Technische Universität München
Department of Mathematics

Average Percolation

Perkolation von Mittelwerten

Master's Thesis in Cooperation with
The University of California, Berkeley
by Michael Preischl

Supervisor:	Prof. Nina Gantert (TUM)
Advisors:	Prof. Nina Gantert (TUM)
	Prof. David Aldous (UCB)
Submission Date:	12/17/2015

Acknowledgements

In the making of this thesis I experienced a lot of support and helpfulness, at this point I would like to mention the people I feel most indebted to.
I want to thank Professor David Aldous from the UC Berkeley not only for accepting my request to be my supervisor during the time in California in this exchange program but also for all the effort he put into my project and the very open-hearted way he included me in his group of fellow researchers. To him, I owe the topic of this thesis and many valuable pieces of advice.
In the same way, I want to thank my supervisor here at the TUM, Professor Nina Gantert for all the time and work she invested to answer all my questions, make helpful suggestions and explain every concept I struggled with. Her support and also her lectures, which I found important sources of knowledge in many chapters formed the final state of this thesis.
I would also like to thank the international coordinators of our faculty, especially Carola Jumpertz, for making this exchange program happen. Her professional counselling helped me gain experiences to cherish for a lifetime.
Last but not least I want to express my heartfelt gratitude towards my family who always supported me in any situation thinkable.

I hereby declare that this thesis is my own work and that no other sources have been used except those clearly indicated and referenced.

Munich, December 17th

Michael Preischl

Contents

1 Introduction

Combining randomness with the spatial structure of graphs has been one of the most interesting and fruitful branches of probability theory in the past decades. Whether we think of the classical Erdős Rényi random graph, electrical networks or something like the frog model, this combination usually results in a very rich mathematical theory. Furthermore, most probabilistic models on graphs provide relevant real-world applications. One of the most fascinating and thus also one of the most well-studied of those models is the so called *percolation*. It was first introduced by Broadbent and Hammersley [6] in 1957 and has spawned a series of very interesting questions, many of which are still open today. In this thesis, we want to focus on *percolation of averages* which is quite similar to percolation but in some aspects also fundamentally different. This model is due to Aldous, who first studied it in an article in 1998 [1]. Our aim was to give a comprehensive overview of the research that has already been done on the topic as well as applying the model in slightly different variations to different underlying graphs. In Chapter 3 we want to present Aldous' paper [1] and reconstruct his results. Chapter 2 considers percolation of averages on tree models. More precisely, a phase transition for a critical parameter value is shown and bounds on this critical threshold are stated. This particular case has not been studied yet and although some methods are based on Aldous' work, most of the results in Chapter 2.1 have not been known before. Chapter 2.2 wraps up Aldous' and Bandyopadhyay's survey about recursive distributional equations [3] whereas Chapter 2.4 shows the connection between percolation of averages and Branching Random Walks. Finally in Chapter 4 the concept of *scaling behaviour* as well as the notion of *universality* from percolation theory are introduced. We then cite a paper about the scaling behaviour of percolation of averages by Aldous [2] and also present some very recent research on this topic by Ding [7] respectively Ding and Goswami [8].

In the following, we want to give a short overview of the idea behind percolation as well as some fundamental concepts. Concerning the huge amount of research that has been done on percolation, this introduction is only able to give a very rudimentary glimpse of the topic. For more details, we refer e.g. to Grimmett's book [12].

1.1 Classical Percolation Theory

We start off with an arbitrary graph $\mathbb{G} = (V, E)$, which consists of a vertex set V and edges $E \subset V \times V$. Next, a probability $p \in [0, 1]$ is fixed and every edge $e \in E$ is independently of all other edges defined to be "open" with probability p and hence "closed" with probability $1 - p$. Now roughly speaking, classical percolation theory is concerned with investigating the formation and size of *open clusters* that is, connected components consisting of open edges. Usually, the underlying graph is of infinite or potentially infinite size like the $\mathbb{Z}^2$ lattice or a Galton Watson tree. Then, a naturally interesting question would be, whether there exists an open cluster of infinite size. In particular, we usually fix one vertex x_0 in the graph which we call the *origin* and then define $C(x_0)$ to be the number of edges in the biggest open cluster containing the origin. Next define the *percolation function* $p \mapsto \theta(p)$ by

$$\theta(p) := \mathbb{P}_p(C(x_0) = \infty). \tag{1.1}$$

For convenience, in percolation theory, the underlying graph is in most cases supposed to be *transitive* which means that for every $x, y \in V$ there is a bijection $\pi : V \to V$ with $\pi(x) = y$ and

$$(\pi(u), \pi(v)) \in E \Leftrightarrow (u, v) \in E \qquad \forall u, v \in V.$$

In transitive graphs, we often write C instead of $C(x_0)$ since the percolation function is then independent of the choice of x_0.

Trivially, it holds that $\theta(0) = 0$ and $\theta(1) = 1$, hence we define the *critical value*

$$p_c := \inf\{p : \theta(p) > 0\} \tag{1.2}$$

with $0 \leq p_c \leq 1$. So intuitively speaking, p_c is the smallest value we can choose as the probability to have an open edge such that we still see an infinite open cluster. This abrupt change of behaviour in the sense of

- $p < p_c \Rightarrow$ There is almost surely no infinite open cluster.
- $p > p_c \Rightarrow$ There is an infinite open cluster with probability 1.

is called a *phase transition*. Note that by Kolmogorov's 0-1 law, the probability to have *any* infinite open cluster is 0 or 1, so if the probability to have an infinite open cluster around the origin is positive, we know that there existst almost surely some infinite open cluster.

Even though the percolation function and the critical value seem to be very natural concepts, they turn out to be rather hard to analyze rigorously. In fact, the critical value is only known for some special cases of underlying graphs (including the $\mathbb{Z}^2$ lattice and the Galton Watson tree) and also the exact behaviour of $\theta(p)$ is unknown in most

cases. Especially the question whether θ is continuous at the critical value and the exact scaling behaviour of $\theta(p)$ for $p \downarrow p_c$ are major points of interest in current research.

Percolation theory is however not just a theoretical construct inside mathematics but has some very useful applications. For example, in statistical physics the connection between the Ising model and percolation on the $\mathbb{Z}^2$ lattice is used to handle occuring phase transitions (see e.g. [16]).
Another field of application includes oil drilling. If we interpret the vertices as holes in a rock formation and the edges as intermediate rock respectively soil of differing densitiy, we could assume that, whenever the density of the rock between two holes lies below a certain value, then oil can flow from one hole to the other. Or to translate it in the language of percolation theory: given a graph with independent weights on the edges, we call an edge open, whenever its weight does not exceed some fixed value.

1.2 Percolation of Averages

Instead of the aforementioned classical percolation, we want to study a concept which we call *percolation of averages* or *average percolation*. We start again with some graph $\mathbb{G}$, consisting of vertices V and edges $E \subset V \times V$ but this time we also need a weight function

$$\begin{aligned} \omega : E &\to \mathbb{R}^+ \\ e &\mapsto \omega_e \end{aligned}$$

which assigns a non-negative number to every edge. In particular, we are interested in the case when ω is not a deterministic function but assigns independent and identically distributed (iid) random weights to the edges. If we denote the edge weight distribution by μ, then for any fixed value c, the probability that an arbitrary edge e satisfies $\omega_e \leq c$ is $\mu([0, c])$. We already noted that we get the classical percolation model with 'success' probability $p = \mu([0, c])$ by defining

$$e \text{ open } \Leftrightarrow \omega_e \leq c.$$

So open clusters s around some vertex $x_0 \in V$ are connected subgraphs $s = (V', E')$ of $\mathbb{G}$ with vertex set V' and edge set E' that fulfill

$$x_0 \in s \quad \text{and} \quad \max_{e \in E'} \{\omega_e\} \leq c. \tag{1.3}$$

Now to get to percolation of averages, we just replace the maximum in (1.3) by the mean. Thus an "open" cluster s around x_0 is a connected subgraph $s = (V', E')$ of $\mathbb{G}$ again with vertex set V' and edge set E' that fulfills

$$x_0 \in s \quad \frac{1}{|E'|} \sum_{e \in E'} \omega_e \leq c \tag{1.4}$$

Note that in contrast to classical percolation theory, these open clusters around x_0 with respect to average percolation are no longer uniquely defined.

Similarly as in the classical percolation theory we want to study the maximal size of those open clusters. Just as before, the most interesting question seems to be the existence of infinite open clusters and we want to see whether we can define something analogous to the percolation function in (1.1) and the critical value in (1.2).

Thinking about applications of this model, it might be sensible to look at finite graphs. Then, instead of fixing a certain value c and asking

I) "what is the biggest open cluster with mean weight below c"

we could turn the question around and ask

II) "given a graph and a number m, what is the smallest value c such that we can find a connected subgraph on at least m edges with mean weight not bigger than c."

If we apply question II) to the complete graph on n vertices and set $m = n - 1$, we are dealing with the well-known task of finding minimal spanning trees (MST). So our percolation of averages can not just be seen as a variant of classical percolation but also as some sort of generalization of the MST problem.
In the case that the edge weights follow an exponential distribution with mean n, Frieze [11] showed that the smallest value c from question II) applied to the complete graph converges to $\zeta(3) = \sum_{i=1}^{\infty} i^{-3}$ as $n \to \infty$.
Another interesting question connected with the complete graph on n vertices is the growth rate of the biggest open cluster in average percolation, with fixed threshold c, if we send $n \to \infty$. This is exactly what Aldous investigated in [1] and we will look at his results in greater detail later.

2 Minimal Average Percolation in Tree Models

The reason why we want to put our focus on tree models first is twofold. On the one hand, tree structures are always a natural first choice when examining graph related problems as their spatial structure usually allows for recursions. Indeed, in this case the recursive approach is so effective that, when looking at different graphs later on, we will try to reduce them to the tree case and then apply similar techniques as in this chapter. The second reason is the generality of statements concerning tree models. We can look at regular trees, general trees or even random graphs like the Galton Watson tree and, for most parts, just use the same methods over and over again.

2.1 The d-Regular Tree

Definition 2.1. Let $d \geq 3$ and consider an infinite d-regular tree. That means, a connected graph $t_d = (V, E)$ without cycles, where every vertex has degree d. Furthermore, assign a weight ω_e to each edge $e \in E$. Here the ω_e are independent and follow an exponential distribution with mean d. Denote the resulting weighted graph by T_d. In T_d choose one designated vertex to be "the root" and denote it with x_0. We will adress vertices with a distance of 1 edge to x_0 as the "first generation", vertices with a distance of 2 edges to x_0 as the "second generation" and so on.

Figure 2.1 shows an example for $d = 3$ with a canonical choice of labelling for the vertices. We are now interested in finding a subgraph of T_d which contains the root (we will sometimes call this a "rooted subgraph") and whose average weight is below a certain value c, i.e. we are looking for $t \subset T_d$ with $x_0 \in t$ and

$$\frac{1}{|t|} \sum_{e \in t} \omega_e \leq c. \tag{2.1}$$

Note that expressions like $e \in t$ and $x_0 \in t$ can cause confusion since e denotes an edge, whereas x_0 is a vertex. However we believe that by context and notation (e for edges, x for vertices) it should always be clear what kind of object is meant. By $|t|$ we will

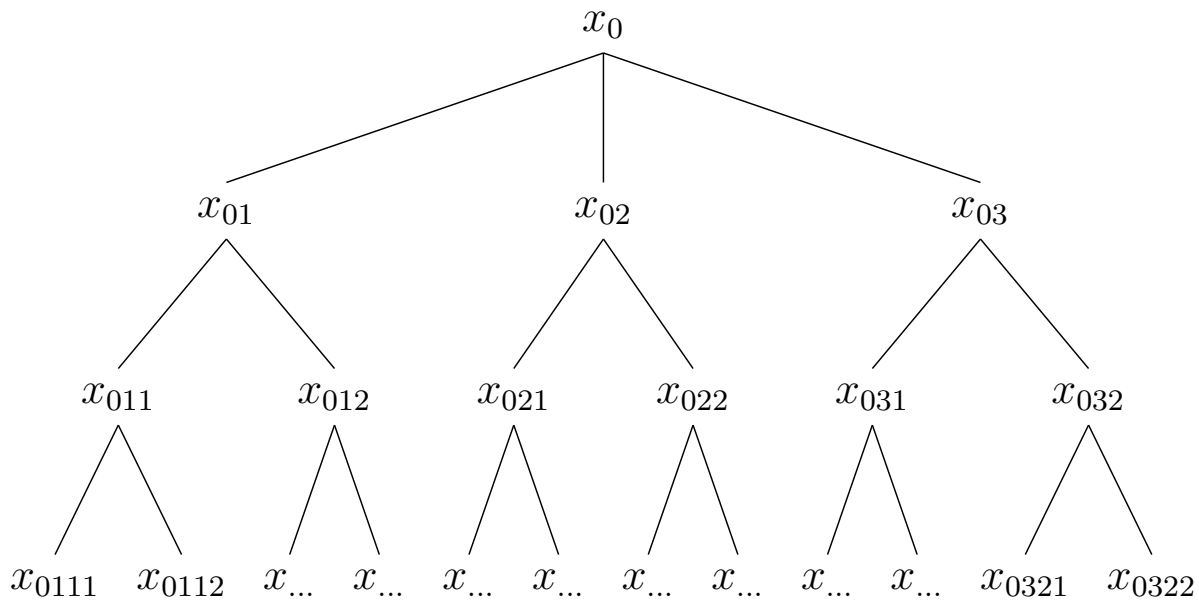


Figure 2.1: The root and the first three generations of a 3-regular tree.

always denote the number of edges of t.

It is now natural to ask for the maximal size of such a subtree. In particular, we want to study the question whether there is an infinite rooted subtree t with mean weight not greater than c. Of course, the answer to this will depend on c. It seems intuitively clear, that for $c = 0$ there will almost surely be no tree at all that fulfills (2.1) (except for the trivial one that consists only of x_0) whereas for $c \to \infty$ it should be relatively easy to find arbitrarily large subtrees that fulfill (2.1). This hints to the existence of a critical value for c, where an infinite component appears. In the following, we want to provide the rigorous arguments that will indeed confirm this conjecture and also provide some more details about the critical value. In order to tackle the aforementioned questions, we will look at a quantity that is somewhat more convenient when passing to the limit $|t| \to \infty$ in (2.1). Let

$$A_n := \min\left\{\frac{1}{n}\sum_{e\in t}\omega_e : t \text{ is a connected subgraph of } T_d,\ |t| = n,\ x_0 \in t\right\}. \tag{2.2}$$

Note that the set of all subtrees with size n that contain the root is finite therefore the definition makes sense. A_n is a random variable, however, we will show that the limit $\lim_{n\to\infty} A_n$ exists and is not random. Since we are always working with connected subgraphs in this chapter, we want to write $t \subset T_d$ for the expression "t is a connected subgraph of T_d".

2.1.1 Existence of a Critical Value

Let $t_{max} := \sup\{|t| : t \subset T_d, x_0 \in t, \frac{1}{|t|}\sum_{e\in t}\omega_e \leq c\}$ As we will see, the probability of the event $\{t_{max} = \infty\}$ depends on the values of c and d. In particular, we want to prove that for every d there is a critical value $c(d)$ such that, almost surely

$$\begin{aligned} c > c(d) \quad &\Rightarrow \quad t_{max} = \infty \\ c < c(d) \quad &\Rightarrow \quad t_{max} < \infty. \end{aligned}$$

To achieve this, we use a *recursive distributional equation (RDE)*, a trick that will be used several times throughout the thesis.

Definition 2.2. Write $\mathcal{M}_1$ for the set of probability measures on $[0,\infty]$ and define $\Gamma_c^* : \mathcal{M}_1 \to \mathcal{M}_1$ by saying that $\Gamma_c^*(\mu)$ is the distribution of $\sum_{i=1}^{d-1}(c - \nu_i + Y_i)^+$, where the Y_i are iid with distribution μ and independent of the ν_i which are iid with an exponential (mean d) distribution. Now we recursively define $\Gamma_c^k(\mu)$ to be the distribution of $\sum_{i=1}^{d}(c - \eta_i + Z_i)^+$ where the Z_i are iid with distribution $(\Gamma_c^*)^{k-1}(\mu)$ and the η_i are again independent and also independent of the Z_i and exponentially distributed with mean d.

The following lemma gives us the existence of a fixed point $c(d)$ of Γ_c that will turn out to be the critical value we are looking for and it is also the limit of A_n for $n \to \infty$. Furthermore, this characterization of $c(d)$ will be useful for numerical simulations later on.

Lemma 2.3. Write δ_x for the Dirac measure at x. There exists a critical value $0 \leq c(d) \leq d$ with the following properties

1. For $c < c(d)$, the map Γ_c has a fixed point μ_c such that $\mu_c([0,\infty)) = 1$ and $\Gamma_c^k(\delta_0) \to \mu_c$ as $k \to \infty$.

2. For $c > c(d)$, the map Γ_c has no fixed point except δ_∞ and $\Gamma_c^k(\delta_0) \to \delta_\infty$ as $k \to \infty$.

Proof. The proof follows from monotonicity arguments. We write $\preceq$ for the 'stochastically less than' partial order on $\mathcal{M}_1$, i.e.

$$\mu_1 \preceq \mu_2 \qquad \Longleftrightarrow \qquad \mu_1[0,x] \geq \mu_2[0,x] \quad \forall 0 \ \leq x \leq \infty.$$

Obviously by definition, Γ_c^* and therefore also Γ_c are increasing maps for any $c > 0$:

$$\mu_1 \preceq \mu_2 \Rightarrow \Gamma_c(\mu_1) \preceq \Gamma_c(\mu_2).$$

By induction, we get that $\Gamma_c^k(\delta_0) \preceq \Gamma_c^{k+1}(\delta_0)$. Hence it follows that there is a limit

$$\Gamma_c^k(\delta_0) \uparrow \mu_c \tag{2.3}$$

for $k \to \infty$. It is easy to see that $\Gamma_c(\lim_{k\to\infty} \Gamma_c^k(\delta_0)) = \lim_{k\to\infty} \Gamma_c^{k+1}(\delta_0)$ holds (note that Γ is defined by a sum over finitely many terms so we can pull the limit out). This means that $\mu_c := \lim_{k\to\infty} \Gamma_c^k(\delta_0)$ is a fixed point. Assume $\mu_c(\infty) = p$. Now we have

$$\begin{aligned}(1-p) &= \mu_c([0,\infty)) = \Gamma_c\mu_c([0,\infty)) \\ &= \mathbb{P}\left(\sum_{i=1}^{d}(c-\eta_i+Z_i)^+ < \infty\right) = \mathbb{P}(Z_i < \infty \quad \forall 1 \le i \le d) \\ &= (1-p)^d.\end{aligned}$$

It follows that $p = 0$ or $p = 1$. This result holds for any fixed point of Γ_c.
Furthermore it is clear that

$$c_1 \le c_2 \Rightarrow \Gamma_{c_1}(\mu) \preceq \Gamma_{c_2}(\mu). \tag{2.4}$$

Hence we can define

$$c(d) := \inf\{c : \mu_c(\infty) = 1\}$$

and trivially $c(d) \ge 0$. To show that $c(d) \le d$, fix $c > d$ and consider independent random variables Y and Z where $Y \sim \mu$ and $Z = c - \nu_1$ with $\nu_1 \sim \exp(\frac{1}{d})$. Write $\mathrm{dist}(X)$ for the distribution of a random variable X, then note that $\Gamma_c(\mu) \succeq \mathrm{dist}(Y+Z)$ and inductively it follows that

$$\Gamma_c^k(\delta_0) \succeq \mathrm{dist}(Z_1 + \cdots + Z_k) \tag{2.5}$$

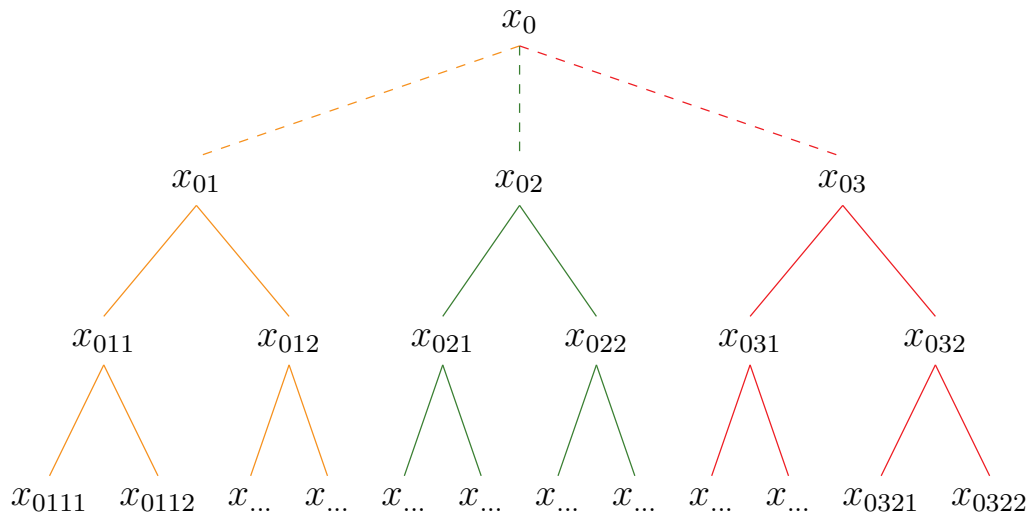
with the Z_i being independent copies of Z. Because of $c > d$, we have $\mathbb{E}[Z] = c - d = a > 0$. With the law of large numbers and (2.5) we conclude $\Gamma_c^k(\delta_0) \overset{k\to\infty}{\longrightarrow} \delta_\infty$. But this means that $c(d) \le d$.

