Weighted distances in the configuration model with conservative edge-weights

Adriaans, E.L.A.

Award date: 2017
Weighted distances in the configuration model with conservative edge-weights

Master Thesis

Erwin Adriaans

Supervisor:
dr. Júlia Komjáthy

first version

Eindhoven, August 2017
Abstract

In this thesis we study first-passage percolation in the configuration model with empirical degree distribution satisfying the bounds of a power-law with exponent $\tau \in (2, 3)$. We assign weights to the edges of the graph with a condition on the distribution of the weights such that the age-dependent branching process described by the degree and weight distribution is conservative. We investigate the shortest weighted distance between two uniformly chosen vertices. We show that this distance divided by a certain sequence converges in probability to 1. This sequence tends to infinity with the amount of vertices. Additionally, this result still holds when the degrees are i.i.d. power-law random variables. At last, we show that the convergence in probability of the weighted distance holds in the erased configuration model as well.
Contents

1 Introduction iv
   1.1 Complex networks ........................................ iv
   1.2 Random Graphs ............................................. v
   1.3 First-passage percolation ................................ vii
   1.4 Configuration model ...................................... vii
   1.5 Universality classes of configuration models ............ viii
   1.6 Structure .................................................. ix

2 Model & Results x
   2.1 Notation .................................................. x
   2.2 The model ................................................ x
   2.3 Discussion and open problems ............................ xii
   2.4 Overview of the proof .................................... xiii

3 Exploration around two vertices xv
   3.1 Coupling of the exploration to a branching process .... xv

4 Percolation of the graph xxiii

5 Proof of the main Theorem xxvi
   5.1 Erased configuration model ............................... xxx

Bibliography xxxiii
Chapter 1

Introduction

Every logistic company wants to be the fastest, cheapest and deliver on time. In order to achieve this, the routes they are driving should be optimal, meaning they should be the least costly and fastest for them in order to be competitive. This does not only hold for logistic companies, also for, for example, companies that design electrical circuits, this needs to be done efficient and cheap but they have the additional condition that the circuits need to be reliable. These type of companies have a bigger influence on the network than a delivery company, since they are designing the circuits themselves. Whereas a delivery company is dependent on the government to supply the road network. A company that makes chips, designs the circuits in those themselves. They want these circuits to be reliable, when a connection breaks down, is the circuit still working? This is also where there is the connection to the logistic company, if they have the best routes but there is a traffic jam or construction work on a certain road than another route becomes the best choice. In these types of networks a broken/congested link is never good, but in other networks this could be valuable. Let us have a look at another model, where we don’t have interconnected electrical components or customers connected via roads, but where the nodes are people and the connections between them are mutual friendships, family relations or colleagues. This model is used among other things to model the spread of a disease. A person sees some of his/her connections more often than others, this can be seen as the time it takes to infect someone once he or she has become infected. The similarity with the road network, it that the time it takes for a person to infect someone else can be seen as the fastest route logistic companies want to find for their trucks. If we know the optimal routes through the human network, we know the time it takes for the disease to spread from one person to another. But what if we could, like a chip company, design the network such that it now would not be reliable, such that a disease cannot spread. So if we can take out nodes and all their connections, by e.g. isolating or vaccinating people, can we stop the spread of a disease after the outbreak? There are some overlapping properties between all these networks that could prove to be valuable, which is why we want to understand the underlying structure of these types of networks.

1.1 Complex networks

Complex networks are widely studied nowadays [2, 40, 43]. They turn out to be everywhere, so it is important to understand their properties and behavior. But what are they exactly or moreover, what is even a network? A definition gives the following statement for network: “a group or system of interconnected people or things”. This definition is very generic but that is because networks appear in many different fields. For example, a social network where the nodes are humans and the connections the social relations between them [31, 34], the Internet where routers and computers are connected through physical and wireless cables or the World-wide web with website connected via hyper links [8], the rail network in which the stations form the nodes and the railways the connections [25], even the human brain, where the nodes represent different specialized regions and weighted distances in the configuration model with conservative edge-weights
the connections represent the communication pathways [19], the network of movie actors that are connected when they acted in the same movie [4]. Since these types of networks have irregular structures and cannot be modeled by regular graphs because there is also randomness involved we call these types of networks complex networks. The organization and topology of a network affect various properties. For example how fast trucks can go through the road network, do parts of the network get isolated if edges break down or the spread of information through a social network.

