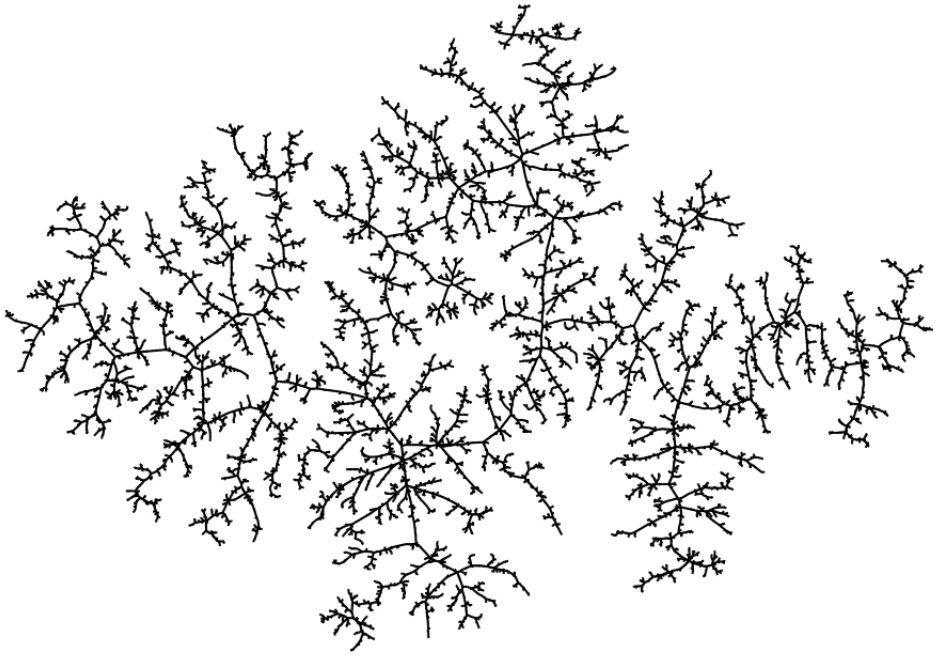
Bachelor-thesis

The Scaling Limit of the Erdős-Rényi Graph



David Geldbach

Supervisor: Prof. Markus Heydenreich
Ludwig-Maximilians-University Munich

Contents

The picture on the titlepage was made by Nicolas Broutin, source: MATHS.OX.AC.UK/NODE/30217.

1 Introduction

This thesis describes the statistical properties of the Erdős-Rényi graph at the critical point. The Erdős-Rényi graph is the most important model for a random graph and simply consists of n vertices which are connected independently with probability p. In the classical paper *"On the evolution of Random Graphs"* [6], it was proven that the Erdős-Rényi graph exhibits a phase transition when we fix $np = 1$ and let n tend to infinity. Thus $p = 1/n$ defines the critical point of the model.

Phase transitions not only have very interesting mathematical properties as discussed in this thesis, they are also very important in the natural sciences, for example when water freezes below 0°C or evaporates above 100°C. A less obvious example for a phase transition is a virus becoming epidemic after infecting a certain threshold of the population. This shows that it is the mathematical concept of the phase transition rather than its realization which is important to understand such transitions. The broad implications of such models have already been commented on in the original paper by Erdős and Rényi [6]:

> "The evolution of random graphs may be considered as a (rather simplified) model of the evolution of certain real communication-nets, e.g. the railway-, road- or electric network system of a country or some other unit, or the growth of structures of anorganic or organic matter, or even the development of social relation. Of course, if one aims at describing such a real situation, our model of a random graph should be replaced by a more complicated but more realistic model."

The authors thus already anticipated how their work was later received in the literature: the model is too simple for real situations yet simple enough so that we can derive mathematical properties.

In this thesis we will focus on one property of their model: how does the Erdős Rényi graph *look like* at the phase transition when n tends to infinity? There are two possible viewpoints to this question, a local and a global perspective, and here we will focus on the global one.

The mathematical tool to describe this property is a *scaling limit*: we change the nature of our object from a graph to a metric space, so that we can rescale it. Thus small local pertubations become trivial in the limit and the global structure emerges. Our goal is to find and characterize a limit object as well as to prove convergence. We will first introduce the Erdős-Rényi graph and the appropriate spaces formally. We then will work out a few properties of the uniform tree which is a related random combinatorical object with the Brownian Continuum Random Tree as scaling limit. Then we will discuss how to encode the Erdős Rényi graph in simpler functions and prove the convergence of those functions before we will introduce the limit object and prove the convergence of the Erdős-Rényi graph.

This thesis is based on the article *"The continuum limit of critical random graphs"* by Louigi Addario-Berry, Nicolas Broutin and Christina Goldschmidt published in *Probability Theory Related Fields* in 2012. The goal of this thesis is to make the article accessible to undergraduate students with basic knowledge in probability theory. Therefore certain arguments will be presented in more detail and more background is added to close the gap between undergraduate knowledge and modern research.

2 Preliminaries

2.1 Erdős-Rényi graph

Throughout this thesis, we will work with graphs usually denoted by $G := G(V, E)$ where V is a discrete set of vertices and $E \subseteq V \times V$ is the set of edges. Sometimes we will denote the edges of G by $E(G)$ and the edges by $V(G)$. We call two vertices $x, y \in V$ neighbours if $(x, y) \in E$. A tuple $(x_1, ...x_k), x_i \in V, 1 \leq i \leq k$ is a path in G if $(x_i, x_{i+1}) \in E$ for all $1 \leq i < k$. As we are only interested in finite graphs, without loss of generality we set $V = [n] = \{1, ..., n\}$ for a graph with n vertices. Furthermore, a graph T is a (discrete) tree if and only if it is connected and has no cycles, i.e. there are no paths of the form $(x_1, x_2, ..., x_k, x_1)$ with $\{x_1, ..., x_k\}$ pairwise different and $k \geq 2$.
The Erdős-Rényi-Graph is a classical and simple model for a random graph: given n vertices, an edge (x, y) exists independently for each edge with probability $p \in [0, 1]$. Formally:

Definition 2.1. *The Erdős-Rényi-Graph $G(n, p), n \in \mathbb{N}, p \in [0, 1]$ is a random variable taking values in the space of graphs on n vertices with its law being defined by:*

$$\mathbb{P}(G(n, p) = G) = p^{|E(G)|}(1 - p)^{\binom{n}{2} - |E(G)|}$$

for any graph G with n vertices.

An important observation about $G(n, p)$ is its growth. When we increase the edge probability p, there should be more edges, hence $G(n, p)$ grows with increasing p in the following sense: for each edge (x, y) sample a uniform variable $U_{x,y}$ on $[0, 1]$ independently and if $U_{x,y} \leq p$, we let $(x, y) \in E_p \subset [n] \times [n]$. The set E_p together with $[n]$ can be seen as random graph. In fact, this graph has the distribution of the Erdős-Rényi graph as $\mathbb{P}((x, y) \in E_p) = \mathbb{P}(U_{x,y} \leq p) = p$. For $p \leq p'$, the set E_p will be contained in $E_{p'}$ and thus this defines a coupling for $\{G(n, p) : p \in [0, 1]\}$. With respect to this coupling, the Erdős-Rényi graph with higher edge probability has more edges than the one with lesser edge probability.

A particular interesting viewpoint on this growth arises when we consider $G(n, \lambda/n)$ with $\lambda \in \mathbb{R}^+$ and vary λ. The expected number of neighbours of a single vertex is $(n - 1)p = (n - 1) \cdot \lambda/n = \lambda + \mathcal{O}(1/n)$, thus if $\lambda > 1$, every vertex has more than one neighbour on average and if $\lambda < 1$ there should be vertices without neighbours. Because of this, it is interesting to calculate how many vertices are connected with each other in a single component in terms of n when n tends to infinity. In fact, the Erdős-Rényi-Graph exhibits a phase transition at $\lambda = 1$, when the graph jumps from being sparsely connected to highly connected. This phase transition

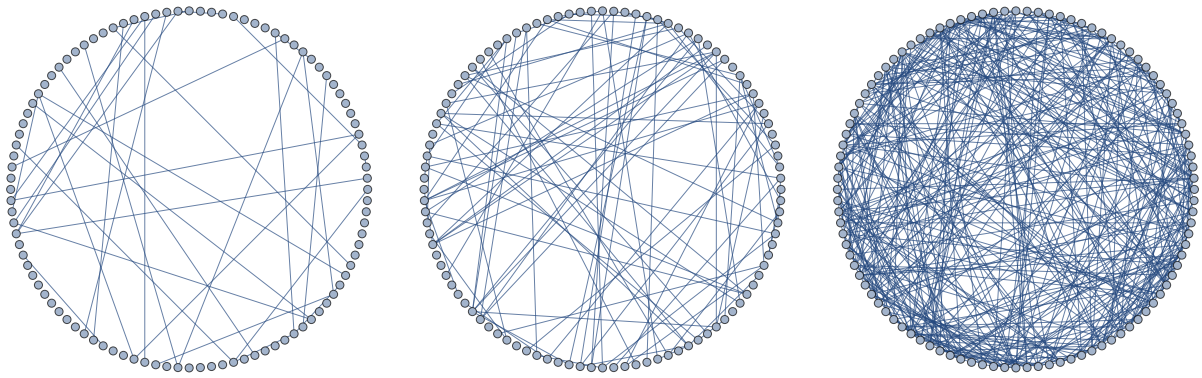


Figure 1: Three realizations of $G(n, p)$ with $n = 100$ and $p \in \{0.01; 0.02; 0.08\}$, simulated using the mathematica interface for igraph.

was first proven by Paul Erdős and Alfréd Rényi in 1960 [6]. The best way to observe this transition is to consider the size of the different components. For a given graph G and a vertex $v \in V(G)$ define the component containing v:

$$C(v) = \{u \in V(G) : \exists \text{ path connecting } u \text{ and } v\} \cup \{v\}$$

Furthermore, denote its size by $|C(v)|$. If G is a random graph like the $G(n,p)$, $|C(v)|$ is clearly a random variable. Define also:

$$|C_{max}| = \max_{v \in [n]} |C(v)|$$

as the size of the maximal component. The following two theorems provide insight on the component sizes below and above the phase transition, respectively.

Theorem 2.2. *Consider $G(n, \lambda/n)$ with fixed $\lambda < 1$. Then for every a big enough there exists $\delta = \delta(a, \lambda) > 0$ such that*

$$\mathbb{P}(|C_{max}| \geq a \log n) = O(n^{-\delta}).$$

Theorem 2.3. *Consider $G(n, \lambda/n)$ with fixed $\lambda > 1$. Then there is $0 < \zeta_\lambda < 1$ such that for every $\nu \in (1/2, 1)$, there exists $\delta = \delta(\nu, \lambda) > 0$ such that*

$$\mathbb{P}(||C_{max}| - \zeta_\lambda n| \geq n^\nu) = O(n^{-\delta}).$$

This means that for $\lambda < 1$ the largest component will be smaller than $\log n$ and for $\lambda > 1$ the largest component will contain a positive proportion of the vertices when $n \to \infty$. Based on this, the cases $\lambda < 1$ and $\lambda > 1$ are called the sub- and supercritical regimes respectively and $\lambda = 1$ is called the critical regime. We will not prove those theorems here - see [7, Theorem 4.4] and [7, Theorem 4.8] for proofs - but we will discuss a link to branching processes [7, Theorem 4.2]:

Proposition 2.4. *Consider $G(n, \lambda/n)$ for fixed λ, then for each $k \geq 1$:*

$$\mathbb{P}(|C(1)| \geq k) \leq \mathbb{P}(T \geq k)$$

where T denotes the total progeny of a Galton-Watson tree with law $Bin(n, \lambda/n)$. The vertex 1 can be replaced by an arbitrary different choice of vertex.

As we know, $Bin(n, \lambda/n)$ converges in distribution to $Poisson(\lambda)$. Furthermore, Galton-Watson trees exhibit a phase transition as well, precisely when the mean of the associated offspring distribution is equal to 1. Thus, it is no surprise that the Erdős-Rényi exhibits a phase transition as well.

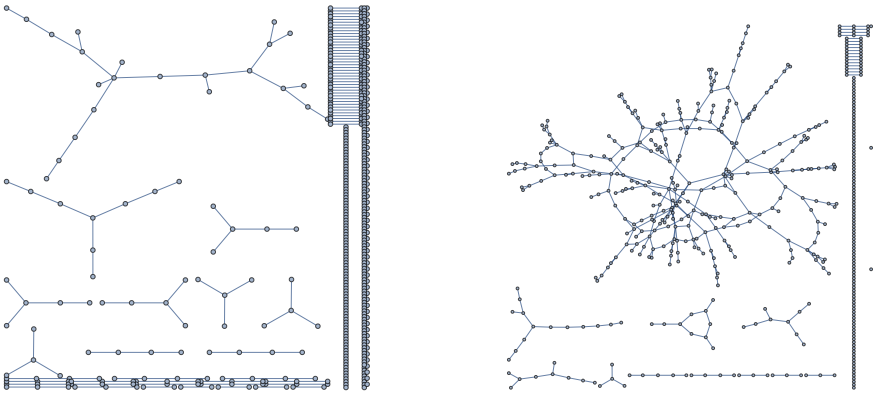


Figure 2: Two realizations of $G(n, \lambda/n)$ with $n = 500$ and $\lambda \in \{1/2; 3/2\}$ illustrating the phase transition, simulated using the mathematica interface for igraph.

Proof of Proposition 2.4. Let $S_i := |\{v \in [n] : d(1,v) = i\}|$, the number of vertices with distance to i to the vertex with label 1, where d is the graph distance, i.e. the length of the shortest path between two vertices. We also call the distance of a vertex to the vertex with label 1 the generation of a vertex. Conditionally on $(S_1, ..., S_{i-1})$, we achieve following estimate for $l \geq 0$:

$$\begin{aligned}\mathbb{P}(S_i \geq l | S_{i-1} = s_{i-1}, ..., S_1 = s_1) &\leq \mathbb{P}(\text{there are at least } l \text{ edges between a vertex at generation } i-1 \\ &\quad \text{and a vertex at generation } i | S_{i-1} = s_{i-1}, ..., S_1 = s_1) \\ &\leq \sum_{l_1+...+l_{s_{i-1}}=l} \mathbb{P}(X_i^1 \geq l_1, ..., X_i^{s_{i-1}} \geq l_{s_{i-1}})\end{aligned}$$

for arbitrary positive integers $s_1, ..., s_{i-1}$. The random variables are X_i^j are independent and identically distributed according to $Bin(n-1-\sum_{j=1}^{i-1} s_j, p)$. Let $Y_i^1, ..., Y_i^{s_{i-1}}$ be *i.i.d.* random variables distributed like $Bin(1+\sum_{j=1}^{i-1} s_j, p)$ independent from X_i^j. Define $X_i^{j,*} = X_i^j + Y_i^j$, a sequence of *i.i.d.* variables distributed like $Bin(n,p)$. Furthermore, we have $X_i^{j,*} \geq X_i^j$ almost surely for any j. Thus:

$$\begin{aligned}\mathbb{P}(S_i \geq l | S_{i-1} = s_{i-1}, ..., S_1 = s_1) &\leq \sum_{l_1+...+l_{s_{i-1}}=l} \mathbb{P}(X_i^{1,*} \geq l_1, ..., X_i^{s_{i-1},*} \geq l_{s_{i-1}}) \\ &= \mathbb{P}(S_i' \geq l | S_{i-1}' = s_{i-1}, ..., S_1' = s_1)\end{aligned}$$

where S_j' counts the offspring of a Galton-Watson tree with offspring distribution $Bin(n,p)$ after j generations. By induction, we can therefore couple S_i and S_i' such that $S_i \leq S_i'$ almost surely for any i. Thus:

$$|C(1)| = 1 + \sum_{i=1}^{n} S_i \leq 1 + \sum_{i=1}^{n} S_i' \leq 1 + \sum_{i=1}^{\infty} S_i' = T$$

which concludes the proof.