□

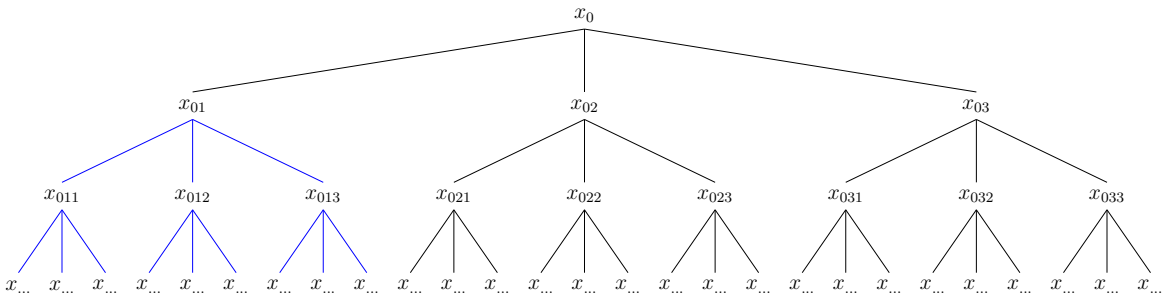
Now we want to establish a connection between the function Γ_c and our tree model. On that account we look at the following definition and lemma.

Definition 2.4. Let T_d^* be a weighted infinite d-ary tree i.e. in contrast to the d-regular tree, every vertex should have d children (which means every vertex except for the root has degree $d+1$) and the edge weights are chosen as iid exponential (mean $d+1$) random variables.

Looking at Figure 2.2a, it is obvious that any weighted d-regular tree T_d as described in Definition 2.1 consists of d independent weighted $(d-1)$-ary trees T_{d-1}^* as described in Definition 2.4 which are all tied together to a common root.
We write $T_{d,m}$ for the d-regular tree up to generation m and, analogously, write $T_{d,m}^*$ for the d-ary tree up to generation m. So $T_{d,m}$, respectively $T_{d,m}^*$ contains all vertices that have a distance of at most m edges to x_0. Note that $T_{d,m}$ and $T_{d,m}^*$ are **finite** trees.



(a) The 3-regular tree T_3. Here the 2-ary trees, starting at x_{01}, x_{02} and x_{03} are drawn in orange, green and red. The edges connecting the individual parts to one common root are the dashed lines.



(b) The root and the first three generations of a 3-ary tree. Note that the blue subtree is itself again a 3-ary tree, starting one generation later. This illustrates the induction argument from the proof of Lemma 2.5.

Figure 2.2: The difference and the connection between d-regular and d-ary trees.

Lemma 2.5. $(\Gamma_c^*)^m(\delta_0)$ is the distribution of $\max\{c|t| - \omega(t) : t \subset T^*_{d-1,m},\, x_0 \in t\}$, where

$$\omega(t) := \sum_{e \in t} \omega_e$$

is the total weight of t.
Consequently, $\Gamma_c^m(\delta_0)$ is the distribution of $\max\{c|t| - \omega(t) : t \subset T_{d,m},\, x_0 \in t\}$.

Proof. Inductively for $(\Gamma_c^*)^m$: The case $m = 0$ is obvious and then, looking at Figure 2.2b, note that the maximal value over height m subtrees can be obtained by considering the maximal value Y_i over height $m-1$ subtrees where each of these trees roots at one of the $d-1$ children of the original root. Each of those $d-1$ subtrees and the corresponding edge to the original root is added if and only if the contribution made by that branch is positive, i.e. if $c - \nu_i + Y_i > 0$. The claim about Γ_c follows from the fact that T_d can be seen as d independent versions of T^*_{d-1}, tied together with independent exponentially (mean d) distributed edges but that is exactly the Definition of Γ_c (See Definition 2.2 for comparison and Figure 2.2a for an illustration). □

Corollary 2.6.

- For $c > c(d)$ we find an infinite, rooted subtree $t \subset T_d$ with average weight less than c with probability 1.
- For $c < c(d)$ we cannot find an infinite, rooted subtree $t \subset T_d$ with average weight less than c with probability 1.

Proof. Let $c > c(d)$ and set
$X := \sup\{c|t| - \omega(t) : t \subset T_d,\, x_0 \in t\}$ and $X_m := \max\{c|t| - \omega(t) : t \subset T_{d,m},\, x_0 \in t\}$. It holds

$$X = \lim_{m\to\infty} X_m = \infty \qquad a.s.$$

since

$$\lim_{m\to\infty} \operatorname{dist} \max\{c|t| - \omega(t) : t \subset T_{d,m},\, x_0 \in t\} = \lim_{m\to\infty} \Gamma_c^m(\delta_0) = \delta_\infty.$$

On the other hand, if $c < c(d)$, we know that

$$X = \lim_{m\to\infty} X_m < \infty \qquad a.s.$$

by the same argument and the claim follows □

A slightly stronger statement, which we already announced in the beginning of this Chapter is given by the following theorem:

Theorem 2.7. *Let $c(d)$ be the critical value from Lemma 2.3, then*

$$A_n \xrightarrow{p} c(d).$$

Proof.
Let $c < c(d)$ and consider a random variable Z_n where

$$Z_n := \max\{c|t| - \omega(t) : t \subset T_{d,n},\, x_0 \in t\}.$$

It clearly holds

$$\begin{aligned} & Z_n \geq c|t| - \omega(t) && \forall t \subset T_{d,n},\, x_0 \in t \\ \iff & \frac{\omega(t)}{|t|} \geq c - \frac{Z_n}{|t|} && \forall t \subset T_{d,n},\, x_0 \in t \\ \Rightarrow & A_n \geq c - \frac{Z_n}{n} \end{aligned}$$

As $c < c(d)$, we can use Lemma 2.3 and Lemma 2.5 to conclude $\lim_{n\to\infty} \mathbb{P}(Z_n \to \infty) = 0$ which then establishes

$$P(A_n \geq c(d) - \varepsilon) \overset{n\to\infty}{\longrightarrow} 1 \qquad \forall \varepsilon > 0. \tag{2.6}$$

Now let $c > c(d)$ and take $n \in \mathbb{N}$ sufficiently large. We shall construct a tree $\widetilde{T}$ with n edges and the desired mean weight.
For $m \in \mathbb{N}$ use again Lemma 2.3 and Lemma 2.5 to choose k such that

$$\Gamma_c^k \delta_0((cm, \infty]) > \frac{1}{2}. \tag{2.7}$$

We use the following algorithm:

1 Set $\widetilde{x}_0 = x_0$ and $\widetilde{T} = \emptyset$. Furthermore let $q := \sum_{i=1}^{k} (d-1)^i$.

2 Starting at $\widetilde{x}_0$, go down one generation and call the first child (at the very left) $\widetilde{x}$.

3 Denote the $d-$regular tree up to generation k with root $\widetilde{x}$ by $T_{d,k}(\widetilde{x})$. Look for a tree $\widetilde{t}$ with at least m edges and average edge weight at most c in $T_{d,k}(\widetilde{x})$. Because of (2.7) the chance to succeed is at least $\frac{1}{2}$. If the search is successful go to 4, otherwise go to 5.

4 Set $\widetilde{T} = \widetilde{T} \cup \widetilde{t}$ and also add the edge $(\widetilde{x}_0, \widetilde{x})$. If $|\widetilde{T}| \geq n - q$ go to step 6, otherwise find the following vertex: starting at $\widetilde{x}$ go down $\widetilde{t}$ by always choosing the most left edge possible. If a dead end is reached, stop and set the last visited vertex to be the new $\widetilde{x}_0$. Now repeat at step 2.

5 Define the new $\widetilde{x}_0$ to be the second child (i.e. the second from left) of the old $\widetilde{x}_0$ and add the edge between the new $\widetilde{x}_0$ and the old $\widetilde{x}_0$ to $\widetilde{T}$. If $|\widetilde{T}| \geq n - q$ go to step 6, otherwise repeat at step 2.

6 Add random edges (that is, edges the algorithm has not looked at yet) until $|\widetilde{T}| = n$.

It remains to show that the tree $\widetilde{T}$ generated by this algorithm has the desired properties. First of all, we note that the maximal number of edges that can be added in step 4 is $q+1$, so using step 6 as soon as $|\widetilde{T}| \geq n-q$ ensures that we don't have to delete edges to reach $|\widetilde{T}| = n$. In order to reach enough edges to begin with step 6, we need at most $\lceil \frac{n-q}{m} \rceil \leq \frac{n-q}{m} + 1$ successful steps. As the success probability is bigger than $\frac{1}{2}$ we won't need more than $3\frac{n-q}{m} + 3$ attempts in the algorithm. In the end we have a tree of size $n - q - 3\frac{n-q}{m} - 3$ with mean weight less than c, q random edges with mean d and one linking edge for each attempt (successful or not). Note that the algorithm has not seen any of the linking edges, therefore they also have mean d. In total this gives:

$$\begin{aligned}\frac{\omega(\widetilde{T})}{|\widetilde{T}|} &= \frac{c(n-q-3\frac{n-q}{m}-3) + qd + (3\frac{n-q}{m}+3)d}{n} \\ &= c + \frac{q(d-c)}{n} + \frac{3(d-c)}{m} - \frac{3q(d-c)}{mn} + \frac{3(d-c)}{n}.\end{aligned}$$

Now choose m big and we get

$$P(A_n \leq c(d) + \varepsilon) \overset{n\to\infty}{\longrightarrow} 1 \qquad \forall \varepsilon > 0. \tag{2.8}$$

Combining (2.6) and (2.8) gives $A_n \xrightarrow{p} c(d)$.

□

Having proved that A_n converges in probability, the question arises whether we can say something about other modes of convergence. For this purpose, note that the sequence (A_n) is uniformly integrable, that is

$$\lim_{k\to\infty} \left(\sup_n \int_{\{|A_n| \geq k} |A_n| d\mathbb{P} \right).$$

To show this, we use the following lemma without proof

Lemma 2.8. Assume that $g : \mathbb{R}_+ \to R_+$ is increasing with $\lim_{x\to\infty} \frac{g(x)}{x} = \infty$ and that $\sup_n \mathbb{E}[g(|X_n|)] < \infty$.
Then (X_n) is uniformly integrable.

Now let $X_1, X_2, \ldots$ be iid exponentially distributed random variables with mean d and write $S_n := \sum_{i=1}^n X_i$.
Choose $g(x) := x^2$, it holds

$$\sup_n \mathbb{E}[g(|A_n|)] \leq \sup_n \mathbb{E}\left[g\left(\frac{S_n}{n}\right)\right] = \sup_n \mathbb{E}\left[\frac{S_n^2}{n^2}\right]. \tag{2.9}$$

Note that

$$\begin{aligned}\mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{\mathrm{Var}(X_1)}{n} &= \mathbb{E}\left[\frac{S_n^2}{n^2}\right] - \mathbb{E}\left[\frac{S_n}{n}\right]^2 \\ \Rightarrow \sup_n \mathbb{E}\left[\frac{S_n^2}{n^2}\right] &= \sup_n \frac{d^2}{n} + d^2 < \infty.\end{aligned} \tag{2.10}$$

Hence, by Lemma 2.8, we can follow from (2.9) and (2.10) that (A_n) is indeed uniformly integrable.

Getting back to the modes of convergence, we look at the following very well-known result

Theorem 2.9. *If a sequence of random variables (X_n) is uniformly integrable and converges in probability to a random variable X, then we have $X_n \to X$ in $\mathcal{L}^1$.*

So since we already proved that A_n converges in probability and also showed that the sequence in uniformly integrable, we can immediately deduce that A_n converges in $\mathcal{L}^1$. A very similar calculation involving $M_{\frac{S_n}{n}-c(d)}$ the moment generating function of $\frac{S_n}{n} - c(d)$ actually shows that $A_n \to c(d)$ in $\mathcal{L}^p$ for any $p > 1$.

Naturally, the next thing to check would be whether A_n also converges almost surely. However, it turns out that this question is quite hard to answer. As a first attempt, one can look at a slightly easier object. Define

$$\widetilde{A}_n := \min\left\{ \frac{1}{|t|} \sum_{e \in t} \omega_e : t \subset T_d,\ |t| \leq n,\ x_0 \in t \right\}$$

and let $\mathcal{F}_n$ be the sigma algebra generated by the edge weights of the first n generations. It is fairly easy to see that $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is a filtration and that

- $\widetilde{A}_n$ is $\mathcal{F}_n$-measurable for all n,
- $\mathbb{E}[|\widetilde{A}_n|] < \infty$ for all n,
- $\mathbb{E}[\widetilde{A}_{n+1}|\mathcal{F}_n] \leq \widetilde{A}_n$.

Hence, $\widetilde{A}_n$ is a supermartingale with respect to the filtration $\mathcal{F}_n$. Recall Doob's forward convergence theorem, which states

Theorem 2.10. *Suppose X_n is a super martingale which satisfies*

$$\sup_n \mathbb{E}[|X_n|] < \infty$$

Then, almost surely $\lim_{n\to\infty} X_n =: X_\infty$ exists and $\mathbb{E}[|X_\infty|] < \infty$.

This is a classical result in martingale theory. For a proof see e.g. [9].
Since $\widetilde{A}_1$ is integrable and $\mathbb{E}[|\widetilde{A}_n|] \leq \mathbb{E}[|\widetilde{A}_m|]$ for $n > m$, it is clear that $\sup_n \mathbb{E}[|\widetilde{A}_n|] < \infty$ holds, hence we know that the sequence $\widetilde{A}_n$ converges almost surely and since it is dominated by $\widetilde{A}_1$ we know it also converges in $\mathcal{L}^p$ for any p.

Unfortunately $\widetilde{A}_n$ does not really give us relevant information about the minimal average edge weight of infinite subtrees in T_d, which is what we are actually interested in and what we described by A_n. Furthermore, we cannot hope to be able to apply the same technique to A_n since A_n is not a (super)martingale.

Another possibility to show almost sure convergence of the sequence A_n would be to look directly at the distribution of A_n and prove that the tails decay sufficiently fast for $n \to \infty$. To do this, define $a_n := nA_n$ or equivalently $\frac{a_n}{n} = A_n$. Now we want to show that for all $\varepsilon > 0$

$$\mathbb{P}\left(\frac{a_n}{n} \geq c(d) + \varepsilon\right)$$

decays exponentially in n. Then, the summation over all n would still be finite, hence we could use the Borel Cantelli lemma to deduce $\lim_{n\to\infty} A_n \leq c(d)$ almost surely. A natural first attempt to achieve this would be using Markov's inequality which yields

$$\mathbb{P}\left(\frac{a_n}{n} \geq c(d) + \varepsilon\right) \leq \mathbb{E}\left[e^{\lambda a_n}\right] e^{-\lambda n(c(d)+\varepsilon)}$$

for any $\lambda > 0$. But unfortunately it is a priori unclear that there is a $\lambda > 0$ such that $\mathbb{E}\left[e^{\lambda a_n}\right] e^{-\lambda n(c(d)+\varepsilon)}$ decays exponentially. The problem here is that the distribution of a_n is quite complicated and hence calculating the exponential moments rigorously seems to be out of reach. It might be possible to find an upper bound on $\frac{a_n}{n}$ that is easier to handle. Obviuosly, for $X_1, X_2, \ldots$ iid exponentially distributed the sum $S_n := \sum_{i=1}^{n} X_i$ would satisfy $\frac{a_n}{n} \leq \frac{S_n}{n}$, however this bound is too coarse.

In conclusion, it can be said that the question for almost sure convergence of A_n is not trivial and might be worth to invest further research into. In this thesis however, we rather tried to gain more information about the critical value $c(d)$ itself, so the next two sections are dedicated to deriving upper and lower bounds on $c(d)$.

2.1.2 An Upper Bound on $c(d)$

We can get an obvious upper bound by considering the following greedy algorithm: Starting from the origin, always choose the edge with the smallest weight.
So in every step we have $d-1$ edges to choose from which means the edge weight on the resulting graph has the distribution of $\min_{i\in[d-1]}(\nu_i)$ where $\nu_i \sim \exp(\frac{1}{d})$ and the ν_i are independent. Recall that the Minimum of an $\exp(a)$ random variable and an independent $\exp(b)$ random variable is again exponentially distributed with parameter $a+b$. So in our case, we know that the edge weight on our subgraph follows an $\exp((d-1)\cdot\frac{1}{d})$ distribution and therefore has mean $\frac{d}{d-1}$ which gives us the bound $c(d) \leq \frac{d}{d-1}$.

We will try to find a better bound by using some percolation theory. First note that

$$P(Y \leq a) = \int_0^a \frac{1}{d} e^{-\frac{1}{d}x} dx = 1 - e^{-\frac{a}{d}}.$$

So if we define in our tree all edges with a weight smaller thanor equal to a to be open, we are in the case of bond percolation on a tree with "open' probability $1 - e^{-\frac{a}{d}}$.

Recall that the critical probability for bond percolation on a $d-$regular tree equals $\frac{1}{d-1}$, so the probability to have an infinite open cluster starting from the origin is positive if and only if

$$1 - e^{-\frac{a}{d}} > \frac{1}{d-1} \iff a > -d \log \frac{d-2}{d-1} \tag{2.11}$$

so choose $a^* := -d \log \frac{d-2}{d-1}$.

In this infinite open cluster, all edges have weight less than or equal to $-d \log \frac{d-2}{d-1}$ so obviously also the average edge weight of this cluster is less than or equal to $-d \log \frac{d-2}{d-1}$. Therefore a^* is an upper bound for $c(d)$. Note that having a positive probability for this rooted open cluster is sufficient since we know by our previous results, that the existence of an infinite, rooted subtree with average weight below a certain value is a 0-1 event. To answer the question if this bound is useful, we look at the upper bound $c(d) < \frac{d}{d-1}$ that we already have by the greedy algorithm. A short calculation shows that $-d \log \frac{d-2}{d-1} \leq \frac{d}{d-1}$. So the new bound is indeed better than the previous one.

We can also investigate the behaviour for $d \to \infty$:

$$\lim_{d \to \infty} -d \log \frac{d-2}{d-1} = \ldots = 1 \tag{2.12}$$

However, we still want to go a little further and try to improve our bound. We found a cluster where *every* edge has a weight below a^*, but in fact we only demand the average of the weights to be below a certain value. To make use of this we introduce the following construction:

Given $a > 0$, set $q := P(Y < a)$ where $Y \sim \exp(\frac{1}{d})$.
Step 1: Construct a bond percolation on the $d-$regular tree with $P(\text{"open"}) = q$.
Step 2: Assign independent weights ω_e to the edges according to the following rule:

- If e is open, then ω_e follows an $\exp(\frac{1}{d})$ distribution conditioned on $\omega_e < a$.
- If e is closed, then ω_e follows an $\exp(\frac{1}{d})$ distribution conditioned on $\omega_e > a$.

