-2\varepsilon/p} = (1 + o(1)) \beta p,
\]

which shows that (64) is verified and finishes the proof of the lemma.

The combination of Lemma 6.5 and Theorem 6.3 can then be used to prove Theorem 1.5. We are required to compare the quantities \(\rho(\kappa)\) given in (65), and \(\gamma\) given in (6), as well as the size of the giant components of \(G_n\) with the size of the giants in \(G_{BD}(1 - \delta_n, n)\) and \(G_{BD}(1 + \delta_n, n)\).

Proof of Theorem 1.5. By Theorem 6.3 it is sufficient to prove that \(\rho(\kappa) = \gamma\). Recall that in our setting \(\mathcal{S} = (0, 1]\), \(d\mu(z) = dz\) and \(\kappa(x,y) = \beta \max\{x,y\}^{-2\varepsilon/p}\). Moreover, the construction of the multi-type Galton-Watson process \(X_U\), where \(U \sim \mu\) is the uniformly chosen root of the branching process, describes the local weak limit of the graph sequence \((G^V(n, \kappa_n), o_n))_{n \geq 1}\), where \(o_n\) is chosen uniformly at random from \(\{x_i\}_{i=1}^{|V_n|}\). Note that in our case both \((G_{BD}(1 - \delta_n, n))_{n \geq 1}\) and \((G_{BD}(1 + \delta_n, n))_{n \geq 1}\) are graph sequences on the same vertex sets and whose kernels both converge to \(\kappa\). Therefore, each of these sequences converges locally to \(X_U\). At the same time, by Theorem 1.3 \((G_n)_{n \geq 1}\) converges locally to \(\mathcal{T}\), and by Lemma 6.2 \(G_{BD}(1 - \delta_n, n) \subseteq G_n \subseteq G_{BD}(1 + \delta_n, n)\) w.h.p. for a suitable coupling of the three graphs. Hence, \(\mathcal{L}(\mathcal{X}) = \mathcal{L}(\mathcal{T})\), and \(\rho(\kappa) = \gamma\) as survival probabilities of the two branching processes.

Remark 6.6. One may construct the process \(\mathcal{T}\) from \(X_U\) by applying the mapping \(x \mapsto x^{2\varepsilon/p}\) to all types of the particles of \(X_U\) (for our purposes it is more convenient to think of the latter process as with a random root \(U\) distributed according to \(\mu\), rather than with a deterministic root \(x \in \mathcal{S}\) and integrating w.r.t. \(d\mu(x)\) at the end). This yields \(a_0 = U^{2\varepsilon/p}\) for the type of the root in \(\mathcal{T}\). We then define the kernel \(\hat{\kappa}(y,z) := \frac{\beta p}{2\varepsilon} (1 - p/(2\varepsilon)) \max\{y, z\}^{-1}\) and observe that a particle \(v\) of \(\mathcal{T}\) with type \(a_v\) produces offspring independently of all other particles according to a Poisson point process on \((0,1]\) with density \(\hat{\kappa}(a_v, z)dz\) (see Definition 1.2). We then obtain for any \(0 \leq a \leq b \leq 1\),

\[
\int_{a^{2\varepsilon/p}}^{b^{2\varepsilon/p}} \kappa(U, w) d\mu(w) = \int_{a^{2\varepsilon/p}}^{b^{2\varepsilon/p}} \frac{\beta}{U^{2\varepsilon/p}} \max\{U^{2\varepsilon/p}/w, w^{2\varepsilon/p}\} dw = \int_{a}^{b} \frac{\beta p}{2\varepsilon} \max\{a_z, z\}^{-1} dz = \int_{a}^{b} \hat{\kappa}(a_0, z) dz.
\]

Hence, the Poisson point process on \((0,1]\) with density \(\kappa(U, w) d\mu(w)\) can be coupled to the Poisson point process on \((0,1]\) with density \(\hat{\kappa}(a_0, z)dz\) by applying the mapping \(x \mapsto x^{2\varepsilon/p}\) to each particle in the former point process. Applying this recursively for each particle in \(X_U\), we obtain \(\mathcal{T}\).
The irreducibility of \( \kappa \) is clear from the definition since \( \inf_{x,y} \kappa(x, y) = \beta > 0 \); however, we need to work a bit to distinguish the subcritical regime from the supercritical. Concerning the operator \( T_\kappa \), for all \( f : (0, 1] \to \mathbb{R} \) and \( x \in (0, 1] \) we have

\[
(T_\kappa f)(x) = \int_{y \in (0,1]} \kappa(x, y) f(y) d\mu(y) = (1 + o(1)) \beta \left( \int_{y \in (0, x]} x^{-2\varepsilon/p} f(y) dy + \int_{y \in [x, 1]} y^{-2\varepsilon/p} f(y) dy \right).
\]

Unfortunately, computing its norm directly leads to an optimisation problem that we are unable to solve explicitly. At best, we can provide bounds for the operator norm, as presented in the next lemma. Note that this result directly implies Proposition 1.7.

**Lemma 6.7.** For every \( p \in (1/2, 1] \) we have

\[
\beta \sqrt{\frac{p(1+4p)}{2-p}} \leq \beta \sup_{t \in (-1/2, \infty)} \sqrt{\frac{(1+2t)(2t^2 + 7t + 4 + 1/p)}{(1 + t)^2(t + 1/p)(2t + 2/p - 1)}} \leq \|T_\kappa\| \leq \beta \sqrt{\frac{p}{1-p}}. \tag{66}
\]

Moreover, the critical parameter \( \beta_c = \beta_c(p) \) is a non-increasing continuous function of \( p \) and satisfies

\[
\max \left\{ \sqrt{\frac{1-p}{p}}, \frac{1}{4} \right\} \leq \beta_c(p) \leq \inf_{t \in (-1/2, \infty)} \left( \frac{(1+2t)(2t^2 + 7t + 4 + 1/p)}{(1 + t)^2(t + 1/p)(2t + 2/p - 1)} \right)^{-1/2} \leq \sqrt{\frac{2}{p(1+4p)}}. \tag{67}
\]

**Proof.** First of all, given any \( f : (0, 1] \to \mathbb{R}^+ \) satisfying \( \|f\|_2 \leq 1 \), the Cauchy-Schwarz inequality implies

\[
\|T_\kappa f\|_2 = \sqrt{\int_{(0,1]} (T_\kappa f)(x)^2 dx} = \sqrt{\int_{(0,1]} \left( \int_{(0,1]} \kappa(x, y) f(y) dy \right)^2 dx} \leq \sqrt{\int_{(0,1]^2} \kappa(x, y)^2 dy dx}.
\]

For \( \varepsilon \in (0, 1/2 \setminus \{1/6\} \), the right-hand side can then be written as

\[
\beta \sqrt{2 \int_{0 < x < y < 1} \max\{x, y\}^{-4\varepsilon/p} dx dy} \leq \beta \sqrt{2 \int_{[0,1]} \frac{1 - x^{-4\varepsilon/p}}{1 - 4\varepsilon/p} dx} = \beta \sqrt{\frac{2}{1 - 4\varepsilon/p} - \frac{2}{(1 - 4\varepsilon/p)(2 - 4\varepsilon/p)}},
\]

which equals \( \beta \sqrt{p/(1-p)} \). Similar computations for \( \varepsilon = 1/6 \) result in \( \beta \sqrt{2} \), and hence \( \|T_\kappa f\|_2 \leq \beta \sqrt{p/(1-p)} \) for all \( \varepsilon \in (0, 1/2) \).

