Random Graphs

Lecture Notes

Peter Mörters

Universität zu Köln
Mathematisches Institut

Wintersemester 2022/23

Contents

Chapter 1

Random graphs: Motivation and Background

1.1 Motivation

Graphs $G = (V, E)$ consist of a finite or countably infinite set V of *vertices* and a set E of *edges*, where an edge is an unordered pair of vertices, or more formally a subset $e \subset V$ with exactly two elements. They are often a tool for a rough description of highly complex situations, for example

- social relations, like friendship. The vertices are individuals and an edge is established if the individuals are friends. This includes formalisation of the concept of friends in social media.
- technical networks like the internet (vertices are computers and edges are wired or wireless links) or telecommunication networks.
- biological networks like the brain (vertices are neurons and edges are established if they interconnect) or brain functional networks (vertices are spots in the brain and edges are established if the electrical activity signals received at the spots are correlated).
- many others like collaboration networks or the world-wide web.

A common feature of our examples is that the graphs are finite (i.e. have a finite number of vertices) but extremely large. We can often only access a very small part of the graph and do not have a complete global description. It is therefore natural to take samples from a randomly chosen part of a graph and look at the distribution of such a random sample. This leads to a a description which is *local* and *random*. We use such a description to build the models we investigate in this lecture.

We consider models that are

- finite random graphs,
- given by local descriptions,
- studied in the limit as the number of vertices goes to infinity.

In particular we are interested in global features that emerge from the local description (for example, whether the graph is connected or if it is, what is the length of the shortest path between two randomly chosen vertices).

Our focus is on *sparse* graphs which means that in the limit as the number of vertices goes to infinity the number of edges goes to infinity with the same speed, more precisely we have a sequence of graphs $G_n = (V_n, E_n)$ such that the limit

$$\rho := \lim_{n\to\infty} \frac{|E_n|}{|V_n|}$$

exists and is finite. It is called the *edge density*.

In Section 1.2 I will give some general thoughts on large graphs before introducing some examples in Section 1.2 and then, in Section 1.3 the general framework for this lecture.

1.2 Sampling from large graphs

Given a finite graph $G = (V, E)$ we can associate to every vertex $u \in V$ its *degree* by

$$d(u) := \big|\{v \in V : \{u, v\} \in E\}\big|.$$

Picking a vertex U from the finite set V uniformly at random we get a random variable $d(U)$, its distribution (a probability measure on $\mathbb{N}_0$) is called the *empirical degree distribution*. The number of edges in the graph is then

$$|E| = \frac{1}{2}\sum_{u\in V} d(u) = \frac{|V|}{2}\,\mathrm{E}d(U).$$

Hence a sequence of graphs $G_n = (V_n, E_n)$ is sparse if for a sequence of uniformly picked vertices $U_n \in V_n$ the expected degrees converge

$$\mathrm{E}d(U_n) \to 2\rho.$$

A sequence of graphs $G_n = (V_n, E_n)$ is *distributionally sparse* if there exists a probability distribution μ on $\mathbb{N}_0$ such that

$$\mathrm{P}\{d(U_n) = k\} \overset{n\to\infty}{\longrightarrow} \mu(k) \quad \text{for all } k \in \mathbb{N}_0.$$

The probability here refers to the random choice of the vertex U_n. When we apply this definition to *random* sequences of graphs the probabilities on the left are random variables and we ask this convergence to hold in probability. We always insist on the *asymptotic degree distribution* μ being deterministic.

Of particular interest are random graph sequences where the asymptotic degree distribution is heavy tailed, i.e. decays polynomially.

Definition 1.1. *A distributionally sparse random graph sequence is called* scale-free *if the asymptotic degree distribution μ satisfies*

$$\lim_{k\to\infty} \frac{\log \mu(k)}{\log(1/k)} = \tau$$

for some $\tau > 0$ called the power-law exponent.

Many of the examples which motivate our theory the graphs are scale-free and the power-law exponent can be measured. The measured values typically lie in the interval $(2, 4)$. Also, we will see that often properties of the graph depend just on the edge density ρ and power-law exponent τ of the graphs. But naturally, this can only be verified rigorously in the framework of particular models and cannot hold in general.

Here is a general result that limits the possible range of power-law exponents.

Lemma 1.1. *If a graph sequence is sparse and scale-free, then $\tau \geq 2$.*

Proof. We first assume that the graphs are not random. Suppose $\tau < 2$. Then there exists $K_0 > 1$ such that

$$\frac{\log \mu(k)}{\log(1/k)} \leq 2$$

and hence $\mu(k) \geq k^{-2}$ for all $k \geq K_0$. Fix $K \geq K_0$, then

$$|E_n| = \frac{|V_n|}{2}\, \mathbb{E}d(U) \geq \frac{|V_n|}{2} \sum_{k=K_0}^{K} k\mathrm{P}\{d(U_n) = k\}.$$

Now we choose n_0 such that for all $n \geq n_0$ and $1 \leq k \leq K$,

$$\mathrm{P}\{d(U_n) = k\} \geq \mu(k) - k^{-3}.$$

Altogether we get

$$\frac{|E_n|}{|V_n|} \geq \tfrac{1}{2} \sum_{k=K_0}^{K} k(\mu(k) - k^{-3}) \geq \tfrac{1}{2} \sum_{k=K_0}^{K} \frac{k-1}{k^2}.$$

As the sum on the right diverges to infinity as $K \uparrow \infty$ we see that the graph sequence is not sparse. In the random case the penultimate inequality holds with high probability (i.e. the probability that it holds converges to one as $n \to \infty$) and the same conclusion can be drawn. □

Exercise (Weak law of large numbers):
Let $X_{ij}^{(n)}$ be independent Bernoulli random variables with expectation $p_{ij}^{(n)}$. Then, if

$$\frac{1}{n}\sum_{i,j=1}^{n} p_{ij}^{(n)} \to c,$$

then

$$\frac{1}{n}\sum_{i,j=1}^{n} X_{ij}^{(n)} \to c \text{ in probability.}$$

1.3 Examples of random graphs

Here are some random graph models of interest that will also fit into our general framework.

1.3.1 The Erdős-Rényi graph.

This is the easiest nontrivial model. We let $G_n = (V_n, E_n)$ with $V_n = [n]$ and we connect each pair of distinct vertices independently with probability c/n. As $n \to \infty$ the total number of edges is the sum of $\binom{n}{2}$ independent Bernoulli random variables with expectation c/n. Divided by n this converges to $c/2$ by the weak law of large numbers, so that the Erdős-Rényi graph is *sparse* with edge density $\rho = c/2$. Then the degree of any vertex is the sum of $n-1$ independent Bernoulli random variables with expectation c/n and hence it is binomially distributed with parameters $n-1$ and c/n. The empirical degree distribution therefore converges to a Poisson distribution with parameter c, hence the graph is *distributionally sparse* and its asymptotic degree distribution μ is a Poisson distribution. As the Poisson distribution has super-exponentially decreasing tails the Erdős-Rényi graph is *not scale-free*.

1.3.2 The stochastic block model.

Assume now that $V_n = V_n^{(1)} \cup V_n^{(2)} = [n]$ and vertices from $V_n^{(1)}$ are of type one, vertices from $V_n^{(2)}$ are of type two. We connect each pair of distinct vertices independently, with probability a/n if they are of the same type, and with probability b/n if they are of different types. This is a special instance of the stochastic block model, the most used model in the statistics of random

graphs. A typical question in statistics would be, from observation of a very large sample graph, whether $a \neq b$ and if so to associate a type to each vertex.

1.3.3 The Chung-Lu graph.

We let $G_n = (V_n, E_n)$ with $V_n = [n]$ and take positive weights $w_1, \ldots, w_n$ either at random or deterministically. We interpret w_i as weight of vertex i. Given the weights we connect two vertices i and j independently with a probability p_{ij} proportional to the product of the weights, more precisely

$$p_{ij} = \frac{w_i w_j}{\ell_n} \wedge 1, \qquad \text{where } \ell_n := \sum_{i=1}^{n} w_i.$$

The case of constant weights corresponds to the Erdős-Rényi graph. If

$$\max_{i=1}^{n} w_i \leq \sqrt{\ell_n} \tag{1.1}$$

we can drop the $\wedge 1$. As twice the expected total number of edges is

$$\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} p_{ij} = \frac{1}{\ell_n} \sum_{i=1}^{n} w_i \sum_{\substack{j=1 \\ j \neq i}}^{n} w_j = \Big(\sum_{i=1}^{n} \frac{w_i}{\sqrt{\ell_n}} \Big)^2 - \sum_{i=1}^{n} \Big(\frac{w_i}{\sqrt{\ell_n}} \Big)^2$$

the graphs are *sparse* with edge density $\frac{1}{2}(m_1^2 - m_2)$ if the limits

$$m_1 = \lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{w_i}{\sqrt{\ell_n}} \text{ and } m_2 = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \Big(\frac{w_i}{\sqrt{\ell_n}} \Big)^2$$

exist. A possible choice is $w_i = (i/n)^{-\gamma}$ for $0 < \gamma < \frac{1}{2}$. Without assuming (1.1) for this choice the sparsity can even be obtained for all $0 < \gamma < 1$.

1.3.4 The simple preferential attachment model.

There is an abundance of models for scale-free networks, but a particularly interesting concept is *preferential attachment*. The idea, popularised 20 years ago by Barabasi and Albert, is that a graph is built by adding new vertices, which connect themselves at random but preferably to powerful vertices. In the classical models the power of a vertex is measured by its current degree, hence vertices arriving early are typically the most powerful. In this lecture we discuss the probably simplest incarnation of preferential attachment.

Vertices arrive one-by-one and vertex n attaches to each vertex $m \in \{1, \ldots, n-1\}$ independently with a probability proportional to $m^{-\gamma}$ for

some parameter $0 \le \gamma < 1$ parametrising the strength of the preference of early vertices. To make the model sparse the proportionality factor is chosen so that the expected number of connections of a vertex is asymptotically constant. As

$$\sum_{m=1}^{n-1} m^{-\gamma} \sim cn^{1-\gamma}$$

the proportionality factor has to be of order $n^{\gamma-1}$. Altogether, the connection probabilities of two distinct vertices with number (or rank) i and j is

$$p_{ij} = \beta(i \vee j)^{\gamma-1}(i \wedge j)^{-\gamma}$$

where $0 < \beta < 1$ is a fixed parameter and all connections are independent.

1.4 Inhomogeneous random graphs

We now present a general framework for random graphs that includes all the examples above and a lot more. The idea is that the individual properties of vertices are expressed by their *type*, which is taken from a separable metric space $\mathcal{S}$. Types can play different roles, for example the weight associated with a vertex or the birthtime of a vertex in a dynamical graph model. The vertex set of $G_n = (V_n, E_n)$ consists of n vertices with types $(x_1, \dots, x_n)$. Given the types of vertices, independently for every pair of distinct vertices with types x_i and x_j we set an edge with probability

$$\frac{1}{n}\kappa(x_i, x_j),$$

where $\kappa\colon \mathcal{S}^2 \to [0, \infty)$ is a symmetric function that encodes the different features that can be built in the model. Note that the existence of potential edges are *conditionally* independent events given the types, but if the types are chosen at random this allows to model *unconditionally* rather strong dependencies between these events.

We now give the formal definition.

Definition 1.2.

(a) A type space *is a pair $(\mathcal{S}, \mu)$ where $\mathcal{S}$ is a separable metric space and μ is a Borel probability measure on $\mathcal{S}$.*

(b) A vertex space *is a type space $(\mathcal{S}, \mu)$ and a sequence $(\mathbf{x}_n)$ of vectors $\mathbf{x}_n = (x_1, \dots, x_n)$ consisting of n types from the type space such that*

$$\frac{1}{n}\big|\{v \in [n]\colon x_v \in A\}\big| \to \mu(A),$$

for every μ-continuity set $A \subset \mathcal{S}$.

(c) A kernel $\kappa\colon \mathcal{S}^2 \to [0,\infty)$ *is a symmetric measurable function.*

(d) The inhomogeneous random graph *associated with a vertex space and sequence $(\kappa_n)_n$ of kernels is the graph sequence $(G_n)_n$ with $G_n = (V_n, E_n)$ where $V_n = \{1,\dots,n\}$ and for vertices $i \neq j$ we have $\{i,j\} \in E_n$ independently with probability*

$$p_{ij} := \frac{1}{n}\Big(\kappa_n(x_i,x_j) \wedge n\Big).$$

If the type space is finite, we say that the associated inhomogeneous graphs are *of finite type*. Not all kernels give good graphs. We formulate some useful conditions.

Definition 1.3. *A kernel κ is called* graphical *if*

(a) $\kappa\colon \mathcal{S}^2 \to [0,\infty)$ is continuous μ-almost everywhere,

(b) $\displaystyle\iint_{\mathcal{S}^2} \kappa(x,y)\,\mu(dx)\mu(dy) < \infty,$

(c) $\displaystyle\frac{1}{n^2}\sum_{1\le i<j\le n}\big(\kappa(x_i,x_j)\wedge n\big) \longrightarrow \frac{1}{2}\iint_{\mathcal{S}^2}\kappa(x,y)\,\mu(dx)\mu(dy) < \infty.$

A sequence $(\kappa_n)_n$ of kernels converges graphically to *κ if κ is a graphical kernel and*

(d) for μ-almost every x, y, we have that $x_m \to x$, $y_m \to y$ imply that

$$\sup_{n\ge m}\big|\kappa_n(x_m,y_m) - \kappa(x,y)\big| \to 0,$$

(e) $\displaystyle\frac{1}{n^2}\sum_{1\le i<j\le n}\big(\kappa_n(x_i,x_j)\wedge n\big) \longrightarrow \frac{1}{2}\iint_{\mathcal{S}^2}\kappa(x,y)\,\mu(dx)\mu(dy) < \infty.$

A useful consequence of this definition is sparsity of the graph.

Lemma 1.2. *The inhomogeneous random graph associated with a sequence $(\kappa_n)_n$ converging graphically to a kernel κ is sparse with edge density*

$$\frac{1}{2}\iint_{\mathcal{S}^2}\kappa(x,y)\,\mu(dx)\mu(dy).$$

Proof. By the weak law of large numbers for Bernoulli variables, it suffices to show convergence of the expectations. This convergence is precisely given in (e). □

We now look at the examples from Section 1.3 and show how they fit into our framework.

The Erdős-Rényi graph. Take $\mathcal{S} = \{1\}$ and $\kappa(1,1) = c$.

The stochastic block model. Take $\mathcal{S} = \{1,2\}$ and

$$\kappa(i,j) = a + (b-a)1_{i \neq j}$$

and $x_i = k$ if $i \in V_n^{(k)}$. We assume that there are numbers $\mu(1), \mu(2)$ with

$$\frac{|V_n^{(k)}|}{|V_n|} \longrightarrow \mu(k) \text{ for } k \in \{1,2\}.$$

The Chung-Lu graph. Take $\mathcal{S} = [0,1]$ with μ the Lebesgue measure. Take F the cummulative distribution function associated with a probability measure P on the positive reals. Let $\psi = (1-F)^{-1}$ the generalized inverse so that $\psi(U)$, for U uniform on $(0,1)$ has the distribution P. Let $x_i = i/n$ for $i = 1, \ldots, n$ and

$$\kappa_n(x,y) = \frac{n}{\ell_n}\psi(x)\psi(y)$$

with $w_i = \psi(x_i)$ and $\ell_n = \sum_{i=1}^n w_i$. This is the Chung-Lu model as before with the given weights. It coincides with the inhomogeneous random graph with the given kernel sequence on the type space $(\mathcal{S}, \mu)$. If $\mathbb{E}\psi(U) < \infty$ then $\ell_n/n \to \mathbb{E}\psi(U)$ and the kernel sequence converges graphically to the kernel

$$\kappa(x,y) = \frac{1}{\mathbb{E}\psi(U)}\psi(x)\psi(y)$$

We have used that that ψ and hence κ are continuous except for at most countably many points.

The simple preferential attachment model. We take $\mathcal{S} = (0,1]$ with the Lebesgue measure μ and types $x_i = i/n$ for all $i \in \{1, \ldots, n\}$. The kernel is given by

$$\kappa(x,y) = \beta(x \vee y)^{\gamma-1}(x \wedge y)^{-\gamma},$$

where $\beta > 0$ is arbitrary. The kernel is graphical if $0 \leq \gamma < 1$. The special case $\gamma = 0$ is also known as the *Dubins model*.

Chapter 2

The degree distribution of inhomogeneous random graphs

In this section we prove the following theorem.

Theorem 1. *Suppose $(\kappa_n)_n$ converges graphically to the kernel κ. Then the sequence of inhomogeneous random graphs $G_n = (V_n, E_n)$ associated with these kernels is distributionally sparse and its asymptotic degree distribution ν is given by*

$$\nu(k) = \int_{\mathcal{S}} \frac{\lambda(x)^k}{k!} e^{-\lambda(x)}\, \mu(dx), \text{ for } k \in \mathbb{N}_0,$$

where

$$\lambda(x) = \int_{\mathcal{S}} \kappa(x, y)\, \mu(dy).$$

Equivalently, if $N_k(n)$ is the number of vertices in G_n with degree k then, in probability,

$$\frac{1}{n} N_k(n) \longrightarrow \int_{\mathcal{S}} \mathbb{P}_{\lambda(x)}\{X = k\} \mu(dx),$$

where X under $\mathbb{P}_\lambda$ is Poisson distributed with mean λ. The distribution on the right hand-side is called the *mixed-Poisson distribution* with mixing distribution $\mu \circ \lambda^{-1}$.

We interpret this as follows: The degree of a vertex of type x is asymptotically Poisson distributed with mean $\lambda(x)$, while the distribution of types is given by μ. This interpretation can be made rigorous for a finite type space, and this is what we show first.

2.1 Finite type space

We suppose

$$\mathcal{S} = \{1, \ldots, m\}.$$

Let $d(u,i)$ be the number of edges linking vertex $u \in G_n$ with vertices of type $i \in \mathcal{S}$. Then

$$d(u) = \sum_{i=1}^{m} d(u,i).$$

Let $N_{i,k}(n)$ be the number of vertices in G_n of type i with degree k. Then

$$n_i(n) := \sum_{k=0}^{n} N_{i,k}(n)$$

is the number of vertices of type i and we have, by definition of the vertex space,

$$\frac{n_i(n)}{n} \to \mu(i).$$

The probability that vertices of type j, i are connected is

$$\frac{1}{n}\Big(\kappa_n(j,i) \wedge n\Big).$$

For a vertex u of type j the random variables $d(u,i)$ with $i = 1, \ldots, m$ are independent and binomially distributed with parameters

$$n_i(n) - 1_{i=j} \text{ and } \frac{1}{n}\big(\kappa_n(j,i) \wedge n\big).$$

This binomial law converges to a Poisson distribution with parameter

$$\lim_{n\to\infty} \frac{n_i(n) - 1_{i=j}}{n}\big(\kappa_n(j,i) \wedge n\big) = \mu(i)\kappa(j,i).$$

Hence the law of $d(u)$ converges to a Poisson distribution with parameter

$$\lambda(j) = \sum_{i=1}^{m} \mu(i)\kappa(j,i) = \int \kappa(j,i)\mu(di).$$

To obtain the convergence of $\frac{1}{n}N_{j,k}(n)$ we look at its expectation and variance. We have

$$\mathbb{E}N_{j,k}(n) = n_j(n)\mathbb{P}\{d(u) = k\} \sim n\mu(j)\frac{\lambda(j)^k}{k!}e^{-\lambda(j)}.$$

For distinct vertices u, v of type j we have

$$\mathbb{P}\{d(u) = k, d(v) = k\} = \mathbb{P}\{d(u) = k\}^2 + o(\frac{1}{n})$$
$$+ \frac{1}{n}\kappa(j,j)\bigg(\Big(\frac{\lambda(j)^{k-1}}{(k-1)!}e^{-\lambda(j)}\Big)^2 - 2\frac{\lambda(j)^{k-1}}{(k-1)!}e^{-\lambda(j)}\frac{\lambda(j)^k}{k!}e^{-\lambda(j)} + \Big(\frac{\lambda(j)^k}{k!}e^{-\lambda(j)}\Big)^2\bigg).$$

and therefore

$$\mathbb{E}N_{j,k}^2(n) = n_j^2(n)\mathbb{P}\{d(u)=k\}^2 + n_j(n)\left(\frac{\lambda(j)^k}{k!}e^{-\lambda(j)} - \left(\frac{\lambda(j)^k}{k!}e^{-\lambda(j)}\right)^2\right) + O(n),$$

and hence $\mathrm{Var}(N_{j,k}(n)) = O(n)$. By Chebyshev's inequality

$$\mathbb{P}\Big(|N_{j,k}(n) - \mathbb{E}N_{j,k}(n)| > \varepsilon n\Big) = \frac{\mathrm{Var}(N_{j,k}(n))}{\varepsilon^2 n^2} \to 0.$$

Summarising,

$$\frac{N_{j,k}(n)}{n} \to \mu(j)\frac{\lambda(j)^k}{k!}e^{-\lambda(j)} \text{ in probability,}$$

from which the result follows in the case of finite type space.

2.2 General case

This is based on an approximation argument, which we carefully prepare. We first construct suitable partitions of the type space $\mathcal{S}$.