This construction assures us to have an open cluster with independent edge weights and it is obvious that we thus constructed T_d with exponential (mean d) edge weights again. In the case that $a \geq -d \log \frac{d-2}{d-1}$, we have $q \geq \frac{1}{d-1}$ and therefore with positive probability an infinite open cluster that contains x_0. We can now calculate the average weight of

this open cluster:

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{S_n}{n}\middle| X_i \leq -d\log\frac{d-2}{d-1} \quad \forall i \in [n]\right] = \int_0^{a^*} x\frac{1}{d}e^{-\frac{1}{d}x}\cdot(d-1)dx$$
$$= \ldots = d(d-2)\left(\log\left(\frac{d-2}{d-1}\right)-1\right)+d^2-d$$

So we get $c(d) \leq d(d-2)\left(\log\left(\frac{d-2}{d-1}\right)-1\right)+d^2-d$. Which is again better than our previous bounds and has the limit $\frac{1}{2}$ for $d\to\infty$.

2.1.3 A Lower Bound on $c(d)$

We will try to find a lower bound for the critical value $c(d)$ by estimating
$\mathbb{E}[|\{t : |t| = n, w(t) < cn\}|]$ where t is a subtree of the d-regular tree which contains the root.
First, we need an estimate for the number of rooted subtrees in T_d that have exactly n edges.

Let a_n denote the number of rooted subtrees of a d-ary tree that contain exactly n edges. Furthermore let b_n be the number of rooted subtrees of a $(d+1)$-regular tree that contain exactly n edges.

Using generating functions, it can be shown that

$$a_n = \frac{1}{n+1}\binom{d(n+1)}{n},$$

see e.g. [21] for a proof of this result. As we consider only d-regular trees for $d \geq 3$, it suffices to look at d-ary trees with $d \geq 2$. Hence, using Stirling's formula to approximate the factorials we get

$$\begin{aligned} a_n &= \frac{1}{n+1}\binom{d(n+1)}{n} \overset{d\geq 2}{<} \frac{1}{n+1}\binom{d(n+1)}{n+1} \\ &\approx \underbrace{\frac{\sqrt{d}}{(n+1)\sqrt{2\pi(d-1)(n+1)}}}_{\leq 1}\left(d\left(1+\frac{1}{d-1}\right)^{d-1}\right)^{n+1} \\ &\leq (de)^{n+1}. \end{aligned} \tag{2.13}$$

Because our basic model is a regular tree, we are interested in b_n. It is not hard to see that

$$b_n = a_n + a_0a_{n-1} + a_1a_{n-2} + \cdots + a_{n-2}a_1 + a_{n-1}a_0,$$

so with (2.13) we get

$$b_n \leq (de)^{n+1} + n(de)^n. \tag{2.14}$$

Now we want to estimate the probability, that a given tree of size n has average edge weight smaller than c. For this we use some large deviations theory, especially *Cramér's Theorem* which says, given a sequence Y_i of iid random variables with law μ and $\mathbb{E}[Y_i] = k$, and some $c < k$, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \leq c\right) \leq -\inf_{x\in(-\infty,c]} \Lambda^*_\mu(x).$$

Here, S_n denotes $\sum_{i=1}^n Y_i$ and Λ^*_μ is the Legendre transformation of μ, i.e.

$$\Lambda^*_\mu(x) := \sup_{\lambda\in\mathbb{R}} \left(\lambda x - \log\left(\int e^{\lambda x} d\mu\right)\right)$$

whenever $\int e^{\lambda x} d\mu$ exists and Λ^*_μ is set to ∞ otherwise. So in our case of exponentially distributed edge weights with mean $d+1$, we get

$$\int e^{\lambda x} d\mu = \int_0^\infty \frac{1}{d+1} e^{-\frac{1}{d+1}x} e^{\lambda x} dx = \frac{1}{1-\lambda(d+1)}.$$

Note that we have to assume $\lambda < \frac{1}{d+1}$ for this integral to be finite.
Now by calculating $\frac{d}{d\lambda}\Lambda^*(x)$ and looking for zero points of this derivative, we find that $\lambda = \frac{1}{d+1} - \frac{1}{x}$ is the only solution. After having checked that this is indeed a maximum, we get

$$\Lambda^*_\mu(x) = \left(\frac{1}{d+1} - \frac{1}{x}\right) x - \log\left(\frac{1}{1-\left(\frac{1}{d+1} - \frac{1}{x}\right)(d+1)}\right) = \frac{x}{d+1} + \log\frac{d+1}{x} - 1.$$

We conclude

$$\mathbb{P}\left(\frac{S_n}{n} \leq c\right) \leq e^{-n\left(\frac{c}{d+1} + \log\frac{d+1}{c} - 1\right)}. \tag{2.15}$$

Putting together (2.14) and (2.15), we find

$$\begin{aligned}\mathbb{E}[|\{t \subset T_{d+1} : |t| = n, w(t) < cn,\ x_0 \in t\}|] &\leq \left((de)^{n+1} + n(de)^n\right) e^{-n\left(\frac{c}{d+1} + \log\frac{d+1}{c} - 1\right)} \\ &= (de+n)\left(\frac{de}{e^{\frac{c}{d+1} + \log\frac{d+1}{c} - 1}}\right)^n.\end{aligned}$$

So it holds

$$\frac{de}{e^{\frac{c}{d+1} + \log\frac{d+1}{c} - 1}} < 1 \qquad \Rightarrow \qquad \mathbb{E}[|\{t \subset T_{d+1} : |t| = n, w(t) < cn,\ x_0 \in t\}|] \to 0$$

which is the case iff $\frac{c}{d+1} - \log(c) + \log\frac{d+1}{d} - 2 > 0$. But note that $c < d+1$, so we are looking for the smaller zero point of $\frac{c}{d+1} - \log(c) + \log\frac{d+1}{d} - 2$.

In conclusion, this means, $c(d+1)$ is greater than the smaller zero point of $\frac{c}{d+1} - \log(c) + \log\frac{d+1}{d} - 2$ or equivalently, $c(d) \geq c^-(d)$, where $c^-(d)$ is the smaller zero point of $\frac{c}{d} - \log(c) + \log\frac{d}{d-1} - 2$.

Again, we calculate the limit for $d \to \infty$:

$$\lim_{d\to\infty} c^-(d) = e^{-2}. \tag{2.16}$$

The Graphics 2.3 and 2.4 show the function $\frac{c}{d} - \log(c) + \log\frac{d}{d-1} - 2$ for $d = 4$ where we can approximately see $c^-(4) = 0.189$.

Figure 2.3: The function $y = \frac{x}{d} - \log(x) + \log \frac{d}{d-1} - 2$ for $d = 4$. The first zero point is $c^-(4)$.

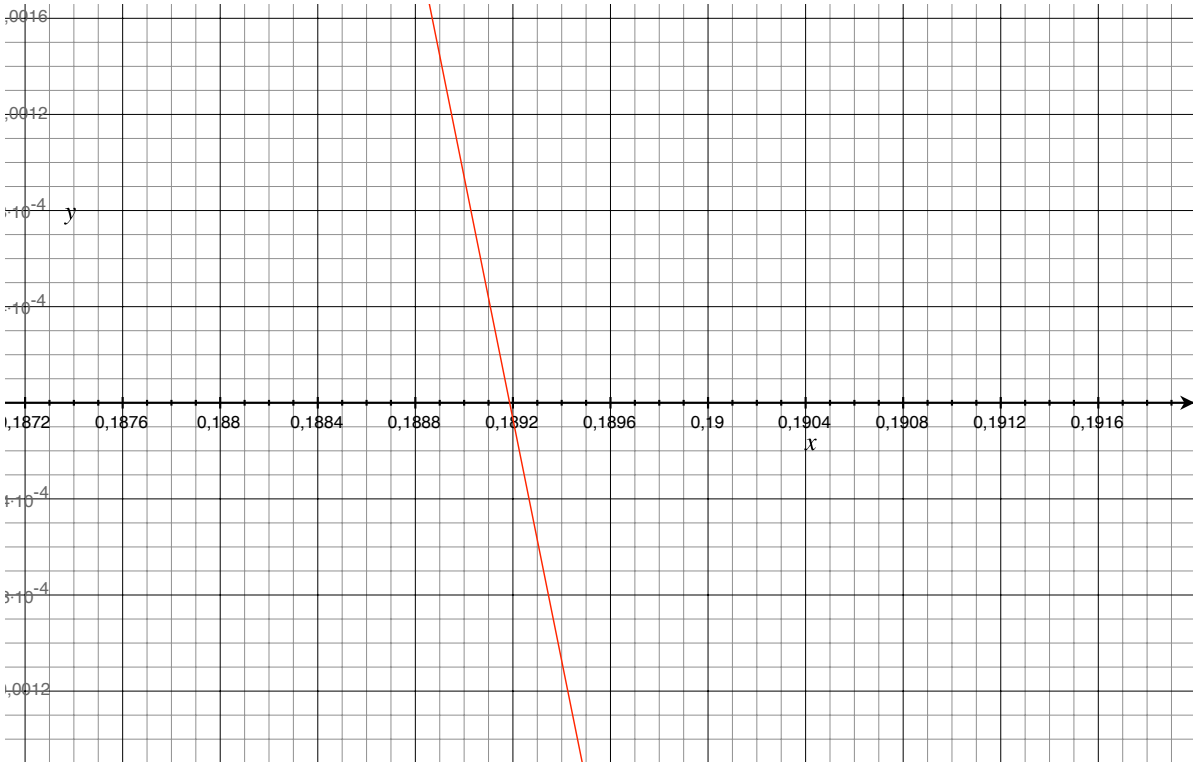


Figure 2.4: The first zero point in detail.

2.1.4 Approximations via Monte Carlo Methods

Whenever the exact value of some constant is hard to find, numerical simulations can be used to find at least a good approximation.

In this particular case, it is actually quite straightforward to think of a Monte Carlo simulation, as all one has to do is simulating the RDE from Definition 2.2. That means, start with a vector of zeros and apply the map Γ_c until the fixed point is (nearly) reached. Since we already proved in Lemma 2.3 that $\Gamma_c^m(\delta_0)$ converges to μ_c for $m \to \infty$, we can be sure to obtain a sample that follows (almost) the distribution μ_c.

While this sounds very easy in theory and is indeed not very hard to implement, it turns out that around 100000 iterations of the map Γ_c are needed, which corresponds to investigating a d-regular tree up to its 100000th generation. Since this is computationally very expensive, we restricted ourselves to the values $d = 4$, $d = 5$ and $d = 6$. The results are shown in Table 2.1 , where we also state the upper and lower bound, as they were calculated in the preceeding chapters.

	lower bound	approximated value	upper bound
$d = 4$	0.189	0.295	0.756
$d = 5$	0.175	0.265	0.685
$d = 6$	0.167	0.249	0.645

Table 2.1: The approximated values of the critical threshold $c(d)$ for $d = 4$, $d = 5$ and $d = 6$ along with the corresponding bounds. These estimates were found by Monte Carlo Simulations of 100000 iterations of $\Gamma_c(\delta_0)$ for varying c. In each step, 500 values were drawn to ensure the empirical distributons are sufficiently close to the actual ones.

2.2 Excursion: Recursive Distributional Equations

At this point we want to say a few words about the method we applied to show the existence of the critical value. We realized that the spatial structure of the tree can be represented by a map Γ_c which was then checked for fixed points and convergence behaviour.

This is a very useful technique and possible applications pop up in quite a lot of different problems, for example in Galton Watson branching processes or statistical physics models on trees. However, somewhat surprisingly, there is relatively few literature on the subject and it wasn't until 2005 that Aldous and Bandyopadhyay coined the expression *recursive distributional equation*. We want to follow their approach from [3] and

introduce the notion more generally. Let $\mathcal{M}(S)$ be the space of probability distributions on some measurable space (S, σ) and let (T, τ) be another measurable space. Set

$$T^* := T \times \bigcup_{0 \leq m \leq \infty} S^m$$

where S^m is the product space and the union is a disjoint union. Given a measurable function $g : T^* \to S$ and a probability measure ν on $T \times \{0, 1, 2, \ldots; \infty\}$ we can make the following definition.

Definition 2.11. $\Gamma(\mu)$ is the distribution of $g(\eta, X_i, 1 \leq i \leq N)$ with

i The $(X_i, 1 \leq i)$ are independent with distribution μ.

ii (η, N) has distribution ν.

iii The families $(X_i, 1 \leq i)$ and (η, N) are independent.

A fixed point equation of the form

$$X \stackrel{d}{=} g(\eta, X_i, 1 \leq i \leq N) \tag{2.17}$$

is called a *recursive distributional equation (RDE).*

In our previous example, $N = d - 1$ was fixed and η was a sequence of length N of independent random variables but note that in particular N does not need to be finite and can even be random, we will indeed encounter both of these cases later.

Notice that the case $N = 1$ corresponds to the well studied task of finding a stationary distribution for a discrete time Markov Chain. Unfortunately for general N the standard techniques from the MC setting don't apply, thus we have to think of other ways to identify the limits (if existent) and to state the basin of attraction.
In many cases, existence can be deduced quite easily from monotonicity, as we did in the previous section. Concerning attractiveness the case isn't quite that obvious but a very useful criterium can be gained via the *contraction method* [3]:

Lemma 2.12. Let M be a subset of $\mathcal{M}_1(S)$ such that Γ maps M into M and let d be some complete metric on M. Further suppose that

$$\sup_{\mu_1 \neq \mu_2 \in M} \frac{d(\Gamma(\mu_1), \Gamma(\mu_2))}{d(\mu_1, \mu_2)} < 1$$

Then Γ has a unique fixed point μ in M whose domain of attraction is all of M.

However, as this result won't be of immediate importance for our thesis, we don't want to state a proof here and also not stress on attractiveness but instead look at another topic that we will give more attention later: the question for scaling behaviour.

Example 2.13. Let $S = \mathbb{R}^+$ and let η be $\mathbb{R}$-valued with $\mathbb{E}[\eta] = \beta$. Fix $c \in \mathbb{R}$ and consider the RDE

$$X \stackrel{d}{=} \max(0, X + \eta - c).$$

There is a solution X_c on $\mathbb{R}^+$ iff $c > \beta$ and it is not hard to show that this solution is given by

$$X_c \stackrel{d}{=} \max_{j \geq 0} \sum_{i=1}^{j} (\eta_i - c)$$

where the η_i are independent copies of η. Now a possible question to ask would be how $\mathbb{E}[X_c]$ behaves for $c \to \beta$. As it turns out, for $\mathrm{Var}(\eta) \in (0, \infty)$ it holds

$$\mathbb{E}[X_c] \sim \frac{\mathrm{Var}(\eta)}{2(c - \beta)} \qquad \text{as } c \downarrow \beta \tag{2.18}$$

which can be shown by the fact that random walks converge weakly to Brownian motions with drift. What we will be aiming to do later is to gain a statement of the form (2.18) for our mean weight model as well.

Although we won't actually *solve* any of the occuring RDEs, they do not only give us an existence statement about the critical value but also provide a very easy and computationally reasonable way of simulating our models as we have seen in chapter 2.1.4.

For a much more detailed look at the subject we refer to [3], a very nice application of RDEs in the context of the frog model can be found in [14].

2.3 The Galton Watson Tree

Instead of using a deterministic d-regular tree as underlying graph structure, we could also work on a random graph and thus add a little more complexity and generality to our model. As we still want to make use of a tree structure, we focus on random trees, in particular on Galton Watson trees.

A Galton Watson tree consists of one starting vertex, again called "the root" and denoted by x_0, and a given offspring distribution. That means the number of children of x_0 is a random variable, denoted by X_0, with

$$P(X_0 = j) =: p_j \ \forall j \in \{0, 1, 2, \dots\} \qquad \text{and} \qquad \sum_{j=0}^{\infty} p_j = 1.$$

Throghout this chapter, we assume that the number of children of one single vertex is almost surely bounded by some number N, which means $p_n = 0$ for all $n > N$. We

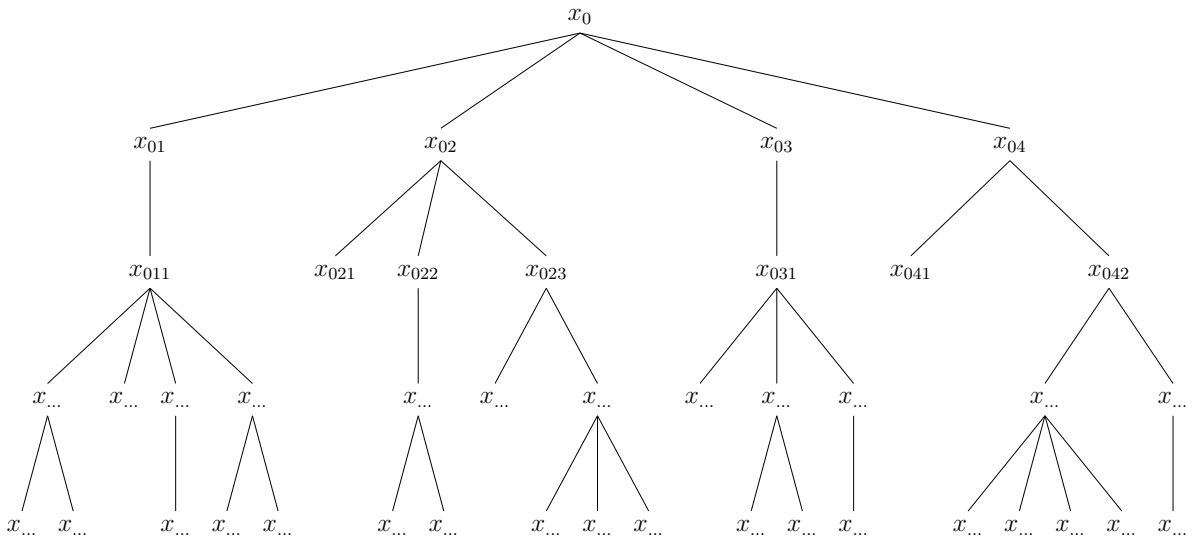


Figure 2.5: The root and the first four generations of a Galton Watson tree.

denote the children of the root (if there are any) by $x_{01}, \ldots, x_{0X_0}$ Now each child of the root also gives birth to a random number of children X_{0i}, $i \in \{1, \ldots, X_0\}$ (which are then grandchildren of the root), again according to the same offspring distribution. We will use the notation $x_{011}, \ldots, x_{01X_{01}}$ for the children of the first child of the root, $x_{021}, \ldots, x_{02X_{02}}$ for the children of the second child of the root and so on. This construction is continued at each new vertex. Furthermore, the offspring of any vertex is assumed to be independent of all other vertices. The resulting tree is often interpreted as a family tree, which is why we call the children of the root the "first generation" the grandchildren of the root the "second generation" and in general all vertices with a distance of n edges to the root the n-th generation. The root itself constitutes the zeroth generation. Obviously, a d-ary tree is just a special case of a Galton Watson tree with the offspring distribution $p_j = 1$ for $j = d$ and $p_j = 0$ otherwise.

Definition 2.14. Let $u := \mathbb{E}[X] = \sum_{j=1}^{\infty} j p_j$ be the expected offspring for any vertex. We now assign weights to all edges of the Galton Watson tree. The weights are assumed to be independent of each other, independent of the offspring distributions at any vertex and they should follow an exponential distribution with mean u. We call the resulting model the weighted Galton Watson tree and denote it by $\mathcal{G}_u$. See Figure 2.5 for an illustration of a Galton Watson tree.

We always want to assume $p_1 \neq 1$, since this would obviously result in a very trivial tree. Note that for $p_0 > 0$ it is possible that the tree "dies out" i.e. that after a finite number of generations there is no offspring anymore. A standard result about Galton Watson trees states that this extinction probability, denoted by q, satisfies

- For $u \leq 1$, it holds $q = 1$.

- For $u > 1$, it holds $0 \leq q < 1$.

To avoid the (for our purpose) uninteresting case of almost sure extinction, we always assume $u > 1$, if not otherwise stated. A Galton Watson tree with $u > 1$ is also called *supercritical*.

Just as before, we now want to examine the biggest rooted subgraphs that have an average weight below some value c. Similarly as in the d-regular case, we consider an RDE.

Definition 2.15. Define $\Gamma_c^G : \mathcal{M}_1 \to \mathcal{M}_1$ by saying that $\Gamma_c^G(\mu)$ is the distribution of $\sum_{i=1}^{X_0}(c - \nu_i + Y_i)^+$. Here the Y_i are iid with distribution μ and independent of the ν_i which are also iid and are exponentially distributed with mean u. Furthermore, X_0 follows the offspring distribution of the Galton Watson tree.