This made people wonder if these networks are fundamentally random or if there is some design structure in these network. Then, a few parallel developments initiated the research in complex networks. At first there was a desire to understand and find the underlying architecture of complex networks. This shift in interest was helped by the rise of computers, along with the increased computing power, which handed the possibility to analyze large networks of millions of nodes. By which the realization came that they appeared in a lot of different fields. Using this increased computational power the topology was empirically analyzed to find statistical properties such as degree distributions and path lengths. A lot of complex networks turn out to share some common properties, i.e. there are large universality classes of networks. A common property is that there are many nodes of relatively high degrees, as first noted in [24]. In particular the degree distribution follows a power law and these types of networks were later on called scale free [3]. This means that the amount of nodes with $k$ connections, for large $k$, is proportional to $k^{-\tau}$ for some $\tau \geq 1$. This is encountered in the world-wide web [8], the movie-actor collaboration network [7], the network of citations of scientific publications [42]. The second property was popularized by Millgram [37] as: “everyone on this planet is separated from anyone else by only six people”. Despite the fact that there are 7 billion people, we are all just six others away from each other, this property was later shown to hold for the Facebook network as well [6]. This phenomena is known for networks as the small-world property and some networks even are ultra-small worlds. A network exhibits the small-world property if the minimal amount of connections to go from one node to another is of order $\log(n)$ or $\log\log(n)$ for ultra-small worlds, with $n$ the amount of nodes in the network. This effect is not only seen in social networks, but also in neurological networks like the brain [1, 19] or food webs [38]. A third common property is clustering, which can be explained as correlation between edges which was first pointed out by Watts and Strogatz [46]. Complex networks, like real-networks, can form communities. High clustering means that two vertices in the graph are more likely to be connected to one another when they have a common neighbor. This is a common feature in for example social networks, two friends of someone are likely to be friends as well mutually. The amount of clustering in a graph is measured by the clustering coefficient $C$, which is the empirical probability that two neighbors of a vertex are connected to each other. These big networks require a different analytical approach, since many questions that were solved in small networks are not relevant in much larger networks. Like the consequences of the breakdown of one connection, since these complex networks are so large that one connection has no influence in the connectedness of the network [20, 21]. Since the 1950s large-scale networks with no apparent design principles were described as random graphs, proposed as the simplest and most straightforward realization of a complex network. We use random graphs to create models of networks that can help us to understand the meaning of these properties, how they came to be as they are, and how they interact with one another. The extensive study in these models is done at a local level since they are generally too large to be studied globally. In the upcoming section we first introduce a graph and then give some random graph models.

1.2 Random Graphs

The most natural way to model a network from a mathematical point of view is to see this as a graph. Then, moreover, a random graph is in some way an extension to the common graph by adding randomness. The group of people or things we are modeling are the vertices and the connections or links between them are the edges in the graph. A graph $G = (V, E)$ is a collection of a set of vertices $V$ and a set of edges $E$. An edge is a pair of two vertices and generally undirected, also the graph model we investigate in this thesis uses undirected edges. Traditionally, these
networks used to be analyzed by graph theory in which the modeled medium would represent the
nodes and the mutual connections the edges. Regular graphs are studied in combinatorics and the
study of graphs is much older than to random graphs. The first paper that studied the topology
of a network, that is an analysis on the vertices and edges, was a paper by Euler in 1736 on the
seven bridges of Königsberg [16]. The big difference is that where graph theory focuses on regular
graphs, these large scale networks were described as random graphs that do not have any apparent
design structure. Regular can not well enough describe the structure of the complex networks as
described in the previous section. Thus new models were needed that could better describe this
complex behavior, which are nowadays called random graphs.
The study of random graphs is believed to have started in 1960 with a paper of Erdős and Rényi [23],
in which they were the first to show threshold phenomena in a graph with no apparent structure.
The model they used considers \( n \) vertices where any possible edge exists with fixed probability
\( p \) independently of the other edges, now known as the Erdős–Rényi random graph model. This
meant the start of the by now widely studied subject of complex networks. It was the first time
anyone showed properties of a graph without actually creating a graph. This basically captures
the essence of random graph theory, they describe universality classes of complex networks and
then find properties of these random graphs that also hold for the modeled complex networks. In
that time, by the absence of data on large networks, this graph was rarely tested in the real world.
As said in the section before, a commonly shared property of these models is that the degrees
follow a power-law distribution. Their model results in the degrees having only 'simple' Poissonian
distributions, but the great majority of real-world networks are complex and have therefore more
complex architectures. Then there was a period in which the research was relatively on a low level,
until the realization that many complex networks share common properties and the interest in
these rose. Despite that the Erdős–Rényi graph model has proven to give a great insight into the
structure of networks it also had some shortcomings, for example the degrees, as just noted, did
not represent most real-world networks. There were multiple models introduced that have some of
these properties of which we quickly introduce some to give some insight.