Lemma 2.1. *For every $m \in \mathbb{N}$ there exists a partition $\mathcal{P}_m$ of $\mathcal{S}$ into $M = M(m)$ Borel sets $A_1^{(m)}, \dots, A_M^{(m)}$ such that*

- *each set $A_i^{(m)}$ is a μ-continuity set,*
- *$\mathcal{P}_{m+1}$ refines $\mathcal{P}_m$, i.e. each $A_i^{(m)}$ is a union of sets in $\mathcal{P}_{m+1}$,*
- *if $i_m(x)$ is the unique index such that $x \in A_{i_m(x)}^{(m)}$ we have*

  $$\mathit{diam}(A_{i_m(x)}^{(m)}) \to 0 \text{ as } m \to \infty,$$

  for μ-almost every x.

Proof. Let (z_i) be a dense sequence in $\mathcal{S}$. For any z_i the balls

$$B_d(z_i) = \{y \in \mathcal{S} \colon d(y, z_i) \le d\},$$

for $d > 0$ have disjoint boundaries and hence at most countably many of them fail to be μ-continuity sets. Hence, for any $m \ge 1$, we can choose balls $B_{m,i} = B_{d_{m,i}}(z_i)$ that are continuity sets and have radii satisfying

$$\frac{1}{m} < d_{m,i} \le \frac{2}{m}.$$

Then $(B_{m,i})_i$ cover $\mathcal{S}$ for every m. Then

$$B'_{m,i} := B_{m,i} \setminus \bigcup_{j<i} B_{m,j}$$

defines a for each m an infinite partition $(B'_{m,i})_i$ into continuity sets of diameter at most $4/m$. To get a finite partition we choose q_m large so that

$$B'_{m,0} := \bigcup_{i>q_m} B'_{m,i}$$

satisfies $\mu(B'_{m,0}) < 2^{-m}$. Then $(B'_{m,i})_{i=0}^{q_m}$ is a partition into continuity sets with diameter of $B'_{m,i}, i \geq 1$ at most $4/m$.

Finally, let $\mathcal{P}_m$ consist of all intersections

$$\bigcap_{l=1}^{m} B'_{l,i_l}$$

with $0 \leq i_l \leq q_l$. Then this is a partition of $\mathcal{S}$ satisfying the first two bulletpoints. To verify the third bulletpoint we note that

$$\sum_{m=1}^{\infty} \mu(B'_{m,0}) \leq \sum_{m=1}^{\infty} 2^{-m} < \infty,$$

hence by Borel-Cantelli μ-almost every x is in only finitely many sets $B'_{m,0}$, $m \in \mathbb{N}$. Hence $\operatorname{diam}(A^{(m)}_{i_m(x)}) \leq 4/m$ for all sufficiently large m. □

We use Lemma 2.1 to construct approximating kernels of finite type. We define

$$\underline{\kappa}_m(x,y) = \inf\{\kappa(x',y') \colon x' \in A^{(m)}_{i_m(x)}, y' \in A^{(m)}_{i_m(y)}\}.$$

We then have

$$\underline{\kappa}_m(x,y) \leq \kappa(x,y) \text{ on } \mathcal{S} \times \mathcal{S}.$$

Hence we can construct the inhomogeneous random graphs G_n associated with these kernels on a joint probability space such that every edge present in the graph associated with $\underline{\kappa}_m$ is also present in the graph associated with κ.

As $\underline{\kappa}_m$ is constant on the partition sets of $\mathcal{P}_m$ we can reduce the graph with this kernel to a graph where the type space is $\{1, \ldots, M(m)\}$ with the measure μ' given by

$$\mu'(j) := \mu(A^{(m)}_j),$$

and the types are

$$i_m(\mathbf{x}_n) := (i_m(x_1), \ldots, i_m(x_n)).$$

This means in particular that if $N_k^{(m)}(n)$ is the number of vertices with degree k in G_n with kernel $\underline{\kappa}_m$ we get from the finite type case that

$$\lim_{n\to\infty} \frac{N_k^{(m)}(n)}{n} \to \sum_{j=1}^{M(m)} \mu(A_j^{(m)}) \frac{\lambda_j^k}{k!} e^{-\lambda_j} \text{ in probability,}$$

where, for any $x \in A_j^{(m)}$, we have

$$\lambda_j = \int_{\mathcal{S}} \underline{\kappa}_m(x,y)\mu(dy).$$

It remains to show that for large m the left had side approximates

$$\lim_{n\to\infty} \frac{N_k(n)}{n}$$

and the right hand side approximates

$$\int_{\mathcal{S}} \frac{\lambda(x)^k}{k!} e^{-\lambda(x)}\, \mu(dx), \text{ where } \lambda(x) = \int_{\mathcal{S}} \kappa(x,y)\, \mu(dy).$$

We will do this now in the slightly more general context of kernels κ_n converging graphically to some κ.

Proof of Theorem 1. Suppose (κ_n) converges graphically to κ. Define

$$\underline{\kappa}_m(x,y) = \inf\{(\kappa \wedge \kappa_n)(x',y') \colon x' \in A_{i_m(x)}^{(m)}, y' \in A_{i_m(y)}^{(m)}, n \geq m\}.$$

Then $\underline{\kappa}_m \leq \kappa_n$, for all $n \geq m$, and $\underline{\kappa}_m \leq \kappa$. For μ-almost every x, y, we have

$$\sup_{\substack{n\geq m \\ x'\in A_{i_m(x)}^{(m)}, y'\in A_{i_m(y)}^{(m)}}} \left|\kappa_n(x',y') - \kappa(x,y)\right| \to 0.$$

Hence $\underline{\kappa}_m(x,y) \to \kappa(x,y)$ for μ-almost every x, y and, by dominated convergence, there exists, for every $\varepsilon > 0$, an m_0 such that for all $m \geq m_0$,

$$\iint \underline{\kappa}_m(x,y)\mu(dx)\mu(dy) > \iint \kappa(x,y)\mu(dx)\mu(dy) - \varepsilon.$$

Fix $k \in \mathbb{N}_0$, $\varepsilon > 0$ and m as above. We couple the graph associated with $\underline{\kappa}_m$ to the graphs associated with κ_n, $n \geq m$, so that every edge in the former also exists in the latter. Let $E_{n,m}$ be the set of edges in G_n with kernel $\underline{\kappa}_m$. Then

$$E_{n,m} \subset E_n,$$

and, recalling Lemma 1.2,

$$\begin{aligned}\frac{1}{n}\big|E_n \setminus E_{n,m}\big| &= \frac{1}{n}\big|E_n\big| - \frac{1}{n}\big|E_{n,m}\big| \\ &\overset{n\to\infty}{\longrightarrow} \frac{1}{2}\iint_{\mathcal{S}^2} \kappa(x,y)\,\mu(dx)\mu(dy) - \frac{1}{2}\iint_{\mathcal{S}^2} \underline{\kappa}_m(x,y)\,\mu(dx)\mu(dy) \\ &< \frac{\varepsilon}{2}.\end{aligned}$$

Let $N_k^{(m)}(n)$ be the number of vertices with degree k in G_n with kernel $\underline{\kappa}_m$. Then, for large n,

$$\Big|\frac{N_k(n)}{n} - \frac{N_k^{(m)}(n)}{n}\Big| \le \frac{2}{n}\big|E_n \setminus E_{n,m}\big| \le 2\varepsilon, \tag{2.1}$$

with high probability. As explained above, the finite type calculation gives that

$$\lim_{n\to\infty} \frac{N_k^{(m)}(n)}{n} \to \mathbb{P}\{D^{(m)} = k\}, \tag{2.2}$$

where the law of $D^{(m)}$ is the mixed Poisson distribution where the parameter equals

$$\lambda_j = \inf_{x\in A_j^{(m)}} \int_{\mathcal{S}} \underline{\kappa}_m(x,y)\mu(dy)$$

with probability $\mu(A_j^{(m)})$. As $\underline{\kappa}_m \le \kappa$ we have $\lambda_j \le \lambda(x)$ for $x \in A_j^{(m)}$ and hence we can couple $D^{(m)}$ and D such that $D^{(m)} \le D$. Then

$$\begin{aligned}\mathbb{P}\{D \neq D^{(m)}\} = \mathbb{P}\{D - D^{(m)} \ge 1\} &\le \mathbb{E}[D - D^{(m)}] \\ &= \iint_{\mathcal{S}^2} \kappa(x,y)\,\mu(dx)\mu(dy) - \iint_{\mathcal{S}^2} \underline{\kappa}_m(x,y)\,\mu(dx)\mu(dy) \\ &< \varepsilon. \end{aligned} \tag{2.3}$$

Combining (2.1), (2.2), (3) and choosing first m and then n large yields

$$\Big|\frac{N_k(n)}{n} - \mathbb{P}\{D = k\}\Big| \le 4\varepsilon,$$

and the result follows as $\varepsilon > 0$ was arbitrary. □

2.3 Scale-free graphs and networks

Recall that the asymptotic degree distribution of an inhomogeneous random graph is, by Theorem 1, a mixed Poisson distribution with mixing measure $\mu\circ\lambda^{-1}$. While the Poisson distribution itself has light tails, the mixed Poisson distribution can have polynomially decaying tails if the mixing distribution has such tails. We now state one (of several possible) results of this nature.

Theorem 2. *Suppose (κ_n) converges graphically to κ and assume*

$$c\lambda^{-\tau} \leq \frac{\mu \circ \lambda^{-1}(d\lambda)}{d\lambda} \leq C\lambda^{-\tau} \text{ for all } \lambda > \lambda^*,$$

where $\tau > 2$ and

$$\lambda \colon \mathcal{S} \to [\lambda^*, \infty), \quad \lambda(x) = \int_{\mathcal{S}} \kappa(x, y)\, \mu(dy).$$

Then the inhomogeneous random graph associated with the sequence (κ_n) is scale-free with power-law exponent τ.

Proof. We even show the stronger statement,

$$(c + o(1))k^{-\tau} \leq \int_{\mathcal{S}} \frac{\lambda(x)^k}{k!} e^{-\lambda(x)}\, \mu(dx) \leq (C + o(1))k^{-\tau}.$$

Recall the following well-known property of the Gamma function

$$\frac{\Gamma(k + x)}{\Gamma(k)} \sim k^x \text{ as } k \to \infty, \text{ for all } x \in \mathbb{R}.$$

By assumption,

$$\begin{aligned} \int_{\mathcal{S}} \frac{\lambda(x)^k}{k!} e^{-\lambda(x)}\, \mu(dx) &= \int_{\lambda^*}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda}\, \mu \circ \lambda^{-1}(d\lambda) \\ &\leq C\, \frac{1}{k!} \int_0^{\infty} \lambda^{k-\tau} e^{-\lambda}\, d\lambda = C\, \frac{\Gamma(k + 1 - \tau)}{\Gamma(k + 1)} \sim Ck^{-\tau}. \end{aligned}$$

Similarly,

$$\int_{\mathcal{S}} \frac{\lambda(x)^k}{k!} e^{-\lambda(x)}\, \mu(dx) \geq c\frac{1}{k!}\Big(\int_0^{\infty} \lambda^{k-\tau} e^{-\lambda}\, d\lambda - (\lambda^*)^k \Big) \sim ck^{-\tau},$$

which completes the proof. □

We now look at our examples. It is easy to see that graphs with a finite type space are never scale-free. In that case the measure $\mu \circ \lambda^{-1}$ on $[0, \infty)$ has finite support and vanishing tails. The more interesting cases arise when $\mathcal{S}$ is infinite.

The Chung-Lu graph. Recall that $\mathcal{S} = [0, 1]$ with μ the Lebesgue measure. Take a cumulative distribution function F satisfying, for some $\tau > 2$,

$$F(x) = 1 - x^{-\tau+1} \text{ for all } x > 1.$$

Then $\psi(u) = u^{1/(1-\tau)}$. Then

$$\kappa_n(x,y) = \frac{n}{\ell_n}(xy)^{1/(1-\tau)}$$

with

$$\ell_n/n \to \mathbb{E}\psi(U) = \int_0^1 u^{1/(1-\tau)}\,du = \frac{\tau-1}{\tau-2}.$$

As the kernel sequence converges graphically to the kernel

$$\kappa(x,y) = \frac{\tau-2}{\tau-1}(xy)^{1/(1-\tau)}$$

we get

$$\lambda(x) = x^{1/(1-\tau)}.$$

Hence, for $r > 1$,

$$\mu\{x \in \mathcal{S}\colon \lambda(x) > r\} = r^{-\tau+1}$$

and taking the derivative, for $\lambda > \lambda^* = 1$,

$$\frac{\mu \circ \lambda^{-1}(d\lambda)}{d\lambda} = (\tau-1)\lambda^{-\tau}.$$

We infer from Theorem 2 that the Chung-Lu graph with the given F is scale-free with power-law exponent τ.

The simple preferential attachment model. We take $\mathcal{S} = (0,1]$ with the Lebesgue measure μ and types $x_i = i/n$ for all $i \in \{1,\dots,n\}$ kernel

$$\kappa(x,y) = \beta(x \vee y)^{\gamma-1}(x \wedge y)^{-\gamma},$$

where $\beta > 0$ is arbitrary and $0 < \gamma < 1$,

$$\begin{aligned}
\lambda(x) &= \beta \int_0^1 (x \vee y)^{\gamma-1}(x \wedge y)^{-\gamma}\,dy \\
&= \beta \int_0^x x^{\gamma-1}y^{-\gamma}\,dy + \beta \int_x^1 y^{\gamma-1}x^{-\gamma}\,dy \\
&= \frac{\beta}{1-\gamma}x^{\gamma-1}x^{1-\gamma} + \frac{\beta}{\gamma}x^{-\gamma}(1-x^{\gamma}) \\
&= \frac{\beta}{\gamma}x^{-\gamma} + \frac{\beta}{1-\gamma}.
\end{aligned}$$

Hence

$$\mu\{x \in \mathcal{S}\colon \lambda(x) > r\} = \mu\{x \in \mathcal{S}\colon x^{-\gamma} > \tfrac{\gamma}{\beta}r - \tfrac{\gamma}{1-\gamma}\} = \left(\tfrac{\gamma}{\beta}r - \tfrac{\gamma}{1-\gamma}\right)^{-1/\gamma},$$

and taking the derivative, for $\lambda > \lambda^* = \frac{\beta}{\gamma(1-\gamma)}$, we get

$$\frac{\mu \circ \lambda^{-1}(d\lambda)}{d\lambda} = \frac{1}{\beta}\left(\frac{\gamma}{\beta}\lambda - \frac{\gamma}{1-\gamma}\right)^{-(1+1/\gamma)}.$$

From Theorem 2 we infer that the simple preferential attachment model is scale-free with power-law exponent $\tau = 1 + \frac{1}{\gamma}$. Observe that in the special case $\gamma = 0$, the Dubins model, is not scale-free.

Chapter 3

Local convergence of inhomogeneous random graphs

The asymptotic degree distributions tell us about the number of neighbours of a randomly chosen vertex in G_n. We can go further and 'explore' the graph: Starting from a random vertex we look at its neighbours, then the neighbours of these neighbours, and so fourth. This way we get a connected random graph with a marked vertex called the root. If we take a limit in distribution under random choice of the vertex at which we start the exploration, the limit will again be a random graph with a root vertex. This graph may be easier to study than the original one.

Suppose for example that G_n does not have many short cycles. Then we have a good chance to explore for some (fixed) time without revisiting a vertex that we have seen before. In this case, due to the nature of the limit we take, the limiting graph will be a tree, i.e. a simpler object than the one we started with. In the case of inhomogeneous random graphs (and many other cases) this tree can be studied using the theory of branching processes. From the behaviour of the branching process we will get useful information about the underlying graph sequence. This is our programme for this and the following chapter. We first formalize the concept of weak local limit (Section 3.1), then learn some branching process theory (Section 3.2) and show existence of weak local limits for inhomogeneous random graphs given in terms of branching processes (Section 3.3). In Section 3.4 and also in Chapter 4 we apply the theory thus developed to an in-depth study of inhomogeneous random graphs.

3.1 Weak local limits: General theory

3.1.1 The space $\mathscr{G}$ of rooted graphs

Our local limits will be probability measures on the metric space $\mathscr{G}$ of locally finite, rooted graphs. Elements (G, o) of $\mathscr{G}$ are graphs G with a finite or countably infinite vertex set and a distinguished vertex $o \in G$ called the root, such that every vertex has a finite degree. We need to clarify when we consider two rooted graphs to be the same, so formally we will now define an equivalence relation and then $\mathscr{G}$ is really the space of equivalence classes, though we will still see them as rooted graphs.

For $(G, o) \in \mathscr{G}$ and $n \in \mathbb{N}$ we denote by $G \wedge n$ the embedded finite subgraph consisting of all vertices in G that can be reached from o by a path with no more than n edges. Two elements $(G_1, o_1), (G_2, o_2) \in \mathscr{G}$ have metric distance $d(G_1, G_2) = \frac{1}{N+1}$ where

$$N = \max\big\{n\colon \exists \text{ bijection } \phi\colon G_1 \wedge n \to G_2 \wedge n \text{ with } \phi(o_1) = o_2 \text{ and}$$

$$\{u, v\} \text{ is an edge in } G_1 \text{ iff } \{\phi(u), \phi(v)\} \text{ is an edge in } G_2\big\}.$$

We identify (G_1, o_1) and (G_2, o_2) if $N = \infty$. In particular, we identify a rooted graph with the connected component of its root. We call a rooted graph *infinite* if it is equivalent to an infinite connected graph. Note that this means it is not equivalent to a finite graph.

Lemma 3.1. *$\mathscr{G}$ is a complete, separable metric space.*

Proof. It is easy to check that $\mathscr{G}$ is a metric space. To check separability we build a sequence containing all finite rooted graphs. This sequence is dense as for every rooted graph (G, o) the finite graphs $G \wedge n$ appear in the sequence and

$$d(G, G \wedge n) \leq \frac{1}{n+1} \longrightarrow 0.$$

To see completeness we take a Cauchy sequence (G_n) in $\mathscr{G}$. We take $G_n^0 = G_n$. Suppose, for $k \in \mathbb{N}$ a sequence $(G_n^{k-1})_n$ has been constructed such that, for all $j \leq k-1$,

- $(G_n^{j-1})_n$ is a subsequence of $(G_n^j)_n$
- $(G_n^j \wedge j)_n$ is constant.

Then $(G_n^{k-1} \wedge k)_n$ is a sequence of graphs which contains a constant subsequence, as by the Cauchy property there is n_0 such that, for all $n, m \geq n_0$,

$$d(G_n^{k-1}, G_m^{k-1}) \leq \frac{1}{k+1}.$$

We take $(G_n^k)_n$ to be this sequence. To find the convergent subsequence we now pick the diagonal sequence $(G_n^n)_n$ which has the property that $(G_n^n \wedge k)_n$ is constant for $n \geq k$ and hence convergent to the unique graph with the property $G_n^n \wedge n = G \wedge n$. □

Remark 3.1. *Actually $\mathscr{G}$ is even an ultrametric space, as two balls in $\mathscr{G}$ are either disjoint or one is contained in the other.*

Exercise: $\mathscr{G}$ is not compact.

Lemma 3.2. *A closed set $K \subset \mathscr{G}$ is compact iff for every $k \in \mathbb{N}$ it can be covered by a finite number of pairwise disjoint balls of radius $\frac{1}{k+1}$.*

Proof. Suppose K is compact and fix $k \in \mathbb{N}$. The collection of balls around x of radius $1/(1+k)$ with $x \in K$ is an open cover of K and hence has a finite subcover. If two balls in this cover are intersecting, then one is contained in the other and hence can be removed from the collection.