This definition is obviously analogous to Definition 2.2 and it is also easy to see that $(\Gamma_c^G)^m(\delta_0)$ is the distribution of $\max\{c|t| - \omega(t) : t \subset \mathcal{G}_{u,m}\}$, where $\mathcal{G}_{u,m}$ is the finite weighted Galton Watson tree up to the m-th generation.

Since in the Galton Watson case extinction will occur with positive probability, we cannot hope to find a limit distribution μ_c^G of $(\Gamma_c^G)^m(\delta_0)$ for $m \to \infty$ that fulfills $\mu_c(\infty) = 0$ or 1. Therefore, we first want to assume $p_0 = 0$, i.e. every vertex has at least one child. Looking at the definition of $(\Gamma_c^G)^m$

$$(\Gamma_c^G)^m(\delta_0) = \text{dist} \sum_{i_1=1}^{X_0} \left(c - \nu_{i_1} + \sum_{i_2=1}^{X_{0i_1}} \left(c - \nu_{i_2}^{(1)} + \cdots + \sum_{i_m=1}^{X_{0i_1\ldots i_{m-1}}} (c - \nu_{i_m}^{(m)})^+ \ldots \right)^+ \right)^+$$

we see that by interpreting X_a as the number of children of vertex x_a with some indexing sequence a, every realisation of $(\Gamma_c^G)^m$ uniquely defines a realisation of a Galton Watson tree up to generation m.

To prove the existence of a critical threshold, we are seeking to establish a similar result as Lemma 2.3.

Lemma 2.16. For $p_0 = 0$ exists a critical value $0 \leq c_G(u)$ with the following properties:

1. For $c < c_G(u)$, the map Γ_c^G has a fixed point μ_c^G such that $\mu_c^G([0, \infty)) = 1$ and $(\Gamma_c^G)^m(\delta_0) \to \mu_c^G$ as $m \to \infty$.

2. For $c > c_G(u)$, the map Γ_c^G has no fixed point except for δ_∞ and $(\Gamma_c^G)^m(\delta_0) \to \delta_\infty$ as $m \to \infty$.

Proof. It is easy to see that we still have monotonicity in both ways as before:

$$\mu_1 \preceq \mu_2 \Rightarrow \Gamma_c^G(\mu_1) \preceq \Gamma_c^G(\mu_2)$$

and

$$c_1 \leq c_2 \Rightarrow \Gamma_{c_1}^G(\mu) \preceq \Gamma_{c_2}^G(\mu).$$

Therefore, we can deduce that there exists an increasing limit

$$(\Gamma_c^G)^m(\delta_0) \uparrow \mu_c^G \qquad \text{as } m \to \infty.$$

To establish

$$\Gamma_c^G(\lim_{m\to\infty}(\Gamma_c^G)^m(\delta_0)) = \lim_{m\to\infty}(\Gamma_c^G)^{m+1}(\delta_0) \tag{2.19}$$

we can use monotone convergence.

Now let $p := \mu_c^G(\infty)$. As μ_c^G is a fixed point, we have

$$p = \mu_c^G(\infty) = \Gamma_c^G(\mu_c^G)(\infty) = \lim_{m\to\infty}(\Gamma_c^G)^m(\infty).$$

Furthermore we can be sure that the number of summands in the definition of $\lim_{m\to\infty}(\Gamma_c^G)^m$ is infinite because we know that the underlying Galton Watson tree does not die out. This obviously means $p > 0 \Rightarrow \Gamma_c^G(\mu_c^G)(\infty) = 1$ which implies $p = 0$ or 1.

Finally we define $c_G(u) := \inf\{c : \mu_c^G(\infty) = 1\}$. □

Indeed, we could even drop the restriction $p_0 = 0$. However the statement we get then gives slightly less information.

Lemma 2.17. $(\Gamma_c^G)^m(\delta_0)$ converges for $m \to \infty$ to a probability measure μ_c^G. Furthermore, there exists a critical value $0 \leq c_G(u)$ with the following properties:

1. For $c < c_G(u)$, the measure μ_c^G fulfills $\mu_c^G([0,\infty)) = 1$.

2. For $c > c_G(u)$, the measure μ_c^G fulfills $\mu_c^G(\infty) > 0$.

The proof is basically verbatim the same as the proof of Lemma 2.16, using monotonicity in two ways to show the existence of the limit. It is, however not possible to reuse the former proof in order to show that μ_c^G is indeed a fixed point of Γ_c^G and also the assumption that $\mu_c^G(\infty) = 0$ or q might seem very plausible but actually needs further justification. Hence we can just define $c_G(u) := \inf\{c : \mu_c^G(\infty) > 0\}$ in the case of possible extinction.

Of course an immediate question is, how the Galton Watson case is connected with the deterministic trees we considered earlier. We noted that a Galton Watson tree with offspring distribution δ_d is just a d-ary tree but how about a Galton Watson tree with offspring distribution, say $\lambda = \frac{1}{2}\delta_2 + \frac{1}{2}\delta_3$? Intuitively, it seems obvious that the critical value $c_\lambda^G(2.5)$ of this Galton Watson tree has to lie between $c(2)$ and $c(3)$. The reason why this is true is, because we can compare the distributions λ and δ_2 respectively λ and δ_3 via the partial ordering $\preceq$. Then, with $\delta_2 \preceq \lambda \preceq \delta_3$ we can use the same monotonicity

arguments as in the proof of Lemma 2.16 to deduce $c(2) \leq c_\lambda^G(2.5) \leq c(3)$. However, finding an explicit formula for $c_\lambda^G(2.5)$ in terms of $c(2)$ and $c(3)$ seems to be very hard to find.
Unfortunately, $\preceq$ is just a partial ordering, which means not every pair of distributions λ and μ has to be comparable. In the case of Galton Watson trees with non-comparable offspring distributions λ and μ we cannot use monotonicity and it is a priori unclear what role the means, variances and other characteristica of the two distributions play. We tried to adress this problem in Section 2.4.2

2.4 Paths Instead of Trees

Up to now, we considered subgraphs of weighted trees and examined the formation of clusters with small average weight. In this section we want to look only at specific subgraphs, namely paths starting at the root. For the sake of generality, we will use the weighted, supercritical Galton Watson tree $\mathcal{G}_u$ as underlying graph again.

In analogy to (2.2), we define

$$B_n := \inf \left\{ \frac{1}{n} \sum_{e \in \pi} \omega_e : \pi \text{ is a path in } \mathcal{G}_u \text{ which starts at } x_0,\ |\pi| = n \right\}. \tag{2.20}$$

Our aim is to get a statement about the behaviour of B_n for $n \to \infty$. Interestingly, this particular model has been researched before. In literature this setting is known as *branching random walk* and was first studied by Hammersley [13].

In the following, we want to use some celebrated results from the field of branching random walks to tackle the problem we are interested in. The proofs require mostly large deviations theory and are close to those in Peres' book [19], where in Chapter 18, the question for the *maximal* displacement is adressed, whereas we are interested in B_n which corresponds to the minimal displacement in the language of branching random walks.

Recall that in $\mathcal{G}_u$, the weights ω_e on the edges are iid exponentially distributed with mean u. For any vertex x, let π_x be the unique path from x to the root x_0. Obviously,

$$S_x := \sum_{e \in \pi_x} \omega_e,$$

the weight of the path from x to x_0 is distributed as the sum of iid exponential random variables.
Next, we want to apply Cramér's Theorem in a more general form than what we used in Chapter 2.1.3.

Theorem 2.18 (Cramér's Theorem)**.**
Let $Y_1, Y_2, \dots$ be iid with law $\mu \in \mathcal{M}_1(\mathbb{R})$. Write $S_n := \sum_{i=1}^n Y_i$.
For $A \subseteq \mathbb{R}$ closed, we have

$$\limsup_n \frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} \in A\right) \leq -\inf_{x \in A} \Lambda_\mu^*(x).$$

For $U \subseteq \mathbb{R}$ open, we have

$$\liminf_n \frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} \in U\right) \geq -\inf_{x \in U} \Lambda_\mu^*(x).$$

Here, the rate function

$$\Lambda_\mu^*(x) := \sup_{\lambda \in \mathbb{R}} \left(\lambda x - \log\left(\int e^{\lambda x} d\mu\right)\right)$$

is again the Legendre transformation of the distribution μ. By the same calculation as in Chapter 2.1.3, we find for the particular case when μ is an exponential distribution with mean u:

$$\Lambda_\mu^*(x) = \frac{x}{u} + \log\frac{u}{x} - 1. \tag{2.21}$$

Since we will not look at any other cases than the exponential distribution, we just want to write Λ^* instead of Λ_μ^*. We are now ready to state the main theorem of this section.

Theorem 2.19. *Let $\mathcal{G}_u$ be a weighted Galton Watson tree with mean $u > 1$ and iid weights ω_e that are exponentially distributed with mean u. Then, conditioned on the event that $\mathcal{G}_u$ survives*

$$\lim_{n \to \infty} B_n = s_*$$

with $s_* := \inf\{s : \Lambda^*(s) \leq \log u\}$.

This Theorem is actually just a special case of a much stronger statement which was established as the result of a series of papers by Hammersley [13], Kingman [17] and Biggins [5]. The stronger Theorem, along with the details of the more general model of branching random walks, that also involves non-independent edge weights, can be found in a very nice introduction by Shi [20].

Proof. We first show the lower bound, that is, we want to prove that

$$\lim_{n \to \infty} \frac{1}{n} \min_{x : |\pi_x| = n} S_x \geq s_*. \tag{2.22}$$

Let $\varepsilon > 0$ and let $Y_1, Y_2, \dots$ be iid exponentially distributed with mean u. Write

$$S_n := \sum_{i=1}^n Y_i. \tag{2.23}$$

Applying Theorem 2.18 to the closed set $\{x \in \mathbb{R}^+ : x \leq n(s_* - \varepsilon)\}$ yields

$$\mathbb{P}(S_n \leq n(s_* - \varepsilon)) \leq 2e^{-n(\inf_{0 \leq x \leq (s_* - \varepsilon)} \Lambda_\mu^*(x))}.$$

By the definition of s_*, we know that there is a $\delta > 0$ such that $\Lambda^*(s_* - \varepsilon) > \log u + \delta$. Hence

$$\mathbb{P}(S_n \leq n(s_* - \varepsilon)) \leq 2e^{-n(\log u + \delta)} = u^{-n}e^{-n\delta}. \tag{2.24}$$

Now a standard result about Galton Watson trees states that the number of vertices in the n-th generation divided by u^n, the expected offspring to the n-th power, is a martingale. Hence, if Z_n denotes the number of vertices in the n-th generation, we have $\mathbb{E}[Z_n] = u^n$. Combining this with (2.24) and using σ-subadditivity leads to

$$\mathbb{P}(\underbrace{S_x \leq n(s_* - \varepsilon) \text{ for some } x \text{ with } |\pi_x| = n| \text{ non-extinction}}_{=:C_n}) \leq \frac{u^n}{1-q} u^{-n} e^{-n\delta}.$$

Therefore $\sum_n \mathbb{P}(C_n) < \infty$. Using the Borel Cantelli lemma, we can conclude $\mathbb{P}(C_n \text{ holds for infinitely many } n) = 0$ which means (2.22) holds almost surely.

For the upper bound, choose $a > s_*$ sufficiently close to s_*. Using (2.21) it is easy to check that $\Lambda^*(x)$ is strictly decreasing in an open neighbourhood of s_*. Hence we can choose $\varepsilon > 0$ such that $\Lambda^*(a) + \varepsilon < \log u$.
Now, for any $k \geq 1$ and $M \in [1, \infty]$ consider the following embedded branching process $\widetilde{B}(k, M)$:
Start at x_0 and let the set of offspring of any vertex x be all its descendants w in $\mathcal{G}_u$ that satisfy

- $|\pi_w| = |\pi_x| + k$.
- $S_w < S_x + ka$.
- $S_u < S_x + M(|\pi_u| - |\pi_x|)$ for all u on the path from x to w.

So by the first two requirements, the offspring of a vertex x are all w that are k generations further down the Galton Watson tree than x and for which the mean edge weight of the path connecting x to w is less than a. Roughly speaking, the third requirement demands that the deviation of the mean weight along the path should be bounded. Note that for $M = \infty$, the last requirement is always satisfied. Figure 2.6 respectively Figure 2.7 shows a possible realisation of this embedded branching process $\widetilde{B}(k, M)$ for $k = 2$ and how it evolves from $\mathcal{G}_u$.

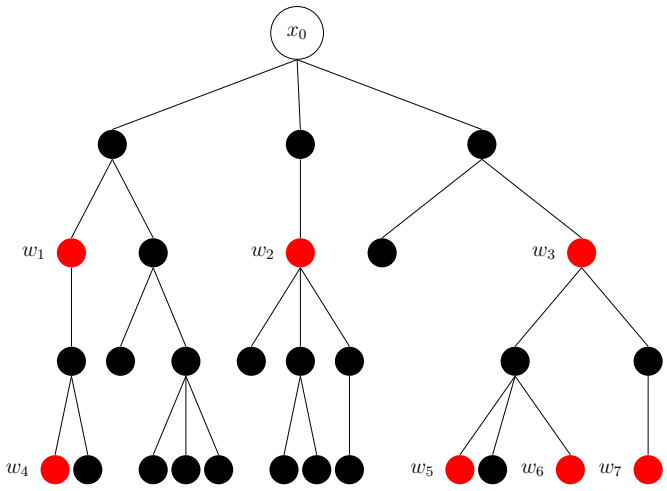


Figure 2.6: The root and the first 4 generations of a Galton Watson tree. All vertices that fulfill the requirements of $\widetilde{B}(2, M)$ are drawn in red.

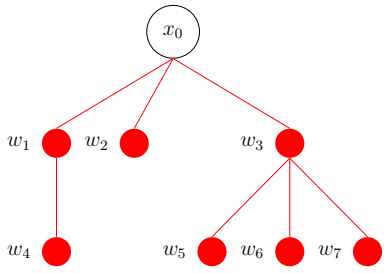


Figure 2.7: Here, $\widetilde{B}(2, M)$ is exracted as a Galton Watson tree of its own.

Obviously, $\widetilde{B}(k, M)$ is again a Galton Watson process, which means the number of children of any two vertices x, y are independent and identically distributed. Write $\mathbb{E}\left[|\widetilde{B}(k, M)|\right]$ for the expected offspring of any vertex in $\widetilde{B}(k, M)$. Let S_n be again the sum defined in (2.23). By Theorem 2.18, applied to the open set $\{x \in \mathbb{R}^+ : x < a\}$, it holds

$$\mathbb{E}\left[|\widetilde{B}(k, \infty)|\right] = u^k \mathbb{P}(S_k < ka) \geq u^k \frac{1}{2} e^{-k \inf_{x<a} \Lambda^*(x)}.$$

Since Λ^* is continuous, we have

$$u^k \frac{1}{2} e^{-k \inf_{x<a} \Lambda^*(x)} \geq \frac{1}{2} u^k e^{-k(\Lambda^*(a)+\varepsilon)}.$$

Hence

$$\mathbb{E}\left[|\widetilde{B}(k, \infty)|\right] \geq \frac{1}{2} u^k e^{-k(\log u - \varepsilon)} > 2$$

for k large enough. So by choosing M sufficiently large, we get that the embedded branching process $\widetilde{B}(k, M)$ satisfies

$$\mathbb{E}\left[|\widetilde{B}(k, M)|\right] > 1$$

which means $\widetilde{B}(k, M)$ is supercritical and therefore it survives with a positive probability. Now the crucial step is to realize that whenever $\widetilde{B}(k, M)$ does not die out, there is a path π in $\mathcal{G}_u$, starting at x_0 which has an average edge weight of at most a, or in other words $\lim_{n\to\infty} B_n \leq a$ with B_n from (2.20).

To conclude the proof of Theorem 2.19 we need the notion of *inherited properties*. A

property A of a Galton Watson tree $\mathcal{G}_u$ is called *inherited* if all finite trees have property A and the subtrees $\mathcal{G}_u(1), \dots, \mathcal{G}_u(X_0)$ which are the Galton Watson trees rooted at all the children of x_0, have property A whenever $\mathcal{G}_u$ has property A. Inherited properties are for example {The tree is finite} or {The number of vertices in the n-th generation grows polynomially in n}. The following is a well known result about Galton Watson trees. A proof can be found in Chapter 3 of [19].

Theorem 2.20. *Let $\mathcal{G}_u$ be a supercritical Galton Watson tree. If A is inherited, then*

$$\mathbb{P}(A|\ \textit{non-extinction of}\ \mathcal{G}_u) \in \{0, 1\}.$$

It is fairly straightforward that for any a the property
$A := \{\mathcal{G}_u$ is finite or $\lim_{n\to\infty} B_n \leq a\}$ is inherited. Theorem 2.20 together with the fact that $\widetilde{B}(k, M)$ survives with a positive probability implies that

$$\mathbb{P}(\lim_{n\to\infty} B_n \leq a|\ \text{non-extinction}) = 1$$

which proves the upper bound. □

2.4.1 Another Upper Bound on $c(d)$

At this point it is worth noting that every path is, of course, also a tree. If we compare Definitions 2.2 and 2.20, we see that $A_n \leq B_n$. Hence, the value s_* from Theorem 2.19 is actually another upper bound for the critical value $c(d)$ which we tried to estimate in Section 2.1.2 since we already mentioned that a d-ary tree is just a special case of a Galton Watson tree with the offspring distribution δ_d. Note that in T_d^* as a "deterministic Galton Watson tree" extinction is not possible, we do not have to consider conditional probabilities. We now want to give the calculation of $\lim_{n\to\infty} B_n = s_*$ in detail.
Recall that s_* was defined as $s_* := \inf\{s : \Lambda^*(s) \leq \log u\}$ with

$$\Lambda^*(x) = \frac{x}{u} + \log\frac{u}{x} - 1.$$

Now in the case that $\mathcal{G}_u = T_d*$, the expected number of offspring is obviously d. Hence, an upper bound for the critical calue $c^*(d)$ of the d-ary tree is the smallest solution of

$$\frac{x}{d} + \log\frac{d}{x} - 1 - \log d.$$

Note that we are talking about d-ary trees here. In order to get a result about the d-regular tree model we studied in Chapter 2.1 one would have to consider a weighted Galton Watson tree with expected offspring $d-1$ and edge weights that are iid exponentially distributed with mean d. So the same calculation for d-regular trees yields, that an upper bound for $c(d)$ from Chapter 2.1 is given by the smallest solution of

$$\frac{x}{d} + \log\frac{d}{x} - 1 - \log(d-1).$$

Figure 2.8 and 2.9 show the case for $d = 4$ and indicate an upper bound of $0,565$ for $c(4)$. Compared to the value from Section 2.1.2, which calculates as $0,756$, this is an improvement.

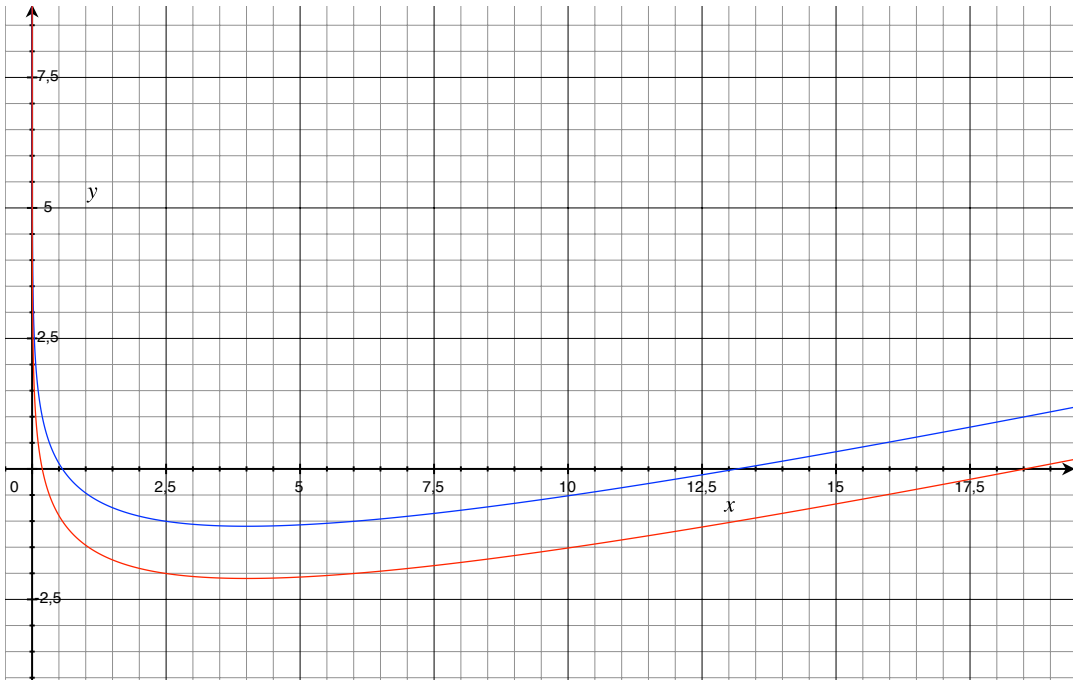


Figure 2.8: In blue: the function $y = \frac{x}{d} + \log \frac{d}{x} - \log(d-1) - 1$ for $d = 4$. The first zero point is an upper bound on $c(4)$. For comparison, we also plotted in red the function $y = \frac{x}{d} - \log(x) + \log\left(\frac{d}{d-1}\right) - 2$ from Section 2.1.3. Recall that the first zero point of the red function was $c^-(4)$, the lower bound on $c(4)$.