In 1998 Strogatz and Watts introduced their small-world network [46]. They noticed that many
real-world networks are a combination between regular and random graphs. Their model starts
with \( n \) vertices on a ring that are all connected to their \( k \) nearest neighbors, this is done to create
clustering in the graph. Then any edge is with probability \( p \) rewired to any other uniformly chosen
vertex. For \( p = 0 \) this gives a completely regular graph while for \( p = 1 \) this graph is completely
random. The main advantage of this model is that it combines high clustering with short path
lengths. The high clustering is an effect of each vertex being connected to its \( k \) nearest neighbors
and the rewiring of some edges to other vertices at random results in a graph structure where
many nodes are reached in a few steps.

The 1999 paper [24] showed that the empirical degrees follow a power-law distribution on
the Internet and the world-wide web and later studies found that this is also the case in other
varieties of complex networks. So the above described models did not describe these types of
complex networks well enough. Additionally, to cope with the fact that most real-world networks
are dynamic, Barabasi and Albert [2] introduced the BA-model or also known as the preferential
attachment model. In this graph vertices are added one by one and \( m \) edges are attached to
vertices already in the graph with a probability proportional to their degree. This means that the
vertices with a high degree have a higher probability of getting new edges, this principle is known
as ‘rich get richer’.

What all these models have in common is that they realize a degree sequence that follows
a certain distribution. The opposite to this would be a generalized random graph model that
generates random networks with a given degree sequence. The first to study a prescribed degree
sequence were Bender and Canfield [12] and later called the configuration model by Bollobás [18].
This is a model to describe large static networks with a prescribed degree sequence. The biggest
advantage of the configuration model is that we can use any degree distribution or sequence to
model a network. We studied this model, so we explain it in more detail after an introduction to
edge weights.
1.3 First-passage percolation

In the models just described the distance that two vertices are apart can be investigated, which is the minimal path length that connects these two vertices and is called the graph distance. In that case any edge represents a length of 1. In a real-life network, a distance of 1 could represent the user’s connections on LinkedIn or friends on Facebook, in this case an edge indicates these two people know each other. LinkedIn even shows how far away someone is from the user, where 2nd indicates that this person is connected with someone the user is connected to and 3rd means there is an additional connection necessary to reach this person etc. When we would be modeling the spread of information over time in a social network, simply a connection cannot tell us how long it takes before information reaches a certain person. To model time, money or distances in random graphs we add edge weights to the edges.

In a graph with edge weights there are two additional metrics defined on the graph, next to the graph distance we have the weighted distance and the hopcount. The weighted distance is the 'shortest' weighted path from one vertex to another and the hopcount is the number of edges on the shortest weighted path. First-passage percolation is used to study these three metrics and was first introduced by Hammersley and Welsh [26]. This can be seen as a flow, starting from a vertex, flowing through the edges at a rate equal to the respective edge-weight, the distance from this vertex to another vertex is the time it takes the front of the flow to reach the respective vertex.

There have been written multiple papers about first-passage percolation in random graphs and research to these metrics has produced surprising results. Already in 1999 Janson showed that in a complete graph the weighted distance between two given points is asymptotically \( \log n/n \) [32]. Additionally he shows that the hopcount is of order \( \log n \), which is quite surprising considering that the graph distance equals 1 for any pair of vertices.

The focus of our study is the configuration model with edge weights and we study the last universality class for the setting where the degrees follow a power-law. First we introduce the configuration model in more detail and later we give more results about first-passage percolation, but on this model.