□

Naturally, one wants to analyze what happens in the case $\lambda = 1$. David Aldous proved the following theorem:

Theorem 2.5 ([2, Thm. 3])**.** *Consider $G(n, \lambda/n)$ with $\lambda = 1$. There exists a non-trivial random variable Z such that:*

$$n^{-2/3}|C_{max}| \xrightarrow{d} Z.$$

He even provides an explicit description of Z: Consider the process

$$W^0(t) := W(t) - \frac{t^2}{2}$$

where $(W(t), t \geq 0)$ is a standard Brownian motion. Then Z is distributed like the length of the longest excursion of the reflected process $W^0(t) - \min_{0 \leq s \leq t} W^0(s)$. In fact, the original result was stronger: he proved that not only the largest component is of size order $n^{2/3}$, but the whole sequence of component sizes. But because we only want to look at the largest component it is enough to know its size. The way we want to look at the critical Erdős Rényi graph is that it is almost a tree, on the one hand it is not connected - but there are individual big components - and on the other hand each component has very few cycles, if any at all. Thus we can hypothesize that the limit object of the Erdős-Rényi graph looks like an $\mathbb{R}$ tree with a few added cycles. We will later see this to be true.

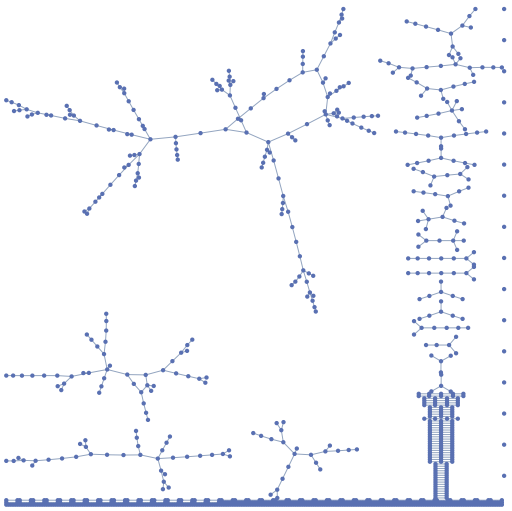


Figure 3: A realization of the critical $G(n, 1/n)$ with $n = 1000$, simulated using the mathematica interface for igraph.

2.2 Real trees and the Gromov-Hausdorff distance

To describe the limit of the Erdős-Renyi graph which we will construct later, we need some more ingredients: a continuous equivalent of graphs and a topology that allows us to compare different graphs. Therefore we define $\mathbb{R}$-trees and Gromov-Hausdorff distance.

Definition 2.6. *A metric space (X, d) is an $\mathbb{R}-$tree if it is geodesic and acyclic, where:*

1. *(X, d) geodesic $\Leftrightarrow$ for all $x, y \in X$ there is an isometric embedding $f : [0, d(x, y)] \to X$ such that $f(0) = x$ and $f(d(x, y)) = y$.*

2. *(X, d) acyclic $\Leftrightarrow$ there are no embedded circles in X, i.e. there are no injective continuous maps of the form $f : \mathbb{S}^1 \to X$ where $\mathbb{S}^1$ is the unit circle.*

To see why is this a continuous generalization of trees, we explain how to construct an $\mathbb{R}$-tree from a tree: if given a tree T, identify each edge with a line segment of length 1 and connect the ends of the line segments if there is a vertex joining the corresponding edges. This yields a space X_T and we define the metric d^T on X_T as the length of the shortes path connecting two points. Naturally, the two conditions of a graph being a tree - connected and no cycles - translate into our conditions for a metric space to be an $\mathbb{R}$-tree: geodesic and acyclic. The second point of the definition can also be substituted by requiring the geodesics between two points to be unique for all points, just like there is a unique shortest path between any two vertices in a discrete tree.

Constructing an $\mathbb{R}$-tree from a discrete tree is only one way to obtain an $\mathbb{R}$-tree, another context in which $\mathbb{R}$-trees appear naturally is the following: Consider a continuous function $g : [0, 1] \to [0, \infty[$ with $g(0) = g(1) = 0$. We define a pseudo-distance on $[0, 1]$:

$$d_g(s, t) := g(s) + g(t) - 2 \inf_{r \in [s \wedge t, s \vee t]} g(r) \qquad \text{for all} \quad s, t \in [0, 1]$$

Using this, define an equivalence relation:

$$s \sim t \Leftrightarrow d_g(s, t) = 0$$

such that d_g is a distance on $T_g := [0, 1]/\sim$. We denote the canonical projection by $\rho_g : [0, 1] \to T_g$.

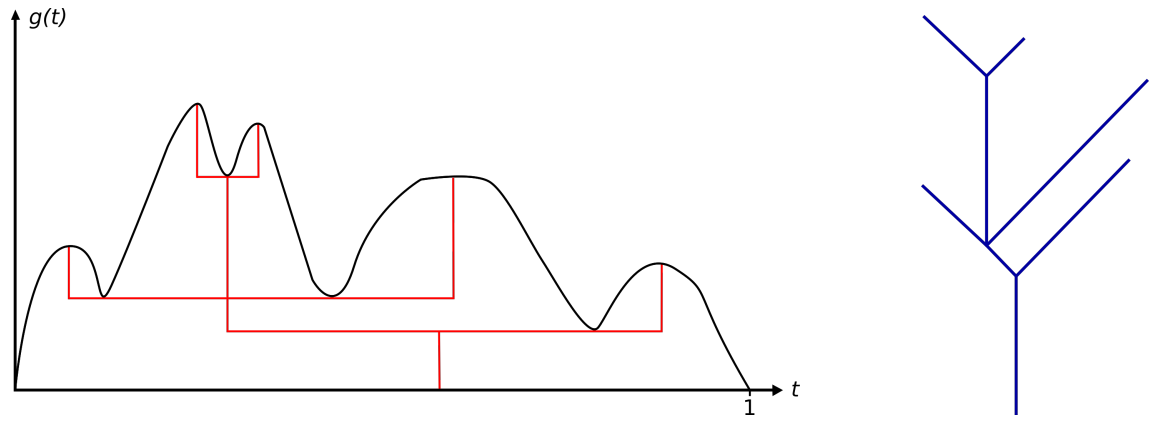


Figure 4: The black function g on the left encodes the real tree on the right. The horizontal red lines are each identified to a single point where as the vertical red lines become the branches of the tree. Furthermore, local maxima become leafs and local minima branching points.

Theorem 2.7 ([9, Thm. 3.1])**.** *The metric space (T_g, d_g) is a real tree.*

This is perhaps the most important way to construct an $\mathbb{R}$-tree. Furthermore, we can replace the interval $[0,1]$ by $[0,n], n \in \mathbb{N}$ and the resulting space will still be an $\mathbb{R}$-tree. By an appropiate choice of function, we can also encode an $\mathbb{R}$-tree which corresponds to a discrete tree. This function is called contour function and we will introduce it later. See Figure 4 for an example.

So now that we have a way to generalize trees we need a way to compare them. To do that, one can define a distance on the space of compact metric spaces. But first, we need a distance for compact sets in the same metric space. Recall:

Definition 2.8. *(Hausdorff distance)*
For a metric space (X,d) and compact subsets $K, K' \subseteq X$, define the Hausdorff distance between K and K':

$$d_H(K,K') := \inf\{\epsilon > 0 | K \subseteq F_\epsilon(K') \text{ and } K' \subseteq F_\epsilon(K)\}$$

where

$$F_\epsilon(K) := \{x \in X | dist(x,K) \leq \epsilon\}.$$

Now that we can measure the distance between compact sets, we want to measure the distance between two spaces by embedding them into any bigger metric space in a way that preserves the topology of the initial spaces. Then, the spaces become subspaces and we can use the notion of Hausdorff distance again. The optimal distance attainable by choice of embedding is then called the Gromov-Hausdorff distance. Formally:

Definition 2.9. *(Gromov-Hausdorff distance)*
For two compact metric spaces $(X,d), (X',d')$ define the Gromov-Hausdorff distance:

$$d_{GH}(X,X') := \inf_{(Y,\delta),\phi,\phi'} \delta_H(\phi(X), \phi'(X'))$$

where (Y,δ) is any metric space and ϕ, ϕ' are isometric embeddings of X, X' into Y. Also, δ_H is the induced Hausdorff distance by δ.

Claim: The Gromov-Hausdorff distance is in fact a metric on the set of compact spaces, if we identify X, X' if they are isometrically isomorphic.

Proof. Proof of Claim. Let $(X_1,d_1), (X_2,d_2), (X_3,d_3)$ be compact metric spaces:

- $d_{\mathrm{GH}}(X_1, X_2) = 0 \implies \exists$ isometric isomorphism $X_1 \to X_2 \implies X = X'$ by identification.
- $d_{\mathrm{GH}}(X_1, X_2) = d_{\mathrm{GH}}(X_2, X_1)$ by symmetry of inf and δ_H in the definition.
- Assume we have space (Y_{12}, δ_{12}) such that we can embed X_1, X_2 into Y_{12} isometrically and another space (Y_{23}, δ_{23}) such that we can embed X_2, X_3 into Y_{23} isometrically. Define $\delta_{13} := \inf_{x_2 \in X_2}\{\delta_{12}(x_1, x_2) + \delta_{23}(x_2, x_3)\}$ for $x_1 \in X_1, x_3 \in X_3$. Define:

$$\delta(x, x') = \begin{cases} d_1(x, x') & \text{if } x, x' \in X_1 \\ d_3(x, x') & \text{if } x, x' \in X_3 \\ \delta_{13}(x, x') & \text{if } x \in X_1, x' \in X_3 \end{cases}$$

  this defines a metric on the set $X_1 \cup X_3$. (positive definite and symmetry clear, triangle inequality inherited by the other distances) Hence, $X_1 \cup X_3$ is a metric space, such that we can embed X_1, X_3 isometrically. We conclude:

$$\delta_{12,H}(X_1, X_2) + \delta_{23,H}(X_2, X_3) \geq \delta_H(X_1, X_3) \geq d_{\mathrm{GH}}(X_1, X_3)$$

  Taking the infimum on the left side yields the triangle inequality.

□

The disadvantage of the definition of Gromov-Hausdorff distance is that the infimum ranges over all metric spaces (as long as there are suitable embeddings) and even finding a space that comes close to realizing the infimum may be very counterintuitive. Luckily, there is a characterisation for Gromov-Hausdorff distance that only relies on measuring distances in the spaces themselves without embedding them into a larger space:

Proposition 2.10. *Let $(X, d), (X', d')$ be two metric spaces, then:*

$$d_{GH}(X, X') = \frac{1}{2} \inf_{R \text{ correspondence}} dis\ R$$

where $R \subset X \times X'$ is a correspondence if for every $x \in X$ there is an $x' \in X'$ such that $(x, x') \in R$ and symmetrically for X' and

$$dis\ R = \sup_{(x_1, x_1'), (x_2, x_2') \in R} |d(x_1, x_2) - d'(x_1', x_2')|$$

is the distortion of the correspondence R.

Proof. We complete the proof of [5, Theorem 7.3.25].
Assume $r > d_{\mathrm{GH}}(X, X')$. Then there is a metric space (Z, δ) such that X, X' are isometric to subspaces of Z and $\delta_H(X, Y) < r$. Define $R \subset X \times X'$:

$$R = \{(x, x') \in (X, X') : \delta(x, x') < r\}$$

and R is a correspondence: assume the opposite, then there would be (without loss of generality) $x_0 \in X$ such that $\delta(x_0, x') \geq r$ for all $x' \in X'$. But this contradicts $\delta_H(X, X') < r$, thus R is a correspondence. Observe:

$$\begin{aligned} dis\ R &= \sup_{(x_1, x_1'), (x_2, x_2') \in R} |d(x_1, x_2) - d'(x_1', x_2')| = \sup_{(x_1, x_1'), (x_2, x_2') \in R} |\delta(x_1, x_2) - \delta(x_1', x_2')| \\ &\leq \sup_{(x_1, x_1'), (x_2, x_2') \in R} \delta(x_1, x_1') + \delta(x_2, x_2') < 2r \end{aligned}$$

Taking the infimum over $r > d_{\mathrm{GH}}(X, X')$ yield the inequality

$$d_{\mathrm{GH}}(X, X') \geq \frac{1}{2} \inf_{R \text{ correspondence}} dis\ R.$$

On the other hand, let R be any correspondence and let $dis\ R = 2r$. Consider the disjoint union $X \sqcup X'$ and define a metric δ on $X \sqcup X'$:

$$\delta(a,b) = \begin{cases} d(a,b) & \text{if} \quad a,b \in X \\ d'(a,b) & \text{if} \quad a,b \in X' \\ \inf_{(x,x')\in R}\{d(a,x) + r + d'(b,x')\} & \text{if} \quad a \in X, b \in X' \end{cases}$$

Assume $r > 0$ without loss of generality. The symmetry of δ and $\delta(a,b) \geq 0$ are clear. Assume $\delta(a,b) = 0$: if a, b are both elements of X, then $a = b$, because d is a metric, the same holds if a, b are both elements of X'. If $a \in X$ and $b \in X'$, then $\delta(a,b) \geq r > 0$, thus $a \neq b$. For the triangle inequality, it is enough to look at $a, c \in X, b \in X'$, then:

$$\begin{aligned} \delta(a,b) + \delta(b,c) &= \inf_{(x,x'),(y,y')\in R} d(a,x) + d(c,x) + 2r + d'(b,x') + d'(b,y') \\ &\geq \inf_{(x,x'),(y,y')\in R} d(a,x) + d(c,x) + 2r + d'(x',y') \\ &\geq \inf_{(x,x'),(y,y')\in R} (d(a,c) - d(c,x)) + (d(x,c) - d(x,y)) + d'(x',y') \\ &\geq d(a,c) + 2r - \sup_{(x,x'),(y,y')\in R} |d'(x',y') - d(x,y)| \\ &= d(a,c) + 2r - dis\ R \\ &= \delta(a,c) \end{aligned}$$

by applying the triangle inequality in X and X'. The other cases follow from symmetry and the fact that d, d' are metrics. Furthermore, identify X and X' with the subspaces of $X \sqcup X'$ isometric to X and X'. Then for any $x \in X$:

$$\begin{aligned} dist(x, X') = \inf_{x'\in X'} \delta(x,x') &= \inf_{x'\in X'} \inf_{(x_0,x_0')\in R} r + d(x,x_0) + d'(x',x_0') \\ &\leq r + \inf_{(x,x_0')\in R} \inf_{x'\in X'} d(x,x) + d(x',x_0') = r \end{aligned}$$

because there is a $x_0' \in X'$, such that $(x, x_0') \in R$. A symmetric argument holds for all $x' \in X'$ and thus we conclude that $\delta_H(X, X') \leq r$. Thus, $d_{\mathrm{GH}}(X, X') \leq r = \frac{1}{2} dis\ R$. With the reverse inequality we showed earlier, we can derive the claim. □

With the notation of Theorem 2.7 we derive the following corollary:

Corollary 2.11 ([9, Cor. 3.7])**.** *Let $g, g' : [0,1] \to [0, \infty[$ be two continuous functions with $g(0) = g(1) = g'(0) = g'(1) = 0$ and $T_g, T_{g'}$ their induced $\mathbb{R}$-trees, then:*

$$d_{GH}(T_g, T_{g'}) \leq 2\|g - g'\|$$

where $\|g - g'\| = \sup_{t\in[0,1]} |g(t) - g'(t)|$ is the supremum norm throughout the thesis.