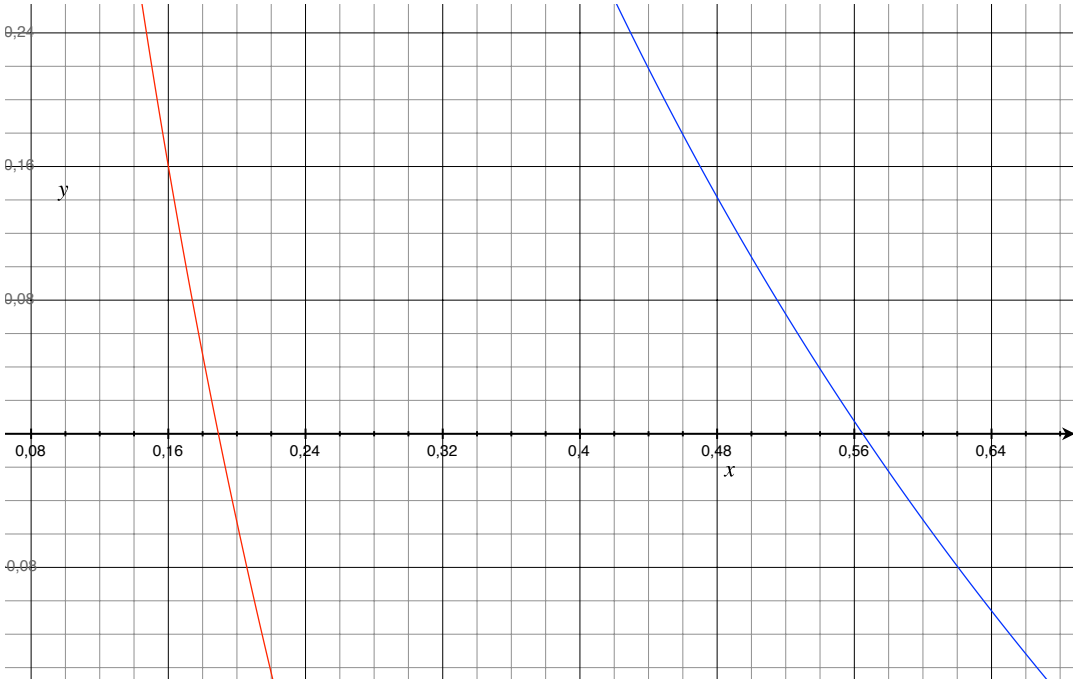


Figure 2.9: The two zero points in detail. The critical value $c(4)$ has to be in between the blue and the red line.

2.4.2 Average Percolation and Polar Sets

Besides finding a formula to explicitly calculate the critical value for paths in the sense of (2.20) for any given offspring distribution, it might be interesting to look at average percolation from a more abstract point of view. Doing so, we also hope to gain insights about the critical value for subtrees $c_G(u)$.

We already noted that the intuitive notion of "infinite subtrees respectively paths with bounded mean weight" is mathematically not quite clear, hence we want to present a slightly different interpretation. Instead of looking at B_n as introduced in (2.20)

$$B_n := \inf \left\{ \frac{1}{n} \sum_{e \in \pi} \omega_e : \pi \text{ is a path in } \mathcal{G}_u \text{ which starts at } x_0,\ |\pi| = n \right\}$$

and then taking the limit $n \to \infty$, we now look at the whole infinite tree and ask whether there is a path $\pi = (x_0, x_1, x_2, \dots)$ starting at the root that satisfies

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n} \omega_{(i,i+1)} \leq c \tag{2.25}$$

where $\omega_{(i,i+1)}$ denotes the weight of the edge connecting x_i and x_{i+1}. The difference between the two concepts is subtle: when we are trying to prove $\lim_{n\to\infty} B_n \leq c$ we just demand that for any (big) length n, we find **some** rooted path π whose mean weight is arbitrarily close to c. When we are working with the notion of (2.25) however, we want to find **one specific** rooted path whose mean weight is, in the limes superior, bounded by c.

Even so this new interpretation seems to be, at first glance, harder to study than B_n, it can be tackled by the concept of so called *polar sets* for random trees. The idea goes back to Evans [10] and attempts to characterize random trees by special sets. The results we are interested in were published in a work by Pemantle and Peres [18] which also gives a nice introduction to the theory of polar sets that we want to outline briefly.

Definition 2.21. Let T be an infinite, possibly random but locally finite tree with root x_0 and further vertices $x_1, x_2, \dots$ labelled in breadth-first order from left to right. Furthermore let $\{X_i\}$ be a collection of iid random variables, indexed by all the vertices of the tree except for the root. In our case, we imagine X_i to be the weight of the unique edge that connects x_i to its parent. Now for any closed set $B \subset \mathbb{R}^\infty$ let $\mathbb{P}(T; B)$ denote the probability that there exists an infinite, non intersecting path $x_0, x_{i_1}, x_{i_2}, \dots$ for which $(X_{i_1}, X_{i_2}, \dots) \in B$. Obviously, $\mathbb{P}(T; B)$ depends on the tree T as well as on the set B and the distribution of the X_i which we want to denote by μ. A set B with $\mathbb{P}(T; B) = 0$ is called *polar* for the tree T with respect to μ. Now two trees T_1 and T_2 are called *equipolar* with respect to μ if for every closed set $B \subset \mathbb{R}^\infty$ holds

$$B \text{ is polar for } T_1 \text{ with respect to } \mu \quad \Longleftrightarrow \quad B \text{ is polar for } T_2 \text{ with respect to } \mu.$$

The following theorem tells us that Galton Watson trees with the same mean have the same polar sets, which is the main statement (and even the title) of [18].

Theorem 2.22. *Let p and p' be the offspring distributions of two Galton Watson trees that satisfy $\sum_n np_n = u = \sum_n np'_n$ and $p_0 = p'_0 = 0$. If the variances of the offspring distributions, $\sum_n n^2 p_n$ and $\sum_n n^2 p'_n$ are both finite, then the $\mathbb{P}_p \times \mathbb{P}_{p'}$ probability of picking two equipolar trees is* 1.

In our case, we consider the closed sets

$$B_c := \left\{ (a_1, a_2, \dots) : \limsup \frac{1}{n} \sum_{i=1}^{n} a_i \leq c \right\}$$

which encodes exactly the notion of (2.25). Next, we define the critical value $\widetilde{c}$:

$$\widetilde{c} := \inf\{c : \mathbb{P}(\mathcal{G}_u; B_c) > 0\}.$$

Now by Theorem 2.22, we know that any two Galton Watson trees with the same mean have the same critical value $\widetilde{c}$, as long as their offspring distributions have finite variance and satitisfy $p_0 = 0$ respectively $p'_0 = 0$.

Although this result only gives information about the critical value $\widetilde{c}$ for the interpretation from (2.25), it seems promising to try and adapt the concept of polar sets to fit the notions of (2.2) and (2.20). Anyway, principles of equivalence of Galton Watson trees have been vivid fields of interest in the past decades and exploring the stand of average percolation in this context might be worth some further research.

3 The Complete Graph Case

After having studied the largest trees respectively paths with small average weight on tree structures, we want to shift our focus to a slightly different model. Instead of an infinite tree we now look at C_n, the complete graph on the vertex set $\{1, 2, \ldots, n\}$. Again, we assign independent weights to the edges, this time following an exponential distribution with mean n. Call the resulting graph with random weights $\mathcal{W}_n$ and for $c \in \mathbb{R}$ define

$$L_\pi(n, c) := \max\{|\pi| : \pi \text{ is a path in } \mathcal{W}_n, |\pi|^{-1}\omega(\pi) \leq c\} \tag{3.1}$$

and

$$L_t(n, c) := \max\{|t| : t \text{ is a tree in } \mathcal{W}_n, |t|^{-1}\omega(t) \leq c\}. \tag{3.2}$$

Probably the most immediate question related to these quantities is how they behave in the limit $n \to \infty$. At a first glance this looks very similar to what we did in the previous chapter, however the fact that the edge weight distribution depends on n adds some complexity to the problem. This is also the model that Aldous first studied in his 1998 paper [1] which started the interest in this kind of stochastic mean field model altogether. There he was able to show that, for both path and tree, there exists a critical value $\tilde{c}(0)$ respectively $c(0)$ such that $L_\pi(n, c)$ respectively $L_t(n, c)$ makes the transition from $o(n)$ to $\Theta(n)$. Furthermore, [1] provides the exact critical value for the path case which in turn has lead to a series of studies about the scaling behaviour near that critical value. In this chapter, we want to present the central results concerning the complete graph case and also look at some implications thereof.

3.1 The PWIT as a local limit of $\mathcal{W}_n$

We start this section by introducing the so called *Poisson weighted infinite tree* or short *PWIT*. It will serve us as a local approximation of $\mathcal{W}_n$ in a sense that we will define shortly. This is the crucial idea for dealing with the complete graph case, as it turns out that the tree structure provided by the PWIT makes it much easier to study the questions we are interested in.

Start with a vertex x_0 and call it the root. Attach to the root an infinite series of children $(x_{0i})_{i\in\mathbb{N}}$. Assign weights to the edges $0 < \omega(x_0, x_{0,1}) < \omega(x_0, x_{02}) < \ldots$ which

are the points of a Poisson rate 1 process. That means $\omega(x_0, x_{01})$ is the time of the first occurence of an exp(1) event, $\omega(x_0, x_{02})$ is the time of the second occurence and so on. At each child of the root, again attach infinitely many vertices with the same weight distribution on the edges, independently of all other edges and continue this construction at each new vertex. The resulting infinite tree is, for obvious reasons, called Poisson weighted infinite tree (PWIT) and we will denote it by $\mathcal{T}^\infty$.
Furthermore we will use $\mathcal{T}_k^\infty$ to denote the subtree that contains only the first k generations and ${}_L\mathcal{T}_k^\infty$ the subtree that consists of the first k generations and inlcudes only the first L children of every vertex. Clearly ${}_L\mathcal{T}_k^\infty$ is a finite tree. What we will do now is to construct a map that establishes the connection between ${}_L\mathcal{T}_k^\infty$ and $\mathcal{W}_n$. So first of all look at the vertex set of ${}_L\mathcal{T}_k^\infty$ which we will denote by ${}_LN_k$. It is not hard to see that

$$ {}_LN_k = \{x_0\} \cup \{x_{0i} : i \in \cup_{j=1}^k \{1, 2, \dots, L\}^j\}.$$

For n large enough, we can define $\varphi : {}_LN_k \to \{1, 2, \dots, n\}$ inductively as follows:

- Set $\varphi(x_0) := 1$
- For any $x_{0\dots ij}$ in ${}_LN_k$ with parent $x_{0\dots i}$ define $\varphi(x_{0\dots ij}) := \arg\min_y \omega_{(\varphi(x_{0\dots i}),y)}$ where the minimum is taken over all vertices in $\mathcal{W}_n$ that have not yet been the image of any pervious vertex under φ.

As a result, we get a weighted tree ${}_L\mathcal{T}_k^{(n)}$ in $\mathcal{W}_n$ on the same vertex set as ${}_L\mathcal{T}_k^\infty$. Example 3.1 might help to visualize this construction.

Example 3.1. For this example, we set $k = L = 2$ and look at $\varphi : {}_2N_2 \to \{1, \dots, 6\}$. Because $|\{1, \dots, 6\}| < |{}_2N_2|$, we won't be able to map every point in ${}_2N_2$ and thus only see the first steps of the construction. However, as complete graphs tend to look rather messy for bigger n, we decided to map to $\{1, \dots, 6\}$ and hope this example makes it clear how to apply the mapping for bigger n. The edge weights in K_6, the complete graph on 6 vertices shall be given as

$$\begin{array}{ccccc} \omega_{(1,2)} = 3,97 & \omega_{(1,3)} = 3,14 & \omega_{(1,4)} = 2,01 & \omega_{(1,5)} = 7,30 & \omega_{(1,6)} = 1,12 \\ \omega_{(2,3)} = 2,48 & \omega_{(2,4)} = 4,23 & \omega_{(2,5)} = 3,13 & \omega_{(2,6)} = 1,98 & \omega_{(3,4)} = 5,55 \\ \omega_{(3,5)} = 2,01 & \omega_{(3,6)} = 4,20 & \omega_{(4,5)} = 1,07 & \omega_{(4,6)} = 0,96 & \omega_{(5,6)} = 8,80. \end{array}$$

Now by definition, we set $\varphi(x_0) = 1$. Next we look at the neighbours of 1 and see that $(1, 6)$ is the edge with the minimal weight among all edges incident to 1 (we say "6 is the cheapest neighbour"). Hence, we set $\varphi(x_{01}) = 6$. For $\varphi(x_{02})$, we look for the cheapest neighbour again, but as 6 has already been used, we choose the second cheapest neigbour and set $\varphi(x_{02}) = 4$. Since we are mapping from ${}_2N_2$, the next vertex waiting to be mapped is x_{011}. Again, we look for the cheapest neigbour, that hasn't previously been used and get $\varphi(x_{011}) = 2$. Applying the same procedure again, we find $\varphi(x_{012}) = 3$. For the resulting subtree ${}_2\mathcal{T}_2^{(6)}$ (or the first part of it) see Figure 3.1.

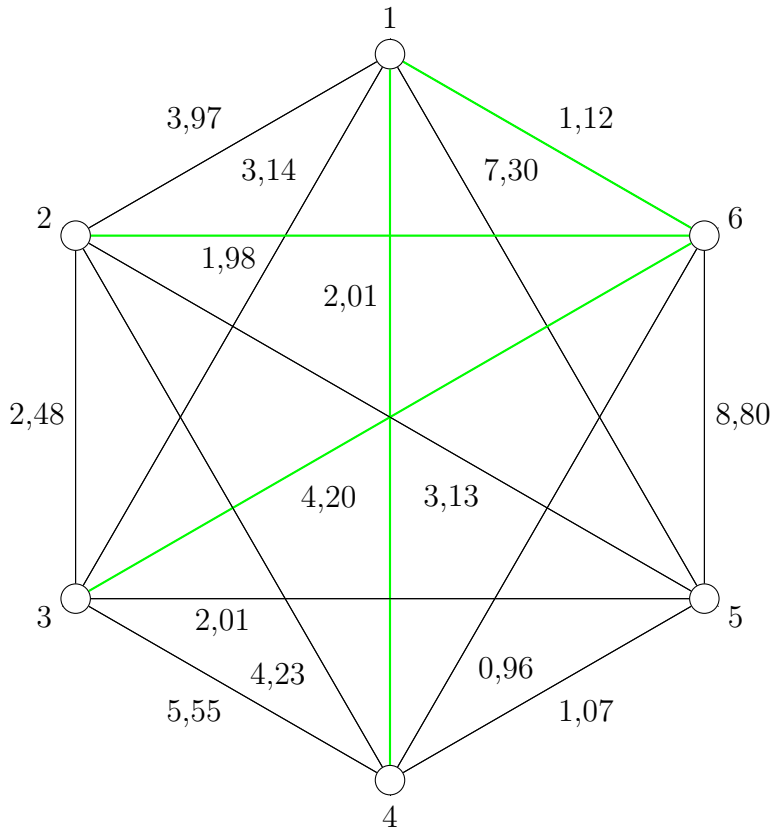


Figure 3.1: $\mathcal{W}_6$, the weighted, complete graph on 6 vertices. The edges of ${}_2\mathcal{T}_2^{(6)}$ are drawn in green.

Obviously the edge weights of any vertex in ${}_L\mathcal{T}_k^{(n)}$ to its children are distributed as the order statistics of L exponential mean n distributed random variables $\eta_1^n < \eta_2^n < \cdots < \eta_L^n$. But it is not hard to show that

$$(\eta_1^n, \eta_2^n, \ldots, \eta_L^n) \xrightarrow{d} (\xi_1, \xi_2, \ldots, \xi_L)$$

for $n \to \infty$. Where $\xi_1 < \xi_2 < \cdots < \xi_L$ are the points of a rate 1 Poisson process.

So in conclusion we see that, for fixed k and L, the edge weights of ${}_L\mathcal{T}_k^{(n)}$, which is a tree in $\mathcal{W}_n$ converge in distribution to those of ${}_L\mathcal{T}_k^{\infty}$, which is a subtree of $\mathcal{T}^{\infty}$. We call this sense of convergence *local weak convergence* because it means that for big n we can identify any finite neighbourhood around a vertex in $\mathcal{W}_n$ with a finite neighbourhood around the root of $\mathcal{T}^{\infty}$.

3.2 Looking for Paths in $\mathcal{W}_n$

Equipped with the local approximation for $\mathcal{W}_n$ we can now begin to study the quantities defined at the beginning of this chapter. We want to start with the path case as it is slightly easier and also because we can prove that the exact critical value equals e^{-1}.

Theorem 3.2. *Let L_p be defined as in* (3.1)*. It holds*

- *For all $c < e^{-1}$ we have*

$$\lim_{n\to\infty} \mathbb{P}(L_p(n,c) > \varepsilon n) = 0$$

  for all $\varepsilon > 0$.

- *For all $c > e^{-1}$ there is a $\delta(c) > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}(L_p(n,c) > \delta(c) n) = 1.$$

Proof. : For the subcritical case we use a counting argument, i.e. we want to determine $\mathbb{E}\left[|\{\pi : |\pi| = k, \omega(\pi) \leq ck\}|\right]$ where π denotes a path in $\mathcal{W}_n$. Obviously, a path of length k (in edges) is defined by an ordered tuple of vertices $(v_0, \ldots, v_k)$ and there are precisely $n \cdot (n-1) \cdot \ldots \cdot (n-k)$ different vertex-tuples of length $k+1$ in $\mathcal{W}_n$. Noting that we do not care about the orientation of the path we can identify any two tuples that are the reverse of each other. Denote the number of paths of length k in $\mathcal{W}_n$ by N_π and we arrive at

$$N_\pi = \frac{n!}{2(n-k-1)!}.$$

Notice that we could also allow paths where the starting vertex equals the ending vertex. The number of those "rings" only differs by the last factor and leads to the same asymptotic behaviour. We will not carry out the exact calculation for the rings but bear in mind that our results are equally valid in the ring case.