Suppose now, for every k, the set K can be covered by a finite number of pairwise disjoint balls of radius $\frac{1}{k+1}$. From every sequence (G_n) in $\mathscr{G}$ we can take a subsequence such that all elements are in the same ball. Doing this successively for every $k \in \mathbb{N}$ and taking the diagonal sequence we arrive at a subsequence (G'_n) such that for any n all $G'_m, m \geq n$ are contained in a ball of radius $\frac{1}{n+1}$. Hence (G'_n) is a Cauchy sequence and therefore convergent. If K is closed the limit is in K and therefore K is compact. □

Recall that a sequence (μ_n) of probability measures is called *tight* iff for every $\varepsilon > 0$ there is a compact set $K \subset \mathscr{G}$ such that $\mu_n(K) > 1 - \varepsilon$ for all $n \in \mathbb{N}$. We say that (μ_n) *converges weakly* to a probability measure μ if, for all $f \colon \mathscr{G} \to \mathbb{R}$ continuous and bounded,

$$\lim_{n\to\infty} \int f \, d\mu_n = \int f \, d\mu.$$

Lemma 3.3. *If the sequence (μ_n) of probability measures is tight, then there exists a weakly convergent subsequence.*

Proof. Let $\varepsilon > 0$ and $A > 1$. Pick a compact set K such that $\mu_n(K) > 1 - \frac{\varepsilon}{8A}$ for all n. Let $f \colon \mathscr{G} \to \mathbb{R}$ continuous and bounded by A. As f is uniformly continuous on K there exists m such that

$$|f(x) - f(y)| \leq \frac{\varepsilon}{16A} \text{ for all } x, y \in K \text{ with } d(x,y) < \frac{1}{1+m}.$$

We take the finitely many disjoint balls of radius $\frac{1}{m+1}$ needed to cover K and suppose we have a sequence (μ'_n) such that for each of those balls B the

sequence $(\mu'_n(B))$ converges to a limit $m(B)$. For every probability measure μ with $\mu(B) = m(B)$ we get

$$\begin{aligned}\lim_{n\to\infty} |\int f\,d\mu'_n - \int f\,d\mu| &\le \lim_{n\to\infty} \sum_B |\int_B f\,d\mu'_n - \int_B f\,d\mu| \\ &\qquad + \lim_{n\to\infty} \int_{K^c} |f|\,d\mu'_n + \int_{K^c} |f|\,d\mu \\ &\le \lim_{n\to\infty} \sum_B A|\mu'_n(B) - m(B)| + 4\frac{\varepsilon}{16A} + 2A\frac{\varepsilon}{8A} \\ &\le \varepsilon.\end{aligned}$$

Given (μ_n) we can use the diagonal procedure to construct a subsequence (μ'_n) such that for every ball B the sequence $(\mu'_n(B))$ converges to a limit $m(B)$. Our consideration implies that if there is a probability measure μ such that $\mu(B) = m(B)$ for every ball, then (μ'_n) converges weakly to this measure μ. We use Caratheodory's theorem for the construction of such a μ. By ultrametricity the finite disjoint unions of the set difference of a ball with finitely many balls contained in it form an algebra. We can easily extend m to a finitely additive measure on this algebra. We now use tightness to see that m is a σ-additive measure on this algebra. Indeed, if $B = \bigcup B_k$ is a ball represented as a union of pairwise disjoint balls, then

$$\begin{aligned}m(B) &= \lim_{n\to\infty} \mu_n\Big(\bigcup_{k=1}^{\infty} B_k\Big) = \lim_{n\to\infty} \sum_{k=1}^{\infty} \mu_n\big(B_k\big) \ge \lim_{n\to\infty} \sum_{k=1}^{M} \mu_n\big(B_k\big) \\ &= \sum_{k=1}^{M} m\big(B_k\big) \to \sum_{k=1}^{\infty} m\big(B_k\big).\end{aligned}$$

On the other hand, given $\varepsilon > 0$ we find a compact K with $\mu_n(K^c) < \varepsilon$. The balls (B_k) form an open cover of the compact set $B \cap K$ and we can assume $B_1, \dots, B_M$ are a finite subcover. Then

$$\begin{aligned}m(B) &= \lim_{n\to\infty} \mu_n\Big(\bigcup_{k=1}^{\infty} B_k\Big) \le \lim_{n\to\infty} \mu_n\Big(\bigcup_{k=1}^{\infty} B_k \cap K\Big) + \varepsilon \\ &= \lim_{n\to\infty} \mu_n\Big(\bigcup_{k=1}^{M} B_k \cap K\Big) + \varepsilon \le \lim_{n\to\infty} \sum_{k=1}^{M} \mu_n\big(B_k\big) + \varepsilon \\ &\le \sum_{k=1}^{\infty} m\big(B_k\big) + \varepsilon.\end{aligned}$$

By Caratheodory's theorem this m can be extended to a probability measure μ on the Borel σ-algebra. Hence (μ'_n) converges weakly to the probability measure μ. □

3.1.2 Definition of weak local limit

We take a sequence of graphs $(G^{(N)})_N$ such that $G^{(N)}$ has $a_N \to \infty$ vertices. We say that $(G^{(N)})_N$ *converges weakly locally* to a random rooted graph (G, o) if, for every bounded continuous $h\colon \mathscr{G} \to \mathbb{R}$, we have

$$\frac{1}{a_N} \sum_{o \in G^{(N)}} h\big(G^{(N)}, o\big) \overset{N\to\infty}{\longrightarrow} \mathbb{E}\big[h(G, o)\big]. \tag{3.1}$$

This is just a weak limit (or limit in distribution) where the deterministic graph $G^{(N)}$ is turned into a random rooted graph by picking the root uniformly at random from the vertices of $G^{(N)}$.

If the graphs $(G^{(N)})_N$ themselves are random, the objects on the left hand side of (3.1) are random variables and the convergence is supposed to hold in probability.

3.1.3 Criteria and examples for weak local convergence

The following result is useful for the identification of weak local limits.

Lemma 3.4. *The sequence $(G^{(N)})_N$ converges weakly locally to a random rooted graph (G, o) iff, for every finite rooted graph (H, r) and $k \in \mathbb{N}$, we have in probability,*

$$\frac{1}{a_N} \sum_{o \in G^{(N)}} 1_{\{(G^{(N)}, o) \wedge k = (H, r)\}} \overset{N\to\infty}{\longrightarrow} \mathbb{P}\big\{(G, o) \wedge k = (H, r)\big\}.$$

Proof. As $(G, o) \mapsto 1_{\{(G,o)\wedge k=(H,r)\}}$ is the indicator of a set which is both closed and open in $\mathscr{G}$, it is a continuous and bounded function. Hence the condition of the lemma follows from weak local convergence.

We prove the converse for deterministic sequences $(G^{(N)})_N$, as the case of random graphs follows by picking a first a subsequence such that the convergence holds for all k and (H, r) almost surely. Under the conditions of the lemma we have to show (a) for every subsequence the existence of a subsubsequence such that the weak local limit exists and (b) that this limit equals (G, o). To see (a) we need to show for every k tightness of the laws of $G_N \wedge k$ with $G_N := (G^{(N)}, o_N)$, where o_N is uniformly chosen.

Let $\varepsilon > 0$ and denote by μ the law of (G, o). Then there exists a compact set K with $\mu(K) \geq 1 - \varepsilon$. Indeed, we can take the dense sequence (g_n) in $\mathscr{G}$ and observe

$$\mathscr{G} = \bigcup_n B\big(g_n, \tfrac{1}{m+1}\big).$$

Hence

$$1 = \lim_{N\to\infty} \mu\Big(\bigcup_{n=1}^{N} B\big(g_n, \tfrac{1}{m+1}\big)\Big).$$

We choose N_m so large that

$$\mu\Big(\bigcup_{n=1}^{N_m} B\big(g_n, \tfrac{1}{m+1}\big)\Big) > 1 - 2^{-m}\varepsilon.$$

Then

$$K := \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{N_m} B\big(g_n, \tfrac{1}{m+1}\big)$$

is a compact set with $\mu(K) \geq 1 - \varepsilon$.

Fix $k \in \mathbb{N}$ and consider

$$\mathscr{G} \wedge k := \{G \wedge k \colon G \in \mathscr{G}\}.$$

As $\phi\colon \mathscr{G} \to \mathscr{G} \wedge k, G \mapsto G \wedge k$ is a continuous map, the set $K \wedge k$ is compact in the space $\mathscr{G} \wedge k$. As this space is discrete, probability measures ν_n on $\mathscr{G} \wedge k$ converge to a probability measure ν on $\mathscr{G} \wedge k$ iff $\nu_n(G) \to \nu(G)$ for every $G \in \mathscr{G} \wedge k$. Therefore our assumption implies

$$\lim_{N\to\infty} \frac{1}{a_N} \sum_{o\in G^{(N)}} 1_{\{\phi(G^{(N)},o)\in K\wedge k\}} = \mathbb{P}(\phi(G,o) \in K\wedge k) = \mu\circ\phi^{-1}(K\wedge k) \geq 1-\varepsilon,$$

as required to show tightness and hence (a).

(b) follows because if (G', o') is any of the limits constructed in (a) then, for every finite rooted graph (H, r) and $k \in \mathbb{N}$,

$$\mathbb{P}\big\{(G', o') \wedge k = (H, r)\big\} = \mathbb{P}\big\{(G, o) \wedge k = (H, r)\big\},$$

because they are the limits of the same subsequence. This implies that the laws of (G', o') and (G, o) agree on a $\cap$-stable generator of the Borel σ-algebra and hence are the same. □

Lemma 3.5. *Suppose (G, o) is a random element of $\mathscr{G}$ and $\mathscr{T} \subset \mathscr{G}$ a Borel set with $\mathbb{P}((G, o) \in \mathscr{T}) = 1$. Then the sequence $(G^{(N)})_N$ converges weakly locally to the random rooted graph (G, o) if, for every finite rooted graph $(H, r) \in \mathscr{T} \wedge k$ and $k \in \mathbb{N}$, we have*

$$\frac{1}{a_N} \sum_{o\in G^{(N)}} 1_{\{(G^{(N)},o)\wedge k=(H,r)\}} \overset{N\to\infty}{\longrightarrow} \mathbb{P}\big\{(G, o) \wedge k = (H, r)\big\},$$

in probability.

Proof. Again we can focus on a deterministic sequence $(G^{(N)})_N$. As $\mathscr{T} \wedge k$ is countable and $\mathbb{P}((G,o) \in \mathscr{T}) = 1$, for every $\varepsilon > 0$ there exists m and a subset $\mathscr{T}(k,m)$ of size at most m such that $\mathbb{P}\{(G,o) \wedge k \in \mathscr{T}(k,m)\} \geq 1-\varepsilon$. Then, with o_N a uniform random vertex from $G^{(N)}$ we get

$$P((G^{(N)}, o_N) \wedge k \notin \mathscr{T} \wedge k\} \leq 1 - P((G^{(N)}, o_N) \wedge k \in \mathscr{T}(k,m)\}.$$

Therefore

$$\limsup_{N\to\infty} P((G^{(N)}, o_N) \wedge k \notin \mathscr{T} \wedge k\} \leq 1 - \mathbb{P}\{(G,o) \wedge k \in \mathscr{T}(k,m)\} \leq \varepsilon.$$

This implies, for $(H,r) \notin \mathscr{T} \wedge k$, that

$$\frac{1}{a_N} \sum_{o \in G^{(N)}} 1_{\{(G^{(N)},o) \wedge k = (H,r)\}} \overset{N\to\infty}{\longrightarrow} 0 = \mathbb{P}\big\{(G,o) \wedge k = (H,r)\big\},$$

as required. □

Example 3.1. Regular trees. *In a d-regular tree every vertex has $d-1$ children and a parent, except for the root, which has d children and no parent. We look at the graph sequence $(G^{(N)})_N$, which consists of all vertices in a d-regular tree up to generation N. The graph $G^{(N)}$ has*

$$1 + d + (d-1)d + \cdots + (d-1)^{N-1}d$$

vertices. We identify its weak local limit.

Take $G^{(N)}$ and place a temporary root at any of the $(d-1)^{N-1}d$ leaves. Then let $N \to \infty$. This gives an infinite rooted graph G, which has a unique infinite self-avoiding path starting at the root. We let $o_0, o_1, o_2, \ldots$ be the vertices along this path. We take the random rooted tree (G,o) with

$$\mathbb{P}\{o = o_l\} = (d-2)(d-1)^{-(l+1)}, l \in \mathbb{N}_0.$$

This random rooted tree is the weak local limit of the sequence $(G^{(N)})$.

Proof. We use the criterion in the lemma above. Let O_N be a random vertex from $G^{(N)}$. We have, for $N \geq k > l$,

$$(G^{(N)}, O_N) \wedge k = (G, o_l) \wedge k$$

iff O_N is precisely l steps from the nearest leaf.

There are $d(d-1)^{N-l-1}$ vertices satisfying this, and

$$1 + d + (d-1)d + \cdots + (d-1)^{N-1}d \sim \frac{d(d-1)^N}{d-2}$$

vertices altogether. Thus,

$$P\big((G^{(N)}, O_N) \wedge k = (G, o_l) \wedge k\big) \sim (d-2)(d-1)^{-l-1},$$

as required. □

3.2 Multitype branching processes

We now take a finite vertex space $\mathcal{S} = \{1, \ldots, m\}$ and the distribution of a random matrix

$$L = (L_{i,j} \colon i, j \in \mathcal{S})$$

with entries in $\mathbb{N}_0$. We denote $m_{i,j} = \mathbb{E}L_{i,j}$ and let $M = (m_{i,j} \colon i, j \in \mathcal{S})$. We build a random tree from independent copies of the random matrix L, say

$$L^{(k,n)} = (L_{i,j}^{(k,n)} \colon i, j \in \mathcal{S}), \text{ for all } k \in \mathbb{N}, n \in \mathbb{N}_0.$$

We start with one vertex in generation zero of arbitrary type. Given the vertices $v_1, \ldots, v_k$ in generation $n \geq 0$ we build the next generation using the matrices

$$L^{(1,n)}, \ldots, L^{(k,n)}.$$

If vertex v_ℓ has type x_ℓ, then for every j it has exactly $L_{x_\ell,j}^{(\ell)}$ children of type j. Hence the total number of individuals of type j in generation $n+1$ is

$$\sum_{\ell=1}^{k} L_{x_\ell,j}^{(\ell,n)},$$

and if the number of individuals of each type in generation n is given as

$$\mathcal{Z}_n = (Z_n^{(1)}, \ldots, Z_n^{(m)}),$$

then

$$Z_{n+1}^{(j)} = \sum_{i=1}^{m} \sum_{k=1}^{Z_n} 1_{x_k=i} L_{i,j}^{(k,n)},$$

where

$$Z_n := Z_n^{(1)} + \ldots + Z_n^{(m)}.$$

The process $(\mathcal{Z}_n)$ is a *multitype Galton-Watson process* and the associated tree a *multitype Galton-Watson tree*.

3.2.1 Multitype Galton-Watson trees

We now explore properties of the multitype Galton-Watson tree. A lot can be read from the matrix M with nonnegative entries m_{ij} denoting the expected number of children of type j by a parent of type i.

Lemma 3.6. *If $\mathbb{E}\mathcal{Z}_0 = v$ we have $\mathbb{E}\mathcal{Z}_n = vM^n$.*

Proof. Argue by induction. The case $n = 0$ is clear, suppose the result holds for n. Then

$$\mathbb{E}\big[Z_{n+1}^{(j)}\big|\mathcal{Z}_n\big] = \sum_{i=1}^{m}\sum_{k=1}^{Z_n} 1_{x_k=i} m_{i,j} = \sum_{i=1}^{m} Z_n^{(i)} m_{i,j},$$

and using the induction hypothesis

$$\mathbb{E}\big[Z_{n+1}^{(j)}\big] = \sum_{i=1}^{m} (vM^n)_i m_{i,j} = (vM^{n+1})_j,$$

as required. □

We assume that M is *positive regular*, i.e. there is a power of M with all entries positive. This assumption is a kind of irreducibility and allows us to use the following result from the spectral theory of nonnegative matrices.

Lemma 3.7. *There is an eigenvalue ρ of the matrix M which is single, real and positive and is strictly larger than the absolute value of all other eigenvalues. It is called the* principal eigenvalue. *Corresponding right, resp. left, eigenvectors a, b can be taken to have positive entries summing to one.*

Proof. This is the Perron-Frobenius theorem, which is an exercise. □

Lemma 3.8. *If a and b are right, resp. left, eigenvectors corresponding to the principal eigenvalue of M normalised such that $ba = 1$, then, as $k \to \infty$, we have*

$$M^k \sim \rho^k ab.$$

Proof. By Lemma 3.7 we have for every $v \perp b$ that $(\lambda b + v)M^k \sim \rho^k \lambda b$. Hence M^k/ρ^k converges to the orthogonal projection P onto the space spanned by b. This can be represented as $P = ab$ for some column vector a with $ba = 1$. Then $M^k a \sim \rho^k P a = \rho^k a(ba) = \rho^k a$, hence a is a right eigenvector. □

We additionally assume that $(\mathcal{Z}_n)$ is *nonsingular* in the sense that

$$\mathbb{P}(Z_n = 1) \neq 1 \text{ for some } n \in \mathbb{N}.$$

This assumption does not depend on the type of the initial particle and only rules out a degenerate case. We now look at the event

$$\{\text{extinction}\} := \{Z_n \to 0\} = \{\exists N \text{ such that } Z_n = 0 \text{ for all } n \geq N\}.$$

Lemma 3.9.

$$\mathbb{P}\{\textit{extinction}\} + \mathbb{P}\{Z_n \to \infty\} = 1.$$

Proof. We have to show that, for every N,

$$\mathbb{P}\{0 < Z_n < N \text{ for infinitely many } n\} = 0.$$

Suppose the event has positive probability. Then there exists $z \in \mathbb{N}_0^k$ with $z \neq (0, \ldots, 0)$ such that $\mathcal{Z}_n = z$ for infinitely many n with positive probability. By the strong Markov property, if the process started in z has positive extinction probability, it will become extinct almost surely, a contradiction. Hence the process started in z cannot become extinct. In particular there exists a nonempty set $\mathcal{E}$ of types such that the process started with a particle of that type survives almost surely. Nonextinction is then equivalent to the occurrence of a type $\mathcal{E}$ vertex in the tree.

Using nonsingularity we see that the tree started with a vertex of type $\mathcal{E}$ will eventually branch. When it branches by positive regularity every branch has a positive probability of containing a vertex of type $\mathcal{E}$ and at least one of the branches will survive and hence contain such a vertex. We conclude that the tree has infinitely many disjoint subtrees rooted in elements of $\mathcal{E}$ and hence $Z_n \to \infty$ almost surely, which contradicts the assumption. □

Proposition 3.1. *The event*

$$\{\textit{nonextinction}\} := \{Z_n > 0 \textit{ for all } n \in \mathbb{N}\}$$

has positive probability if and only if $\rho > 1$.

Remark 3.2. *If* $\rho > 1$ *the process* $(\mathcal{Z}_n)$ *is called* supercritical.

Proof. Because of positive regularity the positivity of the nonextinction probability does not depend on the type of the initial vertex. If we start with a particle with random type distributed according to the probability vector b, then $\mathbb{E}\mathcal{Z}_n = bM^n = \rho^n b$. If $\rho \leq 1$ this implies $\mathbb{E}Z_n \leq 1$ and hence $\mathbb{P}\{Z_n \to \infty\} = 0$. By Lemma 3.9 this implies $\mathbb{P}\{\text{extinction}\} = 1$ and hence

$$\mathbb{P}\{\text{nonextinction}\} = 0.$$

Now suppose $\rho > 1$. Take a vertex of type i and look at the expected number of offspring of type i after d generations, by Lemma 3.6 this is

$$(e_i M^d)_i \sim \rho^d (e_i ab)_i = \rho^d a_i b_i.$$

As $a_i b_i > 0$ we can pick d large so that this number exceeds one. Then there is a supercritical single-type Galton-Watson process (X_k) with $Z_{kd}^{(i)} \geq X_k$. Hence $X_k \to \infty$ implies $Z_{kd}^{(i)} \to \infty$. We infer that $\mathbb{P}\{\text{nonextinction}\} > 0$. □

Our next result concerns the asymptotic distribution of the types in a generation. This is given by the left eigenvalue of M which we constructed in Lemma 3.7.

Proposition 3.2. *Almost surely on nonextinction,*

$$\lim_{n\to\infty} \frac{\mathcal{Z}_n}{Z_n} = b.$$

The proof relies on a version of the strong law of large numbers.

Lemma 3.10. *Suppose* (N_k) *are random variables and* $(X_n^{(k)} : n \geq 1, k \geq 1)$ *iid random variables with mean zero. If* $\{\liminf_{k\to\infty} N_k = \infty\}$ *with positive probability, then almost surely on this event*

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} X_n^{(k)} = 0.$$

Proof. By the classical strong law of large numbers there exist (k_N) such that, for fixed $k, \varepsilon > 0$,

$$\mathbb{P}\left\{ \left| \frac{1}{n} \sum_{m=1}^{n} X_m^{(k)} \right| < 2^{-N} \, \forall n \geq k_N \, \forall N \geq 1 \right\} > 1 - \varepsilon 2^{-k}.$$

The assumption ensures that for every N we have $N_k \geq k_N$ for k large enough. The result follows by summing the complementary probabilities over k. □

Remark 3.3. *Fix* $i \in \mathcal{S}$ *and suppose* $\rho > 1$. *As seen in the proof of Proposition 3.1 there exists* d *such that in every subtree emanating from any vertex of the multitype Galton-Watson tree there is a supercritical (single type) Galton-Watson tree embedded in every* d*th generation of the subtree. Such a Galton-Watson process goes to infinity unless it becomes extinct. Hence on survival of the multitype Galton-Watson process for* $N_n = Z_n^{(i)}$ *we have* $\liminf_{k\to\infty} N_k = \infty$.