1.4 Configuration model

Like we noted before, this model is relatively old compared to other models like small-world the network or the BA-model. The first paper written about this topic was in 1978 [12] and later called configuration model graph in [18]. The model uses a prescribed degree sequence \( d = (d_1, \ldots, d_n) \) for its \( n \) nodes. Although, given just a degree distribution \( P(d) \), we can generate the degree sequence for \( n \) nodes by sampling the degrees independently from the distribution \( P(d) \). Then, for \( i = 1, \ldots, n \), we assign \( d_i \) half-edges to node \( i \), then two half-edges are chosen uniformly at random and merged into an edge. In this graph self-loops and multiple-edges are allowed and hence the generated graph is called a multigraph. There is a variation to this model called the erased configuration model, in which after the creation of the normal configuration model, all self-loops are deleted and multiple-edges merged into one edge such that a simple graphs is created. The resulting graph is then called the erased configuration model.

The advantage of the model is that the degree sequence can be of any distribution. Where the Erdős-Rényi model has Poissonian degrees, the small-world has fixed degrees or a mix between fixed and binomial, depending on the model definition and the BA-model has power-law degrees. However, the model was defined before the invention of the clustering coefficient, so this only describes networks that do not show clustering. But this also has an advantage for the analysis of the model, since the probability that for a vertex the neighbors of neighbors are among the original neighbors of the vertex goes as \( n^{-1} \). This is the reason that the clustering coefficient is low, but because of this we can see the exploration around a vertex as a branching process (see e.g. [10, 27]), which turns out to be a good tool to find results regarding distances in the graph.

It is often seen as a disadvantage that this model fails to capture the phenomenon of clustering. However, Newman proposed a slight variation to the configuration model to include clustering [39].

Weighted distances in the configuration model with conservative edge-weights
CHAPTER 1. INTRODUCTION

This model does not start with just one prescribed degree sequence but two, say $d_1$ and $d_2$, where the first represents the number of half-edges and the second the number of ‘wedges’. Like before the half-edges are paired uniformly at random, the wedges are a bit different. As for them we uniformly choose three wedges and then pair the three vertices together to form a triangle. This way there is a positive probability that two of the neighbors are connected mutually. This model is however not studied thoroughly yet.

Since in many real-life networks the degree distributions follow a power-law distribution. We study the setting of the configuration model that matches this property. We investigate a model with power-law degrees with exponent between 2 and 3. That is we assume that the empirical way there is a positive probability that two of the neighbors are connected mutually. This model is  

This model does not start with just one prescribed degree sequence but two, say $d_1$ and $d_2$, where the first represents the number of half-edges and the second the number of ‘wedges’. Like before the half-edges are paired uniformly at random, the wedges are a bit different. As for them we uniformly choose three wedges and then pair the three vertices together to form a triangle. This way there is a positive probability that two of the neighbors are connected mutually. This model is however not studied thoroughly yet.

Since in many real-life networks the degree distributions follow a power-law distribution. We study the setting of the configuration model that matches this property. We investigate a model with power-law degrees with exponent between 2 and 3. That is we assume that the empirical way there is a positive probability that two of the neighbors are connected mutually. This model is  

This model does not start with just one prescribed degree sequence but two, say $d_1$ and $d_2$, where the first represents the number of half-edges and the second the number of ‘wedges’. Like before the half-edges are paired uniformly at random, the wedges are a bit different. As for them we uniformly choose three wedges and then pair the three vertices together to form a triangle. This way there is a positive probability that two of the neighbors are connected mutually. This model is however not studied thoroughly yet.

Since in many real-life networks the degree distributions follow a power-law distribution. We study the setting of the configuration model that matches this property. We investigate a model with power-law degrees with exponent between 2 and 3. That is we assume that the empirical way there is a positive probability that two of the neighbors are connected mutually. This model is however not studied thoroughly yet.  

1.5 Universality classes of configuration models.

This thesis is part of a line of research to find universality classes of configuration model random graphs. This research started by a sequel of three papers by Hofstad et al. [27, 29, 45], that distinguishes between three universality classes and gives results for the typical distances on the graph. The three considered universality classes were models with power-law degrees with exponent $\tau$ such that: the mean and variance of the degree are finite ($\tau > 3$), infinite variance but finite mean $(\tau \in (2,3))$ and infinite mean and variance $(\tau \in (1,2))$. Later was shown in [10] that there are actually two different universality classes for $\tau \in (2,3)$ depending on the weight distribution. For each of the universality classes we first review results on the graph distance and then we move on to results about weighted distances. The results obtained are typically asymptotic results in the sense that the number of nodes in the graph tends to infinity, with i.i.d. degrees satisfying a certain condition (e.g. a power-law with a certain exponent).