Now for P_π, the probability that a path of length k has total weight less than ck, we use the representation in form of a Poisson process. Let η_i be the time between the i-th and $(i-1)$-th occurence of an $\exp(\frac{1}{n})$ distributed event. We want $\eta_1 + \eta_2 + \cdots + \eta_k \leq ck$ so at time ck we want that at least k events have occured. Hence we find

$$P_\pi = \mathbb{P}\left(\text{Poisson}\left(\frac{1}{n} \cdot ck\right) \geq k\right) \leq \frac{\left(\frac{ck}{n}\right)^k}{k!} e^{-\frac{ck}{n}}.$$

If we now combine our results, we find that

$$\mathbb{E}\left[|\{\pi : |\pi| = k, \omega(\pi) \leq ck\}|\right] \leq N_\pi \cdot P_\pi.$$

We use Sterling's formula $n! \approx \sqrt{2\boldsymbol{\pi} n}\left(\frac{n}{e}\right)^n$ to handle the factorials. In order to avoid ambiguities we use the bold letter $\boldsymbol{\pi}$ to denote the constant, whereas π denotes a path. Luckily all the $\boldsymbol{\pi}$'s cancel out very soon. We find

$$\begin{aligned}
k^{-1}\log\mathbb{E}\left[|\{\pi : |\pi| = k, \omega(\pi) \le ck\}|\right] &\le k^{-1}\log\left(\frac{n!}{2(n-k-1)!}\left(\frac{ck}{n}\right)^k \frac{1}{k!}e^{-\frac{ck}{n}}\right) \\
&\overset{\text{Sterling}}{\approx} \frac{1}{k}\log\frac{1}{2} + \frac{1}{k}\log\left(\frac{\sqrt{2\boldsymbol{\pi} n}}{\sqrt{2\boldsymbol{\pi}(n-k-1)}}\left(\frac{n}{e}\right)^n\left(\frac{e}{n-k-1}\right)^{n-k-1}\right) \\
&+ \log\frac{ck}{n} - \frac{c}{n} - \frac{1}{k}\log\left(\sqrt{2\boldsymbol{\pi} k}\left(\frac{k}{e}\right)^k\right) \\
&= \frac{1}{k}\log\frac{n}{e} + \frac{n-k-1}{k}\log\frac{n}{n-k-1} + \log(c) + o(1).
\end{aligned}$$

So we get

$$\mathbb{E}\left[|\{\pi : |\pi| = k, \omega(\pi) \le ck\}|\right] = \frac{n}{e}\left(\frac{n}{n-k-1}\right)^{n-k-1} c^k.$$

Now we can deduce

$$\begin{aligned}
P(L_\pi(c,n) > \varepsilon n) &\le \sum_{\varepsilon n \le k \le n-1} \mathbb{E}\left[|\{\pi : |\pi| = k, \omega(\pi) \le ck\}|\right] \\
&= \sum_{\varepsilon n \le k \le n-1} \frac{n}{e}\left(\frac{n}{n-k-1}\right)^{n-k-1} c^k \\
&\le (n-1)\frac{n}{e}\left(\frac{n}{n-\varepsilon n-1}\right)^{n-\varepsilon n-1} c^{\varepsilon n} \\
&= (n-1)\frac{n}{e}\left(1 + \frac{1+\varepsilon n}{n-\varepsilon n-1}\right)^{n-\varepsilon n-1} c^{\varepsilon n} \\
&\overset{\text{for } n \text{ big}}{\approx} (n-1)\frac{n}{e}e^{1+\varepsilon n}c^{\varepsilon n}.
\end{aligned}$$

So we see that for $c < e^{-1}$ we get

$$\mathbb{P}(L_\pi(c,n) > \varepsilon n) \overset{n\to\infty}{\longrightarrow} 0. \tag{3.3}$$

For the supercritical case we use the local correspondence between $\mathcal{W}_n$ and $\mathcal{T}^\infty$ and therefore want to interpret $\mathcal{T}^\infty$ as a continuous time branching process that looks as follows: Start with some vertex x_0. Every vertex has an infinite number of children, where the "birth" of x_{0i}, the i-th child of x_0 is distributed as ξ_i, the i-th point of a rate 1 Poisson process. This process is obviously equivalent to $\mathcal{T}^\infty$. The continuous process is sometimes called the *Yule branching process*.
So for large n and any given vertex v^* in $\mathcal{W}_n$, we can identify any finite size neighbourhood around v^* with a Yule process starting at v^*.

If we denote the maximal generation up to some time s by G_s, we have the following result from [4]

$$\frac{G_s}{s} \xrightarrow{s\to\infty} e \quad a.s.. \tag{3.4}$$

But how does it translate to our situation concerning paths? Thinking about it, it is clear that we can interpret G_s as $|\pi|$ and s as $\omega(\pi)$. So instead of (3.4) we can write

$$\min\left\{\frac{\omega(\pi)}{|\pi|}, \pi \text{ is a path in } \mathcal{T}^\infty,\ |\pi| = m,\ x_0 \in \pi\right\} \xrightarrow{m\to\infty} e^{-1}. \tag{3.5}$$

For convenience, we want to use the notation $\pi \subset \ldots$ to say "π is a path in…" as we will always use the letter π for paths and t for trees we don't have to fear ambiguities. What we want to do now is to construct a path of order n in $\mathcal{W}_n$. For this define Ψ_c^k to be the distribution of $\max\{c|\pi| - \omega(\pi) : \pi \subset \mathcal{T}^\infty \text{ up to generation } k\}$ and $\Psi_{c,L}^k$ is the same distribution when only the first L children of every vertex are considered. Clearly $\Psi_{c,L}^k \xrightarrow{d} \Psi_c^k$ for $L \to \infty$. From (3.5) we can deduce that

$$\Psi_c^k \xrightarrow{k\to\infty} \delta_\infty \quad a.s. \tag{3.6}$$

for all $c > e^{-1}$.

Lemma 3.3. For $e^{-1} < c < 1$, $a \in \mathbb{N}$ and n sufficiently large, there exists an algorithm on $\mathcal{W}_n$ which takes an arbitrary vertex x_0 and finds vertices $x_1, \ldots, x_{q-1}$. With probability at least $\frac{1}{2}$ the algorithm finds a path $\pi \subset \{x_0, \ldots, x_{q-1}\}$ which contains both x_0 and x_{q-1} and it holds $c|\pi| - \omega(\pi) > ca$. Furthermore, the algorithm will not look at more than the q vertices $x_0, \ldots x_{q-1}$ and will not look at any edges that have no endpoint in the set $\{x_0, \ldots x_{q-1}\}$ and q depends only on c and a.

Proof. of Lemma 3.3: Because of (3.6) we know that for every $a \in \mathbb{R}$ we can choose k such that $\Psi_c^k([ca, \infty]) \geq \frac{1}{2}$. Furthermore we can choose $L \in \mathbb{N}$ such that $\Psi_{c,L}^k([ca, \infty]) \geq \frac{1}{2}$. But this means that under the assumption that for big n we have $\mathcal{W}_n = \mathcal{T}^\infty$ in a local sense, we can fix any vertex x_0 in $\mathcal{W}_n$ and with probability at least $\frac{1}{2}$ find a path π originating at x_0 such that $c|\pi| - \omega(\pi) > ca$. This path obviously has at least a edges and average edge weight at most c. As we restricted ourselves to k generations and L children in $\mathcal{T}^\infty$, we considered at most $q = \sum_{i=0}^k L^i$ vertices and didn't look at any more edges. □

Here is the construction of an order n path in $\mathcal{W}_n$ with average weight at most c:

1 Take $e^{-1} < c < 1$ and fix an arbitrary vertex x_0 in $\mathcal{W}_n$. Take k and L as above such that $\Psi_{c,L}^k([ca, \infty]) > \frac{1}{2}$ and use the algorithm from Lemma 3.3 to look for a path π with $c|\pi| - \omega(\pi) > ca$. Suppose the algorithm succeeds (i.e. it finds a path with

the desired properties), denote the q vertices that were examined during the search with $x_1, \ldots, x_{q-1}$ (if less than q vertices were examined during the search just include arbitrary extra vertices).

2 Look for the vertex $y_0 := \arg\min_{y \in \mathcal{W}_n \setminus \{x_0, \ldots, x_{q-1}\}} \omega(x_{q-1}, y)$. Remove the vertices $x_0, \ldots, x_{q-1}$ from $\mathcal{W}_n$ and apply the algorithm from Lemma 3.3 again but this time starting at y_0 and c replaced by $\frac{n}{n-q}c$. This modification is necessary as we are now looking at a complete graph on $n-q$ vertices but the edge weights still have mean n. Denote the q vertices that were examined during the search with $y_1, \ldots, y_{q-1}$.

3 If the algorithm succeeds add the path and the linking edge (x_{q-1}, y_0) to π and repeat the previous step, that is look for the vertex $z_0 := \arg\min \omega(y_{q-1}, z)$ where the minimum is taken over all $z \in \mathcal{W}_n \setminus (\{x_1, \ldots, x_{q-1}\} \cup \{y_1, \ldots, y_{q-1}\})$ and apply the algorithm from Lemma 3.3 again with starting point z_0 and $\frac{n}{n-2q}c$ instead of c.

4 If the algorithm fails, remove the vertices $y_0, \ldots, y_{q-1}$ from the graph and look for the best neighbour of x_{q-1} in the remaining graph.
That means search for $z_0 := \arg\min \omega(x_{q-1}, z)$ where the minimum is taken over all $z \in \mathcal{W}_n \setminus (\{x_1, \ldots, x_{q-1}\} \cup \{y_1, \ldots, y_{q-1}\})$. Now apply the algorithm from Lemma 3.3 again with starting point z_0 and $\frac{n}{n-2q}c$ instead of c.

5 Repeat this procedure for multiple steps. Note that in step $s+1$ we look for a path with average edge weight at most $\frac{n}{n-sq}$.

We continue for $\frac{\varepsilon n}{q}$ stages with $0 < \varepsilon < 1$. As the success probability was at least $\frac{1}{2}$ we can say that for $n \to \infty$ at least one third of our tries, which is $\frac{\varepsilon n}{3q}$ were successes. In this case, π consists of $\frac{\varepsilon n}{3q}$ single paths of size a and a linking edge between each of those paths, resulting in a total edge number of $\frac{\varepsilon n}{3q}(a+1) - 1$.
All the single paths have an average edge weight of at most

$$c\frac{n}{n - \frac{\varepsilon n}{q}q} = \frac{c}{1-\varepsilon}$$

and if we denote the average weight of the linking edges by γ, we find that the average edge weight of the total path is at most

$$\frac{c}{1-\varepsilon} + \frac{\gamma}{a+1}.$$

As for large n, γ is approximately distributed as the first success of a Poisson process with rate 1, we can say that $P(\gamma \leq 3) \overset{n\to\infty}{\longrightarrow} 1$ which means the average edge weight is asymptotically

$$\frac{c}{1-\varepsilon} + \frac{3}{a+1}.$$

Because ε and a were arbitrary, we can choose ε small and a big and thus get the desired path with length of order n and average weight at most $c + \delta$. □

3.3 Looking for Trees in $\mathcal{W}_n$

Similarly as in the first chapter, we will use a recursive distributional equation to establish the critical value for trees. However, by this method we can only hope to gain insights about $\mathcal{T}^\infty$ and the difficulty will then be to transfer these results to $\mathcal{W}_n$. Let's introduce the relevant RDE first.

Definition 3.4. Let $0 < \xi_1 < \xi_2 < \dots$ be the points of a Poisson rate 1 process. Then, for any $\mu \in \mathcal{M}_1$, define $\widehat{\Gamma}_c(\mu)$ to be the distribution of $\sum_{i=1}^{\infty}(c - \xi_i + Y_i)^+$, where the Y_i are independent with distribution μ. So $\widehat{\Gamma}_c : \mathcal{M}_1 \to \mathcal{M}_1$.

It's not hard to see that we have again

$$(\widehat{\Gamma}_c)^m(\delta_0) = \text{dist } \max\{c|t| - \omega(t) : t \subset \mathcal{T}_m^\infty,\, x_0 \in t\}.$$

Furthermore, we also want to introduce $\widehat{\Gamma}_{c,L} : \mathcal{M}_1 \to \mathcal{M}_1$ by saying that $\widehat{\Gamma}_{c,L}(\mu)$ is the distribution of $\sum_{i=1}^{L}(c - \xi_i + Y_i)^+$ with ξ_i and Y_i as before. Consequently, it holds

$$(\widehat{\Gamma}_{c,L})^m(\delta_0) = \text{dist } \max\{c|t| - \omega(t) : t \subset {}_L\mathcal{T}_m^\infty,\, x_0 \in t\}.$$

Just as we did earlier, we want to define the critical value by using the map $\widehat{\Gamma}_c$. For this, look at the following lemma.

Lemma 3.5. There exists a critical value $0 \le c(0) \le 1$ with the following properties

1. For $c < c(0)$, the map $\widehat{\Gamma}_c$ has a fixed point μ_c such that $\mu_c([0,\infty)) = 1$ and $(\widehat{\Gamma}_c)^m(\delta_0) \to \mu_c$ as $m \to \infty$.
2. For $c > c(0)$, the map $\widehat{\Gamma}_c$ has no fixed point except for δ_∞ and $(\widehat{\Gamma}_c)^m(\delta_0) \to \delta_\infty$ as $m \to \infty$.

Proof. : Analogously to the two similar lemmas before, we use monotonicity: It is quite easy to see, that

$$\mu_1 \preceq \mu_2 \Rightarrow \widehat{\Gamma}_c(\mu_1) \preceq \widehat{\Gamma}_c(\mu_2)$$

and

$$c_1 \le c_2 \Rightarrow \widehat{\Gamma}_{c_1}(\mu) \preceq \widehat{\Gamma}_{c_2}(\mu).$$

So by induction, we know that $(\widehat{\Gamma}_c)^m(\delta_0) \preceq (\widehat{\Gamma}_c)^{m+1}(\delta_0)$, which means that there exists an increasing limit

$$(\widehat{\Gamma}_c)^m(\delta_0) \uparrow \mu_c \quad \text{as } m \to \infty.$$

The fact that

$$\widehat{\Gamma}_c\left(\lim_{m\to\infty}(\widehat{\Gamma}_c)^m(\delta_0)\right) = \lim_{m\to\infty}(\widehat{\Gamma}_c)^{m+1}(\delta_0)$$

follows again from monotone convergence.
It is easy to see, that any measure μ_c which is a fixed point under $\widehat{\Gamma}_c$ must fulfill $\mu_c(\infty) = 0$ or 1. Together with the monotonicity this allows us to define

$$c(0) := \inf\{c : \mu_c(\infty) = 1\}.$$

While $c(0) \geq 0$ is obvious, the claim that $c(0) \leq 1$ is again proven by noting that for independent random variables Y and Z with $Y \sim \mu$ and $Z \sim c - \xi_1$, we have $\text{dist}(Z + Y) \preceq \widehat{\Gamma}_c(\mu)$. hence, for independent copies $Z_1, \ldots, Z_m$ of Z it holds $\text{dist}(Z_1 + \cdots + Z_m) \preceq (\widehat{\Gamma}_c)^m(\delta_0)$. Obviously, $\mathbb{E}[Z] = c - 1$ so if $c > 1$ the law of large numbers tells us that $(\widehat{\Gamma}_c)^m(\delta_0) \overset{m\to\infty}{\longrightarrow} \delta_\infty$ which means that $c(0) \leq 1$. □

We shall now present the main result of this section.

Theorem 3.6. *Let $c(0)$ be the critical value from Lemma 3.5. It holds*

- *For $c < c(0)$ we have $\lim_{n\to\infty} \mathbb{P}(L_t(n, c) > \varepsilon n) = 0$ for all $\varepsilon > 0$.*
- *For $c > c(0)$ there is $\varepsilon(c) > 0$ such that $\lim_{n\to\infty} \mathbb{P}(L_t(n, c) > \varepsilon(c)n) = 1$.*

So in other words: there is a critical threshold $c(0)$, where the quantity $L_t(n, c)$ makes the transition from $o(n)$ to $\Theta(n)$.

In the interests of clarity, we will divide the proof into two parts.

3.3.1 Proof of the Subcritical Behaviour

We start by stating a helpful lemma

Lemma 3.7. Let $(t_n)_{n\in\mathbb{N}}$ be a sequence of trees in $\mathcal{W}_n$ such that each t_n contains an edge e_n between two vertices $v_1^{(n)}$ and $v_2^{(n)}$ and the edge weight $\omega(e)$ is the s_n-th highest among the edge weights of all edges incident to $v_1^{(n)}$. If $\lim_{n\to\infty} s_n = \infty$ then $\omega(t_n) \to \infty$ as $n \to \infty$.

Proof. : Let $q(m, a, s_n, n)$ be the probability that there exists a path $\pi = v_0, v_1, \ldots, v_j$ with $j \leq m$ in $\mathcal{W}_n$, such that

i) $\max_{i\leq j} \omega_{v_{i-1},v_i} \leq a$.

ii) $|\{v : \omega_{v_{j-1},v} < \omega_{v_{j-1},v_j}\}| \geq s_n$.

Now fix m and a. The edge weights follow an exponential distribution with mean n, so the probability for one edge weight to be smaller than a is given by

$$\int_0^a \frac{1}{n} e^{-x\frac{1}{n}} dx = 1 - e^{-\frac{a}{n}} \leq \frac{a}{n}.$$

In a complete graph on n vertices the number of paths on i edges is bounded by n^i, hence the expected number of paths that fulfill $i)$ in the definition of q can be bounded by

$$\sum_{i=0}^{m-1} n^i \left(\frac{a}{n}\right)^i = \sum_{i=0}^{m-1} a^i.$$

If we now consider condition $ii)$, we see that

$$\left|\{v : \omega_{v_{j-1},v} < \omega_{v_{j-1},v_j}\}\right| \leq \left| \left\{ v : \omega_{v_{j-1},v} < \underbrace{\max_{i \leq j} \omega_{v_{i-1},v_i}}_{=a} \right\} \right|.$$

And $\mathbb{P}(|\{v : \omega_{v_{j-1},v} < a\}| \geq s_n) \overset{n\to\infty}{\longrightarrow} \mathbb{P}(\xi_{s_n+1} \leq a)$ where ξ_i is again the Poisson process with rate 1. So for each path, the probability that there is some vertex v_j fulfilling condition $ii)$ tends to $\mathbb{P}(\xi_{s_n+1} \leq a)$. We conclude that

$$\limsup_{n\to\infty} q(m, a, s_n, n) \leq \mathbb{P}(\xi_{s_n+1} \leq a) \sum_{i=0}^{k-1} a^i = 0.$$

Because obviously $\lim_{n\to\infty} \mathbb{P}(\xi_{s_n+1} \leq a) = 0$. Looking at the sequence of trees $(t_n)_{n\in\mathbb{N}}$, we see that each of them contains a path (the edge e_n between $v_1^{(n)}$ and $v_2^{(n)}$) which fulfills condition $ii)$. By what we have just shown, the probability that this edge weight is bounded by some arbitrary number a tends to 0, therefore $\omega(e_n)$ and hence also $\omega(t_n)$ tends to ∞. □

The statement we just proved seems rather technical and indeed we will only need it to show the next lemma which will then lead to the subcritical behaviour.

Lemma 3.8. Let $c < c(0)$. Take $x > 0, m \geq 1$ and let $N(c, m, x, n)$ be the set of all vertices v of $\mathcal{W}_n$ such that there is some tree t containing v with the following properties

- $|t| \leq 3m$.
- $\omega(t) \leq c|t| - x$.

Then for any $\varepsilon > 0$, it holds

$$\limsup_{n\to\infty} \mathbb{P}(|N(c, m, x, n)| > \varepsilon n) \leq \varepsilon^{-1} \mu_c([x, \infty))$$

where $\mu_c = \lim_{m\to\infty} (\widehat{\Gamma}_c)^m(\delta_0)$.

Proof. : Let $\varepsilon > 0$. By the construction of the PWIT (see Section 3.1 for details) we have

$$\text{dist } \max\{c|t| - \omega(t) : t \subset {}_L\mathcal{T}_{3m}^{(n)},\, x_0 \in t\} \to (\widehat{\Gamma}_{c,L})^{3m}(\delta_0) \qquad \text{as } n \to \infty$$

so, obviously

$$\limsup_{n\to\infty} \text{dist } \max\{c|t| - \omega(t) : t \subset {}_L\mathcal{T}_{3m}^{(n)},\, x_0 \in t,\, |t| \leq 3m\} \preceq (\widehat{\Gamma}_{c,L})^{3m}(\delta_0). \tag{3.7}$$

It is clear that $(\widehat{\Gamma}_{c,L})^{3m}(\delta_0) \preceq (\widehat{\Gamma}_c)^{3m}(\delta_0)$, hence we claim

$$\limsup_{n\to\infty} \text{dist } \max\{c|t| - \omega(t) : t \subset \mathcal{W}_n,\, 1 \text{ is in } t,\, |t| \leq 3m\} \preceq (\widehat{\Gamma}_c)^{3m}(\delta_0). \tag{3.8}$$

And here is why: Assume that (3.8) was not true, then, because (3.7) holds for any L we know that the sequence of t_n that yield the maximum in (3.8) satisfies the following

- Each t_n contains vertices $v_1^{(n)}, v_2^{(n)}$ such that the edge e_n that links $v_1^{(n)}$ to $v_2^{(n)}$ has the l_n-th smallest edge weight among all edges incident to $v_1^{(n)}$ where l_n is some sequence with $\limsup_{n\to\infty} l_n = \infty$.