Proof of Proposition 3.2 (Kurtz et al (1997)). Fix i, j. By the strong law of large numbers and the following remark,

$$\lim_{n\to\infty} \frac{1}{Z_n^{(i)}} \sum_{k=1}^{Z_n} 1_{x_k=i} \big(L_{i,j}^{(k,n)} - m_{i,j} \big) = 0,$$

almost surely on survival. Averaging over the types i gives

$$\lim_{n\to\infty} \frac{1}{Z_n} \Big(Z_{n+1}^{(j)} - \sum_{i=1}^{m} Z_n^{(i)} m_{i,j} \Big) = \lim_{n\to\infty} \sum_{i=1}^{m} \frac{Z_n^{(i)}}{Z_n} \left(\frac{1}{Z_n^{(i)}} \sum_{k=1}^{Z_n} 1_{x_k=i} \big(L_{i,j}^{(k,n)} - m_{i,j} \big) \right) = 0.$$

Write

$$v_n := \mathcal{Z}_n/Z_n, A := M/\rho, \gamma_{n+1} := Z_{n+1}/(\rho Z_n).$$

We have shown that, with $|\cdot|$ denoting the sum of entries of a vector,

$$\lim_{n\to\infty} \left|\gamma_{n+1}v_{n+1} - v_n A\right| = 0.$$

Fix $k \in \mathbb{N}$. As

$$v_n \prod_{\ell=0}^{k-1} \gamma_{n-\ell} - v_{n-k}A^k = \sum_{r=0}^{k-1} (\gamma_{n-r}v_{n-r} - v_{n-r-1}A)A^r \prod_{\ell=r+1}^{k-1} \gamma_{n-\ell},$$

and the product on the right hand side is bounded from zero and infinity by random constants, the triangle inequality gives, almost surely on non-extinction,

$$\lim_{n\to\infty} \left|v_n \prod_{\ell=0}^{k-1} \gamma_{n-\ell} - v_{n-k}A^k\right| = 0.$$

As $A^k \to ab$ for a suitable right eigenvalue a, we can choose k large enough such that

$$\limsup_{n\to\infty} \left|v_n \prod_{\ell=0}^{k-1} \gamma_{n-\ell} - v_{n-k}ab\right|$$

is arbitrarily small, which implies that

$$\limsup_{n\to\infty} \left|v_n - (v_{n-k}a \prod_{\ell=0}^{k-1} \gamma_{n-\ell}^{-1})b\right|$$

is arbitrarily small. As both v_n and b are nonnegative vectors summing to one we infer that $v_n \to b$ almost surely on nonextinction, as required. □

3.2.2 Poisson Galton-Watson trees

We now move to Galton-Watson trees with infinite type space, as they will appear in our study of inhomogeneous random graphs. We specialize to the case of Poisson Galton-Watson trees, when the children of a vertex form a Poisson process in the type space. In the finite case this means that the random variables $L_{i,j}$, for $i, j \in \mathcal{S}$ are independent Poisson distributed random variables with mean $m_{i,j}$. The role of the matrix M will now be taken by a linear operator, given by a kernel.

In the general case, let $\mathcal{S}$ be a separable metric space and ν a locally finite Borel measure on $\mathcal{S}$, i.e. all balls have finite measure. A *Poisson process* with intensity ν is a random Borel measure P such that

- for every Borel set $\mathcal{A}$ with $\nu(\mathcal{A}) < \infty$ the random variable $P(\mathcal{A})$ is Poisson distributed with mean $\nu(\mathcal{A})$,
- for disjoint Borel sets $\mathcal{A}, \mathcal{B}$ the random variables $P(\mathcal{A})$ and $P(\mathcal{B})$ are independent.

A Poisson process exists for every intensity ν. The intuition is that $P(\mathcal{A})$ counts the number of points that have landed in the set $\mathcal{A}$. Given that $P(\mathcal{A}) = n$ the points $X_1, \ldots, X_n$ falling in $\mathcal{A}$ are independent uniformly distributed random variables with distribution $\frac{1}{\nu(\mathcal{A})}\nu|_{\mathcal{A}}$.

Let now $(\mathcal{S}, \mu)$ be a type space and take an integrable and almost everywhere continuous function

$$\kappa \colon \mathcal{S} \times \mathcal{S} \to [0, \infty),$$

which need not be symmetric. Then the *Poisson Galton-Watson tree* associated with κ is given as follows:

- the offspring of every vertex in a given generation is independent,
- each vertex of type x has children given by a Poisson process with intensity

  $$\kappa(x, y)\mu(dy),$$

  which means that for every Borel set $\mathcal{A}$ the number of children with type in $\mathcal{A}$ are Poisson distributed with intensity

  $$\int_{\mathcal{A}} \kappa(x, y)\mu(dy).$$

If the root has a random type sampled according to μ the Poisson Galton-Watson tree is called *unimodular*. We denote this process by $(\mathcal{Z}_n)$ and interpret the state $\mathcal{Z}_n$ at generation n as either a finite random (multi-)set of elements in $\mathcal{S}$, or as a random measure on the set $\mathcal{S}$ taking values in $\mathbb{N}_0$.

Let us quickly match this definition with the finite case. Suppose a particle of type i has a Poisson number of type j children with mean m_{ij}, independently for all j. We take a positive probability vector μ on $\mathcal{S}$ as the law of the initial particle and let

$$\kappa(i, j) := m_{i,j}/\mu(j)$$

If μ satisfies the detailed balance equation

$$\mu(i)m_{i,j} = \mu(j)m_{j,i},$$

then κ is a symmetric kernel, but for the branching process this is not necessary. Then the number of children of type in $\mathcal{A}$ is indeed Poisson with mean

$$\sum_{j\in\mathcal{A}} \kappa(i,j)\mu(j) = \int_{\mathcal{A}} \kappa(i,j)\mu(dj),$$

And for disjoint sets $\mathcal{A}$ and $\mathcal{B}$ the number of particles with corresponding types is independent. Hence the multitype Galton-Watson tree with independent Poisson offspring numbers is a Poisson Galton-Watson tree with kernel κ.

We define the operator T_κ acting on measurable functions $f\colon \mathcal{S} \to \mathbb{R}$ by

$$(T_\kappa f)(x) = \int_{\mathcal{S}} \kappa(x,y) f(y) \mu(dy).$$

This is defined if $f \geq 0$ with $T_\kappa f$ possibly taking the value ∞, or if f is bounded with $T_\kappa f \in L^1(\mathcal{S},\mu)$. We define the norm of the operator as

$$\|T_\kappa\| = \sup\{|T_\kappa f|\colon f \geq 0, |f| \leq 1\},$$

where $|f| = \sqrt{\int f^2\, d\mu}$. The following result is easy to check.

Lemma 3.11. *$\|T_\kappa\|$ is finite iff T_κ is a bounded linear operator on $L^2(\mathcal{S},\mu)$.*

The operator T_κ plays a similar role as the matrix M in the case of finite type space, for example in describing the average state of the process.

Lemma 3.12. *For every $f\colon \mathcal{S} \to [0,\infty)$ we have*

$$\mathbb{E} \sum_{x\in\mathcal{Z}_n} f(x) = \int_{\mathcal{S}} (T_\kappa^n f)(x) \mu(dx).$$

Proof. This can be shown by induction. It trivially holds for $n = 0$. Suppose it holds for $n-1$. With every $x \in \mathcal{Z}_{n-1}$ we associate an independent Poisson process P_x and note that

$$\mathbb{E} \sum_{y\in P_x} f(y) = \int \kappa(x,y) f(y)\, \mu(dy) = (T_\kappa f)(x).$$

This follows by an application of the monotone class theorem to the definition. We conclude that

$$\mathbb{E} \sum_{x\in\mathcal{Z}_n} f(x) = \mathbb{E} \sum_{x\in\mathcal{Z}_{n-1}} \sum_{y\in P_x} f(y) = \mathbb{E} \sum_{x\in\mathcal{Z}_{n-1}} (T_\kappa f)(x),$$

and by the induction hypothesis this equals

$$\int_{\mathcal{S}} (T_\kappa^{n-1} T_\kappa f)(x) \mu(dx) = \int_{\mathcal{S}} (T_\kappa^n f)(x) \mu(dx),$$

as required. □

We henceforth assume that κ is bounded from zero and the Poisson Galton-Watson process is nonsingular (with the same definition as in the finite type case). As the following proposition shows, the norm $\|T_\kappa\|$ plays the role of the principal eigenvalue ρ and characterises supercriticality.

Proposition 3.3. *The Poisson Galton-Watson process associated with κ has a positive survival probability if and only if $\|T_\kappa\| > 1$.*

Proof. If $\|T_\kappa\| > 1$ we find $f > 0$ with $|f| = 1$ such that

$$\int_{\mathcal{S}} \left(\int_{\mathcal{S}} \kappa(x,y) f(y) \mu(dy) \right)^2 \mu(dx) > 1.$$

On a finite type space we can approximate κ from below by a positive regular and nonsingular kernel $\underline{\kappa}$ and the function f by a vector (f_j) with $\sum f_j^2 \leq 1$ such that

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{m} \underline{\kappa}(i,j) \mu(A_j) f_j \right)^2 \mu(A_i) > 1. \tag{3.2}$$

Let $M = (m_{i,j})$ be the matrix given by

$$m_{i,j} = \underline{\kappa}(i,j) \mu(A_j).$$

Then the Poisson Galton Watson process dominates the multitype Galton Watson process with Poisson offspring with expectation matrix M. Then

$$\|M\| \geq \|M\| \, |f| \geq |Mf| > 1,$$

where the last inequality is (3.2) and $\|M\|$ is the spectral norm of M, which is equal to the principal eigenvalue ρ. Hence the multitype Galton Watson process is supercritical, has a positive survival probability and by stochastic domination this also applies to the Poisson Galton Watson process.

If $\|T_\kappa\| \leq 1$ we use an argument as in Lemma 3.9 (see exercises) to see that

$$\mathbb{P}\{\text{extinction}\} + \mathbb{P}\{Z_n \to \infty\} = 1.$$

If we start with a particle with random type distributed according to μ, then, for $|f| = 1$, we have

$$\mathbb{E} \sum_{x \in \mathcal{Z}_n} f(x) = \int_{\mathcal{S}} (T_\kappa^n f)(x) \mu(dx) \leq \|T_\kappa\|^n \leq 1.$$

This implies $\mathbb{E} Z_n \leq 1$ and hence $\mathbb{P}\{Z_n \to \infty\} = 0$. By the observation above this implies $\mathbb{P}\{\text{extinction}\} = 1$ and hence $\mathbb{P}\{\text{nonextinction}\} = 0$. $\square$

3.3 Weak local limits of inhomogeneous random graphs

In this section we prove the following theorem.

Theorem 3. *If (κ_n) is a sequence of kernels converging graphically to a kernel κ, which is bounded from zero. Then the inhomogeneous random graphs (G_n), where G_n is associated with κ_n, converge weakly locally to the unimodular Poisson Galton-Watson tree associated with κ.*

We prove the result first for the case of finite type space and then use approximation to pass to the general case.

3.3.1 Finite type space

A technical problem here is to properly address the vertices in a tree. We use the *Ulam-Harris labelling*, in which the children of any vertex are ordered. $\oslash$ denotes the root, a its ath child, and $a_1 \dots a_r$ with $a_i \in \mathbb{N}$ denotes the a_rth child of $a_1 \dots a_{r-1}$. We write $(\mathbf{t}, \mathbf{q})$ for a finite ordered tree $\mathbf{t}$ and an allocation

$$\mathbf{q}\colon \mathbf{t} \to \mathcal{S}, v \mapsto q(v)$$

of types to its vertices. Given a graph G_n with vertex set $\{1, \dots, n\}$ and type space $\mathcal{S}$ and a root o we now think of $(G_n, o) \wedge k$ as the ordered typed graph, in which vertices are explored breadth first in the order of the vertex set and are given the type they have in the graph. This allows us to compare $(G_n, o) \wedge k$ with $(\mathbf{t}, \mathbf{q})$ and define

$$N_{n,k}(\mathbf{t}, \mathbf{q}) = \sum_{o \in G_n} 1_{\{(G_n,o)\wedge k = (\mathbf{t},\mathbf{q})\}}.$$

In order to prove Theorem 3, by Lemma 3.5 we need to show that

$$\frac{1}{n} N_{n,k}(\mathbf{t}, \mathbf{q}) \longrightarrow \mathbb{P}\big\{T \wedge k = (\mathbf{t}, \mathbf{q})\big\},$$

in probability, where T is the multitype Galton-Watson tree where

- the root has type chosen according to μ,
- a vertex of type i has a Poisson number of children with parameter

$$\lambda(i) = \sum_{j \in \mathcal{S}} \kappa(i, j)\mu(j),$$

- every child independently gets a type according to the probability vector $p = (p_1, \dots, p_m)$ given by

$$p_j = \frac{\kappa(i,j)\mu(j)}{\sum_{j' \in \mathcal{S}} \kappa(i,j')\mu(j')} = \frac{\kappa(i,j)\mu(j)}{\lambda(i)}.$$

This convergence is now proved using the *second moment method*. We first show the convergence of the expectations and then show that the variance is of order $o(n^2)$, so that the result follows from Chebyshev's inequality.

Convergence of expectations

We start by giving a sufficiently explicit formula for the distribution of the Poisson Galton-Watson tree.

Lemma 3.13.

$$\mathbb{P}\big\{T \wedge k = (\mathbf{t}, \mathbf{q})\big\} = \prod_{i=1}^{|\mathbf{t} \wedge k-1|} e^{-\lambda(q(v_i))} \frac{1}{d(v_i)!} \prod_{j=1}^{d(v_i)} \kappa(q(v_i), q(v_i j))\mu(q(v_i j)),$$

where $v_1, \dots, v_{|\mathbf{t} \wedge k-1|}$ are the vertices of $\mathbf{t} \wedge k-1$ in the lexicographic ordering based on the Ulam-Harris labelling.

Proof. Fix a vertex v of type $q(v)$. The probability of seeing a sequence of $d(v)$ children of types $q(v1), \dots, q(vd(v))$ equals

$$e^{-\lambda(q(v))} \frac{\lambda(q(v))^{d(v)}}{d(v)!} \prod_{j=1}^{d(v)} \frac{\kappa(q(v), q(vj))\mu(q(vj))}{\lambda(q(v))}$$

$$= \frac{e^{-\lambda(q(v))}}{d(v)!} \prod_{j=1}^{d(v)} \kappa(q(v), q(vj))\mu(q(vj)).$$

This is true independently for every vertex v with $|v| \leq k-1$. Hence

$$\mathbb{P}\big\{T \wedge k = (\mathbf{t}, \mathbf{q})\big\} = \prod_{|v| \leq k-1} e^{-\lambda(q(v))} \frac{1}{d(v)!} \prod_{j=1}^{d} \kappa(q(v), q(vj))\mu(q(vj)).$$

Rewriting this by enumerating the vertices in $\mathbf{t} \wedge (k-1)$ in the lexicographical order of the Ulam-Harris labelling gives the result. □

Now we look at the inhomogeneous random graph G_n (as always with fixed types and random edges). Fix a vertex v of type $q(v)$ and explore its

neighbourhood. The probability of seeing a sequence of d neighbours of type $(q(v1), \dots, q(vd))$ equals

$$\frac{\prod_{q\in\mathcal{S}} \binom{n_q}{m_q}}{\binom{d}{m_q\colon q\in\mathcal{S}}} \prod_{q\in\mathcal{S}} (1-\tfrac{1}{n}\kappa_n(q(v),q))^{n_q-m_q} \prod_{j=1}^{d} \tfrac{1}{n}\kappa_n(q(v),q(vj)),$$

where

- n_q is the number of vertices of type q in the graph excluding v,
- m_q is the number of vertices of type q in $q(v1), \dots, q(vd)$.

This equals

$$\frac{1}{d!} \prod_{q\in\mathcal{S}} (1-\tfrac{1}{n}\kappa_n(q(v),q))^{n_q-m_q} \prod_{j=1}^{d} \tfrac{1}{n}\kappa_n(q(v),q(vj))(n_{q(vj)} - m_{q(vj)}(j-1)),$$

where

- $m_q(j)$ is the number of type q vertices in $q(v1), \dots, q(vj)$.

As $n \to \infty$ we have $n_q/n \to \mu(q)$ and the expression above converges to

$$\frac{1}{d!} e^{-\lambda(q(v))} \prod_{j=1}^{d} \kappa(q(v),q(vj))\mu(q(vj)).$$

If we want to iterate this, i.e. explore the neighbourhood of neighbouring vertices just found, we need to take into account the depletion effect: Some vertices already discovered no longer participate in the exploration.

To manage this effect we recall the lexicographic ordering of elements in $\mathbf{t} \wedge (k-1)$ as $v_1, \dots, v_{|\mathbf{t}\wedge(k-1)|}$. For a type q let

- $m_q(i)$ be the number of type q vertices in the set $\{v_1, \dots, v_i\}$,
- $m_q(i,j)$ the number of type q vertices in the set consisting of $v_1, \dots, v_{i-1}$, their children and the first j children of v_i.

Let d_v be the number of children of vertex v in $\mathbf{t}$. Then

$$\begin{aligned}\mathbb{P}\big\{(G_n,o)\wedge k = (\mathbf{t},\mathbf{q})\big\} = \prod_{i=1}^{|\mathrm{t}\wedge(k-1)|} \frac{1}{d_{v_i}!} \prod_{q\in\mathcal{S}} (1-\tfrac{1}{n}\kappa_n(q(v_i),q))^{n_q-m_q(i-1)} \\ \times \prod_{j=1}^{d_{v_i}} \tfrac{1}{n}\kappa_n(q(v_i),q(v_ij))(n_{q(v_ij)} - m_{q(v_ij)}(i,j-1)),\end{aligned}$$

which as above converges to the right hand side in Lemma 3.13.

Bounding the variance

We start by computing

$$\frac{\mathbb{E}N_{n,k}(\mathbf{t},\mathbf{q})^2}{n^2} = \mathbb{P}\{(G_n,o_1)\wedge k = (G_n,o_2)\wedge k = (\mathbf{t},\mathbf{q})\},$$

where the randomness refers to G_n as well as two independently chosen vertices o_1, o_2 from G_n. The following auxilliary result is based on the convergence of expectations shown above.

Lemma 3.14. *For any fixed $k \in \mathbb{N}$ we have, as $n \to \infty$,*

$$\mathbb{P}\{\exists \text{ path of length } \leq k \text{ between } o_1 \text{ and } o_2\} \to 0.$$

Proof. We have

$$\mathbb{P}\{\exists \text{ path of length } \leq k \text{ between } o_1 \text{ and } o_2\} = \frac{1}{n}\mathbb{E}\big|(G_n,o_1)\wedge k\big|.$$

Note that $\frac{1}{n}|(G_n,o_1)\wedge k| \leq 1$ and, by the first part,

$$\frac{1}{n}|(G_n,o_1)\wedge k| \to 0 \text{ in } \mathbb{P}\text{-distribution},$$

and hence also in probability. By dominated convergence we hence get

$$\frac{1}{n}\mathbb{E}\big|(G_n,o_1)\wedge k\big| \to 0,$$

as claimed. □

Hence we have

$$\frac{\mathbb{E}[N_{n,k}(\mathbf{t},\mathbf{q})^2]}{n^2} = \mathbb{P}\{(G_n,o_1)\wedge k = (G_n,o_2)\wedge k = (\mathbf{t},\mathbf{q}), o_2 \notin (G_n,o_1)\wedge 2k\} + o(1).$$

We now condition on the event $(G_n,o_1)\wedge k = (\mathbf{t},\mathbf{q})$ and write

$$\begin{aligned}&\mathbb{P}\{(G_n,o_1)\wedge k = (G_n,o_2)\wedge k = (\mathbf{t},\mathbf{q}), o_2 \notin (G_n,o_1)\wedge 2k\}\\ &= \mathbb{P}\{(G_n,o_1)\wedge k = (\mathbf{t},\mathbf{q}) \mid (G_n,o_2)\wedge k = (\mathbf{t},\mathbf{q}), o_2 \notin (G_n,o_1)\wedge 2k\}\\ &\quad\times \mathbb{P}\{(G_n,o_2)\wedge k = (\mathbf{t},\mathbf{q}), o_2 \notin (G_n,o_1)\wedge 2k\}.\end{aligned}$$

From the first part and Lemma 3.14 we can infer that

$$\mathbb{P}\{(G_n,o_2)\wedge k = (\mathbf{t},\mathbf{q}), o_2 \notin (G_n,o_1)\wedge 2k\} \to \mathbb{P}\big\{T\wedge k = (\mathbf{t},\mathbf{q})\big\}.$$

To study the conditional probability we note that the probability of $\{(G_n,o_1)\wedge k = (\mathbf{t},\mathbf{q})\}$ under the conditioning is the same as if the vertices in $(G_n,o_2)\wedge k$ and adjacent edges are removed from the graph. As the number of these vertices is $o(n)$ the limit of our expectation calculation remains in place. Therefore, also

$$\begin{aligned}&\mathbb{P}\{(G_n,o_1)\wedge k = (\mathbf{t},\mathbf{q}) \mid (G_n,o_2)\wedge k = (\mathbf{t},\mathbf{q}), o_2 \notin (G_n,o_1)\wedge 2k\}\\ &\qquad\to \mathbb{P}\big\{T\wedge k = (\mathbf{t},\mathbf{q})\big\}.\end{aligned}$$

Therefore $\mathrm{Var}(N_{n,k}(\mathbf{t},\mathbf{q})/n) \to 0$ and the proof in the finite case is complete.

3.3.2 General case

The generalisation to arbitrary type space is very similar to the argument given for the convergence of the degree distributions.

Suppose (κ_n) is a sequence of kernels converging graphically to κ. Recasll that, for every $m \in \mathbb{N}$, there exists a partition $\mathcal{P}_m$ of $\mathcal{S}$ into $M = M(m)$ Borel sets $A_1^{(m)}, \ldots, A_M^{(m)}$ such that

- each set $A_i^{(m)}$ is a μ-continuity set,
- $\mathcal{P}_{m+1}$ refines $\mathcal{P}_m$, i.e. each $A_i^{(m)}$ is a union of sets in $\mathcal{P}_{m+1}$,
- if $i_m(x)$ is the unique index such that $x \in A_{i_m(x)}^{(m)}$ we have

  $$\operatorname{diam}(A_{i_m(x)}^{(m)}) \to 0 \text{ as } m \to \infty,$$

  for μ-almost every x.