Now, note that the first inequality in (66) comes from the fact that the left-most expression is equal to \( \sqrt{\frac{(1+2t)(2t^2 + 7t + 4 + 1/p)}{(1 + t)^2(t + 1/p)(2t + 2/p - 1)}} \) when \( t = 0 \). Therefore, we only prove the more precise lower bound. Denote \( f_t : x \in (0, 1] \mapsto \sqrt{1 + 2t} x^t \in \mathbb{R} \) and note that \( \|f\|_2 = 1 \). We get

\[
\|T_\kappa f\|_2 = \beta \sqrt{(1 + 2t) \int_{(0,1]} \left( \int_{(0,1]} \max\{x, y\}^{-2\varepsilon/p} y^t dy \right)^2 dx},
\]

so that splitting the range of the inner integral into \((0, x]\) and \((x, 1]\) yields

\[
\beta \sqrt{(1 + 2t) \int_{(0,1]} \left( \int_{(0,x]} x^{-2\varepsilon/p} y^t dy + \int_{(x,1]} y^{-2\varepsilon/p} dy \right)^2 dx} = \beta \sqrt{(1 + 2t) \int_{(0,1]} \left( x^{1+t - 2\varepsilon/p} \frac{1}{1 + t} + \frac{1 - x^{1+t - 2\varepsilon/p}}{1 + t - 2\varepsilon/p} \right)^2 dx}.
\]

Expanding the squared term and integrating with respect to \( x \) then finally yields

\[
\beta \sqrt{\frac{1 + 2t}{(1 + t - 2\varepsilon/p)^2} \left( \frac{2\varepsilon/p}{1 + t} \right)^2 \frac{1}{3 + 2t - 4\varepsilon/p} - \frac{4\varepsilon/p}{(1 + t)(2 + t - 2\varepsilon/p) + 1}}.
\]
An immediate computation implies that
\[
\left(\frac{2\varepsilon/p}{1 + t}\right)^2 \frac{1}{3 + 2t - 4\varepsilon/p} - \frac{4\varepsilon/p}{(1 + t)(2 + t - 2\varepsilon/p)} + 1 = \frac{(2t^2 + 7t + 6 - 2\varepsilon/p)(1 + t - 2\varepsilon/p)^2}{(1 + t)^2(2 + t - 2\varepsilon/p)(3 + 2t - 4\varepsilon/p)},
\]
which together with the fact that \(2\varepsilon/p = 2 - 1/p\) implies that
\[
||T_\kappa f||_2 = \beta \sqrt{\frac{(1 + 2t)(2t^2 + 7t + 4 + 1/p)}{(1 + t)^2(t + 1/p)(2t + 2/p - 1)}}.
\]

Optimising over \(t \in (-1/2, \infty)\) proves (66), and since \(\beta_c\) is the value for which \(||T_\kappa|| = 1\), (67) follows immediately.

Furthermore, for any \(p_1, p_2 \in (1/2, 1)\) satisfying \(p_1 < p_2\), it is easy to check that the kernel for \(p_2\) dominates the kernel for \(p_1\) as a function over \((0, 1) \times (0, 1)\). From this, we conclude that for every measurable non-negative function \(f\) on \((0, 1)\) we have that \(||T_\kappa f||_2\) is a non-decreasing function of \(p\), and hence \(||T_\kappa||\) is also a non-decreasing function of \(p\) by definition. This yields that \(\beta_c\) is a non-increasing function of \(p\), and thus \(\beta_c(p) \geq \beta_c(1) = 1/4\) for all \(p \in (1/2, 1]\), which yields (67).

Finally, to show that \(\beta_c\) is continuous, we will show the continuity of the norm of the operator \(T_{\kappa,1}\), given by the operator for \(\beta = 1\), as a function of \(p\). Indeed, let \((p_n)_{n \geq 1}\) and \(p\) be real numbers in \((1/2, 1]\) such that \(p_n \to p\) as \(n \to \infty\). Moreover, let \(\mathcal{P} = \{p_n\}_{n \geq 1} \cup \{p\}\) and let \(f\) be a non-negative function in \(L_2((0,1))\). Then, to show that the norm of \(||T_{\kappa,1}f||\) for \(p\) is the limit of the norms of \(||T_{\kappa,1}f||\) for \(p_n\) as \(n \to \infty\), it is sufficient to show that
\[
\int_{(0,1]} \left(\int_{(0,1]} \max\{x,y\}^{-(2p-1)/p} f(y) dy \right)^2 - \left(\int_{(0,1]} \max\{x,y\}^{-(2p_n-1)/p} f(y) dy \right)^2 \right) dx
\]
tends to 0. Since the term in the integral is uniformly bounded by
\[
2\max_{q \in \mathcal{P}} \left(\int_{(0,1]} \max\{x,y\}^{-(2q-1)/q} f(y) dy \right)^2 \leq 2 \left(\int_{(0,1]} \max\{x,y\}^{-1} f(y) dy \right)^2,
\]
which is integrable since
\[
\int_{(0,1]} \left(\int_{(0,1]} \max\{x,y\}^{-1} f(y) dy \right)^2 dx
\]
is at most \(||T_{\kappa,1}||^2\) for \(p = 1\), which is finite. By the dominated convergence theorem, it suffices to show that
\[
\left(\int_{(0,1]} \max\{x,y\}^{-(2p-1)/p} f(y) dy \right)^2 - \left(\int_{(0,1]} \max\{x,y\}^{-(2p_n-1)/p} f(y) dy \right)^2
\]
tends to 0 as \(n \to \infty\) for every \(x \in (0,1]\), or equivalently that
\[
\int_{(0,1]} \left(\max\{x,y\}^{-(2p-1)/p} - \max\{x,y\}^{-(2p_n-1)/p} \right) f(y) dy
\]
tends to 0 as \(n \to \infty\) for every \(x \in (0,1]\). The above expression equals
\[
\int_{(0,x]} (x^{-(2p-1)/p} - x^{-(2p_n-1)/p}) f(y) dy + \int_{(x,1]} (y^{-(2p-1)/p} - y^{-(2p_n-1)/p}) f(y) dy.
\]
While the first term in the sum is a multiple of \( x^{-(2p-1)/p} - x^{-(2p_n-1)/p_n} \), which clearly tends to 0 when \( n \to \infty \) for any \( x \in (0, 1] \), the second term is dominated from above by

\[
\max_{y \in [x, 1]} \{ y^{-(2p-1)/p} - y^{-(2p_n-1)/p_n} \} f(y) \leq (x^{-(2p-1)/p} + x^{-(2p_n-1)/p_n}) f(y),
\]

which is integrable over the interval \((x, 1]\). Applying the dominated convergence theorem once again and using that \( y^{-(2p-1)/p} - y^{-(2p_n-1)/p_n} \) tends to zero as \( n \to \infty \) for every \( y \in (x, 1] \) finishes the proof of the lemma.

\[ \square \]

Finally, note that the limit kernel \( \kappa \) we obtained is unbounded and hence we could not apply Theorems 6.3 and 6.4 in their full generality. Nevertheless, \( \kappa(x, y) \) remains bounded if at least one of \( x \) and \( y \) is bounded away from 0. We use this observation to prove Theorem 1.9.

**Proof of Theorem 1.9.** Denote \( V_n^\lambda = V(G_n^\lambda) \), and let \( G^\lambda_{BD}(1, n) = G_{BD}(1, n)[V_n^\lambda] \) be the graph induced from \( G_{BD}(1, n) \) by the vertex set \( V_n^\lambda \). Define the kernels

\[ \kappa_n^\lambda : (x, y) \in [\lambda, 1] \times [\lambda, 1] \mapsto \kappa_n(x, y). \]

Note that the only difference between \( \kappa_n^\lambda \) and \( \kappa_n \) is the domain of definition. One may readily verify that \( \kappa_n^\lambda \) serves to define the graph \( G^\lambda_{BD}(1, n) \) in the same way as \( \kappa_n \) serves to define \( G_{BD}(1, n) \). Also, with minor modifications the proof of Lemma 6.5 shows that \( (\kappa_n^\lambda)_{n \geq 1} \) converges to \( \kappa^\lambda : (x, y) \in [\lambda, 1] \times [\lambda, 1] \mapsto \kappa(x, y) \). Since \( \kappa^\lambda \) is a bounded kernel, the second point of Theorem 1.9 follows immediately from Part (i) of Theorem 6.3.