Proof. Let $\rho_g, \rho_{g'}$ be the projections from $[0,1]$ into T_g and $T_{g'}$ respectively. Let:

$$R := \left\{(a,a') : \exists t \in [0,1] \text{ such that } a = \rho_g(t) \text{ and } a' = \rho_{g'}(t)\right\} \subset T_g \times T_{g'}$$

which is a correspondence by construction. By Proposition 2.10:

$$
\begin{aligned}
d_{\mathrm{GH}}(T_g, T_{g'}) &\leq \frac{1}{2} dis\ R = \frac{1}{2} \sup_{(a,a'),(b,b') \in R} |d_g(a,b) - d_{g'}(a',b')| \\
&= \frac{1}{2} \sup_{s,t \in [0,1]} |g(s) + g(t) - g'(s) - g'(t) - 2 \inf_{r \in [s \wedge t, s \vee t]} g(r) + 2 \inf_{r' \in [s \wedge t, s \vee t]} g'(r')| \\
&\leq \sup_{t \in [0,1]} |g(t) - g'(t)| + \sup_{s,t \in [0,1]} |\inf_{r \in [s \wedge t, s \vee t]} g(r) - \inf_{r' \in [s \wedge t, s \vee t]} g'(r')| \\
&\overset{g,g' \geq 0}{\leq} \| g - g' \| + \sup_{s,t \in [0,1]} \inf_{r' \in [s \wedge t, s \vee t]} |g(r') - g'(r')| \\
&\leq 2 \|g - g'\|
\end{aligned}
$$

□

Remark: Instead of (ordinary) metric spaces we will also consider rooted metric spaces, i.e. metric spaces (X, d, τ) with a distinguished point $\tau \in X$. Everything from above holds, if the Gromov-Hausdorff distance is changed slightly to accomodate the roots. Under the same assumptions and notation of Definition 2.9, define for two rooted metric spaces $(X, d, \tau), (X', d', \tau')$:

$$
d_{\mathrm{GH}}(X, X') := \inf_{(Y,\delta),\phi,\phi'} \max \left\{ \delta_H(\phi(X), \phi'(X')), \delta(\phi(\tau), \phi'(\tau')) \right\}
$$

Furthermore, for a set R to be a correspondence between (X, d, τ) and (X', d', τ') require additionally that $(\tau, \tau') \in R$.

2.3 Continuum random tree

The goal for this section is to find a scaling limit for uniform trees on $[n]$ vertices. In the previous section, we have seen that a continuous function with certain constraints induces a real tree. Now, we have to reverse this procedure: given a discrete tree T, how can we turn it into a function, such that it commutes with the earlier procedure illustrated in Theorem 2.7. This is achieved in the following way: Embed the tree in the plane and imagine a particle that explores the tree from left to right at unit speed starting at the root[1]. Let $C(s)$ be its distance to the root at time s such that every edge is traversed exactly twice. Clearly, the total time required is $2(n-1)$, as there are $n-1$ edges. The resulting function $(C(s), 0 \leq s \leq 2(n-1))$ is called the contour function. See Figure 5 for an example.

We now want to use contour functions to analyse uniform trees. If we consider μ to be the critical geometric distribution - i.e. $\mu(k) = (\frac{1}{2})^{k+1}$ for $k \in \mathbb{N}_0$ - there is a useful connection between Galton-Watson trees, uniform trees and the symmetric random walk.

Lemma 2.12. *Let T^n be a uniform tree on $[n]$ and $\mathcal{T}_\mu$ a Galton-Watson tree with the critical geometric distribution μ as offspring distribution. Then:*

$$
\mathbb{P}(\mathcal{T}_\mu = T \big| |\mathcal{T}_\mu| = n) = \mathbb{P}(T^n = T)
$$

for any tree T on $[n]$.

Proof. Let T be a tree on $[n]$ and denote the set of all trees on $[n]$ by $\mathbb{T}$. Recall the definition of uniform trees:

$$
\mathbb{P}(T^n = T) = \frac{1}{\sum_{T' \in \mathbb{T}} 1}
$$

[1] This can be defined in a cleaner way once we introduce the depth first ordering of the vertices in the next section.

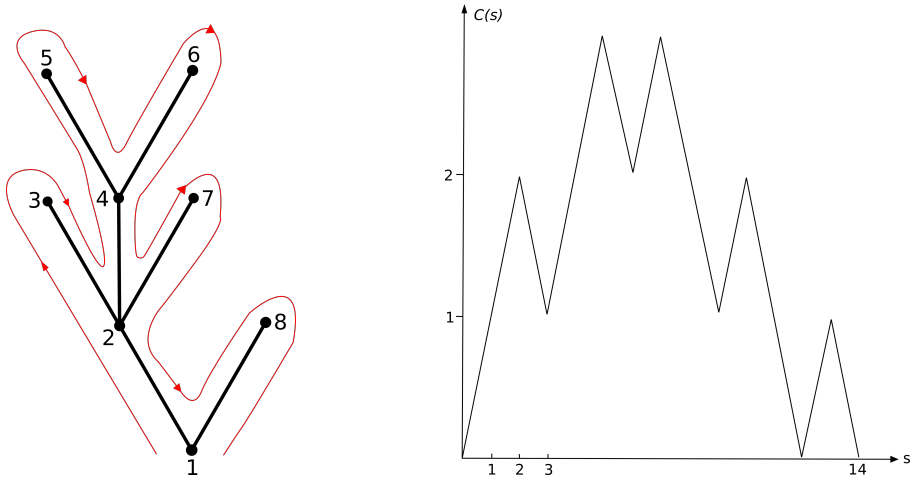


Figure 5: A tree with 8 vertices and its contour function. The red line sympolizes the path of the particle. This is the inverse procedure to the construction used in Theorem 2.7 where we construct an $\mathbb{R}$-tree from a continuous function.

This means every possible tree with n vertices is equally likely. Now let $(k_1^T, ..., k_n^T) \in \mathbb{N}^n$ be such that k_j^T is the number of offspring of vertex j in the tree T. Because every tree with n vertices has $n-1$ edges, we have $\sum_{j=1}^n k_j^T = n-1$, independently of T. Then:

$$\begin{aligned}\mathbb{P}(\mathcal{T}_\mu = T \big| |\mathcal{T}_\mu| = n) &= \frac{\mathbb{P}(\mathcal{T}_\mu = T)}{\sum_{T' \in \mathbb{T}} \mathbb{P}(\mathcal{T}_\mu = T')} = \frac{\prod_{i=1}^n \mu(k_j^T)}{\sum_{T' \in \mathbb{T}} \prod_{i=1}^n \mu(k_j^{T'})} \\ &= \frac{\prod_{i=1}^n (\frac{1}{2})^{k_j^T+1}}{\sum_{T' \in \mathbb{T}} \prod_{i=1}^n (\frac{1}{2})^{k_j^{T'}+1}} = \frac{(\frac{1}{2})^{2n}}{\sum_{T' \in \mathbb{T}} (\frac{1}{2})^{2n}} = \frac{1}{\sum_{T' \in \mathbb{T}} 1} \\ &= \mathbb{P}(T^n = T)\end{aligned}$$

□

Recall the simple random walk on $\mathbb{Z}$: let $(X_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables such that $\mathbb{P}(X_1 = 1) = \frac{1}{2} = \mathbb{P}(X_1 = -1)$. Define $S_0 = 0$ and $S_n = \sum_{i=1}^n X_i, n \geq 1$ and call the sequence $(S_n)_{n \in \mathbb{N}_0}$ the simple random walk. Define a stopping time T as follows:

$$T = \inf\{n \in \mathbb{N} \big| S_n = -1\} < \infty \quad a.s.$$

Then we call $\{S_0, ..., S_{T-1}\}$ an excursion of the simple random walk.

Proposition 2.13 ([9, Prop. 2.6])**.** *The tree corresponding to an excursion of the simple random walk is a Galton-Watson tree with the critical geometric offspring distribution.*

Proof. We complete the proof of [9]. The key to this propositon is the fact that a subtree of a Galton-Watson tree rooted at a descendant of the root is distributed like the Galton-Watson tree itself.

We want to know how many children the root has. For that we need to know how many times S_n *touches* 0 before S_n hits -1. To achieve this formally, define the upcrossing and downcrossing times:

$$U_1 := \inf\{n \geq 1 : S_n = 1\} \quad \text{and} \quad V_1 := \inf\{n \geq U_1 : S_n = 0\}$$

and inductively:

$$U_{j+1} := \inf\{n \geq V_j : S_n = 1\} \quad \text{and} \quad V_{j+1} \inf n \geq U_{j+1} : S_n = 0$$

Let $K := \sup\{j : U_j \leq T\}$ and if the set is empty let $K = 0$. By the way we construct a tree out of a contour function, we clearly have that K is distributed like the number of children of the root. Observe:

$$\mathbb{P}(K = 0) = \mathbb{P}(U_1 > T) = \mathbb{P}(S_1 = -1) = \frac{1}{2}$$

And inductively for $k \geq 1$ by the Markov property:

$$\mathbb{P}(U_k \leq T \big| U_{k-1} \leq T) = \mathbb{P}(S_{U_{k-1}+1} \big| S_{U_{k-1}} = 0) = \mathbb{P}(X_{U_{k-1}+1} = 1) = \frac{1}{2}$$

$$\implies \mathbb{P}(K = k) = \left(\frac{1}{2}\right)^{k+1}$$

Thus, the distribution of the number of offspring of the root is the critical geometric distribution. Conditionally on $\{K = k\}$, there are k excursions above 1 of the form

$$i \in \{1, ..., k\} : \quad S_{U_i+n} - 1, \quad 0 \leq n \leq V_i - U_i - 1$$

These excursions are independent for all $i \in \{1, ..., k\}$ and distributed like an excursion of the simple random walk. Thus, the subtree corresponding to the $i-$th child of the root is distributed like the tree itself. This, together with the fact that the root has the critical geometric offspring distribution characterises the tree uniquely as Galton-Watson tree with the desired offspring distribution. □

To close the gap between discrete trees and discrete excursions, and continuous trees and continuous excursions we need one landmark result of probabilty theory - basically an extension of the central limit theorem to continuous functions[2] - proven by Monroe Donsker in 1952, see for example [4].

Theorem 2.14. *Let $(X_i)_{i\in\mathbb{N}}$ be i.i.d and $\mathbb{P}(X_1 = 1) = \frac{1}{2} = \mathbb{P}(X_1 = -1)$ and consider $W^{(n)} = (S_n^t = \sum_{i=1}^{\lfloor nt \rfloor} X_i + (nt - \lfloor nt \rfloor)X_{\lfloor nt \rfloor + 1}, t \in [0,1])$ and let $W = (W(t), t \in [0,1])$ a standard Brownian motion. Then:*

$$\frac{1}{\sqrt{n}}(W^{(n)}(t), t \in [0,1]) \xrightarrow{d} (W(t), t \in [0,1])$$

in $C([0,1], \mathbb{R})$, equipped with the uniform topology.

This means, that the rescaled and interpolated simple random walk converges in distribution to a Brownian motion. In light of Proposition 2.13 we define the Continuum Random Tree[3]:

Definition 2.15. *Let $(e(t), 0 \leq t \leq 1)$ be a standard Brownian excursion. $\mathcal{T} := T_e$ is called the Continuum Random Tree.*

By standard Brownian excursion we mean a path of the standard Brownian Motion with length 1 such that almost surely: $e(0) = e(1) = 0$, $e(t) > 0$ for $0 < t < 1$ and $e(\cdot)$ is continuous. Theorem 2.7 shows that $\mathcal{T}$ is almost surely an $\mathbb{R}$-tree. This is in fact a well-defined notion, check [9] for reference. The next theorem was again proven by Aldous [3, Theorem 8] and is our key ingredient to analyze the scaling limit of the Erdős-Rényi graph.

Theorem 2.16. *Let T^n be a uniform tree on $[n]$ and $\mathcal{T}$ the Continuum Random Tree. If we view T^n as a compact metric space, then:*

$$n^{-1/2}T^n \xrightarrow{d} \mathcal{T}$$

as compact metric spaces with respect to the Gromov-Hausdorff topology.

[2] We only present the special case of the simple random walk, the theorem holds in fact if the X_i are arbitrary real random variables, as long as they have finite variance.

[3] Also known as Brownian Continuum Random Tree.

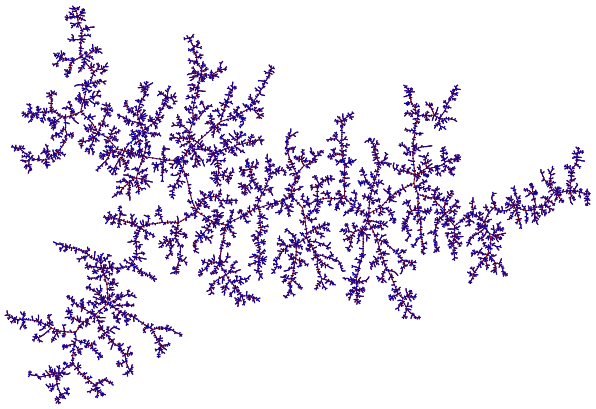


Figure 6: A realization of the Continuum Random Tree, picture by Igor Kortchemski, source: IGOR-KORTCHEMSKI.PERSO.MATH.CNRS.FR

Proof Sketch. The essence of the proof is that a uniform tree can be represented by an excursion of the simple random walk conditioned to become -1 at time n by Proposition 2.13. Then similarly to Theorem 2.14, the interpolated rescaled excursion of the simple random walk converges uniformly to an Brownian excursion and by Corollary 2.11 their corresponding $\mathbb{R}$-trees converge in Gromov-Hausdorff distance. The main technicality of the proof lies in the fact that we need to extract information from zero sets of the standard Brownian motion and that we need to condition on the excursion of the simple random walk. For a full and comprehensive proof see [9, Theorem 3.9]. □

We end this section with a short remark about the properties of the continuum random tree: the continuum random tree is defined by identifying points of the line segment $[0,1]$, denote the projection $[0,1] \to \mathcal{T}$ by ρ. This line segment is naturally equipped with the Lebesgue (probability) measure μ. Thus, we can equip $\mathcal{T}$ with the pushforward measure with respect to ρ. Call $x \in \mathcal{T}$ a leaf if $\mathcal{T} \backslash \{x\}$ is connected.

Proposition 2.17. *Let $\mathcal{L}$ be the set of leaves of $\mathcal{T}$. Then:*

$$\mu(\rho^{-1}(\mathcal{L})) = 1$$

This property is sometimes called $\mathcal{T}$ being "leaf-dense". We can interpret it as follows: if we select a vertex of $\mathcal{T}$ uniformly at random, it will almost surely be a leaf. For a proof of Proposition 2.17, see [3, Theorem 3].