Let us quickly introduce some notation to show the results of some prior research. We write $d_G(u,v)$ for the graph distance, $d_L(u,v)$ for the weighted distance and $d_H(u,v)$ for the hopcount between vertices $u$ and $v$. These are the three main metrics studied on graphs to understand their topology. The graph distance is only dependent on the degree distribution, whereas the hopcount and weighted distance additionally depend on the the distribution of the edge-weights.

The here presented papers all consider the configuration model, with i.i.d. degrees following a power-law with exponent $\tau - 1$. We consider the three universality classes for the different values of $\tau$. First we consider the setting with $\tau \in (1,2)$, for this exponent the mean and variance of the degrees are infinite. The graph distance is with high probability at most 3 i.e., $P(d_H(u,v) = 2) = 1 - P(d_H(u,v) = 3) = p$ [45]. The most studied edge weights are i.i.d. exponential random variables with mean 1, as well as in this model. When these exponential weight are added, the weighted distance $d_L(u,v)$ converges in distribution to $E_1/D_1 + E_2/D_2$ [15], with $E_1$ exponential random variables with mean 1 and $D_1$ following the degree distribution. The hopcount converges in distribution to $2 + H_\infty$ [15], with $H_\infty$ some random variable.

The network in which $\tau > 3$, is essentially the opposite since the mean and variance of the degrees are both finite now. This results in totally different behavior for the considered metrics. In [27] it is shown that the graph distance grows as $\log n$, so this model is a small-world network. Later this result has been generalized to a wide class of random graphs models, including the Erdős-Rényi model [44]. If there are edge-weights added then $d_L(u,v) - c \log n$ converges in distribution to some limiting random variable, given explicitly, and the hopcount satisfies the CLT as $(d_H(u,v) - 1) / \alpha \log n$ converges to a standard normal random variable, for some $\alpha > 1$, these results were first shown for exponential mean 1 edge weights [13] and later for general edge-weights with a density function [14].

This thesis is part of a line of research to find universality classes of configuration model random graphs. This research started by a sequel of three papers by Hofstad et al. [27, 29, 45], that distinguishes between three universality classes and gives results for the typical distances on the graph. The three considered universality classes were models with power-law degrees with exponent $\tau$ such that: the mean and variance of the degree are finite ($\tau > 3$), infinite variance but finite mean $(\tau \in (2,3))$ and infinite mean and variance $(\tau \in (1,2))$. Later was shown in [10] that there are actually two different universality classes for $\tau \in (2,3)$ depending on the weight distribution. For each of the universality classes we first review results on the graph distance and then we move on to results about weighted distances. The results obtained are typically asymptotic results in the sense that the number of nodes in the graph tends to infinity, with i.i.d. degrees satisfying a certain condition (e.g. a power-law with a certain exponent).

Let us quickly introduce some notation to show the results of some prior research. We write $d_G(u,v)$ for the graph distance, $d_L(u,v)$ for the weighted distance and $d_H(u,v)$ for the hopcount between vertices $u$ and $v$. These are the three main metrics studied on graphs to understand their topology. The graph distance is only dependent on the degree distribution, whereas the hopcount and weighted distance additionally depend on the the distribution of the edge-weights.

The here presented papers all consider the configuration model, with i.i.d. degrees following a power-law with exponent $\tau - 1$. We consider the three universality classes for the different values of $\tau$. First we consider the setting with $\tau \in (1,2)$, for this exponent the mean and variance of the degrees are infinite. The graph distance is with high probability at most 3 i.e., $P(d_H(u,v) = 2) = 1 - P(d_H(u,v) = 3) = p$ [45]. The most studied edge weights are i.i.d. exponential random variables with mean 1, as well as in this model. When these exponential weight are added, the weighted distance $d_L(u,v)$ converges in distribution to $E_1/D_1 + E_2/D_2$ [15], with $E_1$ exponential random variables with mean 1 and $D_1$ following the degree distribution. The hopcount converges in distribution to $2 + H_\infty$ [15], with $H_\infty$ some random variable.