Now, we show the third point of Theorem 1.9 for \( \zeta = ||T_n||^{-1} \). By Theorem 6.4 the only thing we need to show is that \( ||T_n\kappa^\lambda|| \) converges to \( ||T_n|| \) as \( \lambda \) decreases to zero. Fix any \( \delta \in (0, 1) \) and let \( f \) be a square integrable positive function over \((0, 1]\) such that \( ||T_n|| \leq (1 + \delta)||T_n f||_2 \). This means that, in particular,

\[
||T_n f||_2^2 = \int_{(0, 1]} \left( \int_{(0, 1]} \kappa(x, y) f(y) dy \right) dx = \int_{[0, 1]} \kappa(x, y_1) \kappa(x, y_2) f(y_1) f(y_2) dy_2 dy_1 dx < \infty.
\]

Now, by the monotone convergence theorem we get that

\[
||T_n\kappa^\lambda||_2^2 = \int_{[\lambda, 1]^3} \kappa^\lambda(x, y_1) \kappa^\lambda(x, y_2) f(y_1) f(y_2) dy_2 dy_1 dx
\]

converges to \( ||T_n f||_2^2 \) when \( \lambda \to 0 \).

Hence, for every sufficiently small \( \lambda > 0 \) we have that \( ||T_n|| \leq (1 + 2\delta)||T_n\kappa^\lambda||_2 \). On the other hand, for every \( \lambda > 0 \) we have that the kernel \( \kappa^\lambda \) may be extended to \( \hat{\kappa}^\lambda : (x, y) \in (0, 1] \times (0, 1] \mapsto \mathbb{1}_{x \in [\lambda, 1]} \mathbb{1}_{y \in [\lambda, 1]} \kappa(x, y) \), which is dominated by \( \kappa \) for any \( \lambda \in (0, 1] \), so \( ||T_n\hat{\kappa}^\lambda|| = ||T_n\kappa|| \leq ||T_n|| \). Hence, \( ||T_n\kappa^\lambda|| \) converges to \( ||T_n|| \) as \( \lambda \to 0 \), which together with Theorem 6.4 proves the third point.

Finally, we show the first point of Theorem 1.9. First of all, note that the local weak limit of \( G_n^\lambda \) exists and it constructed from \((T, 0)\) conditionally on the event \( \{a_0 > \lambda\} \) by pruning the tree by removing all vertices whose type is at most \( \lambda \) together with their descendants.

Let us denote the limit of \( (G_n^\lambda)_{n \geq 1} \) by \((T^\lambda, 0)\) and its survival probability by \( \gamma^\lambda \). Then, for any fixed \( r \geq 1 \), up to minor modifications the proof of Theorem 1.3 shows that the probability that the root 0 of \( \mathcal{T} \) has a vertex with age in the interval \([0, \lambda]\) at distance at most \( r \) tends to 0 as \( \lambda \to 0 \). Therefore, for every finite rooted tree \((T, o)\) we have that

\[
\mathbb{P}\left((T^\lambda, 0) \cong (T, o)\right) \to \mathbb{P}\left((T, 0) \cong (T, o)\right) \text{ as } \lambda \to 0.
\]
Recall $\gamma$ from (6) as well as that $\gamma = \rho(\kappa)$ by the proof of Theorem 1.5. Also, recall that $\mathcal{G}$ is the set of all finite rooted trees. We deduce that

$$|\gamma^\lambda - \gamma| = \left| \sum_{(T,o) \in \mathcal{G}} \mathbb{P}\left((T^\lambda,0) \cong (T,o)\right) - \mathbb{P}\left((T,0) \cong (T,o)\right) \right| \leq \sum_{(T,o) \in \mathcal{G}} \left| \mathbb{P}\left((T^\lambda,0) \cong (T,o)\right) - \mathbb{P}\left((T,0) \cong (T,o)\right) \right|.$$ 

Then, by the reverse Fatou’s lemma,

$$\limsup_{\lambda \to 0} |\gamma^\lambda - \gamma| \leq \sum_{(T,o) \in \mathcal{G}} \limsup_{\lambda \to 0} \left| \mathbb{P}\left((T^\lambda,0) \cong (T,o)\right) - \mathbb{P}\left((T,0) \cong (T,o)\right) \right| = 0,$$

and therefore $\gamma^\lambda \to \gamma = \rho(\kappa)$ as $\lambda \to 0$.

Now, recall the coupling constructed in Lemma 6.2 and the sequence $(\hat{\delta}_n)_{n \geq 1}$. Note that $G^\lambda_{BD}(1 + \hat{\delta}_n,n)$ may be constructed from $G^\lambda_n$ by adding every edge $\{v_i,v_j\}$ satisfying $v_i,v_j \in V_n \setminus V_{\lambda n}$ to $G^\lambda_n$ with probability

$$O\left(\frac{\hat{\delta}_n}{n} \left(\frac{\max\{i,j\}}{|V_n|}\right)^{-2\epsilon/p}\right),$$

while $G^\lambda_{BD}(1 - \hat{\delta}_n,n)$ may be constructed from $G_n$ by removing every edge of $G_n$ with probability $O(\hat{\delta}_n)$. At the same time, for any fixed $r \in \mathbb{N}$, by Theorem 1.3 and the definition of $(T,0)$ w.h.p. the ball of radius $r$ around a uniformly chosen vertex in $G_n$ contains $O(\hat{\delta}_n^{1/2})$ edges. Moreover, since $G^\lambda_n \subseteq G_n$ and $|V^\lambda_n|$ and $|V_n|$ are of the same order, w.h.p. the $r^{th}$ neighbourhood of a uniformly chosen vertex in $V^\lambda_n$ is the same in both $G^\lambda_{BD}(1 - \hat{\delta}_n,n)$ and $G^\lambda_{BD}(1 + \hat{\delta}_n,n)$. We conclude that each of $G^\lambda_{BD}(1 - \hat{\delta}_n,n)$, $G^\lambda_n$ and $G^\lambda_{BD}(1 + \hat{\delta}_n,n)$ converges locally to $(T^\lambda,0)$. With a similar argument as in the proof of Theorem 1.5, we conclude that $\rho(\kappa^\lambda) = \gamma^\lambda$. Hence, for every fixed $\delta' > 0$ we have that, by choosing $\lambda$ sufficiently small, $|\rho(\kappa^\lambda) - \rho(\kappa)| \leq \frac{\delta'}{4\epsilon}$ and therefore w.h.p. $|G^\lambda_1(V^\lambda_n)|/|V^\lambda_n| \in [\rho(\kappa) - \frac{\delta'}{4\epsilon},\rho(\kappa) + \frac{\delta'}{4\epsilon}]$. By combining Lemma 2.1 and Lemma 2.5 applied with $S = 0$, we conclude that $|V^\lambda_n|/n$ converges in probability to $2\epsilon(1 - \lambda^{\epsilon/(2\epsilon)})$. This proves the first part of Theorem 1.9 since $2\epsilon\rho(\kappa) = 2\epsilon\gamma$ and moreover w.h.p.

$$2\epsilon\gamma - \delta' \leq \frac{|V^\lambda_n|}{n} \left(\gamma - \frac{\delta'}{4\epsilon}\right) \leq \frac{|V^\lambda_n|}{n} \left(\gamma + \frac{\delta'}{4\epsilon}\right) \leq 2\epsilon\gamma + \delta'$$

for every sufficiently small $\lambda$. \hfill \Box

### 7 The maximum degree: proof of Theorem 1.10

In this section we study the maximum in-degree, out-degree and degree (that is, in-degree plus out-degree) of the graph $\overline{G}_n$, as stated in Theorem 1.10. Due to the deletion of vertices, the evolution of vertex degrees is quite different from the case $p = 1$ with no vertex deletion, as discussed after Theorem 1.10.