3 Encoding random graphs

In this section we will deal with how to label a tree or a graph in a smart and useful way by using the so called *depth first search*. Using this labeling, we can encode a tree as function in a different way than in the previous chapter. This provides us with a correspondence between trees and certain functions. Furthermore, we will extend the labeling to graphs and in particular connected random graphs. We then look at two ways to construct a graph using the functions obtained by the labeling process and study their distribution as well as the difference between them.

3.1 Depth first search

Assume we are given a rooted and connected graph $G = G(V, E)$ with vertices $V = [m] = \{1, ..., m\}$ where we choose the vertex with label 1 as the root. We will define an algorithm that explores G and thereby encodes the structure of the graph:

Definition 3.1. *Ordered depth first search -* ***oDFS***(G)

Initialization	*Set* $\mathcal{O}_0 = (1), \mathcal{A}_0 = \emptyset$
Step i	$(0 \leq i \leq m-1)$*: Let* v_i *be the first vertex of* $\mathcal{O}_i$ *and let* $\mathcal{N}_i$ *be the neighbors of* v_i*. To construct* $\mathcal{O}_{i+1}$*, remove* v_i *from* $\mathcal{O}_i$ *and add the vertices of* $\mathcal{N}_i \backslash \mathcal{A}_i$ *to the start of* $\mathcal{O}_i$ *in increasing order in terms of vertex label. Furthermore, set* $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \{v_i\}$.

Naturally, the algorithm will end when $\mathcal{O}_m = \emptyset$. Before that, it will visit every vertex exactly once and thereby induce a new labeling on G, i.e. we will label a vertex according to the step when it was visited. Furthermore, **oDFS**(G) induces a spanning tree (a subtree containing all vertices) of G: we connect the two vertices u, v if $v \in \mathcal{N}_i$ given u is explored at step i. We refer to the resulting tree as *ordered-depth first tree* of G. Because the order of vertices traversed is unique, the resulting tree is also unique. Next, we define the *depth-first walk* $(X(i))_{0 \leq i < m}$ of G: let $X(i) = |\mathcal{O}_i| - 1$. This way, $X(i)$ counts the vertices seen, but not yet fully explored at time i. For convenience, we will use a continuous interpolation of $X(i)$; $s \in [0, m-1]$:

$$X(s) := X(\lfloor s \rfloor) + (s - \lfloor s \rfloor) \cdot (X(\lfloor s \rfloor + 1) - X(\lfloor s \rfloor))$$

This seems like a rather complicated construction, but the intuition behind **oDFS**(G) is simple: the algorithm explores G such that it always tries to go one step further away from the root if possible. If that is not possible it will backtrack until it is possible to go deeper again. See figure 7 for an example of this procedure.

In the special case that we deal with a tree T, we will also define the area of the tree:

$$a(T) := \sum_{i=1}^{m-1} X(i)$$

which is also the area under the graph of the depth first walk. Intuitively, the area of a tree measures how much the tree is branched, with bias to branching close to the root.

Now, we ask ourselves how to reconstruct a graph if we are given a tree T as the result of **oDFS**(G). This means, we may add edges to T only in a way, such that the ordered-depth first tree of the resulting graph is still T. Therefore, we define that an egde $(v_i, v_j) \notin E(T)$ is *permitted* by **oDFS**(T) if, during the search procedure on T, there is a k such that $v_i, v_j \in \mathcal{O}_k$.

Lemma 3.2 ([1, Lemma 7])**.** *Given a tree* T *and a connected graph* G *on* $[m]$*, then* T *is the ordered-depth first tree of* $G \iff G$ *can be obtained from* T *by adding edges permitted by* ***oDFS****(*T*).*

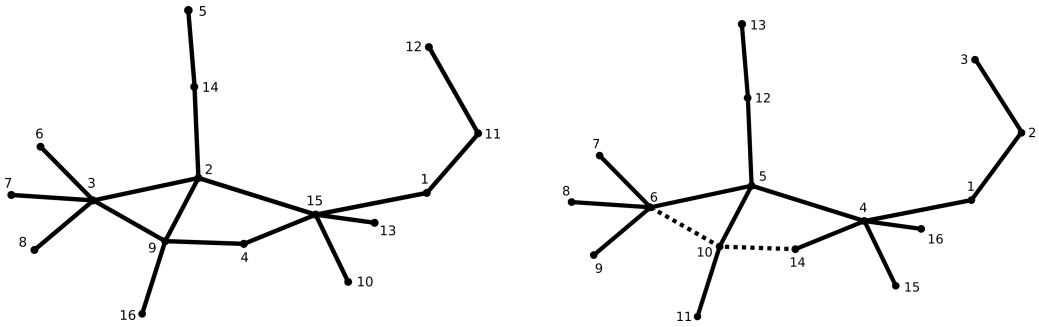


Figure 7: A graph with 16 vertices with arbitrary labeling on the left, and depth first labeling on the right. The dotted edges will not be traveresed during the depth first search.

Proof. $\Longleftarrow$ First of all, T clearly has to be a subgraph of G. We continue by induction over k, the number of edges we need to add to T to obtain G. The case $k = 0$ is trivial, as this means $T = G$. Suppose $k \geq 1$: Assume we would need to add k edges to T to obtain G and the statement is true for all graphs, where we would need to add $k - 1$ edges. Also, let $v_0, ..., v_{m-1}$ be the ordering on $[m]$ obtained by running **oDFS**(T). Let $(v_i, v_j); i < j$ be the first edge of G not contained in T - first in the sense of the vertex ordering, i.e. there is no edge $(v_{i'}, v_{j'}) \in E(G)\backslash E(T)$ such that $i' < i$. By definition, vertex v_i is explored at timestep i of **oDFS**(T) and the behaviour of oDFS(T) and **oDFS**(G) prior to this step is the same. Hence, because (v_i, v_j) is a permitted edge and v_j is explored after v_i, we have $v_j \in \mathcal{O}_i(T) = \mathcal{O}_i(G)$. We conclude, $(v_i, v_j) \notin E(T)$ and using the induction hypothesis we achieve the desired result.

$\Longrightarrow$ Again, let $v_0, ..., v_{m-1}$ be the ordering of $[m]$ obtained from **oDFS**(T). Suppose there is an edge in $E(G)\backslash E(T)$ that is not permitted by **oDFS**(T) and look for a contradiction. Let $(v_i, v_j), i < j$ be the first of those edges (in the same sense as above). Again, the behaviour of **oDFS**(T) and **oDFS**(G) is identical up to step i. Because (v_i, v_j) is not permitted and v_j will be visited after v_i, we have $v_j \notin \mathcal{O}_i(T) = \mathcal{O}_i(G)$. On the other hand, v_j is a neighbour of v_i in G, which means $v_j \in \mathcal{N}_i(G) \rightsquigarrow (v_i, v_j) \in E(T)$ contradicting our assumption. □

We will derive one more lemma in this section: Lemma 3.2 tells us that we can think of a connected graph as the amalgamation of a spanning tree and a set of permitted edges. Thus it is interesting to look at the number of permitted edges in a certain tree:

Lemma 3.3 ([1, Lemma 6])**.** *Let T be a tree on $[m]$. The number of edges permitted by **oDFS**(T) is exactly $a(T)$.*

Proof. We will proceed by induction on m. The cases $m \in \{1, 2\}$ are trivial. Let $m \geq 3$ and denote the neigbours of 1 with $v_1, ..., v_i$ for some $i \geq 1$. Furthermore, if we remove 1 from T, we will receive i trees, each rooted at $v_j, 1 \leq j \leq i$. Denote those trees by $T_1, ..., T_i$ and the depth first walks of $T_1, ..., T_i$ by $X_{T_1}(t), ..., X_{T_i}(t)$. By definition, $X(T)$ starts at 1, uncover $v_1, ..., v_i$ and then run independently on the T_j.

We will now look at the number of permitted edges between the T_j, i.e. edges (x, y) with $x \in T_j, y \in T_{j'}$ with $j \neq j'$: assume $x \in T_j$. Then y can only be one of the roots of the other subtrees with an higher index than j, as the other vertices are either fully explored or not yet uncovered. We conclude that the number of permitted edges of this kind is

$$\sum_{j=1}^{i} (i - j)|T_j|.$$

Furthermore, the number of permitted edges with both ends in a single subtree is exactly the area of that subtree - by induction hypothesis. In the end, the number of permitted edges is

$$\sum_{j=1}^{i}((i-j)|T_j| + a(T_j)) = \sum_{j=1}^{i}\sum_{k=1}^{|T_j|}((i-j) + X_{T_j}(k)) \overset{(*)}{=} \sum_{t=1}^{m} X(t) = a(T)$$

$(*)$: At timestep $t; 0 < t < m-1$ of the **oDFS**(T), the process explores one of the T_j. This means the value of $X(t)$ will be the value of the respective $X_{T_j}(k)$ for the appropriate k plus the number of neigbours of the root that have yet to be fully explored. □

We conclude this subsection with the remark that we can encode a graph without losing any information by its depth first walk and marking the surplus edges below the graph of the walk. See Figure 8 for an example.

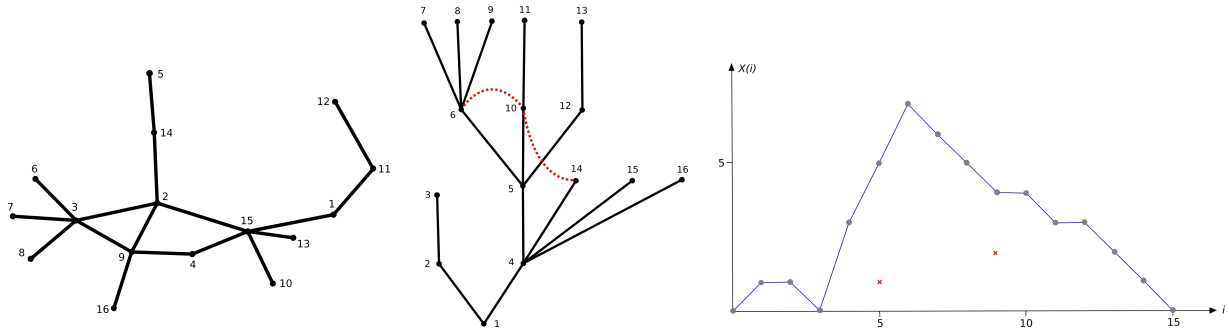


Figure 8: A graph with 16 vertices, its depth first tree and its marked depth first walk.

3.2 Two ways of reconstructing a graph

We want to construct a graph from a pointset $\mathcal{Q} \subset \mathbb{N} \times \mathbb{N}$ and a tree T on $[m]$ with labels in depth first order. First, we define $\mathcal{Q} \cap f$ for a pointset $\mathcal{Q}$ and a continuous function $f : \mathbb{R}^+ \to \mathbb{R}^+$:

$$\mathcal{Q} \cap f := \{(x, y) \in \mathcal{Q} \big| 0 < y \leq f(x)\}$$

This notation will later prove to be useful, as we do not have to restrict ourselves too much on the pointsets, as we can also allow $\mathcal{Q} \subseteq \mathbb{R}^+ \times \mathbb{R}^+$.

We now construct $G^X(T, \mathcal{Q})$ as the inverse procedure from the depth first search in the previous section: By Lemma 3.2, we may add edges that are permitted by **oDFS**(T) to receive a graph whose depth first spanning tree is T. This way, we can obtain any connected graph on m vertices. To construct $G^X(\mathcal{Q}, T)$ add those edges that correspond to the points of $\mathcal{Q} \cap X$, where $(X(t), 1 \leq t < m)$ is the depth first walk on T. By Lemma 3.3 every point in the set $\mathcal{Q} \cap X$ corresponds to exactly one permitted edge.

To construct $G^H(\mathcal{Q}, T)$, we first need to define the height process of a tree: Assume the tree is labeled in depth first order. Define $(H(i), 0 \leq i < m)$ as the height of the $i+1$ vertex - the graph distance between the root and vertex with label $t+1$. Like for the depth first walk $X(t)$, we define a continuous interpolation for $H(s), s \in [0, m-1]$:

$$H(s) := H(\lfloor s \rfloor) + (s - \lfloor s \rfloor) \cdot (H(\lfloor s \rfloor + 1) - H(\lfloor s \rfloor))$$

Now construct the graph $G^H(\mathcal{Q}, T)$: if $(i, j) \in \mathcal{Q}$ and $0 < 2j \leq H(i)$, add an edge between vertex v_i and the vertex at distance $(2j - 1)$ from the root on the path from v_i to the root. Because T is a tree, this path is unique and therefore the added edge is as well. It may happen that the depth first tree of the graph constructed in this manner is not T. See Figure 9 for an example.

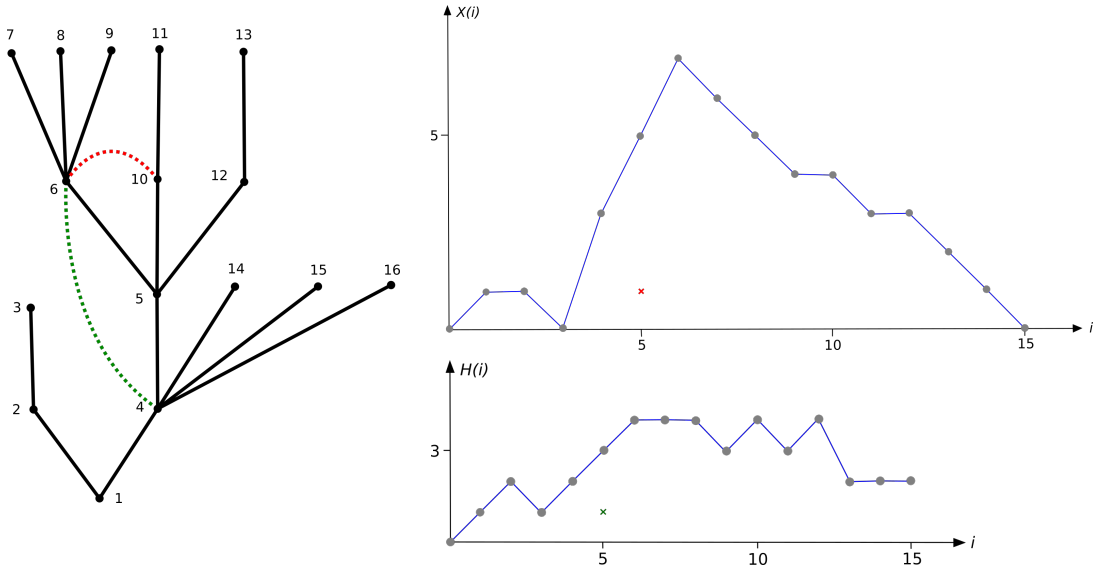


Figure 9: A tree with 16 vertices and its depth first walk $X(i)$ and height function $H(i)$. The same marked point induces different edges, the edge induced in G^X is marked red and the edge induced in G^H is marked green.