The network in which $\tau > 3$, is essentially the opposite since the mean and variance of the degrees are both finite now. This results in totally different behavior for the considered metrics. In [27] it is shown that the graph distance grows as $\log n$, so this model is a small-world network. Later this result has been generalized to a wide class of random graphs models, including the Erdős-Rényi model [44]. If there are edge-weights added then $d_L(u,v) - c \log n$ converges in distribution to some limiting random variable, given explicitly, and the hopcount satisfies the CLT as $(d_H(u,v) - 1) / \alpha \log n$ converges to a standard normal random variable, for some $\alpha > 1$, these results were first shown for exponential mean 1 edge weights [13] and later for general edge-weights with a density function [14].
CHAPTER 1. INTRODUCTION

The last universality class with finite mean and infinite variance ($\tau \in (2, 3)$) is somewhere in the middle between the previous two classes. The graph distance divided by $2 \log \log n$ converges in probability to $1/|\log(\tau - 2)|$ [29], so this model is an ultra-small world. A remarkable change happens when exponential, mean 1, i.i.d. edge weights are added to the graph, the hopcount then satisfies the CLT again as $(d_H(u, v) - \alpha \log n)/\alpha \log n$ but with a different $\alpha$ than before. So where the graph distance is of order $\log \log n$, the amount of edges on the least weighted path becomes of order $\log n$. The weighted distance, $d_L(u, v)$, converges in distribution to some limiting random variable again, but this time without the $\log n$ drift [13]. In [10] a generalization of [13] is given, for the setting where the weights do not satisfy (2.2.5) i.e., the sum is finite. Additionally it is shown that depending on the weight-distribution of the edge-weights, rather different behavior for the weighted distance is possible. The other regime, where the weights satisfy (2.2.5), is further studied in [11]. They prove that when the edge weights are of the form $1 + X$, with $X$ such that the branching process describing the exploration around a vertex is explosive, then both the weighted distance and hopcount minus $2 \log \log n/|\log(\tau - 2)|$ are tight sequences of random variables. In this thesis we further investigate this last case where the edge weights satisfy (2.2.5), in the highest generality, thus providing a full picture of weighted distances in the $\tau \in (2, 3)$ case.

1.6 Structure

In the next section we introduce the studied model, the configuration model, in a mathematical setting along with some assumptions and definitions. When we have these preliminaries we give the main theorem on the shortest weighted distance and additionally the alteration for the erased configuration model. In Section 3 and 4 we proof some lemma’s, necessary for the proof of the theorems. Section 3 considers the lemma’s that are related to the exploration around two uniformly chosen vertices in the graph, which is done in a breadth-first search manner and coupled to branching processes. The section after that we define the degree-dependent percolation and we show that, for some condition on the percolation function, the empirical degree sequence after percolation still satisfies the bounds of a power-law. Having acquired the necessary preliminaries, in the final section, Section 5, we give the proofs of the two theorems. The proof of the main theorem is given using two lemma’s, one for the lower bound and the other for the upper bound. After this proof is completed we can give the proof for the altered theorem for the erased configuration model.
Chapter 2
Model & Results

In this section we introduce the model used together with the assumptions. Then we give the main results of this paper followed by some open problems, but first we introduce some notation that is used in this paper.

2.1 Notation

We say that a sequence of events \((E_n)_{n \in \mathbb{N}}\) holds with high probability (whp) if \(\lim_{n \to \infty} P(E_n) = 1\).

For a sequence of random variables \((X_n)_{n \geq 1}\), we say that \(X_n\) converges in probability to a random variable \(X\) if for all \(\varepsilon > 0\),

\[
\lim_{n \to \infty} P(|X_n - X| > \varepsilon) = 0,
\]

shortly \(X_n \xrightarrow{P} X\).

For a distribution function \(F(x)\) the inverse is defined as

\[
F^{-1}(x) := \inf \{y \in \mathbb{R} : F(y) \geq x\}.
\]

We write lhs and rhs for the left-hand side and right-hand side, respectively.

2.2 The model

We consider the configuration model \(CM_n(d)\) on \(n\) vertices with degree sequence \(d = \{d_1, \ldots, d_n\}\).