Define

$$\underline{\kappa}_m(x,y) = \inf\{(\kappa \wedge \kappa_n)(x',y') \colon x' \in A_{i_m(x)}^{(m)}, y' \in A_{i_m(y)}^{(m)}, n \geq m\}.$$

Then $\underline{\kappa}_m \leq \kappa_n$, for all $n \geq m$, and $\underline{\kappa}_m \leq \kappa$. For μ-almost every x, y, we have

$$\sup_{\substack{n \geq m \\ x' \in A_{i_m(x)}^{(m)}, y' \in A_{i_m(y)}^{(m)}}} \left|\kappa_n(x',y') - \kappa(x,y)\right| \to 0.$$

Hence $\underline{\kappa}_m(x,y) \nearrow \kappa(x,y)$ for μ-almost every x, y.

Lemma 3.15. *Let (κ_m) be a sequence of kernels such that $\kappa_m(x,y) \nearrow \kappa(x,y)$ for μ-almost every x, y. Let T^m be the unimodular Poisson Galton-Watson tree with kernel κ_m, and T be the unimodular Poisson Galton-Watson tree with kernel κ. Then, for every $k \in \mathbb{N}$,*

$$T^m \wedge k \Rightarrow T \wedge k \quad \text{in distribution.}$$

In particular, for μ-almost every x, the probability $P^m(x, \geq k)$ that an individual of type x in T^m has at least k descendants converges monotonically to $P(x, \geq k)$, the corresponding probability in T.

Proof. With $\Delta\kappa_m(x,y) = \kappa_m(x,y) - \kappa_{m-1}(x,y)$ and $\kappa_0(x,y) = 0$ we can write

$$\kappa(x,y) = \sum_{m=1}^{\infty} \Delta\kappa_m(x,y).$$

For every m, define an independent Poisson process given by its intensity $\Delta\kappa_m(x,y)\mu(dy)$ and label the individuals by m. Then, the sum of these Poisson processes determines a Poisson process with intensity $\kappa(x,y)\mu(dy)$ and additionally every individual has a label. We use these processes as offspring distributions in a Poisson-Galton Watson process. Then T^m is the process consisting of the individuals with label at most m, T is the process consisting of all individuals, and $T^m \wedge k \Rightarrow T \wedge k$ because one can choose M such that the probability of $T \wedge k$ containing a label larger than M is arbitrarily small. Finally, we can start the procedure in a fixed vertex of type x and as the indicator of the event that the tree has at most $k-1$ vertices is a continuous and bounded function on $\wedge k$ we note that $P^m(x, \geq k) = 1 - P^m(x, < k) \to 1 - P(x, < k) = P(x, \geq k)$. □

Fix $n \geq m$ as above. We couple the graph associated with $\underline{\kappa}_m$ to the graph associated with κ_n, so that every edge in the former also exists in the latter. Let $E_{n,m}$ be the set of edges in $G_n^{(m)}$ with kernel $\underline{\kappa}_m$ and E_n the set of edges in G_n with kernel κ_n. Then

$$E_{n,m} \subset E_n,$$

and, recalling Lemma 1.2, for given $\varepsilon > 0$, and large enough m, we have

$$\begin{aligned}\frac{1}{n}\big|E_n \setminus E_{n,m}\big| &= \frac{1}{n}\big|E_n\big| - \frac{1}{n}\big|E_{n,m}\big| \\ &\overset{n\to\infty}{\longrightarrow} \frac{1}{2}\iint_{\mathcal{S}^2} \kappa(x,y)\,\mu(dx)\mu(dy) - \frac{1}{2}\iint_{\mathcal{S}^2} \underline{\kappa}_m(x,y)\,\mu(dx)\mu(dy) \\ &< \varepsilon.\end{aligned}$$

Let $N_{k,n}^{(m)}(\mathbf{t},\mathbf{q})$ be the number of vertices $o \in G_n^{(m)}$ with kernel $\underline{\kappa}_m$ such that

$$(G_n^{(m)}, o) \wedge k = (\mathbf{t},\mathbf{q}),$$

and $N_{k,n}(\mathbf{t},\mathbf{q})$ for the corresponding quantity using the kernel κ_n. Denote by K the largest degree occurring in $\mathbf{t}$. Then, for large n,

$$\begin{aligned}\Big|\frac{N_{k,n}^{(m)}(\mathbf{t},\mathbf{q})}{n} - \frac{N_{k,n}(\mathbf{t},\mathbf{q})}{n}\Big| &\\ \leq \sum_{u,v\in G_n^{(m)}} & 1_{\{u\in(G_n^{(m)},v)\wedge k-1,(G_n^{(m)},v)\wedge k=(\mathbf{t},\mathbf{q})\}} 1_{\{d(u)\neq d^{(m)}(u)\}} \\ + \sum_{u,v\in G_n^{(m)}} & 1_{\{u\in(G_n,v)\wedge k-1,(G_n,v)\wedge k=(\mathbf{t},\mathbf{q})\}} 1_{\{d(u)\neq d^{(m)}(u)\}},\end{aligned}$$

because if a vertex $v \in G_n^{(m)}$ has $(G_n^{(m)}, v) \wedge k = (\mathbf{t},\mathbf{q})$ but not $(G_n, v) \wedge k = (\mathbf{t},\mathbf{q})$ then one of the vertices in $(G_n^{(m)}, v) \wedge k - 1$ has a different degree in $G_n^{(m)}$ and in G_n. Similarly, if $v \in G_n^{(m)}$ has $(G_n, v) \wedge k = (\mathbf{t},\mathbf{q})$ but not $(G_n^{(m)}, v) \wedge k = (\mathbf{t},\mathbf{q})$.

For fixed u we have

$$\sum_{v\in G_n^{(m)}} 1_{\{u\in(G_n^{(m)},v)\wedge k-1,(G_n^{(m)},v)\wedge k=(\mathbf{t},\mathbf{q})\}} \le \sum_{l=0}^{k-1} K^l \le \frac{K^k-1}{K-1}.$$

Analogously,

$$\sum_{v\in G_n^{(m)}} 1_{\{u\in(G_n,v)\wedge k-1,(G_n,v)\wedge k=(\mathbf{t},\mathbf{q})\}} \le \frac{K^k-1}{K-1}.$$

Hence

$$\begin{aligned}\Big|\frac{N_{k,n}^{(m)}(\mathbf{t},\mathbf{q})}{n} - \frac{N_{k,n}(\mathbf{t},\mathbf{q})}{n}\Big| &\le 2\frac{K^k-1}{K-1}\times\frac{1}{n}\sum_{u\in G_n^{(m)}} 1_{\{d(u)\neq d^{(m)}(u)\}}\\ &\le 2\frac{K^k-1}{K-1}\times\frac{2}{n}\big|E_n\setminus E_{n,m}\big|,\end{aligned}$$

which we have seen to be arbitrarily small. We have thus seen, for any $\varepsilon>0$ and sufficiently large $n\ge m$ that

$$\Big|\frac{N_{k,n}^{(m)}(\mathbf{t},\mathbf{q})}{n} - \frac{N_{k,n}(\mathbf{t},\mathbf{q})}{n}\Big| \le \varepsilon,$$

and from the finite type case

$$\frac{N_{k,n}^{(m)}(\mathbf{t},\mathbf{q})}{n} \to \mathbb{P}\{T^m\wedge k=(\mathbf{t},\mathbf{q})\}.$$

By Lemma 3.15 we can take m so large that

$$|\mathbb{P}\{T^m\wedge k=(\mathbf{t},\mathbf{q})\} - \mathbb{P}\{T\wedge k=(\mathbf{t},\mathbf{q})\}| < \varepsilon,$$

which finishes the proof of Theorem 3.

3.4 Examples and applications

3.4.1 Exploiting weak local convergence

We first look at some general consequences of the existence of weak local limits. They all come from the choice of suitable continuous, bounded functions $h\colon \mathscr{G}\to\mathbb{R}$. For simplicity we assume, as in all our examples, that $|G_n|=n$. We also assume that the weak local limit (G,o) has a positive probability of being more than the root.

Proposition 3.4. *Let (G_n) be a sequence of random graphs converging weakly locally to a random rooted graph (G, o). Then (G_n) is distributionally sparse and the asymptotic degree distribution μ is given as*

$$\mu(k) = \mathbb{P}\{d(o) = k\}.$$

Proof. Let $g\colon \mathbb{N}_0 \to \mathbb{R}$ be bounded and continuous. Then pick $h\colon \mathscr{G} \to \mathbb{R}$ as $h(G, o) = g(d(o))$, which is also bounded and continuous. Hence, in probability,

$$\frac{1}{n}\sum_{v\in G_n} g(d(v)) = \frac{1}{n}\sum_{v\in G_n} h((G_n, v)) \overset{n\to\infty}{\longrightarrow} \mathbb{E}h((G, o)) = \mathbb{E}g(d(o)),$$

which implies (G_n) is distributionally sparse and $\int g(k)\mu(dk) = \mathbb{E}g(d(o))$. □

We now look at the number of edges in the graph. To get sparsity from weak local convergence we need an extra uniform integrability condition.

Proposition 3.5. *Let (G_n) be a sequence of random graphs converging weakly locally to a random rooted graph (G, o). Assume further that for a uniformly chosen vertex $O_n \in G_n$ the random variables $d(O_n)$ are uniformly integrable. Then*

$$\frac{|E_n|}{n} \longrightarrow \frac{1}{2}\mathbb{E}[d(o)] \quad \textit{in probability,}$$

and hence (G_n) is sparse if $\mathbb{E}[d(o)] < \infty$.

Proof. Note that while $\sum_{v\in G_n} d(v) = 2|E_n|$ we cannot simply apply the weak local limit with $h(G, o) = d(o)$ as this continuous function is unbounded. Here uniform integrability is coming in. For every $\varepsilon, \delta > 0$ we find K such that, for all $k \geq K$ such that

$$\mathbb{P}\Big\{\sum_{v\in G_n} d(v)1_{d(v)>k} > \delta n\Big\} \leq \delta^{-1}\mathbb{E}[d(O_n)1_{d(O_n)>k}] < \varepsilon.$$

If $\mathbb{E}[d(o)] < \infty$ fix $k \geq K$ such that

$$0 \leq \mathbb{E}[d(o)] - \mathbb{E}[d(o)1_{d(o)\leq k}] \leq \delta.$$

Using weak local convergence

$$\frac{1}{n}\sum_{v\in G_n} d(v)1_{d(v)\leq k} \longrightarrow \mathbb{E}[d(o)1_{d(o)\leq k}] \quad \text{in probability,}$$

and noting that $\varepsilon, \delta > 0$ were arbitrary and using the triangle inequality yields the result in the case $\mathbb{E}[d(o)] < \infty$. If $\mathbb{E}[d(o)] = \infty$ an analogous argument shows convergence to infinity. □

Remark 3.4. *The uniform integrability condition is discussed in the exercises. It holds in all our examples. Note that for any inhomogeneous random graph (G_n) with kernels (κ_n) converging graphically to κ we have*

$$\frac{|E_n|}{n} \longrightarrow \frac{1}{2}\mathbb{E}[d(o)] = \frac{1}{2}\iint \kappa(x,y)\mu(dy)\,\mu(dx) \qquad \text{in probability}$$

even without a uniform integrability assumption, see Lemma 1.2 where the weak law of large numbers was used.

We next address the number of connected components in the graph.

Proposition 3.6. *Let (G_n) be a sequence of random graphs converging weakly locally to a random rooted graph (G,o). Let K_n be the number of connected components of G_n. Then, in probability,*

$$\frac{K_n}{n} \longrightarrow \mathbb{E}[\tfrac{1}{|G|}] \geq 0.$$

Proof. This is an exercise. Recall that we identify G with the connected component of the root in (G,o). The proof is based on choosing $h\colon \mathscr{G} \to \mathbb{R}$ as $h(G,o) = |G|^{-1}$ if $|G| < \infty$, and $h(G,o) = 0$ otherwise. This function is bounded and can be checked to be continuous. □

Theorem 4. *For any inhomogeneous random graph (G_n) such that the kernels κ_n converge graphically, there exists a constant $c > 0$ such that the number K_n of connected components satisfies*

$$\frac{K_n}{n} \longrightarrow c > 0 \text{ in probability.}$$

Proof. The conditions of Proposition 3.6 are met and $|G| < \infty$ with positive probability if (G,o) is a unimodular Poisson Galton-Watson tree. □

Now we look at the size of connected components. We are interested in components of fixed size in G_n, but also in the existence of a component of macroscopic size, i.e. with size of order n. We say that (G_n) has a *giant component* if the size S_n of the largest component in G_n satisfies

$$\frac{S_n}{n} \longrightarrow \theta > 0 \text{ in probability.}$$

Proposition 3.7. *Let (G_n) be a sequence of random graphs converging weakly locally to a random rooted graph (G,o). Let $K_n^{(k)}$ be the number of connected components with exactly k vertices in G_n. Then, in probability,*

$$\frac{K_n^{(k)}}{n} \longrightarrow \tfrac{1}{k}\mathbb{P}\{|G| = k\}.$$

In particular, if (G, o) is almost surely finite then

$$\frac{S_n}{n} \longrightarrow 0 \text{ in probability}$$

and (G_n) has no giant component

Proof. The proof is based on choosing $h\colon \mathscr{G} \to \mathbb{R}$ as $h(G, o) = \frac{1}{k} 1_{\{|G|=k\}}$. This function is bounded and continuous. Note that, for the connected component C_v of a vertex v in G_n we have

$$\frac{K_n^{(k)}}{n} = \frac{1}{n} \sum_{v \in G_n} \tfrac{1}{k} 1_{\{|C_v|=k\}} = \frac{1}{n} \sum_{v \in G_n} h((G, v)) \to \mathbb{E}h(G, o) = \tfrac{1}{k}\mathbb{P}\{|G| = k\},$$

as claimed. Now, if (G, o) is almost surely finite, for given $\varepsilon, \delta > 0$ there is k such that $\mathbb{P}\{|G| \geq k\} < \varepsilon\delta$. Then

$$\begin{aligned} \limsup_{n\to\infty} \mathbb{P}\{\exists \text{ component of size} \geq \delta n\} &\leq \limsup_{n\to\infty} \mathbb{E}\frac{1}{\delta n} \sum_{v \in G_n} 1_{\{|C_v| \geq \delta n\}} \\ &\leq \frac{1}{\delta}\mathbb{P}\{|G| \geq k\} < \varepsilon. \end{aligned}$$

This shows that there cannot be a giant component. □

Remark 3.5. *In our examples there is also a converse result, as the quantity $\frac{S_n}{n}$ converges in probability to the probability that (G, o) is infinite. This implies that there is a giant component if and only if (G, o) is infinite with positive probability. This does not follow from the existence of the weak local limit alone and we will look at this for particular models in Chapter 4.*

Theorem 5. *In the inhomogeneous random graph with kernels (κ_n) converging graphically to κ there is no giant component if $\|T_\kappa\| \leq 1$.*

Proof. This follows from Proposition 3.7 and the fact that (G, o) is finite almost surely if and only if $\|T_\kappa\| \leq 1$. □

Next, we look at clustering of (G_n). We define the *clustering coefficient at $v \in G_n$* as

$$C_n(v) = \frac{\Delta(G_n, v)}{\frac{1}{2}d(v)(d(v) - 1)},$$

if $d(v) \geq 2$ and zero otherwise, where $\Delta(G, v)$ is the number of triangles in G that contain $v \in G$. Note that this is the proportion of wedges based at

v that complete to a triangle. The *local clustering coefficient* of G_n is now defined as

$$C_n = \frac{1}{n} \sum_{v \in G_n} C_n(v).$$

The limit of C_n as $n \to \infty$, if it exists, is the *asymptotic local clustering coefficient*.

Proposition 3.8. *Let (G_n) be a sequence of random graphs converging weakly locally to a random rooted graph (G, o). Then*

$$C_n \longrightarrow \mathbb{E}\Big[\frac{\Delta(G,o)}{\frac{1}{2}d(o)(d(o)-1)}\Big] \text{ in probability.}$$

Proof. Let

$$h(G,o) = \frac{\Delta(G,o)}{\frac{1}{2}d(o)(d(o)-1)},$$

if $d(o) \geq 2$ and zero ohterwise. This function is bounded by one and continuous, as it only depends on $(G, o) \wedge 2$. The result follows by taking the weak local limit. □

Remark 3.6. *Note that averaging over local quantities avoids putting too much weight on powerful vertices. By contrast, the* global clustering coefficient *of G_n is defined as*

$$C'_n = \frac{\sum_{v \in G_n} C_n(v)}{\frac{1}{2}\sum_{v \in G_n} d(v)(d(v)-1)},$$

i.e. the overall proportion of wedges that complete to a triangle. If (G_n) is a sequence of random graphs converging weakly locally to a random rooted graph (G, o) and the sequence $(d(O_n)^2)$ of random variables is uniformly integrable, then this quantity converges to

$$\frac{\mathbb{E}\Delta(G,o)}{\frac{1}{2}\mathbb{E}d(o)(d(o)-1)}.$$

But for scale-free random graphs with $\tau < 3$ the uniform integrability typically fails and $C'_n \to 0$, see exercises. In particular the global clustering coefficient is unsuitable to measure clustering for scale-free networks with small τ.

If the weak local limit (G, o) is a tree, then the asymptotic local clustering coefficient is zero. This is a major motivation to study *geometric* random graph models, i.e. spatially embedded graphs, for which weak local limits are not trees and there is clustering.

Theorem 6. *For any inhomogeneous random graph (G_n) with kernels (κ_n) converging graphically to κ the asymptotic local clustering coefficient is zero.*

Finally, we look at the degrees of vertices at the end of a uniformly selected edge. For $k, l \in \mathbb{N}_0$ we let

$$N_n(k,l) = \frac{1}{2|E_n|} \sum_{e \in E_n} 1_{d(\underline{e})=k, d(\overline{e})=l} + 1_{d(\underline{e})=l, d(\overline{e})=k},$$

where $e = \{\underline{e}, \overline{e}\}$ and we observe that this is well-defined. The probability measure N_n on $\mathbb{N}_0 \times \mathbb{N}_0$ is called the *empirical degree-degree distribution.*

Proposition 3.9. *Let (G_n) be a sequence of random graphs converging weakly locally to a random rooted graph (G, o). Assume further that for a uniformly chosen vertex $O_n \in G_n$ the random variables $d(O_n)$ are uniformly integrable. Then, for $k, l \in \mathbb{N}_0$ we have*

$$N_n(k,l) \to \frac{k}{\mathbb{E}[d(o)]} \mathbb{P}\{d(o) = k, d(V) = l\} \quad \text{in probability,}$$

where V is a uniformly chosen neighbour of o in (G, o).

Proof. We rewrite

$$\begin{aligned} \frac{1}{n} \sum_{e \in E_n} 1_{d(\underline{e})=k, d(\overline{e})=l} + 1_{d(\underline{e})=l, d(\overline{e})=k} &= \frac{k}{n} \sum_{u \in G_n} 1_{d(u)=k} \left(\frac{1}{k} \sum_{v \sim u} 1_{d(v)=l} \right) \\ &= kE\big[1_{d(O_n)=k, d(V_n)=l}\big], \end{aligned}$$

where V_n is a uniformly chosen neighbour of O_n in G_n and E refers to the choice of O_n and V_n with fixed G_n. As $h(G, o) = 1_{d(o)=k} P\{d(V) = l\}$ with P referring to V dependis only on $(G, o) \wedge 2$ it defines a bounded and continuous function and

$$E\big[1_{d(O_n)=k, d(V_n)=l}\big] = Eh(G_n, O_n) \to \mathbb{E}h(G, o) = \mathbb{P}\{d(o) = k, d(V) = l\},$$

in probability. A reference to Proposition 3.5 completes the proof. □

Remark 3.7. *(a) Summing over all $l \in \mathbb{N}$ gives*

$$\frac{1}{2|E_n|} \sum_{e \in E_n} 1_{d(\underline{e})=k} + 1_{d(\overline{e})=k} \to \frac{k}{\mathbb{E}[d(o)]} \mathbb{P}\{d(o) = k\},$$

the size-biased degree distribution.

(b) Note that N_n is a symmetric measure, and hence the limit must be symmetric, too. We infer that

$$\frac{1}{l}\mathbb{P}\{d(o)=k, d(V)=l\} = \frac{1}{k}\mathbb{P}\{d(o)=l, d(V)=k\},$$

which is a version of the unimodularity property of (G,o) shown in the exercises.

Theorem 7. *For any inhomogeneous random graph (G_n) with kernels (κ_n) converging graphically to κ the empirical degree-degree distribution converges to the probability measure ν given by*

$$\nu(k,l) = c^{-1} \iint \frac{\lambda(x)^k}{(k-1)!} e^{-\lambda(x)} \kappa(x,y) \frac{\lambda(y)^l}{l!} e^{-\lambda(y)} \mu(dx)\, \mu(dy),$$

with $c = \int \kappa(x,y)\mu(dx)\mu(dy)$.