3.3 Distribution of connected random graphs

In this section we want to find a way to construct an object similar to the Erdős-Renyi graph but in terms of the constructions explained in the previous section. To do that, let $\mathbb{T}_{[m]}$ be the set of labeled trees with m vertices and $\mathbb{G}_T$ the set of connected graphs for which their depth first tree is T. The set

$$\{\mathbb{G}_T : T \in \mathbb{T}_{[m]}\}$$

clearly forms a partition for the connected graphs on $[m]$ vertices. We now construct a connected random graph:

Let $0 < p < 1$ and $m \in \mathbb{N}$. Select a tree $\tilde{T}_m^p$ from $\mathbb{T}_{[m]}$ such that for any given tree T we have:

$$\mathbb{P}(\tilde{T}_m^p = T) \propto (1 - p)^{-a(T)}$$

where $a(T)$ is the area of T with respect to the depth first walk X_T (and we write $\propto$ for *proportional to* in the sense that there is a universal constant independent of T). We also need a so called *binomial pointset* $\mathcal{Q}_p \subset \mathbb{N}_0 \times \mathbb{N}_0$. This means each point $(i, j) \in \mathbb{N} \times \mathbb{N}$ exists in $\mathcal{Q}_p$ independently with probability p. We will then use this to construct the graph G^X in the previous section, i.e. define:

$$\tilde{G}_m^p = G^X(\tilde{T}_m^p, \mathcal{Q}_p)$$

because $\tilde{T}_m^p$ and $\mathcal{Q}_p$ are random, so is $\tilde{G}_m^p$. Another way to think of this, is to add each edge permitted by $\mathbf{oDFS}(\tilde{T}_m^p)$ independently with probability p. Furthermore, we write G_m^p for a connected component of $G(n, p); n \geq m$ - the Erdös-Renyi graph like described in the introduction - with size m.

Proposition 3.4 ([1, Prop. 8]). *For any $0 < p < 1$ and $m \leq n$, the distributions of $\tilde{G}_m^p$ and G_m^p are identical.*

Proof. First, we note that the quantity $s(G) = |E(G)| - (m-1)$ for a connected graph on $[m]$ is the exact number of edges we have to remove to make G a tree. We call $s(G)$ the surplus of G. On the other hand, $s(G)$ is the number of edges we need to add to some tree T to obtain G.

We observe for any connected graph G on $[m]$:

$$\begin{aligned}\mathbb{P}(G_m^p = G) &\propto \mathbb{P}(G(m,p) = G) \\ &= \left(p^{m-1}(1-p)^{\binom{m}{2}-(m-1)}\right)\left(p^{s(G)}(1-p)^{-s(G)}\right) \propto p^{s(G)}(1-p)^{-s(G)}\end{aligned}$$

Again, we used $\propto$ when we left out universal constants for all G considered.

On the other hand:

$$\begin{aligned}\mathbb{P}(\tilde{G}_m^p = G) &\propto (1-p)^{-a(T)}\mathbb{P}(\tilde{G}_m^p = G \big| T \text{ is the depth first tree of } G) \\ &= (1-p)^{-a(T)}\left(p^{-a(T)}(1-p)^{a(T)-s(T)}\right) = p^{s(G)}(1-p)^{-s(G)}\end{aligned}$$

We conclude:

$$\mathbb{P}(G_m^p = G) \propto \mathbb{P}(\tilde{G}_m^p = G) \implies \mathbb{P}(G_m^p = G) = \mathbb{P}(\tilde{G}_m^p = G)$$

because probability measures are normalized. □

3.4 Bridging the difference between G^X and G^H

In the previous subsections we have seen that on the one hand G^X is closely related to a random graph, yet on the other hand we will later see that the construction of G^H translates naturally into a continuous limit. Therefore, in this subsection we will derive bounds on the differences between G^X and G^H. We remember that our notion of convergence is the Gromov-Hausdorff distance as described in the previous chapter. Note that all bounds derived are deterministic.

First, we will derive a quick corollary from Proposition 2.10:

Corollary 3.5.

$$d_{GH}(G^X, G^H) \leq \frac{1}{2}\sup_{x,y\in V}|d^H(x,y) - d^X(x,y)|$$

Proof. This follows immediately from Propositon 2.10, for any correspondence R:

$$d_{\mathrm{GH}}(G^X, G^H) \leq \frac{1}{2} dis\ R$$

The claim follow from choosing the natural correspondence:

$$R = \{(k,k) : k = 1, ..., n\} \subset [n] \times [n] = V(G^X) \times V(G^H)$$

□

Now we are in a position to derive a bound on the distance between G^X and G^H. This we can do under the assumption that the pointsets inducing the additional edges are identical. In fact, this does not have to be the case, but for our constructions of random graphs it will be true with high probability.

Lemma 3.6 ([1, Lemma 20]). *Suppose that $\mathcal{Q} \cap X = \mathcal{Q} \cap (H/2)$ and write $k = |\mathcal{Q} \cap X| = |\mathcal{Q} \cap (H/2)|$. Then*

$$d_{GH}(G^X, G^H) \leq k(\|X - H/2\| + 2).$$

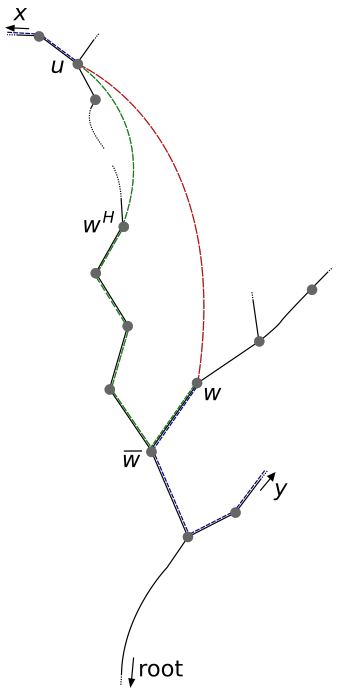


Figure 10: Part of a tree with its induced edges: the red edge is induced in G^X and the green edge is induced in G^H. The blue path together with the red edge is the shortest path between x and y in G^X. The blue path together with the green path is the path we use in G^H instead.

Proof. We want to use the previous Corollary 3.5. Thus, we need to estimate the graph distances in G^X and G^H respectively. Recall that the graph distance between two vertices x and y is the length of the shortest path between x and y. Define some notation for paths in a graph: think of a path as an ordered tupel of vertices, and we can concatenate two paths when the endpoint of one path is the first vertex of the other path. Denote the concatenation operator by $\oplus$. For example:

$$(1, 2, 4, 2) \oplus (2, 7, 1) = (1, 2, 4, 2, 7, 1)$$

Because G^X and G^H are constructed by adding edges to a tree T, a path in G^X or G^H can sometimes traverse only edges in T. Thus, denote the shortest path from a to b in T by $\pi^T(a, b)$. Similarly, denote the shortest path from a to b in G^X or G^H by $\pi^X(a, b)$ and $\pi^H(a, b)$ respectively. Now, we let $x, y \in V$ and let $\pi^X(x, y)$ be the shortest path between them in G^X. Assume there is exactly one edge on the path that was induced by the pointset $\mathcal{Q} \cap X$:

$$\pi^X(x, y) = \pi^T(x, u) \oplus (u, w) \oplus \pi^T(w, y)$$

and the edge (u, w) was induced by $(\xi, \xi') \in \mathcal{Q} \cap X$. Construct a path $\pi(x, y)$ in G^H:

$$\pi(x, y) = \pi^T(x, u) \oplus (u, w^H) \oplus \pi^{X,H}(w^H, w) \oplus \pi^T(w, y)$$

where (u, w^H) was induced by the same point $(\xi, \xi') \in \mathcal{Q} \cap X = \mathcal{Q} \cap (H/2)$ and some other path $\pi^{X,H}$ which we will construct later. Without loss of generality assume that one of the endpoints of this edge is u, otherwise it would be w and we could give use the same argument with reversed roles for u and w. Then we get a bound on the difference of graph distances:

$$d^H(x, y) - d^X(x, y) \leq |\pi^{X,H}(w^H, w)|$$

There is a unique path from u to the root and the vertex w^H lies on that path due to the way we construct G^H. Furthermore, w has to be connected to this path with exactly one edge, as

both u and w were active vertices at the same time during the depth first search. Thus, there is a vertex $\overline{w}$ on the path from u to the root and $d(w, \overline{w}) = 1$. And because $\overline{w}$ and w^H share their path to the root, their distance is determined by their respective height. We can improve our estimate:

$$d^H(x,y) - d^X(x,y) \leq |\pi^{X,H}(w^H, w)| \leq |H(\overline{w}) - H(w^H)| + 1$$

But because (u, w) and (u, w^H) are induced by the same point $(\xi, \xi') \in \mathcal{Q}$, we have $H(w^H) = 2X(w) + 2$:

$$\begin{aligned} d^H(x,y) - d^X(x,y) &\leq |H(\overline{w}) - H(w^H)| + 1 \leq |H(\overline{w}) - 2X(w)| + 3 \\ &\leq |2X(w) - H(w)| + 4 \leq \|2X - H\| + 4. \end{aligned}$$

By reading the argument from bottom to top, we can derive a symmetric bound for reversed roles of G^X and G^H.
Now consider any two vertices $x', y' \in V$. Now the shortest path between x' and y' can contain up to k induced edges, thus we have to repeat the construction above up to k times. We then get the estimate:

$$|d^H(x', y') - d^X(x', y')| \leq k(\|2X - H\| + 4) = 2\left(\|X - (H/2)\| + 2\right).$$

Using Corollary 3.5 we conclude:

$$d_{\mathrm{GH}}(G^X, G^H) \leq \frac{1}{2} \sup_{x,y \in V} |d^H(x,y) - d^X(x,y)| \leq k\left(\|X - (H/2)\| + 2\right).$$

□

4 Convergence of the connected component

In the previous section we have discussed how to encode the Erdős-Rényi graph using the depth first search. The information is encoded in a random function and a random set of points. In this section we will prove convergence of those objects and piece them together to reach our main goal - the scaling limit of the Erdős-Rényi graph.

4.1 Convergence of the depth first walk

In Theorem 2.16 we have seen that the rescaled contour function of a (depth first labeled) uniform tree on $[n]$ converges uniformly to an Brownian excursion. This result can be strenghtened to the height function and to the depth first walk, again the limits are Brownian excursions. Unsurprisingly, those encoding functions are far from independent, in fact they all converge to the same limiting distribution:

Theorem 4.1 ([10])**.** *As $m \to \infty$:*

$$\frac{1}{\sqrt{n}}(X^n(\lfloor n\cdot \rfloor), H^n(\lfloor n\cdot \rfloor), C^n(n\cdot)) \xrightarrow{d} (e(\cdot), 2e(\cdot), 2e(\cdot/2)).$$

Where X^n represents the depth first walk, H^n the height function and C^n the contour function of a uniform tree on $[n]$ with depth first labeling. The convergence takes place in $\mathbb{D}([0,1], \mathbb{R}^+)$ equipped with the Skorohod topology.

Because we do not interpolate X^n and H^n but rather view them as a step functions, we lose continuity and thus we have to work in a different space. The choice is $\mathbb{D}([0,1], \mathbb{R}^+)$, the space of cadlag function, i.e. continuity on the right and limits on the left. Now, those functions can be discontinuous and we let two functions be close if they jump at similar times and are otherwise close in the supremum norm, formally:

Definition 4.2. *A function $f : [0,1] \to \mathbb{R}^+$ is an element of $\mathbb{D}([0,1], \mathbb{R}^+)$ if for every $x \in [0,1]$:*

$$\lim_{t \downarrow 0} f(x+t) = f(x) \qquad \text{and} \qquad \lim_{t \downarrow 0} f(x-t) \quad \text{exists}$$

On this space, define the Skorohod topology, define:

$$\Lambda := \{\lambda : [0,1] \to [0,1] : \lambda \quad \text{is continuous, strictly increasing and} \quad \lambda(0) = 0, \lambda(1) = 1\}$$

Using this, we define a metric on $\mathbb{D}$, let $f, g \in \mathbb{D}([0,1], \mathbb{R}^+)$:

$$d(f,g) = \inf_{\lambda \in \Lambda} \max\{\|f - g \circ \lambda\|_\infty; \|\lambda - id\|_\infty\}$$

where id is the identity map on $[0,1]$. This is a well defined metric and the induced topology is called the Skorohod topology.

Check [4, Chapter 3] for more details.
The goal of this section is to derive an analogous result to Theorem 4.1, but not for uniform trees, but for tilted trees with the probability distribution introduced at the beginning of section 3.3. Yet now we have a slightly different distribution for our trees and thus we need a slightly different distribution of the limiting excursion. Define the space of continuous excursions:

$$\mathcal{E} = \{f \in C(\mathbb{R}^+, \mathbb{R}^+) : f(0) = 0, \exists \sigma \in (0, \infty) \text{ and } f(x) > 0 \ \forall x \in (0, \sigma) \text{ and } f(x) = 0 \ \forall x \geq \sigma\}$$

Definition 4.3. *For $\sigma > 0$, define a tilted excursion of length σ, a random variable on $\mathcal{E}$.*

$$\mathbb{P}\left(\tilde{e}^{(\sigma)} \in \mathcal{B}\right) = \frac{\mathbb{E}\left[\mathbb{1}_{\{e^{(\sigma)} \in \mathcal{B}\}} \exp\left(\int_0^\sigma e^{(\sigma)}(s)\,\mathrm{d}s\right)\right]}{\mathbb{E}\left[\exp\left(\int_0^\sigma e^{(\sigma)}(s)\,\mathrm{d}s\right)\right]}$$

Here, $e^{(\sigma)}$ is a Brownian excursion of length σ, characterized by $(e^{(\sigma)}(t), 0 \le t \le \sigma) \overset{d}{=} (\sqrt{\sigma} e(t/\sigma), 0 \le t \le \sigma)$ and $e(\cdot)$ is a standard Brownian excursion. Write $\tilde{e}$ for $\tilde{e}^{(1)}$.

In the follwing results, we will prove statements only for $\sigma = 1$ for notational simpicity, the proof for arbitrary σ always follows from the fact, that $\tilde{e}$ exhibts a scaling property in the same way e does, i.e.:

$$(\tilde{e}^{(\sigma)}(t), 0 \le t \le \sigma) \overset{d}{=} (\sqrt{\sigma}\tilde{e}(t/\sigma), 0 \le t \le 1)$$

This property is refered to as Brownian scaling.

Theorem 4.4 ([1, Thm. 12])**.** *Let $p = p(m)$ and $mp^{2/3} \to \sigma$ for $m \to \infty$. Then:*

$$\left(\sqrt{\frac{\sigma}{m}} \tilde{X}^m(\lfloor (m/\sigma)t \rfloor), 0 \le t \le \sigma\right) \overset{d}{\to} \left(\tilde{e}^{(\sigma)}(t), 0 \le t \le \sigma\right)$$

in $\mathbb{D}([0,\sigma], \mathbb{R}^+)$.