We let

$$d^+_n(v_i) := \sum_{j=i+1}^n \mathbb{1}_{\{v_i,v_j\} \in \overline{G}_n}, \quad \text{and} \quad d^-_n(v_i) := \sum_{j=1}^{i-1} \mathbb{1}_{\{v_j,v_i\} \in \overline{G}_n},$$

denote the in-degree and out-degree of vertex $v_i$ in $\overline{G}_n$, respectively, and let $d^e_n(v_i) := d^+_n(v_i) + d^-_n(v_i)$ denote the degree of vertex $v_i$. As in the previous section, Lemma 6.2 allows us to “sandwich” the vertex degrees of $\overline{G}_n$ between those of the graphs $\overline{G}^\square_{1,n} := \overline{G}_{RD}(1 - \hat{\delta}_n,n)$ and $\overline{G}^\square_{2,n} := \overline{G}_{BD}(1 + \hat{\delta}_n,n)$. For each of $\square \in \{s,+,-,\}$, we denote by $d^\square_{1,n}(v_i)$ and $d^\square_{2,n}(v_i)$ the degree (if $\square = s$), in-degree (if $\square = +$) or out-degree (if $\square = -$) of the vertex $v_i$ in $\overline{G}^\square_{1,n}$ and in $\overline{G}^\square_{2,n}$, respectively.
The proof of Theorem 1.10 makes use of the following preliminary result, which is similar in spirit to Lemma 1 from [9], combined with the “sandwiching” of the graphs $\tilde{G}_{1,n}$, $\tilde{G}_n$, and $\tilde{G}_{2,n}$. The main conclusion is that, roughly speaking, the maximum in-degree (respectively out-degree) of $\tilde{G}_n$ is approximately $k_n$ if the expected number of vertices of in-degree (respectively out-degree) $k_n$ is about 1.

**Lemma 7.1.** Let $(k_n)_{n \geq 1}$ be a sequence of positive numbers. Then, for $\Box \in \{+, -, \}$,

$$\mathbb{P}\left( \max_{i \in [\|V_n\|]} d^\Box_i(v_i) \geq k_n \bigm| |V_n| \right) = \begin{cases} 0 & \text{if } \lim_{n \to \infty} \sum_{i=1}^{\|V_n\|} \mathbb{P}\left( d^\Box_i(v_i) \geq k_n \bigm| |V_n| \right) = 0, \\ 1 & \text{if } \lim_{n \to \infty} \sum_{i=1}^{\|V_n\|} \mathbb{P}\left( d^\Box_i(v_i) \geq k_n \bigm| |V_n| \right) = \infty. \end{cases}$$

Moreover,

$$\mathbb{P}\left( \max_{i \in [\|V_n\|]} d^\Box_i(v_i) \geq k_n \bigm| |V_n| \right) = 0 \text{ if } \lim_{n \to \infty} \sum_{i=1}^{\|V_n\|} \mathbb{P}\left( d^\Box_i(v_i) \geq k_n \bigm| |V_n| \right) = 0.$$

**Proof.** In this proof everything is done conditionally on $|V_n|$. Firstly, by Lemma 6.2 we know that $\tilde{G}_{1,n} \subseteq \tilde{G}_n \subseteq \tilde{G}_{2,n}$ w.h.p. Hence, proving that the expected number of vertices of degree (or in-degree, or out-degree) $k_n$ in $\tilde{G}_{2,n}$ is $o(1)$ implies that w.h.p. $G_n$ contains no vertex of degree (or in-degree, or out-degree) $k_n$. Secondly, for both $\Box \in \{+, -, \}$ and every $i \in [\|V_n\|]$, define the event $\mathcal{D}^\Box_{i,n} := \{d^\Box_i(v_i) \geq k_n\}$. By independence of the states of the edges in $\tilde{G}_{1,n}$ one may conclude that the events $\mathcal{D}^\Box_{i,n}$ are independent for different values of $i$. Thus, when

$$\lim_{n \to \infty} \sum_{i=1}^{\|V_n\|} \mathbb{P}\left( \mathcal{D}^\Box_{i,n} \bigm| |V_n| \right) = \lim_{n \to \infty} \sum_{i=1}^{\|V_n\|} \mathbb{P}\left( d^\Box_i(v_i) \geq k_n \bigm| |V_n| \right) = \infty$$

holds, it follows that

$$\text{Var}\left( \sum_{i=1}^{\|V_n\|} 1_{\mathcal{D}^\Box_{i,n}} \bigm| |V_n| \right) = \sum_{i=1}^{\|V_n\|} \text{Var}(1_{\mathcal{D}^\Box_{i,n}} \bigm| |V_n|) \leq \mathbb{E}\left[ \sum_{i=1}^{\|V_n\|} 1_{\mathcal{D}^\Box_{i,n}} \bigm| |V_n| \right] \leq \mathbb{E}\left[ \sum_{i=1}^{\|V_n\|} 1_{\mathcal{D}^\Box_{i,n}} \bigm| |V_n| \right]^2.$$  

We conclude via Chebyshev’s inequality that at least one of $(\mathcal{D}^\Box_{i,n})_{i=1}^{\|V_n\|}$ holds w.h.p. and hence $\tilde{G}_{1,n}$ (and thus also $\tilde{G}_n$) contains a vertex of in–degree or out-degree $k_n$ w.h.p. \hfill $\Box$

**Proof of Theorem 1.10.** In this proof everything is done conditionally on $|V_n|$ and the event $\mathcal{Q}_n = \{|\|V_n\| - 2\pi n| \leq n^{2/3}\}$, which holds w.h.p. by Lemma 2.5. We remark that in several places in this proof $1 \pm \delta_n$ is somewhat abusively replaced by $\beta$; note that this abuse does not influence the proof since the limits in (3) and (4) are both continuous as functions of $\beta$. Define $W_0$ as the inverse of the function $f : x \in [-1, \infty) \rightarrow x e^x \in [-1/e, \infty)$ ($W_0$ is also known as the main branch of the Lambert $W$ function). In particular, $e^{W_0(\log n)} = \log n/W_0(\log n)$, and moreover by Theorem 2.7 in [16] we also have that

$$W_0(\log n) = \log \log n - \log \log \log n + o(1),$$

which implies that

$$\frac{\log n}{W_0(\log n)} = \frac{\log n}{\log \log n - \log \log \log n + o(1)} = \frac{\log n}{\log \log n} + \frac{\log n \log \log \log n}{(\log \log n)^2} + o\left( \frac{\log n}{(\log \log n)^2} \right).$$

(69)

Hence, for $\Box \in \{s, +\}$ it suffices to prove that

$$\frac{\max_{i \in [\|V_n\|]} d^\Box_i(v_i) - \log n/W_0(\log n)}{\log n/(\log \log n)^2} \xrightarrow{p} C_{\beta,p} := \left( 1 + \log \left( \frac{\beta p}{1 - p} \right) \right).$$

(70)
to obtain (3). In fact, we show that for every $\mu > 0$ we have that w.h.p.

$$C_{\beta,p} - \mu \leq \frac{\max_{i \in [|V_n|]} d^+_{i,n}(v_i) - \log n/W_0(\log n)}{\log n/(\log \log n)^2}$$

(71)

$$\leq \frac{\max_{i \in [|V_n|]} d^s_{2,n}(v_i) - \log n/W_0(\log n)}{\log n/(\log \log n)^2} \leq C_{\beta,p} + \mu.$$  

(72)

We first show (72). Fix $a_n := \log n/W_0(\log n) + (C_{\beta,p} + \mu) \log n/(\log \log n)^2$. Then, for every $i \in [|V_n|]$ and $t \geq 0$,

$$\mathbb{P}(d^s_{2,n}(v_i) \geq a_n) \leq \mathbb{E}\left[e^{t d^s_{2,n}(v_i)}\right] e^{-t a_n}$$

$$\leq e^{-t a_n} \prod_{j=1}^{i-1} \left(1 + (e^t - 1)(2\varepsilon)^{-(1-p)/p}\beta \left(\frac{i}{n}\right)^{-2e/p} \prod_{j=i+1}^{V_n} \left(1 + (e^t - 1)(2\varepsilon)^{-(1-p)/p}\beta \left(\frac{j}{n}\right)^{-2e/p}\right)\right).$$

Including the term $j = i$ in the first product and using that $1+x \leq e^x$ for all $x \in \mathbb{R}$ and $||V_n| - 2\varepsilon n| = n^{2/3}$ implies that

$$\mathbb{P}(d^s_{2,n}(v_i) \geq a_n) \leq \exp\left(-t a_n + (e^t - 1)\beta \left(\frac{i}{2\varepsilon n}\right)^{1-2e/p} + \frac{1 + O(n^{-1/3})}{(1 - 2\varepsilon/p)(2\varepsilon n)^{1-2e/p}}\right)$$

$$\leq \exp\left(-t a_n + (1 + O(n^{-1/3}))(e^t - 1)\beta \frac{p}{1-p}\right).$$

(73)