Proof. Note that uniform integrability in Proposition 3.9 was only used to ensure sparsity. So it is not required in this example and the result follows by calculating $k\mathbb{P}\{d(o)=k, d(V)=l\}$ for the Poisson Galton-Watson tree. □

3.4.2 The Chung-Lu graph

We now do some calculations for the Chung-Lu graph. In this case the limiting kernel κ has the 'rank one' form

$$\kappa(x,y) = c\,\psi(x)\psi(y),$$

with $c = 1/\mathbb{E}[\psi(U)]$. For such kernels

$$T_\kappa f(x) = c\,\psi(x) \int_{\mathcal{S}} \psi(y) f(y) \mu(dy).$$

By Cauchy-Schwarz for any $f \geq 0$ with $|f| = 1$,

$$T_\kappa f(x) \leq c\,\psi(x)|\psi|,$$

and so

$$\|T_\kappa\| \leq c\,|\psi|^2.$$

Conversely, for $f = \psi$ we get

$$T_\kappa \psi(x) = c\,\psi(x) \int_{\mathcal{S}} \psi^2(y) \mu(dy).$$

Hence ψ is an eigenvector of T_κ with eigenvalue $c|\psi|^2$ and we infer that

$$\|T_\kappa\| = c \int_{\mathcal{S}} \psi^2(y)\mu(dy).$$

Hence in the Chung-Lu model there is no giant component if

$$\int_0^1 \psi^2(y)\, dy \le \int_0^1 \psi(y)\, dy. \tag{3.3}$$

This criterion is sharp, but this requires a more involved proof (given in the cited book by van der Hofstad, Chapter 4 of Volume II). In this course we will instead give a proof for the corresponding result for simple preferential attachment model.

Theorem 8. *For the Chung-Lu model with limiting kernel*

$$\kappa(x,y) = \tfrac{1}{\mathbb{E}[\psi(U)]}\,\psi(x)\psi(y)$$

the giant component exists if and only if the weights $\psi(U)$ satisfy

$$\mathbb{E}[\psi(U)^2] > \mathbb{E}[\psi(U)].$$

For example, look at the case that, for some $\tau > 2$,

$$F(x) = 1 - \beta x^{-\tau+1} \text{ for all } x > \beta^{1/(\tau-1)}.$$

Recall that in this case the graph is scale-free with power law exponent τ. Then $\psi(u) = (u/\beta)^{1/(1-\tau)}$ and

$$\int_0^1 \psi(y)\, dy = \beta^{\frac{1}{\tau-1}} \frac{\tau-1}{\tau-2}.$$

Also

$$\int_0^1 \psi^2(y)\, dy = \beta^{\frac{2}{\tau-1}} \frac{\tau-1}{\tau-3} \text{ if } \tau > 3,$$

and ∞ otherwise. Hence *there is no giant component if $\tau > 3$ and*

$$\beta \le \left(\frac{\tau-3}{\tau-2}\right)^{\tau-1}.$$

The converse is also true: If $\tau \le 3$ there is a giant component no matter what the edge density is (*robust case*), but if $\tau > 3$ there is a giant component if and only if the edge density is large enough

$$\beta > \beta_c := \left(\frac{\tau-3}{\tau-2}\right)^{\tau-1}.$$

3.4.3 The simple preferential attachment graph

Now look at the simple preferential attachment graph, which is the inhomogeneous random graph with kernel

$$\kappa(x,y) = \beta(x \vee y)^{\gamma-1}(x \wedge y)^{-\gamma},$$

where $\beta > 0$ is arbitrary and $0 < \gamma < 1$. Recall that this is a scale-free random network with power law exponent

$$\tau = 1 + \frac{1}{\gamma}.$$

We look at the operator

$$T_\kappa f(x) = \beta\bigg(x^{\gamma-1}\int_0^x y^{-\gamma} f(y)\,dy + x^{-\gamma}\int_x^1 y^{\gamma-1} f(y)\,dy\bigg).$$

Getting an upper bound for arbitrary $|f| = 1$ is nontrivial as we cannot use the Cauchy-Schwarz inequality, or in other words (other than in the case of the Chung-Lu graph) this is not a Hilbert-Schmidt integral operator, as

$$\iint \kappa(x,y)^2 \mu(dx)\mu(dy) = 2\beta^2 \int_0^1 x^{2\gamma-2}\int_0^x y^{-2\gamma} dy dx = \infty.$$

To get some idea for the correct answer we now give a lower bound for $\|T_\kappa\|$. Suppose $\gamma < \frac{1}{2}$ and pick $\gamma < \alpha < \frac{1}{2}$ and $f(x) = cx^{-\alpha}$. Then

$$\int_0^1 f(x)^2\,dx = c^2 \frac{1}{1-2\alpha}$$

and choosing $c = \sqrt{1-2\alpha}$ ensures $|f| = 1$. Then

$$\begin{aligned} T_\kappa f(x) &= c\beta\big(x^{\gamma-1}\tfrac{x^{1-\gamma-\alpha}}{1-\gamma-\alpha} + x^{-\gamma}\tfrac{1-x^{\gamma-\alpha}}{\gamma-\alpha}\big) \\ &= c\beta\big(\tfrac{1}{1-\gamma-\alpha}x^{-\alpha} + \tfrac{1}{\gamma-\alpha}x^{-\gamma} - \tfrac{1}{\gamma-\alpha}x^{-\alpha}\big) \\ &= c\beta\big(\tfrac{1-2\gamma}{(1-\gamma-\alpha)(\alpha-\gamma)}x^{-\alpha} - \tfrac{1}{\alpha-\gamma}x^{-\gamma}\big). \end{aligned}$$

We obtain

$$\begin{aligned} &\int T_\kappa^2 f(x)\,dx \\ &\quad = c^2\beta^2\Big(\big(\tfrac{1-2\gamma}{(1-\gamma-\alpha)(\alpha-\gamma)}\big)^2 \tfrac{1}{1-2\alpha} - 2\tfrac{1-2\gamma}{(1-\gamma-\alpha)(\alpha-\gamma)^2(1-\alpha-\gamma)} + \tfrac{1}{(\alpha-\gamma)^2(1-2\gamma)}\Big) \\ &\quad = \beta^2\Big(\big(\tfrac{1-2\gamma}{(1-\gamma-\alpha)(\alpha-\gamma)}\big)^2 - 2\tfrac{(1-2\alpha)(1-2\gamma)}{(1-\gamma-\alpha)(\alpha-\gamma)^2(1-\alpha-\gamma)} + \tfrac{1-2\alpha}{(\alpha-\gamma)^2(1-2\gamma)}\Big). \end{aligned}$$

Letting $\alpha \uparrow \frac{1}{2}$ we get

$$\|T_\kappa\| \geq \beta\big(\tfrac{1-2\gamma}{(\frac{1}{2}-\gamma)^2}\big).$$

Assuming that this is sharp we conjecture that there is no giant component if $0 \leq \gamma < \frac{1}{2}$ and

$$0 \leq \beta \leq \tfrac{(\frac{1}{2}-\gamma)^2}{1-2\gamma} = \tfrac{1}{4} - \tfrac{\gamma}{2},$$

and that there is one otherwise. This is indeed true and constitutes the following theorem.

Theorem 9. *For the simple preferential attachment model with parameters $\beta > 0$ and $0 < \gamma < 1$ the giant component exists if and only*

$$\gamma \geq \frac{1}{2} \text{ or } \beta > \beta_c := \frac{1}{4} - \frac{\gamma}{2}.$$

In Chapter 4 we will use special properties of the simple preferential attachment graph to prove this result (and further interesting stuff) by purely probabilistic means.

Chapter 4

Simple preferential attachment revisited

4.1 Coupling to a branching random walk

Although the operator T_κ in the case of simple preferential attachment is not a Hilbert-Schmidt operator and therefore does not look very nice at a first glance, it does have a special property that reveals itself when we are transforming the type space. We look at

$$\phi\colon \mathcal{S} = (0,1] \to (-\infty, 0], t \mapsto \log t,$$

which maps types to positions, and first check what happens to the unimodular Poisson Galton-Watson tree in this case.

The type of the root, which was uniform, now becomes position $-X := \phi(U)$, where X is standard exponential. The children of a vertex of type $s \in \mathcal{S}$, which were a Poisson process with intensity

$$\kappa(s,t)\mu(dt) = \beta(s \vee t)^{\gamma-1}(s \wedge t)^{-\gamma}\, dt$$

become the children of a vertex of position $x = \phi(s)$, which is a Poisson process with intensity

$$\kappa(\phi^{-1}(x), \phi^{-1}(y))\mu \circ \phi^{-1}(dy) = \\ \beta(\phi^{-1}(x) \vee \phi^{-1}(y))^{\gamma-1}(\phi^{-1}(x) \wedge \phi^{-1}(y))^{-\gamma}\, d\phi^{-1}(y),$$

which is

$$\beta e^{(\gamma-1)(x\vee y)} e^{-\gamma(x\wedge y)}\, e^y\, dy = \beta(e^{\gamma(y-x)} 1_{y-x>0} + e^{(1-\gamma)(y-x)} 1_{y-x<0})\, dy.$$

The remarkable fact is that this depends only on $y - x$ and not on both x and y. This means that when you follow the positions of the first children in the Poisson-Galton Watson tree they form a random walk — up to the fact that positions to the right of the origin are not allowed.

For a σ-finite continuous measure ν on the reals we formally define the *Poisson branching random walk* starting at $x(<0)$ with offspring intensity ν (killed at the origin) as follows:

- in generation zero there is a single individual at position x,
- given the individuals in generation n and their positions $x_1, x_2, \ldots (<0)$ we sample independent Poisson processes

$$P_1 = \sum_i \delta_{y_i^{(1)}}, P_2 = \sum_i \delta_{y_i^{(2)}}, \ldots$$

  with intensity ν,
- and form the $n+1$st generation of the branching random walk by declaring

$$x_k + y_i^{(k)} \text{ for all } i \text{ (such that } x_k + y_i^{(k)} < 0)$$

  the children of the individual x_k for $k = 1, 2, \ldots$.

The Poisson branching random walk consists of a rooted tree (possibly with countably infinite offspring numbers) together with the positions of the individuals. Here we consider the intensity

$$\pi(dy) = \beta(e^{\gamma y} 1_{y>0} + e^{(1-\gamma)y} 1_{y<0})\, dy$$

and say that the corresponding killed branching random walk is *unimodular* if it is started in $-X$ where X is standard exponentially distributed. Note that for this choice of π every individual in the (not killed) Poisson branching random walk has infinite offspring, but the *killed* process defines a rooted random tree with finite offspring number. This rooted tree is denoted by $\mathfrak{T}$.

The *exploration process* of a rooted graph successively collects information about the graph by defining an order of vertices and checking the existence of edges adjacent to these vertices in that order, starting at the root. We will explain the exploration process below. The exploration process yields a filtration and associated stopping times. For a rooted graph (G, o) we denote by $(G, o) \curlywedge T$ the rooted graph consisting of the vertices and edges seen up to time T. Our aim in this section is to construct a coupling of the simple preferential attachment graph (G_n, O_n) for a uniform random root O_n with the tree $\mathfrak{T}$ defined above such that the exploration processes of the graph and the tree in such a way that up to a stopping time, which is typically large, the explored part of the graph and the tree coincide.

4.1.1 The exploration process

We now define the exploration process of a rooted graph, i.e., we specify the way we collect information about the connected component of a root vertex o. In the first step, we explore all immediate neighbours of o in the graph. To explain a general exploration step we classify the vertices in three categories:

- *veiled vertices*: vertices for which we have not yet found connections to the component of o;
- *active vertices*: vertices for which we already know that they belong to the component, but for which we have not yet explored all its immediate neighbours;
- *dead vertices*: vertices which belong to the cluster and for which all immediate neighbours have been explored.

After the first exploration step the vertex o is marked as dead, its immediate neighbours as active and all the remaining vertices as veiled. In a general exploration step, we choose the *leftmost* active vertex, set its state to *dead*, and explore its immediate neighbours. The newly found *veiled* vertices are marked as *active*, and we proceed with another exploration step until there are no active vertices left.

We define stopping times T_n depending on sequences (n_n) and (c_n) both with values in $\{1, \dots, n\}$ as the first time when in the exploraton of G_n either

(A) the number of dead and active vertices exceeds c_n, or

(B) one vertex in $\{1, \dots, n_n\}$ is activated, or

(C) there are no more active vertices left.

The main result of this section is the following.

Proposition 4.1. *Suppose the sequences $(c_n)_{n\in\mathbb{N}}$ and $(n_n)_{n\in\mathbb{N}}$ satisfy*

$$\lim_{n\to\infty} \frac{c_n}{n_n^{1-\gamma}} = 0 \text{ and } \lim_{n\to\infty} \frac{c_n n^\gamma}{n_n^{\gamma+1}} = 0 \text{ and } \lim_{n\to\infty} \frac{c_n^2}{n_n} = 0.$$

Then each (G_n, O_n) can be coupled to $\mathfrak{T}$ such that with high probability

$$(G_n, O_n) \curlywedge T_n = \mathfrak{T} \curlywedge T_n.$$

4.1.2 A random labelled tree

We prove the coupling in two steps. In the first step we couple the rooted graph (G_n, o_n) to an easier object, namely a random labelled tree $\mathbb{T}(o_n)$. To both of these objects we associate the exploration process as described above.

The tree $\mathbb{T}(o)$ describes a simplified neighbourhood of a vertex $o \in G_n$. Any vertex in the tree is labelled by its *location*, an element of $\{1, \dots, n\}$. The root is given as a vertex with location o. A vertex with location i produces independently descendants in the locations $j \in \{1, \dots, i-1, i+1, \dots n\}$ with probability

$$\beta(i \vee j)^{\gamma-1}(i \wedge j)^{-\gamma}.$$

Note that in this tree different vertices can be labelled by the same location. The link between this labelled tree and the simple preferential attachment graph is given in the following lemma.

Lemma 4.1. *Suppose that $(c_n)_{n\in\mathbb{N}}$, $(n_n)_{n\in\mathbb{N}}$ satisfy*

$$\lim_{n\to\infty} \frac{c_n^2}{n_n} = 0.$$

Then one can couple (G_n, O_n) with the tree $\mathbb{T}(O_n)$ such that with high probability

$$(G_n, O_n) \curlywedge T_n = \mathbb{T}(O_n) \curlywedge T_n,$$

including equality of locations.

We will see that the bad event (E) which leads to G_n being different from $\mathbb{T}(O_n)$ occurs when the descendants of the explored vertex in the labelled tree include a vertex located at a dead or active vertex in the exploration process of the graph. In this case we stop the exploration so that, before stopping, the explored part of G_n is a tree with each node having a unique location. When (E) occurs we say that *the coupling fails*. If we stop the exploration without (E) being the case, we say that the *coupling succeeds*. In this case the veiled parts of the random tree and the network may be generated independently of each other with the appropriate probabilities and we have coupled the random labelled tree and the graph.

Lemma 4.2. *Suppose that $(c_n)_{n\in\mathbb{N}}$, $(n_n)_{n\in\mathbb{N}}$ are sequences of integers such that*

$$\lim_{n\to\infty} \frac{c_n^2}{n_n} = 0.$$

Then the coupling of the exploration processes satisfies

$$\lim_{n\to\infty} \sup_{o\in\{1,\dots,n\}} \mathbb{P}\big(\textit{coupling with initial vertex } o \textit{ ends in } (\mathrm{E})\big) = 0,$$

i.e. the coupling succeeds with high probability.

Proof. We may assume that $o \in \{n_n + 1, \dots, n\}$. To distinguish both exploration processes, we use the term *descendant* for a child in the labelled random tree and the term *immediate neighbour* in the context of the neighbourhood exploration in G_n. In the initial step, we explore the root and find all its immediate neighbours in G_n and all its descendants in $\mathbb{T}(o)$. Both explorations are identically distributed and they therefore can be perfectly coupled. Suppose now that we have performed k steps and that we have not yet stopped the exploration. In particular, we have seen at most c_n vertices, there are still active vertices, both explored subgraphs coincide and any unveiled (i.e. active or dead) elements of the labelled random tree can be uniquely referred to by its location. We now explore the descendants and immediate neighbours of the *leftmost* active vertex, say i.

We sample the descendants of i as in the labelled tree $\mathbb{T}(o)$. Only if these descendants include a dead or active vertex (E) occurs. Otherwise we take the descendants as immediate neighbours of G_n, so that the coupling is perfect. We stop if (A), (B) or (C) occur due to these neighbours and otherwise continue. All that remains is to estimate the probability of the descendants including a dead or active vertex. There are at most c_n such unveiled vertices and their locations are $j \geq n_n + 1$. For each the probability of a connection is

$$\beta(i \vee j)^{\gamma-1}(i \wedge j)^{-\gamma} \leq \beta n_n^{-1}.$$

As there are at most c_n exploration steps until we end in one of the states (A), (B), or (C), the coupling fails due to (E) with a probability bounded from above by

$$\beta \, \frac{c_n^2}{n_n} \to 0 \,,$$

in other words, the coupling succeeds with high probability. □

4.1.3 Coupling to the killed branching random walk

We now map the locations $\{1, \dots, n\}$ of vertices in G_n to positions on the negative halfline such that

$$\varphi(i) = -\sum_{j=i+1}^{n} \frac{1}{j}.$$

Note that the youngest vertex is placed at the origin, and older vertices are placed to the left with decreasing intensity. In particular the position of the particle corresponding to a vertex with fixed location will move to the left as n is increasing.

Looking at a fixed observation window $[a, b]$ on the negative halfline, as $n \uparrow \infty$, we see that the number of vertices in the window is increasing. At the same time the location of the vertex at (or closest to) a fixed position in the window is increasing, which means that the probability of edges between two such vertices is decreasing. As we shall see below, the combination of these two effects leads to convergence of the distribution of offspring locations on the halfline. In particular, thanks to the independence of edges, offspring converges to a Poisson process by the law of small numbers. Our main aim now is to prove the following result. We follow the convention that a sequence of events depending on the index n holds *with high probability* if the probability of these events goes to one as $n \uparrow \infty$.

Lemma 4.3. *Suppose that $(c_n)_{n\in\mathbb{N}}$ and $(n_n)_{n\in\mathbb{N}}$ are sequences of integers with*

$$\lim_{n\to\infty} \frac{c_n}{n_n^{1-\gamma}} = 0 \text{ and } \lim_{n\to\infty} \frac{c_n n^\gamma}{n_n^{\gamma+1}} = 0.$$

Then the tree $\mathbb{T}(O_n)$ can be coupled with $\mathfrak{T}$ such that with high probability

$$\mathbb{T}(O_n) \wedge T_n = \mathfrak{T} \wedge T_n.$$

To establish the relationship between $\mathbb{T}(O_n)$ and $\mathfrak{T}$ we define the projection

$$\pi_n \colon (-\infty, 0] \to \{1, \dots, n\},$$

which maps t onto the smallest $m \in \{1, \dots, n\}$ with $t \le \varphi(m)$. We apply this to each location of $\mathfrak{T}$ and obtain a branching process with location parameters in $\{1, \dots, n\}$, which we call π_n-projected process. We need to show, using a suitable coupling, that when $\mathfrak{T}$ is started with a vertex $-X$, where X is standard exponentially distributed, then up to the stopping time T_n this agrees with a random tree $\mathbb{T}(O_n)$ with high probability.

We first look at the location of the root in both processes.

Lemma 4.4. *Let X be standard exponentially distributed. Then $\pi_n(-X)$ and U Laplace distributed on $\{1, \dots, n\}$ can be coupled such that*

$$\mathbb{P}(\pi_n(-X) \neq U) \to 0.$$

Proof. We have $\pi_n(-X) = i$ iff $\varphi(i-1) < -X \le \varphi(i)$, which happens if

$$\sum_{j=i}^{n} \frac{1}{j} > X \ge \sum_{j=i+1}^{n} \frac{1}{j}.$$

This event has probability

$$\exp(-\sum_{j=i+1}^{n}\tfrac{1}{j}) - \exp(-\sum_{j=i}^{n}\tfrac{1}{j}) = \exp(-\sum_{j=i+1}^{n}\tfrac{1}{j})(1-e^{-1/i})$$

The coupling follows from the fact that

$$\sum_{i=1}^{n}\big|\exp(-\sum_{j=i+1}^{n}\tfrac{1}{j})(1-e^{-1/i}) - \frac{1}{n}\big| \to 0.$$

Further details are an exercise. □

We now start the principal task of coupling the offspring variables by coupling Poisson and Bernoulli variables.

Lemma 4.5. *Let $\lambda \geq 0$ and $p \in [0,1]$, $X^{(1)}$ Poisson distributed with parameter λ, and $X^{(2)}$ Bernoulli distributed with parameter p. Then there exists a coupling of these two random variables such that*

$$\mathbb{P}(X^{(1)} \neq X^{(2)}) \leq \lambda^2 + |\lambda - p|.$$

Proof. This is an exercise. □

Recall that in the π_n-projected process a vertex in position v with $i = \pi_n(v) < 0$ produces a Poissonian number of π_n-projected descendants at the location $1 \leq i-m \leq i-1$ with parameter

$$\lambda_{i-m} := \int_{\varphi(i-m-1)-v}^{\varphi(i-m)-v} \beta(e^{(1-\gamma)y}1_{y<0})\,dy, \tag{4.1}$$

and at the location $n \geq i+m \geq i+1$ with parameter

$$\lambda_{i+m} := \int_{\varphi(i+m-1)-v}^{\varphi(i+m)-v} \beta(e^{\gamma y}1_{y>0})\,dy, \tag{4.2}$$

and at the location i itself with parameter

$$\lambda_{i} := \int_{\varphi(i-1)-v}^{\varphi(i)-v} \beta(e^{\gamma y}1_{y>0} + e^{(1-\gamma)y}1_{y<0})\,dy. \tag{4.3}$$

A vertex in location i in $\mathbb{T}(O_n)$ produces a Bernoulli distributed number of descendants in $i-m$ with success probability

$$p_{i-m} = \beta i^{\gamma-1}(i-m)^{-\gamma}.$$

This needs to be compared with

$$\lambda_{i-m} = \beta \int_{\varphi(i-m-1)-v}^{\varphi(i-m)-v} e^{(1-\gamma)y}\, dy = \tfrac{\beta}{1-\gamma}(e^{(1-\gamma)(\varphi(i-m)-v)} - e^{(1-\gamma)(\varphi(i-m-1)-v)}),$$

which by the mean value theorem equals

$$\beta e^{(1-\gamma)\xi}(\varphi(i-m) - \varphi(i-m-1)),$$

for some $\xi \in (\varphi(i-m-1)-\varphi(i), \varphi(i-m)-\varphi(i-1))$. The following lemma compares these quantities.