Proof. Assume $\sigma = 1$, the general case follows from Brownian scaling like described above. Let T^m be a uniformly chosen tree and $(X^m(i), 0 \le i \le m)$ its depth-first walk. We already now by Theorem 4.1 that

$$\left(m^{-1/2} X^m(\lfloor mt \rfloor), 0 \le t \le 1\right) \overset{d}{\to} (e(t), 0 \le t \le 1)$$

where $e(t)$ is again a standard Brownian excursion. To go from uniform trees to tilted trees, we need to check how the area $a(T^m)$ behaves in the limit $m \to \infty$, observe:

$$m^{-3/2} a(T^m) = \frac{1}{m} \sum_{i=0}^{m-1} \frac{1}{m^{1/2}} X^m(i) = \int_0^1 \overline{X}^m(t)\,\mathrm{d}t \overset{d}{\to} \int_0^1 e(t)\,\mathrm{d}t$$

where $\overline{X}^m(t) = m^{-1/2} X^m(\lfloor mt \rfloor)$ and the last step holds, because the integration operator $h \mapsto \int_0^1 h(t)\,\mathrm{d}t$ is continuous. Furthermore observe:

$$(1-p)^{-a(T^m)} = (1-p)^{-m^{3/2} \int_0^1 \overline{X}^m(t)\,\mathrm{d}t} \overset{d}{\to} \exp\left(\int_0^1 e(t)\,\mathrm{d}t\right)$$

Now let $f : \mathbb{D}([0,1], \mathbb{R}^+) \to \mathbb{R}_0^+$ any bounded function. Then:

$$\begin{aligned}\mathbb{E}\left[f\left(m^{-1/2}\tilde{X}^m(\lfloor m\cdot \rfloor)\right)\right] &= \frac{\mathbb{E}\left[f(\overline{X}^m)(1-p)^{-m^{3/2}\int_0^1 \overline{X}^m(t)\,\mathrm{d}t}\right]}{\mathbb{E}\left[(1-p)^{-m^{3/2}\int_0^1 \overline{X}^m(t)\,\mathrm{d}t}\right]} \\ &\longrightarrow \frac{\mathbb{E}\left[f(e)\exp\left(\int_0^1 e(s)\,\mathrm{d}s\right)\right]}{\mathbb{E}\left[\exp\left(\int_0^1 e(s)\,\mathrm{d}s\right)\right]} = \mathbb{E}\left[f(\tilde{e})\right]\end{aligned}$$

Here we have used the uniform integrability of the family $\{(1-p)^{-\xi a(T_m)}\}_{\xi > 0}$ which is proven in [1, Lemma 14]. Thus we have $\mathbb{E}\left[f\left(m^{-1/2}\tilde{X}^m(\lfloor m\cdot \rfloor)\right)\right] \to \mathbb{E}[f(\tilde{e})]$, which implies $(m^{-1/2}\tilde{X}^m(\lfloor mt \rfloor), 0 \le t \le 1) \overset{d}{\to} (\tilde{e}(t), 0 \le t \le 1)$. □

In light of Theorem 4.1, we cite following related result:

Theorem 4.5 ([1, Thm. 15]). *Let $p = p(m)$ and $mp^{2/3} \to \sigma$ for $m \to \infty$. Then:*

$$\left(\sqrt{\frac{\sigma}{m}}\tilde{H}^m(\lfloor (m/\sigma)t \rfloor), 0 \le t \le \sigma\right) \xrightarrow{d} \left(2\tilde{e}^{(\sigma)}(t), 0 \le t \le \sigma\right)$$

in $\mathbb{D}([0,\sigma], \mathbb{R}^+)$.

4.2 Convergence of the additional edges

In Proposition 3.4 we have seen that the surplus edges of the Erdős-Rényi graph can be encoded via binomial pointsets. Recall that we defined a binomial pointset as subset of $\mathbb{Z}^+ \times \mathbb{Z}^+$ where each point is included independently with probability p. We want to know how the distribution of the surplus edges behaves when we rescale our space. For this we need to define Poisson point processes on $\mathbb{R}^+ \times \mathbb{R}^+$.

Definition 4.6. *A random variable $\mathcal{P}$ taking values in the countable subsets of $\mathbb{R}^+ \times \mathbb{R}^+$ is called Poisson point process if for disjoint measureable bounded sets $A_1, ..., A_k \subset \mathbb{R}^+ \times \mathbb{R}^+, k \in \mathbb{N}$, we have:*

- *$N(A_i) := \#(P \cap A_i), i = 1, ..., k$ are independent random variables.*
- *$N(A_i) \sim Poisson(\mu(A_i))$ for any $A_i, i = 1, .., k$.*

where μ denotes the two-dimensional Lebesgue measure and $\#$ the cardinality of a set.

See [8, Chapter 24] for existence and [8, Korollar 24.9] ensures that our specification of the distribution defines $\mathcal{P}$ uniquely.

A simple fact for this Poisson point process is that $\mathcal{P}$ has almost surely only finitely many points in any compact set and almost surely no points in a any null set. See Figure 11 for illustration.

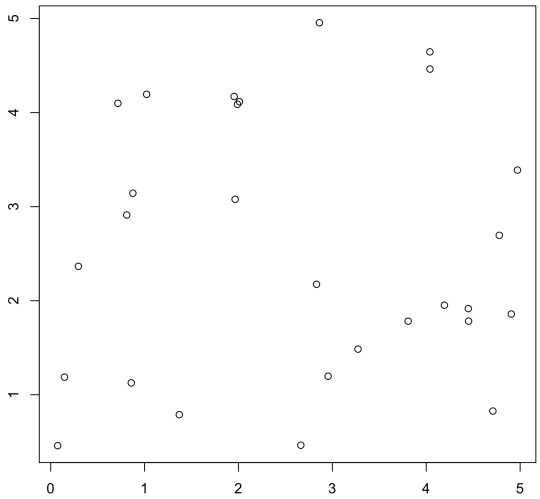


Figure 11: A realization of a Poisson point process with measure μ restricted to $[0,5] \times [0,5]$, simulated using R.

The next lemma is perhaps the most crucial one in this thesis: it classifies the joint convergence of the height process and the additional edges and thus all necessary information contained in the Erdős-Rényi graph:

Lemma 4.7 ([1, Lemma 18]). *Let $p = p(m)$ and $mp^{2/3} \to \sigma$ for $m \to \infty$. Recall that $\tilde{T}_m^p$ is a random tree on $[m]$ and let $\tilde{H}^m$ be its height process. Let $\mathcal{Q}^p \subseteq \mathbb{Z}^+ \times \mathbb{Z}^+$ be a binomial pointset of intensity p. Define:*

$$\mathcal{P}_m = \{((m/\sigma)^{-1}i, (m/\sigma)^{-1/2}j) : (i,j) \in \mathcal{Q}^p\}$$

a rescaled version of $\mathcal{Q}^p$, we will use this notation throughout the rest of the thesis. Then:

$$(((m/\sigma)^{-1/2}\tilde{H}^m(\lfloor (m/\sigma)t \rfloor), 0 \le t \le \sigma), \mathcal{P}_m \cap ((m/\sigma)^{-1/2}\tilde{H}^m(\lfloor (m/\sigma)t \rfloor)/2, 0 \le t \le \sigma)) \xrightarrow{d} (2\tilde{e}^{(\sigma)}, \mathcal{P} \cap \tilde{e}^{(\sigma)})$$

where $\mathcal{P}$ is a Poisson point process like defined above such that $\mathcal{P}$ is independent of $\tilde{e}^{(\sigma)}$. Convergence takes place in $\mathbb{D}([0,\sigma], \mathbb{R}^+)$ in the first coordinate and in the sense of Hausdorff distance in the second coordinate.

Proof. Assume $\sigma = 1$, the general case follows from Brownian scaling like described above. By Theorem 4.5 we have:

$$(m^{-1/2}\tilde{H}^m(\lfloor mt \rfloor), 0 \le t \le 1) \xrightarrow{d} (2\tilde{e}(t), 0 \le t \le 1)$$

in $\mathbb{D}([0,1], \mathbb{R}^+)$. Thus, we will show $\mathcal{P}_m \xrightarrow{d} \mathcal{P}$ and then show that joint convergence holds. Let $k \ge 1$ and $A_1, ..., A_k \subset [0,1] \times \mathbb{R}^+$ be disjoint, measurable and bounded sets. Define

$$N_m(A_i) = \#\{(\lfloor m^{-1/2}x \rfloor, \lfloor my \rfloor) \in \mathcal{Q}^p : (x,y) \in A_i\}$$

the set of lattice points that are both in $\mathcal{Q}^p$ and the rescaled A_i. This is a binomial random variable with parameters $\eta_m(A_i) := \#\{\lfloor m^{-1/2}x \rfloor, \lfloor my \rfloor) : (x,y) \in A_i\}$ and p. Because $\eta_m(A_i)$ approximates A_i by rectangles of size $m^{-3/2}$ we have:

$$m^{-3/2}\eta_m(A_i) \overset{m\to\infty}{\longrightarrow} \mu(A_i)$$

Combined with $m^{3/2}p \to 1$ this implies $N_m(A_i) \xrightarrow{d} Poisson(\mu(A_i))$. Furthermore, the random variables $N_m(A_1), ..., N_m(A_k)$ are independent as they count points of disjoint sets. By [8, Korollar 24.9] this identifies the limit in distribution of $\mathcal{P}_m$ uniquely as $\mathcal{P}$.

Next, we will proof two general claims:
Claim: Assume $f_n : [0,1] \to \mathbb{R}^+$ continuous converges uniformly to some $f : [0,1] \to \mathbb{R}^+$. For any open set $A \subseteq [0,1] \times \mathbb{R}^+$ consider

$$\begin{cases} A_n = \{(x,y) \in A : 0 < y < f_n(x)\} \\ A_* = \{(x,y) \in A : 0 < y < f(x)\} \end{cases}$$

Then $A_n \to A_*$ in the Hausdorff sense.
Proof of Claim: First consider:

$$\begin{cases} B_n = \{(x,y) \in]0,1[\times \mathbb{R}^+ : 0 < y < f_n(x)\} \\ B_* = \{(x,y) \in]0,1[\times \mathbb{R}^+ : 0 < y < f(x)\} \end{cases}$$

Then using the fact that f_n is uniformly bounded and thus all sets B_n, B_* are bounded:

$$d_{\mathrm{H}}(B_n, B) \le d_{\mathrm{H}}(\partial B_n, \partial B_*) \le \|f_n - f\| \overset{n\to\infty}{\longrightarrow} 0$$

Thus $B_n \to B_*$ in the Hausdorff sense. Because $A_n = B_n \cap A$ and $A_* = B_* \cap A$ it follows that $B_n \cap A \to B_* \cap A$. ∎
Claim: Suppose $g_n : \{0, ..., n\} \to \mathbb{Z}^+$, $\|g_n\| \le n$ and $(n^{-1/2}g_n(\lfloor nt \rfloor, 0 \le t \le 1)$ converges to $(g(t), 0 \le t \le 1)$ in $\mathbb{D}([0,1], \mathbb{R}^+)$ such that g is continuous. Then the linear interpolation $\tilde{g}_n : [0,1] \to \mathbb{R}^+$ of $(n^{-1/2}g_n(nt) : t \in \{0, 1/n, ..., 1\})$ converges uniformly to g.

Proof of Claim: By our definition of $\mathbb{D}([0,1],\mathbb{R}^+)$, we find a sequence $(\lambda_n)_n \in \Lambda$ such that $\max\{\|n^{-1/2}g_n(\lfloor n\cdot\rfloor) - g\circ\lambda_n\|_\infty, \|\lambda_n - id\|\} \to 0$. Now:

$$\begin{aligned}\|\tilde{g}_n - g\| &\leq \|\tilde{g}_n - n^{-1/2}g_n(\lfloor n\cdot\rfloor)\| + \|n^{-1/2}g_n(\lfloor n\cdot\rfloor) - g\circ\lambda_n\| + \|g\circ\lambda_n - g\| \\ &= n^{-1/2}\sup_{t\in[0,1]}|g_n(\lfloor nt\rfloor) + n^{-1}(nt - \lfloor nt\rfloor)(g_n(\lfloor nt\rfloor + 1) - g_n(\lfloor nt\rfloor)) - g_n(\lfloor nt\rfloor)| + \\ &\quad \|n^{-1/2}g_n(\lfloor n\cdot\rfloor) - g\circ\lambda_n\| + \sup_{t\in[0,1]}\sup_{|t-t'|\leq\epsilon_n}|g(t) - g(t')| \\ &\leq 2n^{-1/2} + \|n^{-1/2}g_n(\lfloor n\cdot\rfloor) - g\circ\lambda_n\| + \sup_{t\in[0,1]}\sup_{|t-t'|\leq\epsilon_n}|g(t) - g(t')| \overset{n\to\infty}{\longrightarrow} 0\end{aligned}$$

for $\epsilon_n = \|\lambda_n - id\|$. The second term tends to zero by assumption and the third term tends to zero because $\epsilon_n \to 0$ and g is uniformly continuous. ■

We now apply the second claim to the height process $\tilde{H}^m : \{0,...,m\} \to \mathbb{Z}^+$. With the remark above we get that the rescaled linear interpolation of $\tilde{H}^m$ converges uniformly. Thus by the first claim, following convergence holds for $\mathcal{P}_m$ because $\mathcal{P}_m$ is independent of $\tilde{H}^m$:

$$\mathcal{P}_m \cap (m^{-1/2}\tilde{H}^m_{inter}/2(mt), 0 \leq t \leq 1) \overset{d}{\to} \mathcal{P} \cap (\tilde{e}(t), 0 \leq t \leq 1)$$

in the Hausdorff sense and where $\tilde{H}^m_{inter}$ is the interpolated version of $\tilde{H}^m$. We have also used the fact that $\mathcal{P} \cap \{(x,y) \in \mathbb{R}^+ \times \mathbb{R}^+ : x \in [0,1], 2\tilde{e}(x) = y\} = \emptyset$ almost surely. Because $\tilde{H}^m(\lfloor\cdot\rfloor)$ and $\tilde{H}^m_{inter}(\cdot)$ agree on lattice points, we have:

$$\mathcal{P}_m \cap (m^{-1/2}\tilde{H}^m/2(\lfloor mt\rfloor), 0 \leq t \leq 1) \overset{d}{\to} \mathcal{P} \cap (\tilde{e}(t), 0 \leq t \leq 1)$$

in the Hausdorff sense jointly with the convergence of $\tilde{H}^m$. □

4.3 Limit object

Lemma 4.7 suggests that the information needed to construct G^H converges nicely after rescaling. Thus it is natural to define a glueing procedure as similar construction in the continuous case. Given a continuous function $h : [0,\sigma] \to \mathbb{R}^+$ such that $h(0) = h(\sigma) = 0$ and a set $P \subset \mathbb{R}^+ \times \mathbb{R}^+$ such that $P \cap (h/2)$ is finite, define $g(h,P)$: Let $\mathcal{T}$ be the real tree encoded by the height function h with its projection $\rho : [0,\sigma] \to \mathcal{T}$ and root $\rho(0)$ - compare to Theorem 2.7. For every point $(\xi_x,\xi_y) \in P \cap (h/2)$, there is a vertex $\rho(\xi_x) \in \mathcal{T}$ at height $h(\xi_x)$. Because $\mathcal{T}$ is a real tree, there is a unique geodesic of length $h(\xi_x)$ from $\rho(\xi_x)$ to the root. Denote the unique point at height $2\xi_y$ on this geodesic by $\hat{\rho}(\xi_x,\xi_y)$. Identify $\hat{\rho}(\xi_x,\xi_y)$ and $\rho(\xi_x)$ for each $(\xi_x,\xi_y) \in P$ and equip the resulting space with the induced path metric.

Definition 4.8 (Continuum Random Graph)**.** *Let $0 < \sigma < \infty$. Let $\tilde{e}^{(\sigma)}$ be a tilted excursion of length σ and $\mathcal{P}$ a Poisson point process with intensity measure μ - the Lebesgue measure. Then:*

$$\mathcal{M}^{(\sigma)} = g(2\tilde{e}^{(\sigma)}, \mathcal{P})$$

is the Continuum Random Graph with parameter σ. Write $\mathcal{M} := \mathcal{M}^{(1)}$.