By setting $t = \log \left(\frac{(1-p)a_n}{\beta p}\right)$ and using that $\log \left(\frac{1-p}{\beta p}\right) = -\log \left(\frac{\beta p}{1-p}\right)$ we arrive at the upper bound

$$\mathbb{P}(d^s_{2,n}(v_i) \geq a_n) \leq \exp\left(-a_n \log a_n + C_{\beta,p}a_n - \frac{\beta p}{1-p} + o(1)\right).$$

(74)

Moreover, using (68), (69) and the equality $e^{W_0(\log n)} = \log n/W_0(\log n)$ we deduce that

$$a_n \log a_n$$

$$= \left(\frac{\log n}{W_0(\log n)} + \frac{(C_{\beta,p} + \mu) \log n}{(\log \log n)^2}\right) \log \left(\frac{\log n}{W_0(\log n)} + \frac{(C_{\beta,p} + \mu) \log n}{(\log \log n)^2}\right)$$

$$= \left(1 + \frac{(C_{\beta,p} + \mu + o(1)) W_0(\log n)}{\log \log n}\right) \log \left(e^{W_0(\log n)} + \frac{(C_{\beta,p} + \mu + o(1)) W_0(\log n)}{(\log \log n)^2}\right)$$

$$= \log n + \frac{(C_{\beta,p} + \mu + o(1)) \log n}{\log \log n}$$

$$= \log n + (C_{\beta,p} + \mu + o(1)) a_n.$$ (75)

Thus, combining (74), (75) and a union bound over all $|V_n| \leq n$ vertices in $\tilde{G}_{2,n}$, we get that

$$\mathbb{P}\left(\max_{i \in [|V_n|]} d^s_{2,n}(v_i) \geq a_n\right) \leq n \exp\left(-a_n \log a_n + a_n \log \left(\frac{\beta p}{1-p}\right) + a_n - \frac{\beta p}{1-p} + o(1)\right)$$

$$\leq \exp(- (\mu + o(1)) a_n) = o(1).$$

Before showing (71), let us turn our attention to the analogous upper bound in (4). Fix any $\mu > 0$ and set

$$\tilde{C}_{\beta,p} := 1 + \log \beta$$

and $c_n := \log n/W_0(\log n) + (\tilde{C}_{\beta,p} + \mu) \log n/(\log \log n)^2$.

(76)
Note that the out-degree of an arbitrary vertex in $\vec{G}_{2,n}$ is stochastically dominated by $d_{2,n}(v_{|V_n|})$. On the other hand, $d_{2,n}(v_{|V_n|})$ is a binomial random variable with distribution $\text{Bin}(|V_n| - 1, p_{1,|V_n|})$. Observe that on the event $\mathcal{Q}_n$, the mean of $d_{2,n}(v_{|V_n|})$ is $p_{1,|V_n|}|V_n| = (1 + \Omega(n^{-\Omega(1)}))\beta$. We conclude by the stronger version of Chernoff’s inequality (see Lemma 2.3 (i)) that with $\phi(x) = (1 + x) \log(1 + x) - x$,

$$\mathbb{P}\left(d_{2,n}(v_{|V_n|}) \geq c_n\right) \leq \exp \left(- (1 + \Omega(n^{-\Omega(1)}))\beta \phi\left(\frac{c_n}{(1 + \Omega(n^{-\Omega(1)}))\beta}\right)\right). \quad (77)$$

Since

$$\frac{c_n}{(1 + \Omega(n^{-\Omega(1)}))\beta} = \frac{\log n}{\beta \log \log n} + \frac{\log n \log \log \log n}{\beta (\log \log n)^2} + \frac{\hat{C}_{\beta,p} + \mu + o(1)}{\beta} \log n \frac{\log n}{(\log \log n)^2},$$

it follows that

$$\phi\left(\frac{c_n}{(1 + \Omega(n^{-\Omega(1)}))\beta}\right) = \frac{\log n}{\beta} + \frac{\mu + o(1)}{\beta} \log n \frac{\log n}{(\log \log n)}. $$

Together with (77) and a union bound this implies that

$$\mathbb{P}\left(\max_{i \in |V_n|} d_{2,n}(v_i) \geq c_n\right) \leq n \mathbb{P}\left(d_{2,n}(v_{|V_n|}) \geq c_n\right) = o(1). \quad (78)$$

Now, we come back to proving (71). The lower bound in (4) is shown in the same manner, so we mark the necessary modifications along the way. Recall $C_{\beta,p}$ and $\hat{C}_{\beta,p}$ from (70) and (76), respectively, and define

$$b_n^+ := \left[\frac{\log n}{W_0(\log n)} + \frac{(C_{\beta,p} - \mu) \log n}{(\log \log n)^2}\right], \quad b_n^- := \left[\frac{\log n}{W_0(\log n)} + \frac{(\hat{C}_{\beta,p} - \mu) \log n}{(\log \log n)^2}\right], \quad e_n := \left[\frac{\mu \log n}{2(\log \log n)^2}\right].$$

Moreover, fix $i \in |V_n|$ and set $\mathcal{S}_i = \{i + 1, |V_n|\}$ in the lower bound for $\square = +$, and $\mathcal{S}_i = \{i - 1\}$ in the lower bound for $\square = -$. Then, for every $j \in \mathcal{S}_i$, let $I_j$ be the indicator random variable of the event \{v_i, v_j\} $\in \vec{G}_{1,n}$. Then, by (59) we have

$$\lambda_j := \mathbb{P}(I_j = 1) = \begin{cases} (1 + \hat{\delta}_n) \frac{\beta}{|V_n|} \left(\frac{i}{|V_n|}\right)^{-2\varepsilon/\mu} & \text{for } \square = +, \\ (1 + \hat{\delta}_n) \frac{\beta}{|V_n|} \left(\frac{i}{|V_n|}\right)^{-2\varepsilon/\mu} & \text{for } \square = -. \end{cases}$$

Let $(P_j)_{j \in \mathcal{S}_i}$ be independent Poisson random variables where $P_j$ has mean $\lambda_j$. Then, by Lemma 2.2 we can couple $I_j$ and $\mathbb{1}_{\{P_j \leq 1\}} P_j$ such that $I_j \geq \mathbb{1}_{\{P_j \leq 1\}} P_j$ for each $j \in [n]$ almost surely. Setting

$$W_i := \sum_{j \in \mathcal{S}_i} P_j \quad \text{and} \quad Z_i := \sum_{j \in \mathcal{S}_i} \mathbb{1}_{\{P_j > 1\}} P_j,$$

for both $\square \in \{+,-\}$ we get that

$$d_{1,n}^\square(v_i) = \sum_{j \in \mathcal{S}_i} I_j \geq \sum_{j \in \mathcal{S}_i} \mathbb{1}_{\{P_j \leq 1\}} P_j = W_i - Z_i.$$

It then follows that

$$\sum_{i=1}^{|V_n|} \mathbb{P}\left(d_{1,n}^\square(v_i) \geq b_n^\square\right) \geq \sum_{i=1}^{|V_n|} \mathbb{P}(W_i \geq b_n^\square + e_n) - \sum_{i=1}^{|V_n|} \mathbb{P}(Z_i \geq e_n).$$
Now, we show that
\[
\lim_{n \to \infty} \sum_{i=1}^{|V_n|} \mathbb{P} \left( W_i \geq b_n^+ + e_n \right) = \infty, \quad \text{and} \quad \lim_{n \to \infty} \sum_{i=1}^{|V_n|} \mathbb{P} \left( Z_i \geq e_n \right) = 0, \tag{79}
\]
which implies that
\[
\lim_{n \to \infty} \sum_{i=1}^{|V_n|} \mathbb{P} \left( d_{i,n}^w(v_i) \geq b_n^w \right) = \infty \tag{80}
\]
and hence the lower bound by Lemma 7.1.