Lemma 4.6. *There exists a constant $C > 0$ such that the following holds: Let $m \in \mathbb{N}$ and $v \le 0$ with $i := \pi_n(v)$ and define $\lambda_{i\pm m}$ as in (4.1-4.3). Then*

$$\left|\lambda_{i-m} - p_{i-m}\right| \le C i^{\gamma-1}(i-m)^{-\gamma-1} \text{ and } \lambda_{i-m}^2 \le C i^{2\gamma-2}(i-m)^{-2\gamma},$$

and

$$\left|\lambda_{i+m} - p_{i+m}\right| \le C i^{-\gamma-1}(i+m)^{\gamma-1} \text{ and } \lambda_{i+m}^2 \le C i^{-2\gamma}(i+m)^{2\gamma-2},$$

If $m = 0$, a Poisson distributed random variable Υ with parameter λ_i satisfies

$$\mathbb{P}(\Upsilon \neq 0) \le C\,\frac{1}{i}.$$

Proof. We first focus on the minus sign. In this case, as seen above,

$$\lambda_{i-m} = \beta e^{(1-\gamma)\xi} \tfrac{1}{i-m},$$

for some $\xi \in (-\sum_{j=i-m}^{i} \frac{1}{j}, -\sum_{j=i-m+1}^{i-1} \frac{1}{j})$. By the Euler-MacLaurin formula the lower limit of this interval is at least

$$\log \tfrac{i-m}{i} - \tfrac{C}{i-m},$$

and the upper limit at most

$$\log \tfrac{i-m}{i} + \tfrac{C}{i-m}.$$

Hence

$$\lambda_{i-m} \le \beta i^{\gamma-1}(i-m)^{-\gamma} e^{\frac{C}{i-m}} \le p_{i-m} + p_{i-m}(e^{\frac{C}{i-m}} - 1)$$

and

$$\lambda_{i-m} \ge \beta i^{\gamma-1}(i-m)^{-\gamma} e^{-\frac{C}{i-m}} \ge p_{i-m} - p_{i-m}(1 - e^{-\frac{C}{i-m}}).$$

The estimates for the minus sign follow from this.

For the plus sign the argument is analogous. For $m = 0$ we have

$$\begin{aligned}\lambda_i &= \int_{\varphi(i-1)-v}^{0} \beta e^{(1-\gamma)y}\, dy + \int_{0}^{\varphi(i)-v} \beta e^{\gamma y}\, dy \\ &\le \tfrac{\beta}{1-\gamma}(1 - e^{(1-\gamma)(\varphi(i-1)-\varphi(i))}) + \tfrac{\beta}{\gamma}(e^{\gamma(\varphi(i)-\varphi(i-1))} - 1),\end{aligned}$$

which is $O(1/i)$. This implies the second statement of the lemma. ∎

Proof of Lemma 4.1. We look at the exploration of $\mathfrak{T}$. At every exploration step we couple the Poisson variables in the π_n-projected process to the Bernoulli variables in the corresponding labelled tree as in Lemma 4.5. For the exploration of $i = \pi_n(v)$ the probability of failure is bounded by a constant multiple of

$$\frac{1}{i} + i^{\gamma-1}\sum_{j=1}^{i-1} j^{-\gamma-1} + i^{2\gamma-2}\sum_{j=1}^{i-1} j^{-2\gamma} + i^{-\gamma-1}\sum_{j=i+1}^{n} j^{\gamma-1} + i^{-2\gamma}\sum_{j=i+1}^{n} j^{2\gamma-2}.$$

This bound behaves like

$$O(i^{\gamma-1}) + O(n^{\gamma} i^{-\gamma-1}).$$

The probability that the coupling fails before we have seen c_n vertices or a vertex with index smaller than n_n therefore goes to zero if $c_n/n_n^{1-\gamma} \to 0$ and also $c_n n^{\gamma}/n_n^{\gamma+1} \to 0$. □

4.1.4 The component of powerful vertices

We now give a criterion that when the exploration stops because a vertex in $\{1, \dots, n_n\}$ was found, then adding the immediate neighbours of this vertex makes the component of the root larger than c_n with high probability.

Lemma 4.7. *Let $(c_n)_{n\in\mathbb{N}}$ and $(n_n)_{n\in\mathbb{N}}$ be sequences of positive integers such that*

$$\lim_{n\to\infty} \frac{c_n n_n^{\gamma}}{n^{\gamma}} = 0.$$

Suppose that $\mathfrak{F}(T_n)$ is the σ-algebra of the exploration process of (G_n, O_n) at the stopping time T_n. Then, on the event that the process has stopped because a vertex in $\{1, \dots, n_n\}$ was found, we have

$$\mathbb{P}\big(|(G_n, O_n)| < c_n \,\big|\, \mathfrak{F}(T_n)\big) \longrightarrow 0.$$

Proof. Suppose o is the leftmost vertex discovered at the stopping time and I_n the set of vertices seen in the exploration up to T_n. Note that $|I_n| \le c_n$. Then the number of right neighbours of o not already discovered is the sum S_n of independent Bernoulli variables with expectation

$$\sum_{\substack{i=o+1\\ i\notin I_n}}^{n} o^{-\gamma} i^{\gamma-1} \sim \tfrac{1}{\gamma} o^{-\gamma} n^{\gamma}$$

and variance

$$\sum_{\substack{i=o+1\\ i\notin I_n}}^{n} o^{-\gamma} i^{\gamma-1}(1 - o^{-\gamma} i^{\gamma-1}) \sim \tfrac{1}{\gamma} o^{-\gamma} n^{\gamma}.$$

By Chebyshev's inequality

$$\mathbb{P}\big(|(G_n, O_n)| < c_n \,\big|\, \mathfrak{F}(T_n)\big) \le \mathbb{P}(S_n < c_n) \le \mathbb{P}(|S_n - \mathbb{E}S_n| > \tfrac{1}{\gamma} o^{-\gamma} n^\gamma - c_n)$$

$$\le \frac{o^{-\gamma} n^\gamma}{(o^{-\gamma} n^\gamma - \gamma c_n)^2} \le \frac{n_n^\gamma}{n^\gamma}(1 + o(1)) \longrightarrow 0,$$

as required. □

4.2 Survival of the killed branching random walk

We will come back to the coupling and check how to use it to get information about the giant in the third section. But before that we collect information on the limiting object, the killed branching random walk by a probabilistic analysis based on martingale theory.

4.2.1 Martingales associated with branching random walks

We first look at a branching random walk with offspring distribution given by a Poisson process with intensity ν without killing. Given $\alpha > 0$ we define

$$\rho(\alpha) := \int e^{-\alpha t}\, \nu(dt).$$

Although we need the results only for a specific ν (i.e. the π defined above) we will work with the general case assuming just that ν is not degenerate in the sense that there is mass on the left half-axis and there exists an open interval $I \subset (0, \infty)$ (possibly empty) such that $\rho(\alpha) < \infty$ for all $\alpha \in I$ and $\rho(\alpha)$ diverges to infinity when α approaches the boundary points of the interval I.

We describe the nth generation of this process started with a particle in the origin as a sequence of measures $(Y^{(n)}(dx))_n$. With every generation of particles we associate a score

$$X_n := \int e^{-\alpha x}\, Y^{(n)}(dx).$$

This score will be used to define martingales and supermartingales.

Lemma 4.8.

(a) If $\rho(\alpha) < \infty$, then $(\rho(\alpha)^{-n} X_n \colon n \in \mathbb{N})$ is a martingale.

(b) If $\rho(\alpha) \leq 1$, then $(X_n \colon n \in \mathbb{N})$ is a supermartingale.

Proof. With the natural filtration $(\mathscr{F}_n)$ we have

$$\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] = \mathbb{E}\Big[\iint e^{-\alpha(x+y)} P_x(dy) Y^{(n)}(dx) \mid \mathscr{F}_n \Big],$$

where P_x are independent Poisson processes with intensity ν, which are also independent of $\mathscr{F}_n$. Hence this equals

$$\int \mathbb{E}\Big[\int e^{-\alpha(x+y)} P(dy) \Big] Y^{(n)}(dx),$$

where the Poisson process P has intensity ν. Hence

$$\mathbb{E}\Big[\int e^{-\alpha(x+y)} P(dy) \Big] = \int e^{-\alpha(x+y)} \nu(dy) = e^{-\alpha x} \rho(\alpha).$$

Altogether

$$\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] = \rho(\alpha) X_n,$$

from which the claim follows. □

Proposition 4.2. *If there exists $\alpha > 0$ such that $\rho(\alpha) \leq 1$ the killed branching random walk becomes extinct almost surely.*

Proof. Suppose that such an α exists. Then $(X_n \colon n \in \mathbb{N})$ is a supermartingale and thus almost surely convergent. Now fix some $N > 1$, an integer $n \geq 2$ and the state at generation $n-1$. Suppose there is an individual with location $x < N$ in the $(n-1)$st generation. Then there is a positive probability (depending on N but not on n) that $|X_n - X_{n-1}| > 1$ and, as $(X_n \colon n \in \mathbb{N})$ converges, this can only happen for finitely many n. Hence either the branching random walk dies out in finite time or the location of the leftmost particle in $(Y^{(n)}(dx))_n$ diverges to $+\infty$ almost surely. This implies that the killed process dies out almost surely. □

4.2.2 A Kesten-Stigum result

We define a nonnegative martingale by

$$W^{(n)} = \rho(\alpha)^{-n} \int e^{-\alpha x} \, Y^{(n)}(dx),$$

and find a sharp criterion that the almost sure limit W is positive.

Lemma 4.9. *The mapping*

$$\log\rho\colon (0,\infty)\to(-\infty,\infty],\quad \alpha\mapsto\log\rho(\alpha)$$

is convex.

Proof. This is an easy exercise. □

Proposition 4.3 (Biggins (1977))**.** *The martingale* $(W^{(n)})_{n\in\mathbb{N}}$ *converges almost surely on survival to a strictly positive limit* W *if*

$$\log\rho(\alpha)-\frac{\alpha\,\rho'(\alpha)}{\rho(\alpha)}>0\quad\text{and}\quad\mathbb{E}\big[W^{(1)}\log^+W^{(1)}\big]<\infty. \tag{4.4}$$

Moreover, under this assumption the survival probability is positive and almost surely on survival there exist a ray (v_n) *of individuals in the tree with positions* x_n *in generation* n *satisfying*

$$\frac{x_n}{n}\longrightarrow-\frac{\rho'(\alpha)}{\rho(\alpha)}.$$

The result is sharp, as when (4.4) fails, then $W=0$ almost surely, but we will only use the claimed direction. The rest of this subsection is devoted to the proof of this result, we follow Lyons (1995). Given a rooted tree (with distinguishable children) with vertices labelled by displacements on the real line, a *ray* is an infinite line of descent starting at the root. The space of rooted labelled trees gives rise to a filtration $(\mathfrak{F}_n)$, as does the space of rooted labelled trees with a distinguished ray, we denote the latter by $(\mathfrak{F}_n^*)$.

We write (t,X) for a tree and its labels $X\colon t\to\mathbb{R}$. For a vertex $\sigma\in t$ we write $X(\sigma)$ for its label and

$$S(\sigma)=\sum_{o<\tau\le\sigma}X(\tau)$$

for its position. Let

$$W^{(n)}(t,X)=\rho(\alpha)^{-n}\sum_{|\sigma|=n}e^{-\alpha S(\sigma)},$$

which when (t,X) is the labelled tree given by the Poisson branching random walk defines the martingale above. We denote by μ the law of (t,X). We now define an associated law $\hat{\mu}^*$ on the *infinite* rooted labelled trees with a distinguished ray. Let μ_n be the restriction of μ to $\mathfrak{F}_n$. Any $\mathfrak{F}_n^*$-measurable function f can be written as

$$f(t,X,\xi)=\sum_{|\sigma|=n}f_\sigma(t,X)1_{\xi_n=\sigma}.$$

Define μ_n^* as the (not necessarily probability) measure on $\mathfrak{F}_n^*$ given by

$$\int f(t,X,\xi)\,d\mu_n^* = \int \sum_{|\sigma|=n} f_\sigma(t,X)d\mu,$$

and a probability measure $\hat{\mu}_n^*$ on $\mathfrak{F}_n^*$ by

$$\frac{d\hat{\mu}_n^*}{d\mu_n^*}(t,X,\xi) = \rho(\alpha)^{-n}e^{-\alpha S(\xi_n)}.$$

Lemma 4.10. *There exists a (unique) probability measure $\hat{\mu}^*$ such that the restriction to $\mathfrak{F}_n^*$ is $\hat{\mu}_n^*$ for all n.*

Proof. We define $\hat{P}$ as the random variable with law given by the density

$$\frac{\int e^{-\alpha x}P(dx)}{\rho(\alpha)}$$

with respect to the law of the Poisson process P. We generate an iid sequence $\hat{P}_0, \hat{P}_1, \hat{P}_2, \ldots$ of random variables with this law. To define $\hat{\mu}^*$ we start with an individual v_0 at the origin and use $\hat{P}_0$ to generate its offspring. Out of the (nonvanishing) offspring pick an individual $v_1 = v$ with respect to the probability vector

$$\frac{e^{-\alpha X(v)}}{\int e^{-\alpha x}\hat{P}_0(dx)}.$$

We use the positions of all individuals $v \neq v_1$ to start independent branching random walks and use $\hat{P}_1$ to determine the offspring of v_1 and its displacements. Continue by picking $v_2 = v$ using the probability vector

$$\frac{e^{-\alpha X(v)}}{\int e^{-\alpha x}\hat{P}_1(dx)},$$

use the positions of all offspring $v \neq v_2$ to start independent branching random walks, use $\hat{P}_2$ to determine the offspring of v_2 and continue like this ad infinitum.

Now let $\hat{\mu}^*$ be the joint law of the random labelled tree and distinguished ray $(v_0, v_1, v_2, \ldots)$. Then we have

$$\begin{aligned}\frac{d\hat{\mu}_{n+1}^*}{d\mu_{n+1}^*}(t,X,\xi) &= \frac{d\hat{\mu}_n^*}{d\mu_n^*}(t,X,\xi)\frac{\int e^{-\alpha x}P_n(dx)}{\rho(\alpha)}\frac{e^{-\alpha X(\xi_{n+1})}}{\int e^{-\alpha x}P_n(dx)} \\ &= \frac{e^{-\alpha X(\xi_{n+1})}}{\rho(\alpha)}\frac{d\hat{\mu}_n^*}{d\mu_n^*}(t,X,\xi).\end{aligned}$$

The claimed property of $\hat{\mu}_n^*$ follows from this by induction. □

The projection of $\hat{\mu}^*$ onto the space of rooted labelled trees (forgetting which ray has been distinguished) is denoted by $\hat{\mu}$ and satisfies

$$\frac{d\hat{\mu}_n}{d\mu_n}(t, X) = W^{(n)}(t, X), \tag{4.5}$$

for all n and labelled trees (t, X), where $\hat{\mu}_n$ is the restriction of $\hat{\mu}$ to $\mathfrak{F}_n$.

Let $(v_0, v_1, v_2, \ldots)$ be the distinguished ray generated by $\hat{\mu}^*$. Then, for any $k \in \mathbb{N}$, we have

$$\begin{aligned}\int X(v_k)\, d\hat{\mu}^* &= \mathbb{E}\bigg[X(v_k)\, \frac{e^{-\alpha X(v_k)}}{\int e^{-\alpha y} P_{v_{k-1}}(dy)} \frac{\int e^{-\alpha y} P_{v_{k-1}}(dy)}{\rho(\alpha)}\bigg] \\ &= \rho(\alpha)^{-1}\mathbb{E}\big[\int y e^{-\alpha y} P(dy)\big] = -\frac{\rho'(\alpha)}{\rho(\alpha)}.\end{aligned}$$

As under $\hat{\mu}^*$ the labels $(X(v_k)\colon k \in \mathbb{N})$ are iid, the strong law of large numbers gives

$$\frac{S(v_n)}{n} = \frac{1}{n}\sum_{k=1}^{n} X(v_k) \to -\frac{\rho'(\alpha)}{\rho(\alpha)} \qquad \hat{\mu}^*\text{-almost surely.}$$

In particular a ray with

$$\frac{x_n}{n} \longrightarrow -\frac{\rho'(\alpha)}{\rho(\alpha)}$$

exists $\hat{\mu}$-almost surely.

Now denote $\hat{P}_k$ be the random variable used to generate the offspring of distinguished vertex v_k and let $\mathscr{G}$ be the σ-algebra generated by $\hat{P}_0, \hat{P}_1, \hat{P}_2, \ldots$ and $v_1, v_2, \ldots$. Let

$$V_{k+1} = \int e^{-\alpha x} \hat{P}_k(dx).$$

Further let $\hat{\mathbb{E}}^*$ be the the expectation with respect to $\hat{\mu}^*$. Then

$$\begin{aligned}\hat{\mathbb{E}}^*\big[W^{(n)}(t, X)\big|\mathscr{G}\big] &= \sum_{k=0}^{n-1} \frac{e^{-\alpha S(v_k)}}{\rho(\alpha)^{k+1}}\Big(V_{k+1} - e^{-\alpha X(v_{k+1})}\Big) + \frac{e^{-\alpha S(v_n)}}{\rho(\alpha)^n} \\ &= \sum_{k=0}^{n-1} \frac{e^{-\alpha S(v_k)}}{\rho(\alpha)^{k+1}} V_{k+1} - \sum_{k=1}^{n-1} \frac{e^{-\alpha S(v_k)}}{\rho(\alpha)^k}.\end{aligned}$$

The terms $e^{-\alpha S(v_k)}/\rho(\alpha)^k$ decay exponentially by our assumption, while V_{k+1} grow at most subexponentially as

$$\mathbb{E}[\log^+ V_k] = \rho(\alpha)^{-1}\mathbb{E}\big[W^{(1)} \log^+ W^{(1)}\big] < \infty,$$

using Fubini and Borel-Cantelli. Therefore the series on the right both converge $\hat{\mu}^*$-almost surely. By Fatou's lemma we get

$$\liminf_{n\to\infty} W^{(n)}(t,X) < \infty \qquad \hat{\mu}\text{-almost surely.}$$

Now (4.5) implies that $1/W^{(n)}(t,X)$ defines a martingale with respect to $\hat{\mu}$. Hence $(W^{(n)}(t,X))_n$ converges $\hat{\mu}$-almost surely and together with the result on the liminf we get that the limit, denoted $W(t,X)$, is $\hat{\mu}$-almost surely finite.

We conclude the argument by noting that

$$d\hat{\mu}(t,X) = W(t,X)\, d\mu(t,X),$$

which implies that $\mathbb{P}(W > 0)$ is positive. Because of the independence of the offspring one can infer that $W > 0$ almost surely on survival, see exercises for details. In particular the existence of the convergent ray holds μ-almost surely on survival.

4.2.3 A sharp criterion for survival

We now complement Proposition 4.2 with the converse result to give a necessary and sufficient criterion for a positive survival probability.

Proposition 4.4. *If for all $\alpha > 0$ we have $\rho(\alpha) > 1$, then the killed branching random walk survives with positive probability.*

Proof. First, assume that $1 < \rho(\alpha) < \infty$ and the second condition in (4.4) holds for all $\alpha > 0$ and that there exists $\alpha_0 > 0$ such that

$$\rho(\alpha_0) = \min_{\alpha>0} \rho(\alpha) > 1.$$

By convexity and continuous differentiability of $\log \rho$ there exists $\alpha > \alpha_0$ such that

$$\log \rho(\alpha) - \frac{\alpha\, \rho'(\alpha)}{\rho(\alpha)} > 0 \text{ and } \frac{\rho'(\alpha)}{\rho(\alpha)} > 0.$$

Then, by Proposition 4.3, with positive probability the limit W is positive and there exists a ray with

$$\lim_{n\to\infty} \frac{x_n}{n} = -\frac{\rho'(\alpha)}{\rho(\alpha)} < 0.$$

Second, to ensure that the second condition in (4.4) holds we can use a cut-off procedure, and replace the offspring distribution $Y^{(1)}(dx)$ by one that takes only the offspring within distance N to the parent into account. It is

easy to see that, for fixed α and sufficiently large N, we can ensure that the new (finite) value $\rho^{(N)}(\alpha)$ replacing $\rho(\alpha)$ is close to the original one if this was finite, and as large as we wish if it was infinite, i.e.

$$\rho^{(N)}(\alpha) \nearrow \rho(\alpha).$$

By our assumptions on ρ there exists an α_0 such that

$$\rho(\alpha_0) = \min_{\alpha \in I} \rho(\alpha) > 1.$$

Using the fact that a sequence of convex functions, which converges pointwise, converges uniformly on every closed set, we can choose N so that the function $\rho^{(N)}$ takes its minimum in an open interval around α_0 and this minimum is strictly bigger than one, while the cut-off ensures that the second criterion in (4.4) automatically holds. The argument above can now be applied and yields the existence of an ancestral line of particles diverging to $-\infty$, which then automatically also exists in the original branching random walk.