There is a more intuitive description for this object: Take the continuum random tree $\mathcal{T}$ encoded by $2\tilde{e}^{(\sigma)}$. Select a random number N of points on $[0,\sigma]$ with density proportional to $\tilde{e}^{(\sigma)}$. There is a unique point τ in $\mathcal{T}$ corresponding to each random pondom point on $[0,\sigma]$. Identify τ with a uniformly chosen point on the geodesic from τ to the root for each random point. This can be justified formally:

Proposition 4.9 ([1, Prop. 19])**.** *Using the notation above, conditional on $\tilde{e}^{(\sigma)}$, following statements hold for the continuum random graph:*

1. *For any* $(\xi_x, \xi_y) \in \mathcal{P} \cap \tilde{e}^{(\sigma)}$, ξ_x *has density*

$$\frac{\tilde{e}^{(\sigma)}(u)}{\int_0^\sigma \tilde{e}^{(\sigma)}(s)\,\mathrm{d}s}$$

on $[0, \sigma]$. *Furthermore,* $\rho(\xi_x)$ *is almost surely a leaf of* $\mathcal{T}$.

2. *We have* $d(\rho(0), \hat{\rho}(\xi_x, \xi_y)) := 2\xi_y = Ud(\rho(0), \rho(\xi_x))$, *where* U *is a uniform random variable on* $[0, 1]$, *independent of* $\tilde{e}^{(\sigma)}$ *and* ξ_x.

3. *The number of additional vertex identification* N *has a Poisson distribution with mean* $\int_0^\sigma \tilde{e}^{(\sigma)}(s)\,\mathrm{d}s$.

Proof. 1. Again, let μ be the two dimensional Lebesgue measure. Any point $(\xi_x, \xi_y) \in \mathcal{P} \cap \tilde{e}^{(\sigma)}$ is located uniformly in all possible positions, thus for any $[a, b] \subseteq [0, \sigma]$:

$$\mathbb{P}(\xi_x \in [a, b]) = \mathbb{P}((\xi_x, \xi_y) \in ([a, b] \times \mathbb{R}^+) \cap \tilde{e}^{(\sigma)}) = \frac{\int_a^b \tilde{e}^{(\sigma)}(u)\,\mathrm{d}u}{\int_0^\sigma \tilde{e}^{(\sigma)}(s)\,\mathrm{d}s}$$

It follows that ξ_x has the claimed density. Let $\mathcal{L}(\mathcal{T})$ be the leaves of $\mathcal{T}$. Then:

$$\mathbb{P}(\rho(\xi_x) \in \mathcal{L}(\mathcal{T})) = \int_0^\sigma \mathbb{1}_{\{\rho(u) \in \mathcal{L}(\mathcal{T})\}} \frac{\tilde{e}^{(\sigma)}(u)}{\int_0^\sigma \tilde{e}^{(\sigma)}(s)\,\mathrm{d}s}\,\mathrm{d}u = \frac{\int_0^\sigma \tilde{e}^{(\sigma)}(u)\,\mathrm{d}u}{\int_0^\sigma \tilde{e}^{(\sigma)}(s)\,\mathrm{d}s} = 1$$

The distribution of $\tilde{e}^{(\sigma)}$ is absolutely continuous with respect to $e^{(\sigma)}$, this implies that we can use Proposition 2.17 to achieve $\mathbb{1}_{\{\rho(u) \in \mathcal{L}(\mathcal{T})\}} = 1$ almost surely for almost all $\tilde{e}^{(\sigma)}$. Therefore, $\rho(\xi_x)$ is almost surely a leaf.

2. Similar to 1., conditional on ξ_x, ξ_y is distributed uniformly on $[0, \tilde{e}^{(\sigma)}(\xi_x)]$, the statement follows when we identify the interval with the geodesic path from $\rho(\xi_x)$ to the root.

3. The number of vertex identifications is precisely the number of marked points under the graph of $\tilde{e}^{(\sigma)}$, thus:

$$N = |\mathcal{P} \cap \tilde{e}^{(\sigma)}| \overset{d}{=} Poisson\left(\int_0^\sigma \tilde{e}^{(\sigma)}(s)\,\mathrm{d}s\right).$$

□

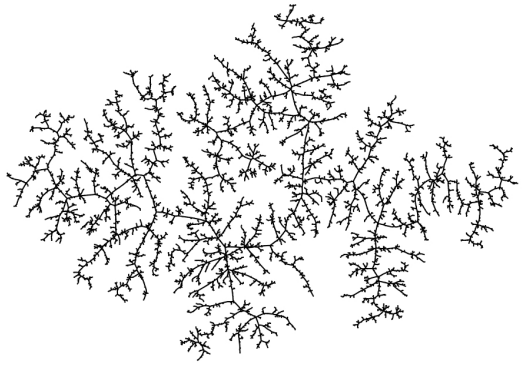


Figure 12: A realization of the Continuum Random Graph. The picture was made by Nicolas Broutin, source: MATHS.OX.AC.UK/NODE/30217.

To finish this section, we cite a lemma that allows us to compare two different metric spaces that were both obtained by the glueing procedure. Let h_1, h_2 be two suitable height functions of length σ and $\mathcal{Q}_1, \mathcal{Q}_2$ two suitable pointsets. Denote $G_1 = g(h_1, \mathcal{Q}_1)$ and $G_2 = g(h_2, \mathcal{Q}_2)$, then:

Lemma 4.10 ([1, Lemma 21]). *Suppose that $k = |\mathcal{Q}_1 \cap (h_1/2)| = |\mathcal{Q}_2 \cap (h_2/2)|$ and $\delta = d_H(\mathcal{Q}_1 \cap (h_1/2), \mathcal{Q}_2 \cap (h/2))$. Then:*

$$d_{GH}(G_1, G_2) \leq \frac{1}{2}(k+1)\left(\delta + 12\|h_1 - h_2\|_\infty + 4 \sup_{|r-r'|\leq\delta} |h_2(r) - h_2(r')|\right)$$

Sketch of proof. The proof is similar to the proof of Lemma 3.6 as we estimate the Gromov Hausdorff distance using Propositon 2.10 and the natural correspondence:

$$d_{\mathrm{GH}}(G_1, G_2) \leq \frac{1}{2} \sup_{t,t'\in[0,1]} |d_1(\rho_1(t), \rho_1(t')) - d_2(\rho_2(t), \rho_2(t'))|$$

where d_1, d_2 denote the metrics in G_1 and G_2 respectively and ρ_1, ρ_2 the projections from $[0, \sigma]$ into G_1 and G_2. The graphs are equipped with the path metric, thus we need to estimate the length of geodesics. Denote the points of $\mathcal{Q}_1 \cap (h_1/2)$ by $\xi_1, ..., \xi_k$ and the points of $\mathcal{Q}_2 \cap (h_2/2)$ by $\eta_1, ..., \eta_k$. Because $d_{\mathrm{H}}(\mathcal{Q}_1 \cap (h_1/2), \mathcal{Q}_2 \cap (h/2)) \leq \delta$, we can relabel those points such that

$$\sup_{i=1,...,k} \|\xi_i - \eta_i\| \leq \delta.$$

Let $x, y \in G_1$ such that $x = \rho_1(t), y = \rho_1(t') : t, t' \in [0, \sigma]$ and $\pi_1(x, y)$ be the geodesic in G_1 connecting x and y. We then construct another path $\pi_2(\rho_2(t), \rho_2(t'))$ in G_2. We can split $\pi_1(x, y)$ up into at most k segements

$$\pi_1(x, y) = \pi(x, u_1) \oplus \bigoplus_{j=1}^{l-1} \pi(u_j, u_{j+1}) \oplus \pi(u_l, y)$$

where $\pi(\cdot, \cdot)$ is a geodesic in the underlying real tree of G_1, $\oplus$ is the concatenation operator and $u_j; j = 1, ..., l \leq k$ is a point where a vertex identification is induced by $\mathcal{Q}_1$. This means that $u_j = \rho_1(\xi_i)$ for every j and some $i = i(j)$. We then construct the path in G_2 by traversing the vertex identification $\rho_2(\eta_i)$ whenever $\rho_1(x_i)$ is traversed. This will contribute a length difference of at most $(\delta + 2\sup_{|r-r'|\leq\delta} |h_2(r) - h_2(r')|)$ for each vertex identification. The geodesic segments in the underlying real tree of G_1 can be approximated by geodesic segments in the underlying real tree of G_2 contibuting a length difference proportional to $\|h_1 - h_2\|_\infty$ according to Corollary 2.11. Using a symmetric argument for switched roles of G_1 and G_2 and

$$\sup_{|r-r'|\leq\delta} |h_1(r) - h_1(r')| \leq 2\|h_1 - h_2\| + \sup_{|r-r'|\leq\delta} |h_2(r) - h_2(r')|$$

to obtain the final estimate. □

4.4 Convergence of the rescaled Erdős-Rényi graph

In this section we will combine our previous results to prove the main theorem which we can finally state:

Theorem 4.11 ([1, Thm. 22]). *Suppose $\sigma > 0$ and $m = m(n) \in \mathbb{Z}^+$ such that $n^{-2/3}m \to \sigma$. Let $p = p(n) \in (0, 1)$ such that $pn \to 1$. Let G_m^p a connected component of $G(n, p)$ conditioned on size m. Then, as $n \to \infty$:*

$$n^{-1/3} G_m^p \xrightarrow{d} \mathcal{M}^{(\sigma)}$$

as metric spaces with respect to the Gromov-Hausdorff distance.

Before we proof the Theorem, we will derive two important corollaries. Using Theorem 2.5 we obtain the following:

Corollary 4.12. *Consider $G(n, 1/n)$. Denote the largest component by G_n^1 and its size by Z_n. Then there is a non-trivial random variable Z such that as $n \to \infty$:*

$$(n^{-2/3}Z_n, n^{-1/3}G_n^1) \xrightarrow{d} (Z, \mathcal{M}^{(Z)})$$

where the convergence in the second component is with respect to the Gromov-Hausdorff distance.

This means that the largest component of the critical Erdős-Rényi graph approaches the continuum random graph when rescaled by factor $n^{-1/3}$.

Furthermore, we can make a statement on the number of cycles in the critical Erdős-Rényi graph. Recall that the surplus $s(G)$ of a connected graph G is the minimal number of edges which we have to remove to obtain a tree.

Corollary 4.13 ([1, Cor. 24])**.** *Suppose that $m = m(n)$ is such that $n^{-2/3}m \to \sigma$ as $n \to \infty$ and let $p = p(n)$ such that $pn \to 1$. Then:*

$$s(G_m^p) \xrightarrow{} Poisson\left(\int_0^\sigma \tilde{e}^{(\sigma)}(s)\,\mathrm{d}s\right).$$

Proof of Corollary: Observe:

$$s(G_m^p) \overset{d}{=} |\mathcal{P}_m \cap ((m/\sigma)^{-1/2}\tilde{X}^m(\lfloor m/\sigma)t\rfloor), 0 \le t \le \sigma)| \xrightarrow{d} |\mathcal{P} \cap \tilde{e}| \overset{d}{=} Poisson\left(\int_0^\sigma \tilde{e}^{(\sigma)}(s)\,\mathrm{d}s\right)$$

The joint convergence of $\mathcal{P}_m$ and $\tilde{X}^m$ follows an analogous argument like Lemma 4.7. □

Proof of Theorem 4.11. Assume $\sigma = 1$, the general case follows from Brownian scaling like described above.
Recall that by Proposition 3.4 G^X and the Erdős-Rényi graph are closely related. We want to make use of G^H and its similarity to the glueing procedure $g(\cdot,\cdot)$. We then want to apply Lemma 3.6 and Lemma 4.10 to obtain estimates on the Gromov-Hausdorff distances. Thus the proof consists of two parts: in the first part we will deal with estimates concering G^H and $g(\cdot,\cdot)$, and in the second part we will deal with estimates concerning G^X and G^H.

Part 1:
We will only sketch this part. While the constructions $G^H(\cdot,\cdot)$ and $g(\cdot,\cdot)$ are very similar, the graph $m^{-1/2}G^H(\tilde{T}_m^p, \mathcal{Q}_p)$ viewed as metric space is not the same as $g(m^{-1/2}\tilde{H}_c^m, \mathcal{Q}^p)$, where $\tilde{H}_c^m$ is the continuous interpolation of $\tilde{H}$. The reason for this is simple: in $m^{-1/2}G^H(\cdot,\cdot)$ we insert edges of length $m^{-1/2}$ where as in $g(\cdot,\cdot)$ we identify points. Furthermore, the function $\tilde{H}_c^m$ does not encode the $\mathbb{R}$-tree $\tilde{T}_m^p$ but the contour function $\tilde{C}^m$ does. Thus we first approximate $m^{-1/2}G^H(\tilde{T}_m^p, \mathcal{Q}_p)$ by $g(\tilde{C}_r^m, \mathcal{Q}_m)$, where $\tilde{C}_r^m(t) = m^{-1/2}\tilde{C}_r^m(2mt), 0 \le t \le 1$ is the rescaled version of $\tilde{C}^m$:

$$d_{\mathrm{GH}}(m^{-1/2}G^H(\tilde{T}_m^p, \mathcal{Q}_p), g(\tilde{C}_r^m, \mathcal{Q}_m)) \le \frac{1}{2}m^{-1/2}k$$

where k is the number of vertex identifications. The reason for this is that both metric spaces have the same underlying real tree and adding an edge of length $m^{-1/2}$ instead of an identification can only change the distortion by at most $m^{-1/2}$, the factor $1/2$ is contributed by Lemma 2.10. Also note, that we replaced $\mathcal{Q}_p$ (or rather $\mathcal{P}_m$) by another set $\mathcal{Q}_m$ - the specific definition does not matter too much - to make up for the fact that we change from the height function to the contour function.

Our next goal is to apply Lemma 4.10. To fulfill the prerequisites of the lemma we apply Skorohod's representation theorem to Lemma 4.7, thus:

$$(m^{-1/2}\tilde{H}^m(\lfloor m\cdot\rfloor), \tilde{C}_r^m(\cdot), \mathcal{P}_m \cap (m^{-1/2}\tilde{H}^m(\lfloor m\cdot\rfloor)/2)) \to (2\tilde{e}, 2\tilde{e}, \mathcal{P}\cap\tilde{e})$$

almost surely in some probability space. Let $\epsilon > 0$ be arbitrary but fixed. Then there is an almost surely random variable Y such that for every $m \geq Y$:

- $|\mathcal{P}_m \cap (m^{-1/2}\tilde{H}^m(\lfloor m\cdot\rfloor)/2)| = |\mathcal{P}\cap\tilde{e}|$,
- $d_{\mathrm{H}}(\mathcal{P}_m \cap (m^{-1/2}\tilde{H}^m(\lfloor m\cdot\rfloor)/2), \mathcal{P}\cap\tilde{e}) \leq \epsilon$,
- $\|\tilde{C}_r^m - 2\tilde{e}\| \leq \epsilon$ and
- $d_{\mathrm{H}}(\mathcal{Q}_m, \mathcal{P}_m \cap (m^{-1/2}\tilde{H}^m(\lfloor m\cdot\rfloor)/2)) \leq \epsilon$.