We start with the first part of (79). Recall that $W_i$ is a sum of independent Poisson random variables with means $\lambda_j$. As a result, $W_i \sim \text{Poi}(\sum_{j \in S_i} \lambda_j)$, so
\[
\mathbb{P} \left( W_i \geq b_n^+ + e_n \right) \geq \mathbb{P} \left( W_i = b_n^+ + e_n \right) = e^{-\sum_{j \in S_i} \lambda_j} \left( \sum_{j \in S_i} \lambda_j \right)^{b_n^+ + e_n} \frac{1}{(b_n^+ + e_n)!}.
\]
By Stirling’s formula $(b_n^+ + e_n)! = \mathcal{O} \left( \left( \frac{b_n^+ + e_n}{e} \right)^{b_n^+ + e_n} \sqrt{2\pi(b_n^+ + e_n)} \right) = \mathcal{O} \left( \left( \frac{b_n^+ + e_n}{e^{1+o(1)}} \right)^{b_n^+ + e_n} \right)$, which yields
\[
\mathbb{P} \left( W_i \geq b_n^+ + e_n \right) = \Omega \left( \left( \frac{b_n^+ + e_n}{e^{1+o(1)}} \sum_{j \in S_i} \lambda_j \right)^{b_n^+ + e_n} \right)
= \exp \left( \left( 1 + \log \left( \sum_{j \in S_i} \lambda_j \right) \right)(b_n^+ + e_n) - (b_n^+ + e_n) \log(b_n^+ + e_n) + o(b_n^+ + e_n) \right). \tag{81}
\]
Now, if $\square = +$, fix an integer $i \in [\sqrt{n}, n/\log n]$. Then,
\[
\sum_{j=i+1}^{|V_n|} \lambda_j = (1 + o(1)) \frac{\beta p}{1-p} \left( 1 - \left( \frac{i}{|V_n|} \right)^{(1-p)/p} \right) = (1 + o(1)) \frac{\beta p}{1-p}, \tag{82}
\]
since $|V_n| = \omega(n/\log n)$ (recall that we work on the event $Q_n$). Combining this with (81) and (82) implies that
\[
\sum_{i=1}^{|V_n|} \mathbb{P} \left( W_i \geq b_n^+ + e_n \right) \geq \sum_{i=[\sqrt{n}]}^{[n/\log n]} \mathbb{P} \left( W_i \geq b_n^+ + e_n \right)
\geq \frac{n}{2\log n} \exp \left( \left( 1 + \log \left( \frac{\beta p}{1-p} \right) \right)(b_n^+ + e_n) - (b_n^+ + e_n) \log(b_n^+ + e_n) + o(b_n^+ + e_n) \right).
\]
Analysis similar to (75) implies that $(b_n^+ + e_n) \log(b_n^+ + e_n) = \log n + (C_{\beta,p} - \mu/2 + o(1)) \log n/\log \log n$, and combined with $(1 + \log(\beta))(b_n^+ + e_n) = (C_{\beta,p} + o(1)) \log n/\log \log n$ this shows the first statement of (79) for $\square = +$.

If $\square = -$, fix an integer $i \in [|V_n| - n/\log n, |V_n|]$. Then, $\sum_{j=1}^{i-1} \lambda_j = (1 + o(1)) \beta$, which together with (81) implies that
\[
\sum_{i=1}^{|V_n|} \mathbb{P} \left( W_i \geq b_n^- + e_n \right) \geq \sum_{i=[|V_n| - n/\log n]}^{|V_n|} \mathbb{P} \left( W_i \geq b_n^- + e_n \right)
\geq \frac{n}{\log n} \exp \left( \left( 1 + \log(\beta) \right)(b_n^- + e_n) - (b_n^- + e_n) \log(b_n^- + e_n) + o(b_n^- + e_n) \right).
\]
Analysis similar to (75) implies that $(b_n^- + e_n) \log(b_n^- + e_n) = \log n + (\bar{C}_{\beta,p} - \mu/2 + o(1)) \log n/\log \log n$, and combined with $(1 + \log(\beta))(b_n^- + e_n) = (\bar{C}_{\beta,p} + o(1)) \log n/\log \log n$ this shows the first statement of (79) for $\square = -$. 

45
We concentrate on the second claim in (79). Fix \( i \in [\|V_n\|] \). Note that \( \max_{i \in [\|V_n\|]} \lambda_i = \max_{i,j \in [\|V_n\|]} p_{ij} = \mathcal{O}(n^{-(1-p)/p}) \), and hence a union bound implies
\[
P \left( \max_{j \in S_i} P_i \geq \left[ \frac{1}{1-p} \right] \right) = \mathcal{O}(\left( \max_{i \in [\|V_n\|]} \lambda_i \right)^{1/(1-p)}) = \mathcal{O}(n^{-1/p}) = o(n^{-1}). \tag{83}
\]
At the same time, for every \( j \in S_i \), using that \( \lambda_j = o(1) \) we get
\[
P (P_j \geq 2) = \sum_{k=2}^{\infty} e^{-\lambda_j} \frac{\lambda_j^k}{k!} \leq \lambda_j^2.
\]
Hence, since \( Z_n \geq e_n \) implies that \( \max_{j \in S_i} P_j \cdot |\{j \in S_i : P_j \geq 2\}| \geq e_n \), and in particular either some of the \( P_j \)'s are at least \( \left[ \frac{1}{1-p} \right] \), or at least \( (1-p)e_n \) of the \( P_j \)'s are at least 2, using (83) we get
\[
P (Z_i \geq e_n) \leq P \left( \max_{j \in S_i} P_j \geq \left[ \frac{1}{1-p} \right] \right) \cup \{|\{j \in S_i : P_j \geq 2\}| \geq (1-p)e_n \}
\leq o(n^{-1}) + P (|\{j \in S_i : P_j \geq 2\}| \geq (1-p)e_n).
\]
Finally, since \( |\{j \in S_i : P_j \geq 2\}| \) is a sum of \( |S_i| \) independent indicators with mean at most
\[
\Lambda_i := \sum_{j \in S_n} \lambda_j^2 = \mathcal{O}(\max_{i,j \in [\|V_n\|]} p_{ij}) = \mathcal{O}(n^{-(1-p)/p}),
\]
by the stronger version of Chernoff’s inequality (see Lemma 2.3i) we get that
\[
P (Z_i \geq (1-p)e_n) \leq o(n^{-1}) + \exp \left( -\Lambda_i \varphi \left( \frac{(1-p)e_n}{\Lambda_i} \right) \right) = o(n^{-1}) + \exp \left( -\Omega(e_n \log(n^{-(1-p)/p})) \right)
\leq o(n^{-1}) + \exp \left( -\Omega \left( \frac{\log n}{\log \log n} \right)^2 \right) = o(n^{-1}),
\]
which finishes the proof of (79) and thus the lower bound in (71). Together with (72) this proves (70) and hence (3).

We finally prove (5) starting with the first statement. Recall \( b_n^+ \) from (78). By Lemma 6.2 it is sufficient so show that for every \( \mu > 0 \), none of the first \( \exp(\log n - 2\mu \log n/\log \log n) \) vertices in \( \overrightarrow{G}_{2,n} \) have degree more than \( b_n^+ \). Indeed, by Lemma 7.1 and (80) we know that \( \overrightarrow{G}_{1,n} \) has maximum in-degree larger than \( b_n^+ \) w.h.p. and moreover \( \overrightarrow{G}_{1,n} \subseteq \overrightarrow{G}_{n} \subseteq \overrightarrow{G}_{2,n} \) w.h.p. At the same time, from the proof of the lower bound in (4) (in particular, replace \( a_n \) with \( b_n^+ \) in (74) and (75)) we have
\[
\max_{i \in [\|V_n\|]} P (d_{2,n}^i(v_i) \geq b_n^+) \leq \exp \left( -\log n + \frac{\mu + o(1)}{\log \log n} \right),
\]
so a union bound over the first \( \exp(\log n - 2\mu \log n/\log \log n) \) vertices in \( \overrightarrow{G}_{2,n} \) implies that w.h.p. \( \log(\min I_n^{\square}) = (1 - o(1)) \log n \) for both \( \square \in \{s, +\} \).