However, we can easily choose a subsequence t_{n_k} such that $\lim_{k\to\infty} l_{n_k} = \infty$. Then we know by Lemma 3.7 that $\omega(t_{n_k}) \to \infty$ but this subsequence cannot yield any maxima for $c|t_{n_k}| - \omega(t_{n_k})$ as we demand $|t_{n_k}| \leq 3m$. So (3.8) is true.

From (3.8) we get that

$$\begin{aligned}&\limsup_{n\to\infty} \mathbb{P}(\max\{c|t| - \omega(t) : t \subset \mathcal{W}_n,\, 1 \text{ is in } t,\, |t| \leq 3m\} \geq x)\\ &= \underbrace{\limsup_{n\to\infty} \mathbb{P}(1 \in N(c,m,x,n))}_{:=p_1} \leq (\widehat{\Gamma}_c)^{3m}(\delta_0)([x,\infty)) \leq \mu_c([x,\infty)).\end{aligned}$$

Furthermore, for n large enough $\frac{|N(c,m,x,n)|}{n}$ can be interpreted as the probability that any specific vertex is in $N(c,m,x,n)$. Because this probability is the same for all vertices in $\mathcal{W}_n$, we have $\frac{|N(c,m,x,n)|}{n} \overset{n\to\infty}{\longrightarrow} p_1$. Using the Markov inequality it follows that

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{P}(|N(c,m,x,n| \geq \varepsilon n) &= \limsup_{n\to\infty} \mathbb{P}\left(\frac{|N(c,m,x,n|}{n} \geq \varepsilon\right)\\ &\leq \limsup_{n\to\infty} \varepsilon^{-1}\mathbb{E}\left[\frac{|N(c,m,x,n|}{n}\right]\\ &\leq \varepsilon^{-1}\mathbb{E}\left[\underbrace{\limsup_{n\to\infty} \frac{|N(c,m,x,n|}{n}}_{=p_1}\right] \leq \varepsilon^{-1}\mu_c([x,\infty))\end{aligned}$$

where we used Fatou's lemma to exchange the limes superior with the expectation. □

Now we are ready to prove the statement about the subcritical regime. Let $m \geq 1$ and $c_1 < c_2 < c_3 < c(0)$. Suppose there is a tree t^* with $|t^*| \geq \varepsilon n$ and $\omega(t^*)|t^*|^{-1} \leq c_1$. If n is big enough such that $\varepsilon n \geq m$, then we can decompose t^* into subtrees which have between m and $3m$ edges each (this is a well known result about trees which is actually pretty easy to see by considering "worst case" trees). Looking at the decomposition of t^* we note that some of the subtrees may have average weight greater than c_2, we call those "heavy trees" but as the average weight of t^* is at most c_1, the number of edges in such heavy trees is at most $\frac{c_1}{c_2}\varepsilon n$. So at least $\left(1 - \frac{c_1}{c_2}\right)\varepsilon n$ edges of t^* lie in "light trees" with average edge weight smaller or less than c_2. A subset of k edges of a tree is incident to at least $k+1$ different vertices, hence there are at least $\left(1 - \frac{c_1}{c_2}\right)\varepsilon n$ vertices in light subtrees with average edge weight at most c_2 and a total number of edges between m and $3m$. We define x by

$$c_2 m = c_3 m - x.$$

It follows that for $m \leq |t| \leq 3m$ we have $c_2|t| \leq c_3|t| - x$. In the next step, we define the events

$$A := L_t(n, c_1) \geq \varepsilon n$$
$$B := |N(c_3, m, x, n)| \geq \left(1 - \frac{c_1}{c_2}\right)\varepsilon n.$$

By the preceding considerations and the definition of $N(c_3, m, x, n)$, we see that $A \Rightarrow B$. Therefore

$$\mathbb{P}(L_t(n, c_1) \geq \varepsilon n) \leq \mathbb{P}\left(|N(c_3, m, x, n)| \geq \left(1 - \frac{c_1}{c_2}\right)\varepsilon n\right).$$

So with the help of Lemma 3.8 we get

$$\limsup_{n\to\infty} \mathbb{P}(L_t(n, c_1) \geq \varepsilon n) \leq \left(1 - \frac{c_1}{c_2}\right)^{-1} \varepsilon^{-1} \mu_{c_3}([m(c_3 - c_2), \infty)).$$

But as m was arbitrary, the subcritical behaviour from Theorem 3.6 follows.

3.3.2 Proof of the Supercritical Behaviour

Similarly as we did in the proof of Theorem 3.2, we will state an algorithm that finds subtrees in $\mathcal{W}_n$ with small average weight.

Lemma 3.9. Let $c(0) < c$, $a \in \mathbb{N}$ and n sufficiently large. Then there is an integer q which depends solely on c and a such that there exists an algorithm on $\mathcal{W}_n$ which takes an arbitrary vertex v_1 and finds vertices $v_2, \dots v_q$. With probability at least $\frac{1}{2}$ the algorithm finds a tree $t \subset \{v_1, \dots v_q\}$ which contains both v_1 and v_q and it holds $c|\pi| - \omega(\pi) > ca$. Furthermore, the algorithm will not look at more than the q vertices $v_1, \dots v_q$ and will not look at any edges that have no endpoint in the set $\{v_1, \dots v_q\}$.

Proof. : The proof is essentially the same as that for Lemma 3.3, replacing "path" with "tree", which is why we chose not to repeat it at this point. □

Using the algorithm from Lemma 3.9 we can construct an order n tree in $\mathcal{W}_n$ with average weight not greater than c for any $c > c(0)$. We do not want to present the construction in detail, as it differs only very slightly from that used to prove the supercritical behaviour in Theorem 3.2. The full argument can be found in Aldous' paper [1].

3.4 Lower and Upper Bounds on c(0)

Unfortunately, our proof for the behaviour of $L_t(n, c)$ did not provide the concrete value of the critical threshold $c(0)$. We have seen the rather crude estimation $0 \leq c(0) \leq 1$, in this section we want to improve on the bounds.

In Section 3.2, we proved that the critical value for $L_\pi(n, c)$ equals e^{-1}. As every path is obviously also a tree, we immediately get the improved lower bound $e^{-1} < c(0)$.

For the upper bound, we use a counting argument. Recall Cayley's formula, which says that the number of trees that can be formed out of n labelled vertices equals n^{n-2}. Furthermore, a tree with k edges must have $k + 1$ vertices, hence we arrive at

$$\begin{aligned}\mathbb{E}\left[|\{t : |t| = k, \omega(t) \leq ck\}|\right] &= \binom{n}{k+1}(k+1)^{k-1}\mathbb{P}(\underbrace{\eta_1 + \cdots + \eta_k}_{iid\ \exp(\frac{1}{n})\ \text{distributed}} \leq ck) \\ &= \binom{n}{k+1}(k+1)^{k-1}\mathbb{P}\left(\text{Poisson}\left(\frac{ck}{n}\right) \geq k\right).\end{aligned}$$

Now take $n \to \infty$ and $k \to \infty$. The asymptotics of the Poisson probability will be given by the term for "$= k$", which is

$$e^{-\frac{ck}{n}}\left(\frac{ck}{n}\right)^k \frac{1}{k!}.$$

Therefore we get

$$\begin{aligned}\frac{1}{k}\log\mathbb{E}\left[|\{t : |t| = k,\, \omega(t) \leq ck\}|\right] \leq &\frac{1}{k}\log\binom{n}{k+1} + \log(k+1) - \frac{1}{k}\log(k-1) \\ &+ \log\left(\frac{ck}{n}\right) - \frac{1}{k}\log k! - \frac{c}{n}.\end{aligned}$$

Using the fact that $\binom{n}{k+1} \leq \frac{n^{k+1}}{(k+1)!}$, we arrive at

$$\begin{aligned}\frac{1}{k}\log\mathbb{E}\left[|\{t : |t| = k,\, \omega(t) \leq ck\}|\right] \leq & \frac{1}{k}\log n + \frac{1}{k}\log\frac{n^k}{(k+1)!} + \log(k+1) \\ & + \log\frac{ck}{n} - \frac{1}{k}\log k! - \frac{1}{k}\log(k-1) - \frac{c}{n} \\ = & \frac{1}{k}\log n + \log(ck(k+1)) - \frac{2}{k}\log(k!) \\ & - \frac{1}{k}\log(k+1) - \frac{1}{k}\log(k-1) - \frac{c}{n}.\end{aligned}$$

We will again use Stirling's formula to get rid of the factorials and thus, for n big, we can write

$$\begin{aligned}&\frac{1}{k}\log\mathbb{E}\left[|\{t : |t| = k,\, \omega(t) \leq ck\}|\right] \leq \\ &\frac{1}{k}\log n + \log(ce^2) + \log\frac{k+1}{k} - \frac{1}{k}\log(2\boldsymbol{\pi}k) - \frac{1}{k}\log(k+1) - \frac{1}{k}\log(k-1) - \frac{c}{n}.\end{aligned}$$

Now fix $\varepsilon > 0$ and look at $\varepsilon n \leq k \leq n$. It is clear that $\log\frac{k+1}{k} - \frac{1}{k}\log(2\boldsymbol{\pi}k) - \frac{1}{k}\log(k+1) - \frac{1}{k}\log(k-1) - \frac{c}{n} = o(1)$ for $\varepsilon n \leq k \leq n$ and $n \to \infty$. So if we choose c such that $ce^2 < \lambda$ for some $\lambda < 1$, then

$$\mathbb{E}\left[|\{t : |t| = k,\, \omega(t) \leq ck\}|\right] \leq n\lambda^k \qquad \varepsilon n \leq k \leq n.$$

In conclusion we get

$$\mathbb{P}(L_t(n,c) > \varepsilon n) \leq \sum_{\varepsilon n \leq k \leq n} \mathbb{E}\left[|\{t : |t| = k,\, \omega(t) \leq ck\}|\right] \leq n^2\lambda^{\varepsilon n} \to 0,$$

which means $c(0) \geq e^{-2}$.

4 Scaling Behaviour Around Criticality

Once the existence of a critical value where a phase transition takes place is proved, it is naturally interesting to investigate the exact behaviour of the model near this critical threshold. In percolation theory this leads to defining so called *scaling exponents* which we want to explain briefly.

4.1 The Idea Behind Scaling Exponents

As stated in the introduction, the important parameter for bond percolation on a graph $\mathbb{G}$ is the probability p for an edge to be open and the critical value p_c is the smallest value of p for which an infinite cluster appears. Therefore we know that $\theta(p) = 0$ for all $p < p_c$ and $\theta(p) > 0$ for all $p > p_c$, here $\theta(p) = \mathbb{P}_p(C = \infty)$ is the chance of an infinite cluster at the origin, see (1.1) for the precise definition. So if we imagine this transition to happen in a "smooth way", we would expect $\theta(p)$ to go to zero for $p \downarrow p_c$. Indeed it is assumed that the behaviour of $\theta(p)$ follows a power law with a *scaling exponent* β in one of the following senses:

1. *logarithmic form*:
$$\lim_{p \downarrow p_c} \frac{\log \theta(p)}{\log(p - p_c)} = \beta, \tag{4.1}$$

2. *bounded-ratios form*: There are constants $0 < c_1 < c_2 < \infty$ such that, uniformly for $p \geq p_c$
$$c_1(p - p_c)^\beta \leq \theta(p) \leq c_2(p - p_c)^\beta, \tag{4.2}$$

3. *asymptotic form*: There is a constant c such that, as $p \downarrow p_c$
$$\theta(p) = c(p - p_c)^\beta(1 + o(1)). \tag{4.3}$$

The equations (4.1) - (4.3) are sometimes summarized by simply writing

$$\theta(p) \sim (p - p_c)^\beta \qquad \text{as} \quad p \downarrow p_c. \tag{4.4}$$

and then noting which of the above forms is meant. Now the somewhat surprising conjecture is that (4.4) is valid for **any** graph $\mathbb{G}$ in one of the three senses defined above. That means the actual value of β may depend on $\mathbb{G}$ but the qualitative behaviour of the model near criticality is quite insensitive to the underlying graph structure. This concept is called *universality* and it has been one of the main fields of interest within percolation theory in the past decades. Note that the assumption of universality comes from numerical observations and considerations within statistical physics, where percolation is a paradigm model. So, besides some special cases, there is no rigorous proof for the existence of a scaling exponent β. Actually even a slightly weaker claim, which is the continuity of $p \mapsto \theta(p)$ has not been proved in general yet. We will look at this "smooth transition" property for mean-weight paths in the next chapter. For a more detailed summary of scaling behaviour and more examples of scaling exponents in classical percolation theory see [15] section 1.2.1.

4.2 Scaling Behaviour in the Complete Graph Case

Now let us, again, look at $\mathcal{W}_n$, the weighted, complete graph on n vertices. The basic question we examined in Chapter 3 was, given a $c \in \mathbb{R}$, what is the maximal size of a path respectively subtree of $\mathcal{W}_n$ such that the average weight is below c. From Theorem 3.2 we know that at the critical point $c = e^{-1}$, the maximal path size makes the transition from $o(n)$ to $\Theta(n)$. Theorem 3.6 gives the same results for trees instead of paths but does not provide the exact critical value $c(0)$.
So it seems natural to define

$$\delta(c) = \lim_{n\to\infty} \mathbb{E}\left[\max\{n^{-1}|\pi| : \pi \text{ is a path in } \mathcal{W}_n,\ |\pi|^{-1}\omega(\pi) \le c\}\right] \tag{4.5}$$

$$\tilde{\delta}(c) = \lim_{n\to\infty} \mathbb{E}\left[\max\{n^{-1}|t| : t \text{ is a tree in } \mathcal{W}_n,\ |t|^{-1}\omega(t) \le c\}\right] \tag{4.6}$$

as deterministic functions in c. Because of the Theorems 3.2 and 3.6, we know that $\delta(c) = 0$ for all $c \le e^{-1}$ and $\tilde{\delta}(c) = 0$ for all $c \le c(0)$. As we expect a "smooth" transition, we would assume that

$$\delta(c) \longrightarrow 0 \qquad \text{as } c \downarrow e^{-1}$$

and

$$\tilde{\delta}(c) \longrightarrow 0 \qquad \text{as } c \downarrow c(0).$$

But we could go even further and assume that these transitions satisfy a power law in the sense of (4.4):

$$\delta(c) \sim (c - e^{-1})^{\beta} \qquad \text{as } c \downarrow e^{-1} \tag{4.7}$$

$$\tilde{\delta}(c) \sim (c - c(0))^{\tilde{\beta}} \qquad \text{as } c \downarrow c(0) \tag{4.8}$$

for some scaling exponents β and $\tilde{\beta}$.
Indeed, Aldous made this conjecture in his paper [2] and also provided a convincing, though non-rigorous numerical argument, which we want to present in the following.

4.2.1 The First Attempt

The approach we will use is known as the *cavity method* and it stems from statistical physics. In their 2005 paper [3] Aldous and Bandyopadhyay give a very nice overview of this method:

Start with a combinatorial optimization problem over some size-n random structure, then follow these steps:

i) Formulate a "size-∞" random structure which is the $n \to \infty$ limit in the sense of local weak convergence.

ii) Define the relevant quantities on the size-∞ structure.

iii) Formulate a corresponding combinatorial optimization problem on the size-∞ structure.

iv) If the size-∞ structure is treelike, observe that the relevant quantities satisfy a problem specific RDE.

v) Solve the RDE and use the unique solution to find the value of the optimization problem on the size-∞ structure.

vi) Show that you can transfer the solution of the size-∞ problem back to a feasible solution of the size-n problem with almost the same cost.

So the main idea is to consider the model $\mathcal{W}_n$ in the limit $n \to \infty$. We already identified this limit model as the Poisson weighted infinite tree $\mathcal{T}^\infty$ and we have already seen that it is somewhat better tractable than $\mathcal{W}_n$. Working within $\mathcal{T}^\infty$ enables us to formulate the search for the maximal size path respectively tree as an optimization problem which can be solved using Lagrange multipliers. Unfortunately this is also one of the reasons why this method is non-rigorous as it seems impossible to find the transformation mentioned in step vi).

We will concentrate on the path case in this thesis. For more details and also a similar treatment of the tree case, we refer to [2].

To begin with, it is convenient to define

$$\varepsilon(\delta) = \lim_{n\to\infty} \mathbb{E}\left[\min\{|\pi|^{-1}\omega(\pi) : \pi \text{ is a path in } \mathcal{W}_n,\ |\pi| \geq \delta n\}\right] \tag{4.9}$$

and to work with $\varepsilon(\delta)$ instead of (4.5). The analogous conjecture to (4.7) is

$$\varepsilon(\delta) \sim \delta^{\alpha} + e^{-1} \qquad \text{as } \delta \downarrow 0 \tag{4.10}$$

with $\alpha = \frac{1}{\beta}$.

Now we can start to follow the steps of the cavity method.

Step i):
Let's again have a closer look at the PWIT model. From Section 3.1 we know the following: Choose an arbitrary vertex of $\mathcal{W}_n$ as the root, denote it with x_0 and fix $r \in \mathbb{R}$. Define the $r-$neighbourhood of x_0 as the configuration of points that are connected to x_0 via a path with a total weight of less than r. Then for $n \to \infty$, the $r-$neighbourhood of x_0 in $\mathcal{W}_n$ converges in distribution to the $r-$neighbourhood of the root in $\mathcal{T}^{\infty}$. This concept of convergence is called local weak convergence.

Step ii):
As we want to study paths π with a length of order n in $\mathcal{W}_n$, we need to know what is their analogon in $\mathcal{T}^{\infty}$. If we fix $r \in \mathbb{R}$ and just consider the $r-$ neighbourhood around the root again, the answer seems intuitively clear: We will see several vertex disjoint fragments of the path crossing through the $r-$window around the root. We will very likely not see any of the two endpoints of the path as this probability goes to 0 with a fixed r. Look at Figure 4.1 for a visualisation.

Step iii):
We are now ready to define the optimization problem on $\mathcal{T}^{\infty}$. We write $\overline{\pi} = \{\pi_1, \pi_2, \dots\}$ for a family of vertex disjoint doubly infinite paths in $\mathcal{T}^{\infty}$ and we want to divide those families in three sets

i $\mathcal{E}_0$ is the set of such families for which no path goes through the root.

ii $\mathcal{E}_2$ is the set of such families for which some path goes through the root.

iii $\mathcal{E}_1$ is the set of such families where, in addition to the doubly infinite paths, there exists exactly one singly infinite path. This path starts at the root.

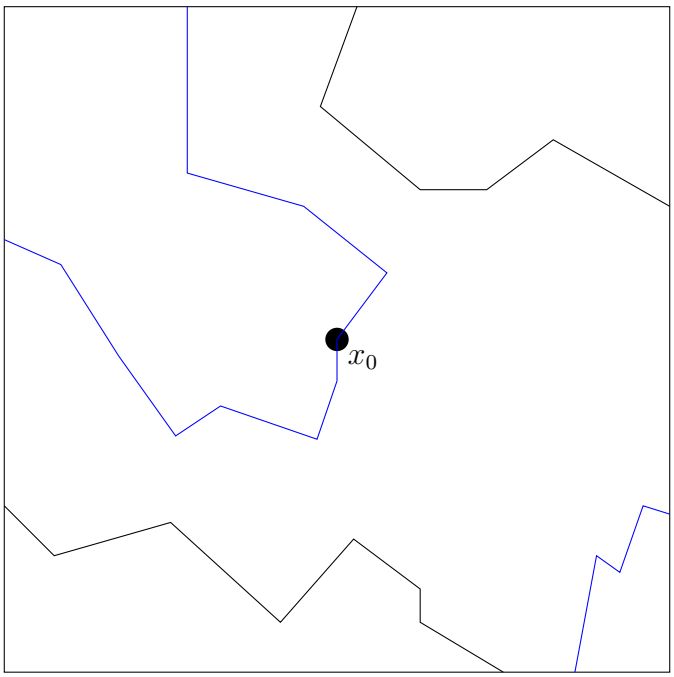


Figure 4.1: A finite window around the root. The black family of paths belongs to $\mathcal{E}_0$, whereas the blue family is in $\mathcal{E}_2$.