We get that in the (not killed) branching random walk with positive probability there exists an ancestral line of particles diverging to $-\infty$. Because ν has mass on the left half-axis, for any $M > 0$, with positive probability the killed branching random walk has a particle in a position $< -M$. These two statements together imply that the killed branching process started in any position on the left half-axis has a positive probability of survival. □

4.3 Existence of the giant component

Recall that we say that *a giant component exists* in the sequence of graphs $(G_n)_{n\in\mathbb{N}}$ if the proportion of vertices in the largest connected component $C_n \subset G_n$ converges, for $n \uparrow \infty$, in probability to a positive number.

Theorem 10 (Existence of a giant component)**.** *The proportion of vertices in the largest component C_n of the simple preferential attachment graph G_n converges, as $n \uparrow \infty$, in probability to*

$$\mathbb{P}(|\mathfrak{T}| = \infty),$$

where $\mathfrak{T}$ is the tree associated with the unimodular killed Poisson branching random walk with intensity π. In particular, a giant component exists if and only if $\mathfrak{T}$ is infinite with positive probability.

We prove this result in several step in the remainder of this section. Theorem 9 follows from this as we have seen that the tree $\mathfrak{T}$ is almost surely finite if there exists $\alpha > 0$ such that $\rho(\alpha) \leq 1$ for

$$\begin{aligned}\rho(\alpha) &= \int e^{-\alpha t}\, d\pi(t) \\ &= \beta \int_0^\infty e^{(\gamma-\alpha)t}\, dt + \beta \int_{-\infty}^0 e^{(1-\gamma-\alpha)t}\, dt \\ &= \frac{\beta}{\alpha-\gamma} + \frac{\beta}{1-\gamma-\alpha},\end{aligned}$$

if $\gamma < \alpha < 1-\gamma$ and infinity otherwise. Only if $\gamma < \frac{1}{2}$ we can pick α with $\rho(\alpha)$ finite and in this case $\alpha = \frac{1}{2}$ minimizes $\rho(\alpha)$ and gives

$$\rho(\tfrac{1}{2}) = \frac{2\beta}{\frac{1}{2}-\gamma},$$

which is no bigger than one if

$$\beta \leq \frac{1}{4} - \frac{\gamma}{2}.$$

4.3.1 Sprinkling

Our proofs, in particular the crucial sprinkling technique, relies on the following continuity property of the survival probability

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

of the killed branching random walk as a function of the edge density parameter β.

Lemma 4.11. *We have*

$$\lim_{\varepsilon \searrow 0} p(\beta - \varepsilon) = p(\beta).$$

Proof. We only need to consider the case where $p(\beta) > 0$, as otherwise both sides of the equation are zero. This assumption implies, by Lemma 4.4, that for all $0 < \alpha < 1$ we have $\rho(\alpha, \beta) := \rho(\alpha) > 1$. As

$$\rho(\alpha, \beta) = \frac{\beta}{\alpha-\gamma} + \frac{\beta}{1-\gamma-\alpha} \text{ for } \gamma < \alpha < 1-\gamma$$

and infinity otherwise, for all sufficiently small $\varepsilon > 0$ we have $\rho(\alpha, \beta-\varepsilon) > 1$ for all $0 < \alpha < 1$. Thus, using again Lemma 4.4, we have $p(\beta - \varepsilon) > 0$.

Now we look at the killed branching random walk started with one particle in position t, constructed using edge intensity parameter $\beta-\varepsilon$. We denote by $E(\varepsilon,t)$ the event this process survives forever, and by $V(\varepsilon,t,\kappa)$ the probability that a particle reaches a site $<\kappa$. Then we have

$$\lim_{\kappa\to-\infty}\inf_{t\le\kappa}\mathbb{P}\big(E(\varepsilon,t)\big)=1.$$

For fixed $\kappa<0$ and $0\le\varepsilon\le\varepsilon_0$ we have

$$\mathbb{P}\big(E(\varepsilon,t)\big)\ge\mathbb{P}\big(V(\varepsilon,t,\kappa)\big)\,\mathbb{P}\big(E(\varepsilon_0,\kappa)\big)\xrightarrow{\varepsilon\downarrow0}\mathbb{P}\big(V(0,t,\kappa)\big)\,\mathbb{P}\big(E(\varepsilon_0,\kappa)\big).$$

Note that the first probability on the right is greater or equal to $p(\beta)$ and that the second probability tends to one, as κ tends to $-\infty$. □

The crucial tool of this section is the 'sprinkling' argument in Proposition 4.5, which shows that the extra edges generate when increasing β by ϵ suffice to connect large components to a giant. Recall that a sequence of events depending on the index n holds *with high probability* if the probability of these events goes to one as $n\uparrow\infty$.

Proposition 4.5 (Sprinkling argument)**.** *Suppose that $(c_n)_{n\in\mathbb{N}}$ is a sequence of integers with*

$$\frac{c_n^2}{n}\to0,\quad c_n-\log n\to\infty$$

and that, for the graphs $(\bar G_n)_{n\in\mathbb{N}}$ with density $0<\bar\beta<\beta$ and $\kappa>0$ we have

$$\sum_{i=1}^{n}1\{|\bar C_n(i)|\ge2c_n\}\ge\kappa n\qquad\text{with high probability,}$$

where $\bar C_n(i)$ is the connected component of the vertex i in $\bar G_n$. Then there exists a coupling of the graph sequences $(G_n)_{n\in\mathbb{N}}$ and $(\bar G_n)_{n\in\mathbb{N}}$ such that $\bar G_n\le G_n$ and all connected components of $\bar G_n$ with at least $2c_n$ vertices belong to one connected component in G_n with at least κn vertices, with high probability.

Proof. Let $\varepsilon=\beta-\bar\beta>0$. Note that we can couple $\bar G_n$ and an independent Erdős-Rényi graph G_n^{ER} with edge probability ε/n with G_n such that

$$\bar G_n\le\bar G_n\vee G_n^{ER}\le G_n.\tag{4.6}$$

Here, $\bar G_n\vee G_n^{ER}$ denotes the graph in which all edges are open that are open in at least one of the two graphs, and $G'\le G''$ means that all edges that are open in G' are also open in G''. We denote by V_n' the vertices in $\bar G_n$ that belong to components of size at least $2c_n$ and write V_n' as the disjoint union $C_1\cup\cdots\cup C_M$, where $C_1,\dots,C_M$ are sets of vertices such that,

- $|C_j| \in [c_n, 2c_n]$ and
- C_j belongs to one component in $\bar{G}_n$, for each $j = 1, \dots, M$.

Recall (4.6), and note that given $\bar{G}_n$ and the sets $C_1, \dots, C_M$, the Erdős-Rényi graph G_n^{ER} connects two distinct sets C_i and C_j with probability at least

$$p_n := 1 - \big(1 - \tfrac{\varepsilon}{n}\big)^{c_n^2} \geq 1 - e^{-\frac{\varepsilon}{n} c_n^2} \sim \frac{\varepsilon}{n}\, c_n^2.$$

By identifying the individual sets as one vertex and interpreting the G_n^{ER}-connections as edges, we obtain a new random graph. Certainly, this dominates an Erdős-Rényi graph with M vertices and success probability p_n, which has edge intensity Mp_n. By assumption, $\frac{1}{2}\frac{\kappa n}{c_n} \leq M \leq n$ with high probability. Hence $M \to \infty$ and $Mp_n - \log M \to \infty$ in probability as $n \uparrow \infty$. We have seen in an exercise that the new Erdős-Rényi graph is connected with high probability. Hence, all vertices of V_n' belong to one connected component in G_n, with high probability. □

4.3.2 The variance of the cluster size

In this section we provide the second moment estimate needed to show that our key empirical quantity, the number of vertices in connected components of a given size, concentrate asymptotically near their mean.

Proposition 4.6. *Suppose that $(c_n)_{n\in\mathbb{N}}$ and $(n_n)_{n\in\mathbb{N}}$ are sequences of integers satisfying $1 \leq c_n, n_n \leq n$. Then, for a constant $C > 0$ we have*

$$\begin{aligned} &Var\Big(\frac{1}{n}\sum_{v=1}^{n} 1\{|C_n(v)| \geq c_n\}\Big) \\ &\quad \leq 2\,\mathbb{P}\big(|C_n(O_n)| < c_n \text{ and } C_n(O_n) \cap \{1, \dots, n_n\} \neq \emptyset\big) \\ &\qquad\qquad + \frac{c_n}{n} + C\, c_n^2 n_n^{-1} + C c_n n_n^{\gamma-1}, \end{aligned}$$

where O_n is independent of G_n and uniformly distributed on $\{1, \dots, n\}$.

Proof. Let v, w be two distinct vertices of G_n. We start by exploring the neighbourhood of v similarly as before. As before we classify the vertices as veiled, active and dead, and in the beginning only v is active and the remaining vertices are veiled. In one exploration step we pick the leftmost active vertex and consecutively (from the left to the right) explore its immediate neighbours in the set of veiled vertices only. Newly found vertices are activated and the vertex explored is set to dead after the exploration. We immediately stop the exploration once one of the events

(A) the number of unveiled vertices in the cluster reaches c_n,

(B) one vertex in $\{1, \dots, n_n\}$ is activated, or

(C) there are no more active vertices left,

happens. Note that when we stop due to (A) or (B) the exploration of the last vertex might not be finished. In that case we call this vertex *semi-active*.

We proceed with a second exploration process, namely the exploration of the cluster of w. This exploration follows the same rules as the first exploration process, where we treat the vertices that remained active or semi-active at the end of the first exploration as veiled. In addition to the stopping in the cases (A), (B), (C) we also stop the exploration once a vertex is unveiled which was also unveiled in the first exploration, calling this event (D). We consider the following events:

E^v : the first exploration started with vertex v ends in (A) or (B);

$E_1^{v,w}$: w is unveiled during the first exploration (that of v);

$E_2^{v,w}$: w remains veiled in the first exploration and the second exploration ends in (A) or (B) but not in (D);

$E_3^{v,w}$: w remains veiled in the first exploration and the second exploration ends in (D).

We have

$$\sum_{v=1}^{n}\sum_{w=1}^{n}\mathbb{P}(|C_n(v)| \geq c_n, |C_n(w)| \geq c_n) \leq \sum_{v=1}^{n}\sum_{w=1}^{n}\sum_{k=1}^{3}\mathbb{P}(E^v \cap E_k^{v,w}) = \sum_{v=1}^{n}\mathbb{P}(E^v)\sum_{k=1}^{3}\sum_{w=1}^{n}\mathbb{P}(E_k^{v,w} \mid E^v).$$

As the first exploration immediately stops once one has unveiled c_n vertices, we conclude that, for fixed v,

$$\sum_{w=1}^{n}\mathbb{P}(E_1^{v,w} \mid E^v) = \mathbb{E}\Big[\sum_{w=1}^{n} 1_{E_1^{v,w}} \Big| E^v\Big] \leq c_n. \tag{4.7}$$

To analyse the remaining terms, we fix distinct vertices v and w and note that the configuration after the first exploration can be formally described by an element $\mathfrak{k}$ of

$$\{\text{open, closed, unexplored}\}^{E_n},$$

where $E_n := \{(i,j) \in \{1, \dots, n\}^2 : i < j\}$ denotes the set of possible edges. We pick a feasible configuration $\mathfrak{k}$ and denote by $\mathcal{E}_\mathfrak{k}$ the event that the first

exploration ended in this configuration. On the event $\mathcal{E}_{\mathfrak{k}}$ the status of each vertex (veiled, active, semi-active or dead) at the end of the first exploration is determined. Suppose $\mathfrak{k}$ is such that w remained veiled in the first exploration, which means that $\mathcal{E}_{\mathfrak{k}}$ and $E_1^{v,w}$ are disjoint events. Next, we note that

$$\mathbb{P}(E_2^{v,w} | \, \mathcal{E}_{\mathfrak{k}}) \leq \mathbb{P}(E^w). \tag{4.8}$$

Indeed, if in the exploration of w we encounter an edge which is open in the configuration $\mathfrak{k}$, we have unveiled a vertex which was also unveiled in the exploration of v, the second exploration ends in (D) and hence $E_2^{v,w}$ does not happen. Otherwise, the event $\mathcal{E}_{\mathfrak{k}}$ does not influence the exploration of w and hence we obtain (4.8).

Finally, we analyse the probability $\mathbb{P}(E_3^{v,w} | \, \mathcal{E}_{\mathfrak{k}})$. If the second exploration process ends in state (D) we have discovered an edge connecting the second exploration to an active or semi-active vertex from the first exploration. Recall that there are at most c_n such vertices and at most one of them is in $\{1, \dots, n_n\}$. For each of these we have to test the existence of edges no more than c_n times. Hence we find $C > 0$ such that

$$\mathbb{P}\big(E_3^{v,w} \big| E^v\big) \leq C c_n^2 n_n^{-1} + C c_n n_n^{\gamma - 1}.$$

Summarising our steps, we have

$$\begin{aligned}
&Var\Big(\frac{1}{n}\sum_{v=1}^{n} 1\{|C_n(v)| \geq c_n\}\Big) \\
&\quad \leq \mathbb{E}\Big[\frac{1}{n^2}\sum_{v=1}^{n}\sum_{w=1}^{n} 1\{|C_n(v)| \geq c_n, |C_n(w)| \geq c_n\}\Big] - \frac{1}{n^2}\sum_{v=1}^{n}\sum_{w=1}^{n}\mathbb{P}(E^v)\mathbb{P}(E^w) \\
&\qquad + 2\,\frac{1}{n}\sum_{v=1}^{n}\mathbb{P}\big(|C_n(v)| < c_n \text{ and } C_n(v) \cap \{1, \dots, n_n\} \neq \emptyset\big) \\
&\quad \leq 2\,\mathbb{P}\big(|C_n(O_n)| < c_n \text{ and } C_n(O_n) \cap \{1, \dots, n_n\} \neq \emptyset\big) \\
&\qquad\qquad + \frac{c_n}{n} + C\,c_n^2 n_n^{-1} + C c_n n_n^{\gamma - 1},
\end{aligned}$$

as required to complete the proof. □

4.3.3 Proof of Theorem 10

We start by proving the lower bound. Suppose therefore that $p(\beta) > 0$, fix $\delta > 0$ arbitrarily small and use Lemma 4.11 to choose $\varepsilon > 0$ such that the survival probability of $\bar{\beta} = \beta - \varepsilon$ is larger than $p(\beta) - \delta$. We denote by $(\bar{G}_n)_{\in \mathbb{N}}$ a sequence of random graphs with edge intensity parameter $\bar{\beta}$ and

let $\bar{C}_n(v)$ the connected component of v in $\bar{G}_n$. Suppose a vertex V is chosen uniformly at random from $\{1,\dots,n\}$. We choose

$$c_n := \lceil (\log n)^2 \rceil$$

and

$$n_n := \lceil n^{\frac{\gamma}{\gamma+1}} (\log n)^5 \rceil.$$

Observe that

$$\lim_{n\to\infty} \frac{c_n}{n_n^{1-\gamma}} = 0 \text{ and } \lim_{n\to\infty} \frac{c_n n^\gamma}{n_n^{\gamma+1}} = 0 \text{ and } \lim_{n\to\infty} \frac{c_n^2}{n_n} = 0.$$

and

$$\lim_{n\to\infty} \frac{c_n n_n^\gamma}{n^\gamma} = 0.$$

As, by Proposition 4.1 and Lemma 4.7,

$$\begin{aligned}\mathbb{E}\Big[\frac{1}{n}\sum_{v=1}^{n} 1\{|\bar{C}_n(v)| \geq c_n\}\Big] &= \mathbb{P}\{(\bar{G}_n, O_n) \curlywedge T_n \text{ stops in (A) or (B)}\} + o(1)\\ &= \mathbb{P}\{\bar{\mathfrak{T}} \curlywedge T_n \text{ stops in (A) or (B)}\} + o(1) \longrightarrow \mathbb{P}\{|\bar{\mathfrak{T}}| = \infty\}, \qquad (4.9)\end{aligned}$$

where $\bar{\mathfrak{T}}$ is the unimodular branching random walk using intensity parameter $\bar{\beta}$. We infer from this that

$$\lim_{n\to\infty} \mathbb{E}\Big[\frac{1}{n}\sum_{v=1}^{n} 1\{|\bar{C}_n(v)| \geq c_n\}\Big] \geq p(\beta) - \delta,$$

as n tends to infinity. By Proposition 4.6 we have

$$\begin{aligned}Var\Big(\frac{1}{n}\sum_{v=1}^{n} 1_{|\bar{C}_n(v)|\geq c_n}\Big) &\leq 2\,\mathbb{P}\big(|\bar{C}_n(O_n)| < c_n \text{ and } \bar{C}_n(O_n) \cap \{1,\dots,n_n\} \neq \emptyset\big)\\ &\quad + \frac{c_n}{n} + C\, c_n^2 n_n^{-1} + C c_n n_n^{\gamma-1}.\end{aligned}$$

The first summand goes to zero by Lemma 4.7 and so do the remaining terms, by the choice of our parameters. Hence

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1\{|\bar{C}_n(v)| \geq c_n\} \geq p(\beta) - \delta \qquad \text{in probability,}$$

and as

$$\frac{c_n^2}{n} \to 0, \quad c_n - \log n \to \infty,$$

Proposition 4.5 implies that, with high probability, there exists a component comprising at least a proportion $p(\beta)$ of all vertices, proving the lower bound.

To see the upper bound we work with the original intensity β. In analogy to (4.9) we obtain

$$\lim_{n\to\infty} \mathbb{E}\Big[\frac{1}{n}\sum_{v=1}^{n} 1\{|C_n(v)| \geq c_n\}\Big] = p(\beta).$$

As in the lower bound, the variance goes to zero, and hence we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1\{|C_n(v)| \geq c_n\} = p(\beta) \qquad \text{in probability.}$$

From this we infer that, in probability, the size of the largest component C_n satisfies

$$\limsup_{n\to\infty} \frac{|C_n|}{n} \leq \limsup_{n\to\infty} \frac{c_n}{n} \vee \Big(\frac{1}{n}\sum_{v=1}^{n} 1\{|C_n(v)| \geq c_n\}\Big) \leq p(\beta)$$

proving the upper bound.

Corollary 4.1. *The proportion of vertices of G_n, which are not in the largest component but in components of size at least $(\log n)^2$ goes to zero in probability. In particular the proportion of vertices in the second largest component of G_n vanishes asymptotically.*

Proof. Note that we have seen, with c_n as before, that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{v=1}^{n} 1\{|C_n(v)| \geq c_n\} = p(\beta) \qquad \text{in probability,}$$

so that, with high probability, the proportion of vertices in clusters of size $\geq c_n$ is asymptotically equal to the proportion of vertices in the giant component. Hence the proportion of vertices, which are not in the giant component but in components of size at least c_n goes to zero in probability. □

4.4 The phase transition: A closer look at cluster sizes

For the simple preferential attachment one can ask further questions, that may be suitable topics for master thesis research. The following questions (and many others) come to mind:

- When $\gamma < \frac{1}{2}$, how quickly does the size of the giant component go to zero when $\beta \downarrow \beta_c > 0$? This has been studied for a more sophisticated model by Eckhoff, Mörters and Ortgiese.

- When $\gamma < \frac{1}{2}$ and $\beta < \beta_c$, how quickly does

$$\lim_{n\to\infty} \mathbb{P}\{|(G_n, O_n)| \geq k\}$$

  decay as $k \to \infty$? There is a useful recent paper on the corresponding problem for killed Poisson branching random walks, see Aidekon, Hu and Zindy.

- When $\gamma < \frac{1}{2}$ and $\beta > \beta_c$ what is the size of the second largest component? Can this be related to $\mathbb{P}\{k \leq |\mathfrak{T}| < \infty\}$ for $k \to \infty$ and can we find its asymptotic behaviour?

- When $\gamma < \frac{1}{2}$ and $\beta < \beta_c$, what is the size of the largest component? Here the branching process approximation cannot be used directly, which makes this harder. The answer is known for the Chung-Lu model, see van der Hofstad II, Theorem 3.22.

- When $\gamma \geq \frac{1}{2}$, how does the size of the giant component behave as $\beta \downarrow 0$? The paper by Eckhoff, Mörters and Ortgiese might have some clues in this case, too, but the answer is less clear.

- When $\gamma \geq \frac{1}{2}$, what is the size of the second largest component? Same problem as before and equally interesting and challenging problem.

Chapter 5

Literaturliste

Aidekon, Hu and Zindy "The precise tail behavior of the total progeny of a killed branching random walk". Ann. Probab. 41: 3786-3878 (2013).

Dereich and Mörters "Random networks with sublinear preferential attachment: The giant component". Ann. Probab. 41: 329-384 (2013).

Eckhoff, Mörters and Ortgiese "Near critical preferential attachment networks have small giant components". J. Stat. Phys. 173: 663-703 (2018).

van der Hofstad "Random graphs and complex networks", Volume I and II. Cambridge University Press

Kurtz "Almost sure convergence of the type distribution for a supercritical branching process". Unpublished (1973).

Lyons "A simple path to Biggins martingale convergence for branching random walk". In: Classical and modern branching processes (1995).