For the second statement in (5), fix any \( \delta \in (0, 2\varepsilon) \). Then, for every \( i \in [\delta n, |V_n|] \) and \( t > 0 \), similarly to (73) we have
\[
P \left( d_{2,n}^i(v_i) \geq b_n^+ \right) \leq \exp \left( -tb_n^+ + (1 + \mathcal{O}(n^{-1/3}))(\varepsilon^t - 1) \beta p \frac{1 - \varepsilon}{1 - p} \left( \frac{\delta}{2\varepsilon} \right)^{1-2\varepsilon/p} \right).
\]
Then, by choosing
\[
t = \log b_n^+ - \log \left( \frac{\beta p}{1 - p} \left( 1 - \frac{2\varepsilon}{p} \left( \frac{\delta}{2\varepsilon} \right)^{1-2\varepsilon/p} \right) \right),
\]
conducting similar steps as in (74) and (75) and using a union bound over all \( |V_n| \leq n \) many vertices, we arrive at
\[
P \left( \max_{i \in [\delta n, |V_n|]} d_{2,n}^i(v_i) \geq b_n^+ \right) \leq (1 - \delta) \exp \left( b_n^+ \left( \mu + \log \left( 1 - \frac{2\varepsilon}{p} \left( \frac{\delta}{2\varepsilon} \right)^{1-2\varepsilon/p} \right) \right) + o(1) \right) + \mathcal{O}(1).
\]
As the logarithm is strictly negative for any choice of \( \varepsilon \in (0,1/2) \) and \( \delta \in (0,2\varepsilon) \), it follows that there exists a sufficiently small \( \mu \) such that the upper bound converges to zero with \( n \), which implies the second statement in (5).

Finally, we concentrate on the final claim in (5). Fix any \( \delta \in (0,2\varepsilon) \). As in the proof of the first claim of (5), using that the events \( \{ \overrightarrow{G}_{1,n} \subseteq \overrightarrow{G}_n \subseteq \overrightarrow{G}_{2,n} \} \) and \( \{ \max_{i\in[V_n]} d_{i,n}(v_i) \geq b_n \} \) hold w.h.p. (the latter event holds w.h.p. for any \( \mu > 0 \) as in the definition of \( b_n \) in (78)), we get

\[
\mathbb{P}\left( \mathcal{I}_n^- \cap \{ v_1, \ldots, v_{\lceil \delta n \rceil} \} \neq \emptyset \right) \\
\leq \sum_{i=1}^{\lceil \delta n \rceil} \mathbb{P}\left( d_{2,n}(v_i) \geq b_n \right) + \mathbb{P}\left( \max_{i\in[V_n]} d_{i,n}(v_i) < b_n \right) + \mathbb{P}\left( \{ G_{1,n} \subseteq G_n \subseteq G_{2,n} \}^c \right).
\]  

(84)

The last two terms both tend to 0 as \( n \to \infty \) by (4) and Lemma 6.2. Also, on the one hand, for every \( i \in [\lceil \delta n \rceil] \) we have that \( d_{2,n}(v_i) \) is stochastically dominated by \( d_{2,n}(v_{\lceil \delta n \rceil}) \). On the other hand, \( d_{2,n}(v_{\lceil \delta n \rceil}) \) has binomial distribution \( \text{Bin}(\lceil \delta n \rceil - 1, p_{1,[\delta n]}(1 + \delta)) \) with mean

\[ (1 + o(1))(\delta) \cdot (1-\beta)/p \frac{1}{n} \delta^{-2\varepsilon}/p = (1 + o(1))\beta \delta (1-p)/p \leq \beta. \]

Hence, the stronger version of Chernoff's inequality (computation done in (77) and the two equations thereafter with \( c_n \) instead of \( b_n \)) shows that for every sufficiently small \( \mu > 0 \) we have that \( \mathbb{P}\left( d_{2,n}(v_i) \geq b_n \right) = o(n^{-1}) \) for every \( i \in [\lceil \delta n \rceil] \). Then, a union bound shows that the sum in (84) also tends to 0 as \( n \to \infty \), which proves the last statement in (5) and concludes the proof.

\[ \square \]

8 Conditional concentration of Lipschitz-type statistics: proof of Theorem 1.11

In this section we prove Theorem 1.11 by applying Azuma's inequality (Lemma 2.4 (ii)) to some suitable martingales with bounded differences. For the first part, we set up our martingale conditionally on \( |V_n| \) as well as the out-degree sequence of \( \overrightarrow{G}_n \), denoted by \( (d_{i,n}^{-}(v_i))_{v_i\in[V_n]} \) (note that in this section, we work with the directed version of \( G_n \)). Then, at each step, the second end of one unmatched edge is revealed. For the second part, conditionally on \( (V_i)_{i=1}^n \) and \( |E_n| = |E(\overrightarrow{G}_n)| \), we reveal the edges of \( G_n \) one by one, and with a suitable probability (which is different for different edges, and measurable in terms of the process \( (V_i)_{i=1}^n \)). Our results are inspired by Theorem 2.19 in [28]; however, despite the fact that the proof method is similar, we do not condition on the entire degree sequence of \( G_n \) as opposed to [28].

Recall that a function is \( L \)-Lipschitz if for any two (directed or undirected) graphs \( G_1 \) and \( G_2 \) that differ in only one edge (that is, \( |E(G_1) \setminus E(G_2)| + |E(G_2) \setminus E(G_1)| \leq 1 \)) we have \( |f(G_1) - f(G_2)| \leq L \). Theorem 1.11 follows immediately from the following two propositions.

Proposition 8.1. For every \( L \)-Lipschitz function \( f \) defined on the set of directed graphs and every \( t \geq 0 \),

\[
\mathbb{P}\left( \left| f(\overrightarrow{G}_n) - \mathbb{E}\left[ f(\overrightarrow{G}_n) \mid |V_n|, (d_{i,n}^{-}(v_i))_{v_i\in[V_n]} \right] \right| \geq t \mid |V_n|, (d_{i,n}^{-}(v_i))_{v_i\in[V_n]} \right) \leq 2 \exp\left( -\frac{t^2}{8|E_n|L^2} \right).
\]

Proof. Order the out-going half-edges of \( \overrightarrow{G}_n \) in increasing order with respect to the label of their starting endvertex, ties being broken arbitrarily. Then, one by one, connect each of these out-going half-edges to a vertex with smaller label in such a way that no double edges are formed. Observe that, conditionally on \( |V_n| \) and \( (d_{i,n}^{-}(v_i))_{v_i\in[V_n]} \), this stochastic algorithm outputs the graph distributed as \( \overrightarrow{G}_n \); for this, it is sufficient to point out that for the vertex with label \( j \in [|V_n|] \), the set, obtained by sampling \( k \leq j - 1 \) vertices consecutively without repetition, is a uniformly random subset of \( [j-1] \) of size \( k \). Let \( (F_i)_{i=1}^{|E_n|} \) be the
natural filtration associated to above stochastic algorithm conditionally on \(|V_n|\) and \((d^-_i(v_i))_{i=1}^{V_n}\). We define the martingale \(\langle X_i \rangle_{i=1}^{|E_n|}\) by setting \(X_i = \mathbb{E}[f(G_n) \mid \mathcal{F}_i]\). Then, we show that

\[
\forall i \in [|E_n| - 1], |X_{i+1} - X_i| \leq 2L. \tag{85}
\]

Note that this is sufficient to conclude by Azuma’s inequality.

To show (85), fix any \(i \in [|E_n| - 1]\), and let \(e_i^+\) be the outgoing edge from a vertex \(u_i\) that is matched at step \(i\) of the algorithm. Also, let \((v_j^i)_{j=1}^{t_i}\) be the vertices, to which \(e_i^+\) may be matched, conditionally on \(\mathcal{F}_i\). Finally, for every \(j \in [t_i]\), let \(\vec{G}_n^j\) be the partially constructed directed graph up to step \(i\) and denote by \(\mathcal{S}_j\) the family of graphs, containing \(\vec{G}_n^j \cup \{u_iv_j\}\). Then, for every pair \((j_1, j_2)\) of different positive integers among \([t_i]\), \(|\mathcal{S}_{j_1}| = |\mathcal{S}_{j_2}|\) since at every step, the number of choices for attaching the next out-going half-edge is fixed. Moreover, the map from \(\mathcal{S}_{j_1} \setminus \mathcal{S}_{j_2}\) to \(\mathcal{S}_{j_2} \setminus \mathcal{S}_{j_1}\) that deletes the edge \(u_iv_{j_1}^i\) and instead constructs \(u_iv_{j_2}^i\) is bijective. We conclude that, since \(f\) is \(L\)-Lipschitz, the average of \(f\) over different classes is the same up to \(2L\) (the factor 2 comes from the fact that one edge is deleted and another is constructed). We conclude that

\[
|X_{i+1} - X_i| \leq \max_{j_1, j_2 \in [t_i]} |\mathbb{E}[f(\vec{G}_n) \mid \vec{G}_n \in \mathcal{S}_{j_1}] - \mathbb{E}[f(\vec{G}_n) \mid \vec{G}_n \in \mathcal{S}_{j_2}]| \leq 2L,
\]

which finishes the proof.