Considering the following objective function

$$b(\overline{\pi}) = c|\{v : v \text{ is a vertex in some } \pi_i\}| - \sum_{e:e \text{ is an edge in some } \pi_i} \omega_e$$

with the corresponding optimization problem

$$\text{maximize } b(\overline{\pi}) \text{ with } \overline{\pi} \in \mathcal{E}_0 \cup \mathcal{E}_2 \tag{4.11}$$

will lead to the optimization problem on $\mathcal{T}^\infty$ that, hopefully, yields information about the $\mathcal{W}_n$ case.
To see the motivation behind this particular objective function it is helpful to notice that our actual goal, which is finding $\varepsilon(\delta)$ or in other words

$$\min_{\pi:\pi \text{ path in } \mathcal{W}_n} \frac{\omega(\pi)}{|\pi|} \qquad \text{subject to } |\pi| \geq \delta n$$

can be reformulated, including the Lagrange multiplier c, as

$$\max_{\pi:\pi \text{ path in } \mathcal{W}_n} \frac{|\pi|}{n} c - \frac{\omega(\pi)}{|\pi|}. \tag{4.12}$$

Now (4.11) is clearly just the corresponding problem in the PWIT. Equation (4.12) has a random solution $\pi_n(c)$ and we suppose that

$$n^{-1}\mathbb{E}\left[|\pi_n(c)|\right] \overset{n\to\infty}{\longrightarrow} \delta(c) \tag{4.13}$$

and

$$\mathbb{E}\left[\frac{\omega(\pi_n(c))}{\pi_n(c)}\right] \overset{n\to\infty}{\longrightarrow} \varepsilon(c). \tag{4.14}$$

Step iv):
Unfortunately, working directly with (4.11) is difficult, as $b(\overline{\pi})$ is the difference of two sums, each having value ∞. Instead we want to consider the following quantities

$$X = \max_{\overline{\pi}\in\mathcal{E}_1} b(\overline{\pi}) - \max_{\overline{\pi}\in\mathcal{E}_0} b(\overline{\pi}) \tag{4.15}$$

$$Z = \max_{\overline{\pi}\in\mathcal{E}_2} b(\overline{\pi}) - \max_{\overline{\pi}\in\mathcal{E}_0} b(\overline{\pi}) \tag{4.16}$$

and see if they follow any RDE. To do so, we want to note an important property of $\mathcal{T}^\infty$. For each child x_{0i} of the root, the subtree T_i which consists of x_{0i} and all descendants thereof has the same distribution as $\mathcal{T}^\infty$ itself. Furthermore the different subtrees T_i are independent for all i.
As a consequence, the random pairs (X_i, Z_i) which are defined exactly as (4.15) and (4.16) but on T_i are distributed as (X, Z) for all i and independent of each other.

Lemma 4.1. The distribution of X from (4.15) satisfies the recursive distributional equation

$$X \overset{d}{=} \max_i \left(c - \omega_{e_i} + X - Z^+\right) \tag{4.17}$$

where e_i denotes the edge between x_0 and x_{0i}. At this point, we recall that the weights of the edges $(\omega_{e_i}, 1 \le i < \infty)$ originating at x_0 are the points of a Poisson process with rate 1.

Proof. : Consider $\overline{\pi}_1$ and $\overline{\pi}_0$, the families in $\mathcal{E}_1$ and $\mathcal{E}_0$, where the maxima in the definition of (4.15) are attained. Suppose $\overline{\pi}_1$ contains an edge from x_0 to x_{0i}. It follows that $\overline{\pi}_0$ and $\overline{\pi}_1$ are identical on all subtrees T_j with $j \neq i$ or in other words, everywhere except for T_i and the edge e_i which links x_0 to x_{0i}. So looking at the tree structure, we realise that

$$X = \max_{\overline{\pi}\in\mathcal{E}_1(i)} b(\overline{\pi}) - \max_{\overline{\pi}\in\mathcal{E}_2(i)\cup\mathcal{E}_0(i)} b(\overline{\pi})$$

here $\mathcal{E}_1(i)$, $\mathcal{E}_2(i)$ respectively $\mathcal{E}_0(i)$ denote the sets $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_0$ on T_i. But it is also obvious that by definition

$$\begin{aligned} X_i &= \max_{\overline{\pi}\in\mathcal{E}_1(i)} b(\overline{\pi}) - \max_{\overline{\pi}\in\mathcal{E}_0(i)} b(\overline{\pi}) \\ Z_i^+ &= \max_{\overline{\pi}\in\mathcal{E}_2(i)\cup\mathcal{E}_0(i)} b(\overline{\pi}) - \max_{\overline{\pi}\in\mathcal{E}_0(i)} b(\overline{\pi}). \end{aligned}$$

So taking into account the edge e_i, we see that $X \overset{d}{=} \max_i(c - \omega_{e_i} + X_i - Z_i^+)$ and since we already noted that (X_i, Z_i) has the same distribution as (X, Z), the lemma follows. □

The next lemma tells us that we can find a similar recursion for Z.

Lemma 4.2. The distribution of Z from (4.16) satisfies the recursive distributional equation

$$Z \stackrel{d}{=} \max_i \left(c - \omega_{e_i} + X - Z^+\right) + \max_i{}^{[2]} \left(c - \omega_{e_i} + X - Z^+\right) \tag{4.18}$$

where $\max_i^{[2]}$ denotes the second maximum.

Proof. : If we interpret the doubly infinite path through the root as two singly infinite paths starting in the root, we can use the same argumentation as in the proof of Lemma 4.1 except that there is obviously just one maximum, which is why the second path has to be chosen as the second maximum. □

From (4.13) we see that $\delta(c)$ is simply the proportion of vertices in the optimal solution to (4.11). If we consider again the procedure that took us from $\mathcal{W}_n$ to the $\mathcal{T}^\infty$, we realise that the root was chosen at random in $\mathcal{W}_n$, so the root is also a typical vertex of $\mathcal{T}^\infty$. This means we can identify expectations at the root with averages over all vertices. Hence $\delta(c)$ is just the probability that the optimal $\overline{\pi}$ in (4.11) contains a path trough the root. This happens if and only if $Z > 0$, which means that

$$\delta(c) = P(Z > 0). \tag{4.19}$$

To derive an equation for $\varepsilon(c)$, we note that, again by the "typical vertex" property the mean edge length over all edges in the optimal family equals the mean edge length over the edges at the root in the optimal family, conditioned on the root being used. As already mentioned, using the root means $Z > 0$ and in that case the lengths of the two edges in the root are $\omega(e_I)$ and $\omega(e_J)$, where

$$I = \arg\max_i (c - \omega_{e_i} + X_i - Z_i^+)$$
$$J = \arg\max_i{}^{[2]} (c - \omega_{e_i} + X_i - Z_i^+).$$

In conclusion we get

$$\varepsilon(c) = \frac{\mathbb{E}\left[\frac{1}{2}(\omega_{e_I} + \omega_{e_J})\mathbb{1}_{Z>0}\right]}{\delta(c)}. \tag{4.20}$$

Step v):
Solving RDEs is a quite hard task. However, a pretty good numerical approximation can be achieved relatively easy by using a bootstrap Monte Carlo method, similarly as we did in Chapter 2.1.4. Just as Aldous did in [2], we represented the distribution of (X, Z) by 1000000 points and iterated the RDE 1000 times while considering only the first 20 points of the Poisson process $(\omega_{e_i}, 1 \leq i < \infty)$. The results of our simulation are shown in Table 4.1. In Figure 4.2, we plotted the δ-ε pairs from our simulations and in red added a plot of the function $\varepsilon = 0.31\delta^{\frac{1}{3}} + e^{-1}$ and looking at the data, it seems as

the claim of the scaling behaviour in Equation (4.10) and hence also Equation (4.7) is validated with $\beta = 3$ and a constant of proportionality of 0.31.

We noted before that the transformation back to the size-n problem was missing. That means, we didn't actually proof the existence of the postulated scaling exponent. Indeed, despite the promising results from this non-rigorous method, the claim about the scaling behaviour of paths is wrong as we shall see in the next section.

c	δ	ε
0.5	0.0250	0.4613
0.48	0.0171	0.4491
0.46	0.0105	0.4361
0.44	0.0053	0.4219
0.43	0.0035	0.4152
0.42	0.0020	0.4068
0.41	0.0011	0.3999

Table 4.1: The simulated values of δ and ε, as well as the corresponding values of c.

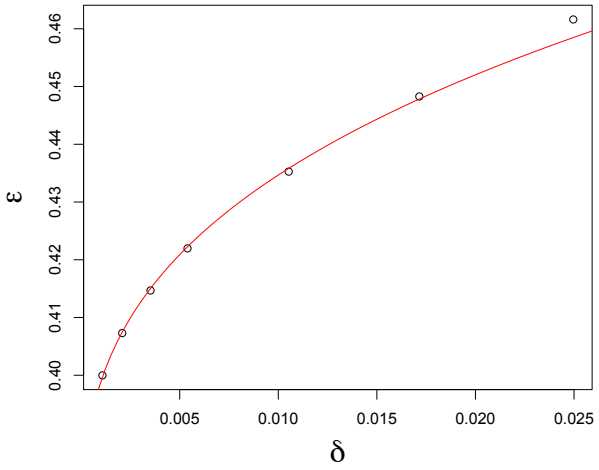


Figure 4.2: The simulated values (circles) and the postulated scaling behaviour (red line).

4.2.2 Recent Results on the Scaling Window and Phase Transitions

As already mentioned, the conjectures about the scaling behaviour of longest paths in $\mathcal{W}_n$ are not correct. Indeed there is a lot more complexity involved and it even turns out that there is yet another phase transition in the supercritical regime. In the following we want to present the results of Ding about the critical regime [7] (2013) as well as the results of Ding and Goswami about the near-supercritical regime [8] (2015).

While the function δ we examined previously incorporates the behaviour of $n^{-1}L_\pi(n, c)$, Ding chose a different scaling and decided to look at $\log(n)^{-3}L_\pi(n, c)$ instead. Then he was able to show (see [7] Theorem 1.1) that if $c(n)$ is chosen in a window of $\pm a \log(n)^{-2}$ around e^{-1} for a certain constant $a > 0$, the probability that $\log(n)^{-3}L_\pi(n, c(n))$ is in some interval $[k, K]$ for constants $k, K > 0$, converges to 1 as $n \to \infty$. At this point we want to remark that we use the notation $c(n)$ to indicate that we are not dealing with an apriori fixed constant but rather choose $c(n)$ dynamically, i.e. it varies as n gets large.

The proof looks at typical paths of fixed length and small average weights and then uses a second moment method to control the maximal deviation. For more details, we refer to Ding's work [7].

So we know that $L_\pi(n, c(n))$ is of order $\log(n)^3$ for $c(n)$ in an interval of length $a2(\log n)^{-n}$ around e^{-1}. But interestingly, there is also another constant $b > a$ and $K' \in \mathbb{R}^+$ such that

$$\mathbb{P}(n^{\frac{1}{4}} \leq L_\pi(n, c(n)) \leq K'n(c(n) - e^{-1})) \overset{n\to\infty}{\longrightarrow} 1$$

for $c(n) \geq e^{-1} + b(\log n)^{-2}$ (see [7] Theorem 1.2 for the statement and also the corresponding proof). In particular, this means that, in addition to the critical point e^{-1} where $L_\pi(n, c)$ makes the transition from $o(n)$ to $\Theta(n)$, there is another critical point somewhere in the shrinking interval $e^{-1} + [a(\log n)^{-2}, b(\log n)^{-2}]$ where $L_\pi(n, c)$ makes the transition from $\Theta((\log n)^3)$ to $\Theta(n)$.

We see that in the critical regime, with the "dynamical" threshold $c(n)$, the search for the longest paths with small average weight offers a pretty rich and diversified scaling behaviour. Now let's focus on the near-supercritical scaling laws again. In the preceding Section 4.2.1 we have seen a non-rigorous argumentation for the existence of a scaling exponent β such that

$$\frac{L_\pi(n, c)}{n} \sim (c - e^{-1})^\beta.$$

This conjecture however, was corrected by Ding and Goswami (see [8] Theorem 1.1) by proving the existence of absolute constants $k^* > 0$, $K^* > 0$ and $\eta^* > 0$ such that for $c = e^{-1} + \eta$ with $\eta \leq \eta^*$ it holds

$$\lim_{n\to\infty} \mathbb{P}(ne^{\frac{-K^*}{\sqrt{\eta}}} \leq L_\pi(n, c) \leq ne^{\frac{-k^*}{\sqrt{\eta}}}) = 1$$

or in other words

$$\frac{L_\pi(n, c)}{n} \sim e^{\frac{-K^*}{\sqrt{\eta}}}.$$

To summarize the above, in total we have the following results for the scaling behaviour of $L_\pi(n, c)$, the longest path length with average edge weight below c in $\mathcal{W}_n$:

- There is a phase transition at e^{-1}, which means for $c < e^{-1}$ fixed, $L(n, c)$ is of order $o(n)$ whereas for $c > e^{-1}$ fixed, $L(n, c)$ is of order $\Theta(n)$.
- There are absolute constants $a, k, K > 0$ such that if we choose $c(n)$ inside the shrinking intervall $e^{-1} \pm a(\log n)^{-2}$ we will have $L_\pi(n, c(n)) \in [k(\log n)^3, K(\log n)^3]$ with probability 1 as $n \to \infty$.
- There are absolute constants $b > a$, K' such that if we choose $c(n) > e^{-1} + b(\log n)^{-2}$ we will have $L_\pi(n, c(n)) \in \left[n^{\frac{1}{4}}, K'n(c(n) - e^{-1})\right]$ with probability 1 as $n \to \infty$.

- There are absolute constants k^*, K^*, η^* such that if we choose $c = e^{-1} + \eta$ fixed with $0 < \eta < \eta^*$ we will have $L_\pi(n, c) \in \left[ne^{\frac{-K^*}{\sqrt{\eta}}}, ne^{\frac{-k^*}{\sqrt{\eta}}}\right]$ with probability 1 as $n \to \infty$.

A somewhat more visual overview of the different regimes is given in Figure 4.3.

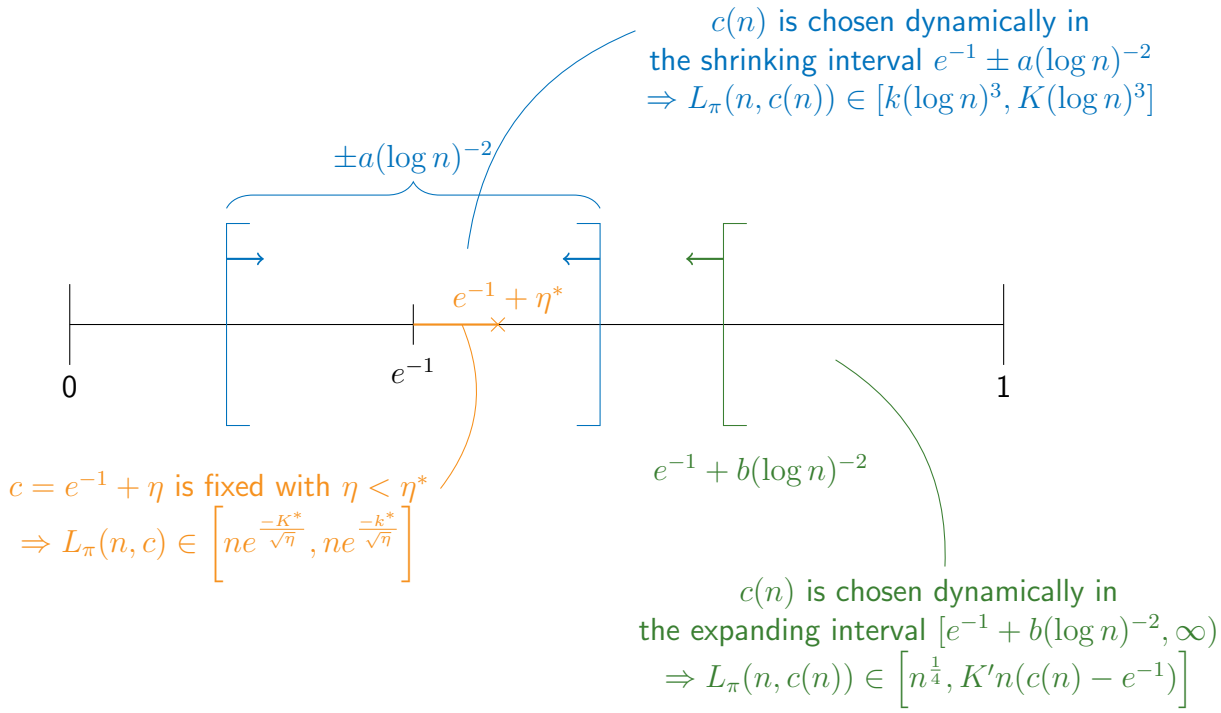


Figure 4.3: An illustration of the scaling behaviour in the different regimes. Note that in the green interval it is possible and in the blue interval even necessary that we choose $c(n)$ arbitrarily close to e^{-1} as n goes to infinity.

Bibliography

[1] Aldous, D. : On the critical value for 'percolation' of minimum-weight trees in the mean-field distance model. In: *Combinatorics, Probability and Computing* 7 (1998), Nr. 01, S. 1–10

[2] Aldous, D. : Percolation–like scaling exponents for minimal paths and trees in the stochastic mean field model. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* Bd. 461 The Royal Society, 2005, S. 825–838

[3] Aldous, D. ; Bandyopadhyay, A. : A survey of max-type recursive distributional equations. In: *Annals of Applied Probability* (2005), S. 1047–1110

[4] Aldous, D. ; Pitman, J. : The asymptotic speed and shape of a particle system. In: *Probability Statistics and Analysis* (1983), S. 1–23

[5] Biggins, J. : The first-and last-birth problems for a multitype age-dependent branching process. In: *Advances in Applied Probability* (1976), S. 446–459

[6] Broadbent, S. ; Hammersley, J. : Percolation processes. In: *Mathematical Proceedings of the Cambridge Philosophical Society* Bd. 53 Cambridge Univ Press, 1957, S. 629–641

[7] Ding, J. : Scaling window for mean-field percolation of averages. In: *The Annals of Probability* 41 (2013), Nr. 6, S. 4407–4427

[8] Ding, J. ; Goswami, S. : Percolation of averages in the stochastic mean field model: the near-supercritical regime. In: *arXiv preprint arXiv:1501.03579* (2015)

[9] Doob, J. : *Classical Potential Theory and Its Probabilistic Counterpart: Advanced Problems*. Bd. 262. Springer Science & Business Media, 2012

[10] Evans, S. : Polar and nonpolar sets for a tree indexed process. In: *The Annals of Probability* (1992), S. 579–590

[11] Frieze, A. : On the value of a random minimum spanning tree problem. In: *Discrete Applied Mathematics* 10 (1985), Nr. 1, S. 47–56

[12] Grimmett, G. : *What is Percolation?* Springer, 1999

[13] Hammersley, J. : Postulates for subadditive processes. In: *The Annals of Probability* (1974), S. 652–680

[14] Hoffman, C. ; Johnson, T. ; Junge, M. : Recurrence and transience for the frog model on trees. In: *arXiv preprint arXiv:1404.6238* (2014)

[15] Hofstad, R. van d.: Percolation and random graphs. In: *New perspectives in stochastic geometry* (2010), S. 173–247

[16] Hu, C.-K. : Percolation, clusters, and phase transitions in spin models. In: *Physical Review B* 29 (1984), Nr. 9, S. 5103

[17] Kingman, J. : The first birth problem for an age-dependent branching process. In: *The Annals of Probability* (1975), S. 790–801

[18] Pemantle, R. ; Peres, Y. : Galton-Watson trees with the same mean have the same polar sets. In: *The Annals of Probability* (1995), S. 1102–1124

[19] Peres, Y. : Probability on trees: an introductory climb. In: *Lectures on probability theory and statistics*. Springer, 1999, S. 193–280

[20] Shi, Z. : Random walks and trees. In: *ESAIM: Proceedings* Bd. 31 EDP Sciences, 2011, S. 1–39

[21] Stanley, R. : *Enumerative combinatorics. Vol. 2, volume 62 of Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999