We then also get that $d_{\mathrm{H}}(\mathcal{Q}_m, P\cap\tilde{e}) \leq 2\epsilon$. This means that for big enough m, the prerequisites of Lemma 4.10 are fulfilled and we apply it to $g(\tilde{C}_r^m, \mathcal{Q}_m)$ and $g(2\tilde{e}, \mathcal{P})$:

$$d_{\mathrm{GH}}(g(\tilde{C}_r^m, \mathcal{Q}_m), g(2\tilde{e}, \mathcal{P})) \leq \frac{1}{2}(|\mathcal{P}\cap\tilde{e}|+1)\left(2\epsilon + 12\epsilon + 4\sup_{|t-s|\leq 2\epsilon}\left|2\tilde{e}(t) - 2\tilde{e}(s)\right|\right)$$

Switching back to G^H, we get:

$$d_{\mathrm{GH}}(m^{-1/2}G^H(\tilde{T}_m^p, \mathcal{Q}_p), g(2\tilde{e}, \mathcal{P})) \leq \frac{1}{2}((1+m^{-1/2})|\mathcal{P}\cap\tilde{e}|+1)\left(14\epsilon + 8\sup_{|t-s|\leq 2\epsilon}\left|\tilde{e}(t) - \tilde{e}(s)\right|\right)$$

Levy's Modulus of Continuity Theorem (check [11, Thm. 2.7] for a proof) tells us:

$$\mathbb{P}\left(\sup_{|t-s|\leq 2\epsilon}|e(t)-e(s)| \geq \epsilon^{1/4}\right) \to 0$$

as $\epsilon \to 0$ for a standard Brownian excursion e. Because the distribution of $\tilde{e}(\cdot)$ is absolutely continuous with respect to $e(\cdot)$, the same holds for $\tilde{e}(\cdot)$. Furthermore:

$$\mathbb{P}(|\mathcal{P}\cap\tilde{e}| \geq \epsilon^{-1/8}) \leq \mathbb{P}\left(|\mathcal{P}\cap\tilde{e}| \geq \epsilon^{-1/8}\middle|\int_0^1 \tilde{e}(s)\,\mathrm{d}s \leq \epsilon^{-1/16}\right) + \mathbb{P}\left(\int_0^1 \tilde{e}(s)\,\mathrm{d}s \leq \epsilon^{-1/16}\right)$$

We estimate both terms using Markov's inequality:

$$\mathbb{P}\left(\int_0^1 \tilde{e}(s)\,\mathrm{d}s \leq \epsilon^{-1/16}\right) \leq \mathbb{E}\left[\int_0^1 \tilde{e}(s)\,\mathrm{d}s\right]\epsilon^{1/16} = L\epsilon^{1/16}$$

$$\mathbb{P}\left(|\mathcal{P}\cap\tilde{e}| \geq \epsilon^{-1/8}\middle|\int_0^1 \tilde{e}(s)\,\mathrm{d}s \leq \epsilon^{-1/16}\right) \leq \epsilon^{1/8}\mathbb{E}[Z] = \epsilon^{1/16}$$

where Z is distributed like a $Poisson(\epsilon^{-1/16})$ random variable and $L = \mathbb{E}\left[\int_0^1 \tilde{e}(s)\,\mathrm{d}s\right]$ is finite. Because ϵ was arbitrary and recall the definition $\mathcal{M} = g(2\tilde{e}, \mathcal{P})$. We obtain

$$d_{\mathrm{GH}}(m^{-1/2}G^H(\tilde{T}_m^p, \mathcal{Q}^p), \mathcal{M}) \xrightarrow{\mathbb{P}} 0$$

completing the first part of the proof.

Part 2:

By Proposition 3.4 we know that a connected component of $G(n,p)$ conditioned on having m vertices is distributed like $G^X(\tilde{T}_m^p, \mathcal{Q}_p)$. Thus it is enough to show that G^X and G^H converge

to the same limit after rescaling. By applying Skorohod's representation theorem to Theorem 4.4 and Lemma 4.7, there is a probability space such that almost surely:

$$\left(\left(\frac{\tilde{H}^m(\lfloor mt\rfloor)}{\sqrt{m}}, 0\le t\le 1\right), \left(\frac{\tilde{X}^m(\lfloor mt\rfloor)}{\sqrt{m}}, 0\le t\le 1\right), \mathcal{P}_m \cap \left(\frac{\tilde{H}^m(\lfloor mt\rfloor)}{2\sqrt{m}}, 0\le t\le 1\right)\right)$$

$$\xrightarrow{d} \big((2\tilde{e}(t), 0\le t\le 1), (\tilde{e}(t), 0\le t\le 1), \mathcal{P}\cap(\tilde{e}(t), 0\le t\le 1)\big)$$

Furthermore by [1, Lemma 16], we may assume that $m^{-1/2}\|\tilde{X}^m - \tilde{H}^m/2\|_\infty \to 0$ almost surely. Consider the integer valued random variable

$$D_m = |(\mathcal{Q}^p \cap \tilde{X}^m) \triangle (\mathcal{Q}^p \cap (\tilde{H}^m/2))|$$

the size of the symmetric difference. D_m is stochastically dominated by a Binomial random variable D'_m with parameters $m\lceil\|\tilde{X}^m - \tilde{H}^m/2\|_\infty\rceil$ and $m^{-3/2}$ in the sense that for any $l\in\mathbb{N}_0$:

$$\mathbb{P}(D_m \ge l) \le \mathbb{P}(D'_m \ge l)$$

Because $m^{-3/2}m\lceil\|\tilde{X}^m - \tilde{H}^m/2\|_\infty\rceil \le m^{-1/2}\|\tilde{X}^m - \tilde{H}^m/2\|_\infty + m^{-1/2} \to 0$, we have that

$$D'_m \xrightarrow{\mathbb{P}} 0.$$

This can be improved to $D'_m \xrightarrow{a.s.} 0$ since D'_m is integer valued. Furthermore:

$$\lim_{m\to\infty}\mathbb{P}(D_m \ne 0) \le \lim_{m\to\infty}\mathbb{P}(D'_m \ne 0) = 0$$

In other words: the probability that the prerequisite of Lemma 3.6 is fulfilled tends to 1 as $m\to\infty$. By Lemma 3.6 we have:

$$d_{\mathrm{GH}}(G^X(\tilde{T}^p_m, \mathcal{Q}_p), G^H(\tilde{T}^p_m, \mathcal{Q}_p)) \le |\mathcal{Q}^p \cap \tilde{X}^m| \cdot \left(\|\tilde{X}^m - \tilde{H}^m/2\| + 2\right)$$

Thus, we can estimate for any $\epsilon > 0$:

$$\mathbb{P}\left(d_{\mathrm{GH}}(G^X(\tilde{T}^p_m, \mathcal{Q}_p), G^H(\tilde{T}^p_m, \mathcal{Q}_p)) \ge \epsilon m^{1/2}\right) \le \mathbb{P}\left(|\mathcal{Q}^p \cap \tilde{X}^m| \cdot \left(\|\tilde{X}^m - \tilde{H}^m/2\| + 2\right) \ge \epsilon m^{1/2}\right) + \mathbb{P}\left(\mathcal{Q}^p \cap \tilde{X}^m \ne \mathcal{Q}^p \cap (\tilde{H}^m/2)\right)$$

We split the first event on the righthand side into two parts:

$$\left\{|\mathcal{Q}^p \cap \tilde{X}^m| \cdot \left(\|\tilde{X}^m - \tilde{H}^m/2\| + 2\right) \ge \epsilon m^{1/2}\right\} \subset \left\{|\mathcal{Q}^p \cap \tilde{X}^m| \ge m^{1/8}\right\} \cup \left\{\|\tilde{X}^m - \tilde{H}^m/2\| \ge \epsilon m^{3/8} - 2\right\}$$

Now we can estimate their probabilities separately:

$$\mathbb{P}\left(|\mathcal{Q}^p \cap \tilde{X}^m| \ge m^{1/8}\right) \le m^{-1/8}\mathbb{E}[|\mathcal{Q}^p \cap \tilde{X}^m|]$$

by Markov's inequality. And by [1, Lemma 16]:

$$\mathbb{P}\left(\|\tilde{X}^m - \tilde{H}^m/2\| \ge \epsilon m^{3/8} - 2\right) \le 2K\epsilon^{-1/6}m^{-1/16}$$

for a universal constant K and m big enough. Furthermore, this term tends to 0 when $m\to\infty$. Observe:

$$\mathbb{E}\left[|\mathcal{Q}^p \cap \tilde{X}^m|\right] \longrightarrow \mathbb{E}[|\mathcal{P}\cap\tilde{e}|] = \mathbb{E}\left[\int_0^1 \tilde{e}(s)\,\mathrm{d}s\right] < \infty$$

This implies that $m^{-1/8}\mathbb{E}\left[|\mathcal{Q}^p \cap \tilde{X}^m|\right] \longrightarrow 0$. Combining all of the estimates above, we can conclude:

$$\mathbb{P}\left(d_{\mathrm{GH}}(G^X(\tilde{T}_m^p, \mathcal{Q}_p), G^H(\tilde{T}_m^p, \mathcal{Q}_p)) \geq \epsilon m^{1/2}\right) \longrightarrow 0$$

as $m \to \infty$. Finally:

$$\begin{aligned} d_{\mathrm{GH}}(m^{-1/2}G^X(\tilde{T}_m^p, \mathcal{Q}^p), \mathcal{M}) &\leq m^{-1/2} d_{\mathrm{GH}}((G^X(\tilde{T}_m^p, \mathcal{Q}^p), (G^H(\tilde{T}_m^p, \mathcal{Q}^p)) \\ &\quad + d_{\mathrm{GH}}(m^{-1/2}G^H(\tilde{T}_m^p, \mathcal{Q}^p), \mathcal{M}) \\ &\xrightarrow{\mathbb{P}} 0 \end{aligned}$$

where the first term converges to 0 in probability by the reasoning above and the second terms by part 1. The result now follows because $m^{-1/2}G^X(\tilde{T}_m^p, \mathcal{Q}^p) \stackrel{d}{=} m^{-1/2}G_m^p$ and $m^{-1/2}n^{1/3} \to 1$. □

4.5 Diameter of the largest component

While the continuum random graph $\mathcal{M}$ is an interesting object on its own, its most important property may be that it turns up as scaling limit of the Erdős-Rényi graph. The reason for this is that it is very easy to derive asymptotics of the Erdős-Rényi graph from the continuum random graph. For example, consider the diameter of a graph:

$$diam\ G = \sup_{x,y \in [n]} d(x, y)$$

for a connected graph G with n vertices. This notion coincides with the diameter of metric spaces when we view a graph as metric space. Furthermore, we have the estimate:

$$\frac{1}{2}\left|diam\ X_1 - diam\ X_2\right| \leq d_{\mathrm{GH}}(X_1, X_2)$$

for two connected metric spaces X_1 and X_2. Thus the diameter is a continuous function from the space of connected metric spaces equipped with the Gromov-Hausdorff topology to the real numbers. This has an important probabilistic consequence: when $X_n \xrightarrow{d} X$ as metric spaces, we have:

$$diam(X_n) \xrightarrow{d} diam(X)$$

Thus we can apply this to Theorem 4.11:

Corollary 4.14 ([1, Thm. 5])**.** *Consider $G(n, 1/n)$. Denote the largest component by G_n^1. Then there is a non-zero random variable D with finite expectation such that:*

$$n^{-1/3} diam(G_n^1) \xrightarrow{d} D.$$

In other words: the diameter of the largest component of the critical Erdős-Rényi graph grows proportional to $n^{1/3}$.

5 Conclusion and outlook

This thesis started from the observation that the Erdős-Rényi graph, which is the simplest and therefore most important model for a random graph, exhibits a phase transition at $\lambda = 1$, where the graph jumps from being sparsely connected to highly connected. In order to understand the nature of the critical graph in more detail, we were using a global scaling approach.

To do so, we made use of the extremely powerful relation between graphs and certain functions. In this thesis, we discussed three different encoding functions that can be defined starting from a given graph: the contour function, the height function and the depth first walk. The last case uses depth first search, an algorithm which allows us to select a spanning tree of a connected graph in a canonical way. We also proved that the surplus edges of the Erdős-Rényi graph can be tracked by marked points under the graph of the depth first walk.

We then discussed how to construct a graph out of these functions and marked points in two different ways: one of those methods was closely related to the Erdős-Rényi graph and the other method translated well into a similar construction which we used to define the continuum random graph. The continuum random graph is a random metric space and restricted to sufficiently small neighbourhoods the continuum random graph exhibts a tree-like structure in the sense of $\mathbb{R}$-trees. We showed that both the functions and the marked points converge after rescaling from which we concluded that the largest component of the critical Erdős-Rényi graph converges in Gromov-Hausdorff distance to the continuum random graph.

Lastly, we have proven a corollary concering the diameter of the critical Erdős-Rényi graph. There we can clearly see the strength of the approach with scaling limits: we easily derived that the diameter grows proportionately to $n^{1/3}$.

We also want to note that many results that we derived for the largest component like or Theorem 4.11 or Corollary 4.14 still hold for the $i-$th largest component. The reason for this is that an equivalent statement to Theorem 2.5 holds not only for the largest component but also for the i-th largest component. Furthermore, Theorem 2.5 can be generalized in another way: if we consider $p = 1/n + \theta/n^{4/3}$ with $\theta \in \mathbb{R}$ instead of $p = 1/n$, the theorem still holds[4]. In all the other proofs we only needed that $np \to 1$ and therefore all results hold. Thus, [1, Theorem 25] - which is the more general version of Theorem 4.11 in a sense - concerns itself not only with a single component but with the whole sequence of connected components and with the more general case of $p = 1/n + \theta/n^{4/3}$.

[4]These values for p are called the critical window of the Erdős-Rényi graph.

6 References

[1] L. Addario-Berry, N. Broutin, C. Goldschmidt; The continuum limit of critical random graphs; *Prob. Theory Related Fields*, 152, 367-406, 2012.

[2] D. Aldous; Brownian Excursions, Critical Random Graphs and the Multiplicative Coalescent; *Ann. Prob.* 25(2): 812–854,1997.

[3] D. Aldous; The Continuum Random Tree I; *Ann. Prob.*, 19(1):1–28, 1991.

[4] P. Billingsley; *Convergence of Probability Measures*; Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1968.

[5] D. Burago, Y. Burago, S. Ivanov; *A Course in Metric Geometry*; Graduate Studies in Mathematics Volume 33, American Math. Soc., Providence, 1964.

[6] P. Erdős, A. Rényi; On the Evolution of Random Graphs; *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[7] R. Hofstad; *Random Graphs and Complex Networks I*; Cambridge University Press, Cambridge Series in Statistical and Probabilistic Mathematics Volume 43, Cambridge University Press, Cambridge, 2017.

[8] A. Klenke; *Wahrscheinlichkeitstheorie*; Springer Verlag, Berlin Heidelberg, 2006.

[9] J. Le Gall, G. Miermont; *Scaling Limits of Random Trees and Planar Maps*; Clay Mathematics Proceedings Volume 15, Clay Math Inst. and American Math. Soc., Providence, 2012.

[10] J.-F. Marckert, A. Mokkadem; The Depth First Processes of Galton Watson Trees Converge to the same Brownian Excursion; *Ann. Prob.*, 31:1655–1678, 2003.

[11] D. Revuz, M. Yor; *Continous Martingales and Brownian Motion*; Springer Verlag, Berlin Heidelberg, 2004.

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vor liegende Bachelorarbeit selbstständig und nur mit den angegebenen Hilfsmitteln verfasst habe. Alle Passagen und Bilder, die ich wörtlich oder sinngemäß aus der Literatur übernommen habe, habe ich deutlich als Zitat mit Angabe der Quelle kenntlich gemacht. Darüber hinaus versichere ich, dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist.

Ort und Datum	Unterschrift