The upper bound in Proposition 8.1 has the disadvantage of being rather constrained in terms of the graph structure as the entire out-degree sequence is exposed. Our next proposition avoids conditioning on the structure of the graph at the price of having complete information about the birth-death process \((V_i)_{i=1}^n\).

**Proposition 8.2.** For any \(L\)-Lipschitz function \(f\) defined on the set of directed graphs and every \(t \geq 0\),

\[
\mathbb{P} \left( \left| f(\vec{G}_n) - \mathbb{E} \left[ f(\vec{G}_n) \mid (V_i)_{i=1}^n, |E_n| \right] \right| \geq t \mid (V_i)_{i=1}^n, |E_n| \right) \leq 2 \exp \left( -\frac{t^2}{8|E_n|L^2} \right).
\]

**Proof.** Let us consider the following process conditionally on \((V_i)_{i=1}^n\) and \(|E_n|\). We start with an empty graph on \(V_n\) and, at each of \(|E_n|\) rounds, we add one edge to the graph. More precisely, for every \(i \in [|E_n|]\), conditionally on the set \(\mathcal{E}_{i-1} := \{e_1, \ldots, e_{i-1}\}\) of already exposed edges, we add any edge \(e_i = u_iv_i \notin \mathcal{E}_{i-1}\) (where \(u_i < v_i\)) with probability proportional to \(\beta/|V_n|\). Setting \(X_i = \mathbb{E}[f(G_n) \mid \mathcal{E}_i]\) for all \(i \in [|E_n|]\), we have that \((X_j)_{j=0}^{|E_n|}\) is a martingale. Moreover, one shows that for all \(i \in [|E_n|]\) we have \(|X_{i-1} - X_i| \leq 2L\) similarly to the second part of the proof of Proposition 8.1 by replacing the families \((\mathcal{S}_j)\) by the families

\[
\mathcal{S}_e := \left\{ \vec{G} \subseteq \{0, 1\}^{\binom{|V|}{2}} : \{e_1, \ldots, e_{i-1}, e\} \subseteq E(\vec{G}) \right\}
\]

for all \(e \notin \mathcal{E}_{i-1}\), and we conclude again by Azuma’s inequality. \(\Box\)

### 9 Conclusion

In this paper, we defined and studied in depth a new model of a dynamic random graph with vertex removal. Some questions remain open.

- Perhaps the most interesting question that we did not answer is to determine the first term in the expression of the diameter of \(G_n\). By Lemma 6.2 and Theorem 6.4 we know that w.h.p. \(\text{diam}(G_n) = \Omega(\log n)\).

- We do not know the expected distance between two vertices of \(G_n\) chosen uniformly at random. It is natural to believe that it satisfies Part 3 of Theorem 1.9 with the same constant \(\zeta\).
• Another natural question is to determine the order of the largest component in $G_n$ when $\beta < \beta_c$. Note that the fact that $\kappa$ is unbounded over $(0,1)^2$ does not allow us to use further results from [5] to establish more precise expression for $|C_1(G_n)|$. We conjecture that $|C_1| = \Theta(\log n)$ w.h.p.

• It would also be interesting to provide a better description of the sets of vertices $I_n$ defined in Theorem 1.10. Although our proof shows bounds that are a bit stronger than the ones stated in Theorem 1.10, we do not provide any details about the size of these sets as well as the positions of the vertices in the bulk (in case $|I_n| = \omega(1)$).

Acknowledgements.

We are thankful to Dieter Mitsche for useful discussions and suggestions.

References

[1] Sayan Banerjee and Shankar Bhamidi. Persistence of hubs in growing random networks. *Probability Theory and Related Fields*, 180(3):891–953, 2021.

[2] Albert Barabási and Reka Albert. Emergence of scaling in random networks. *Science*, 289:509–512, 1999.

[3] Albert L. Barabási, Chaoming Song, and Dashung Wang. Quantifying long-term scientific impact. *Science*, 342(6154):127–132, 2013.

[4] Noam Berger, Christian Borgs, Jennifer Chayes, and Amin Saberi. Asymptotic behavior and distributional limits of preferential attachment graphs. *The Annals of Probability*, 42(1):1–40, 2014.

[5] Béla Bollobás, Svante Janson, and Oliver Riordan. The phase transition in the uniformly grown random graph has infinite order. *Random Structures & Algorithms*, 26(1-2):1–36, 2005.

[6] Fan Chung and Linyuan Lu. Coupling online and offline analyses for random power law graphs. *Internet Mathematics*, 1(4):409–461, 2004.

[7] Colin Cooper, Alan Frieze, and Juan Vera. Random deletion in a scale-free random graph process. *Internet Mathematics*, 1(4):463–483, 2004.

[8] Steffen Dereich and Peter Mörters. Random networks with sublinear preferential attachment: the giant component. *The Annals of Probability*, 41(1):329–384, 2013.

[9] Luc Devroye and Jiang Lu. The strong convergence of maximal degrees in uniform random recursive trees and dags. *Random Structures & Algorithms*, 7(1):1–14, 1995.

[10] Ken A. Dill, Michael J. Hazoglu, Vivek Kulkarni, and Steven S. Skiena. Citation histories of papers: sometimes the rich get richer, sometimes they don’t. *arXiv preprint arXiv:1703.04746*, 2017.

[11] Sergey N. Dorogovtsev, José F. F. Mendes, and Alexander N. Samukhin. Anomalous percolation properties of growing networks. *Physical Review E*, 64(6):066110, 2001.

[12] Rick Durrett. Rigorous result for the CHKNS random graph model. *Discrete Mathematics & Theoretical Computer Science*, 2003.

[13] Paul Erdős and Alfred Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[14] Alessandro Garavaglia, Remco van der Hofstad, and Gerhard Woeginger. The dynamics of power laws: Fitness and aging in preferential attachment trees. *Journal of Statistical Physics*, 168(6):1137–1179, 2017.
[15] Edward N. Gilbert. On random graphs. *The Annals of Mathematical Statistics*, 4(30):1141–1144, 1959.

[16] Abdolhossein Hoorfar and Mehdi Hassani. Inequalities on the Lambert W function and hyperpower function. *J. Inequal. Pure and Appl. Math*, 9(2):5–9, 2008.

[17] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*. John Wiley & Sons, 2000.

[18] Steven Kalikow and Benjamin Weiss. When are random graphs connected? *Israel journal of mathematics*, 62(3):257–268, 1988.

[19] David A. Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

[20] Torgny Lindvall. *Lectures on the coupling method*. Courier Corporation, 2002.

[21] Tiffany Y. Y. Lo. Weak local limit of preferential attachment random trees with additive fitness. *arXiv preprint arXiv:2103.00900*, 2021.

[22] Bas Lodewijks. Location of maximum degree vertices in weighted recursive graphs with bounded random weights. *arXiv preprint arXiv:2110.00522*, 2021.

[23] Bas Lodewijks. On joint properties of vertices with a given degree or label in the random recursive tree. *arXiv preprint arXiv:2204.09032*, 2022.

[24] Michael Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, 1:226–251, 2001.

[25] Justin Salez. *Some implications of local weak convergence for sparse random graphs*. PhD thesis, Université Pierre et Marie Curie – Paris VI; Ecole Normale Supérieure de Paris, 2011.

[26] Larry Shepp. Connectedness of certain random graphs. *Israel Journal of mathematics*, 67(1):23–33, 1989.

[27] Remco van der Hofstad. *Random Graphs and Complex Networks*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2017.

[28] Nicholas C. Wormald. Models of random regular graphs. *London Mathematical Society Lecture Note Series*, pages 239–298, 1999.