Let \(H_n := \sum_{v \in [n]} d_v\), the sum of the degrees with \([n] := \{1, 2, \ldots, n\}\). To every vertex \(v \in [n]\) we assign \(d_v\) half-edges, then all these half-edges are paired uniformly at random and the resulting graph is denoted by \(CM_n(d)\). If the sum of degrees is not even we add an additional half-edge to vertex \(n\), this does not further influence the analysis. We assume that the empirical distribution function, defined as

\[
F_n(x) := \frac{1}{n} \sum_{v \in [n]} 1\{d_v \leq x\},
\]

satisfies the conditions for a power-law distribution, as given in the following assumption.

**Assumption 2.2.1** (Power-law tail behavior). For some \(\gamma \in (0, 1), C > 0, \tau \in (2, 3)\) and \(x \in [0, n^{(1-\varepsilon)/(\tau-1)}]\)

\[
\frac{1}{x^{\tau-1}} e^{-C(\log x)^{\gamma}} \leq 1 - F_n(x) \leq \frac{1}{x^{\tau-1}} e^{C(\log x)^{\gamma}}. \tag{2.2.1}
\]

In addition we assume.

**Assumption 2.2.2** (Minimal degree at least 2). \(\min_{v \in [n]} d_v \geq 2\).

Under Assumptions 2.2.1 and 2.2.2 by [28, Theorem 4.1] whp almost all vertices lie in the same connected component and the giant component is of size \(n(1 - o(1))\). For a random variable with distribution function \(F_n\) we write \(D_n\), which is the degree of an uniformly chosen vertex in \([n]\). We define \(B_n\) as the (size biased version of \(D_n\)-1).

\[
P(B_n = k) := \frac{k + 1}{H_n} \sum_{v \in [n]} 1\{d_v = k+1\} = \frac{k + 1}{\mathbb{E}[D_n]} P(D_n = k + 1). \tag{2.2.2}
\]
We write $F_{B_n}$ for the distribution function of $B_n$. As shown in [30], $F_{B_n}$ also satisfies a similar bound as (2.2.1) but with a different constant $C^*$ and exponent $\tau - 2$,

$$
\frac{1}{x^{\tau-2}}e^{-C^*(\log x)\gamma} \leq 1 - F_n(x) \leq \frac{1}{x^{\tau-2}}e^{C^*(\log x)\gamma}.
$$

(2.2.3)

To be able to state results later we use an additional assumption for distribution functions $F_n$ and $F_{B_n}$.

**Assumption 2.2.3 (Convergence to limiting distributions).** There exist distribution functions $F_D(x), F_B(x)$ such that for some $\kappa > 0$,

$$
\max\{d_{TV}(F_n, F), d_{TV}(F_{B_n}, F_B)\} \leq n^{-\kappa},
$$

where $d_{TV}(F, G) := \frac{1}{2} \sum_{x \in \mathbb{N}} |F(x+1) - F(x) - (G(x+1) - G(x))|$ is the total variation distance between two (discrete) probability measures.

We denote the random variables following the distribution $F_D$ and $F_B$ of Assumption 2.2.3 by $D$ and $B$. We now know that $D_n \xrightarrow{d} D$ and $B_n \xrightarrow{d} B$ since convergence of total variation distance implies weak convergence for discrete random variables. In addition $F_D$ and $F_B$ satisfy (2.2.1) with possibly different constants.

In this paper we study the shortest distance between two arbitrary vertices. For an edge $e$ the length is denoted by $L_e$ and we write $E$ for the set of all edges.

**Definition 2.2.4 (Shortest distance).** Let $u$ and $v$ be two vertices in the graph, then the graph-distance $d_G(u, v)$ is the number of edges used by the shortest path between $u$ and $v$. The shortest $L$-distance from $u$ to $v$ is defined as

$$
d_L(u, v) := \min_{\pi \in \mathbb{E}} \sum_{e \in \pi} L_e,
$$

(2.2.4)

where the minimum is taken over all paths connecting $u$ to $v$, so $\pi \subset E$. $d_L(u, v) = 0$ if $u = v$ and $d_L(u, v) = \infty$ if $u$ and $v$ are not connected. The hopcount, the number of edges on the shortest $L$-distance path, is denoted as $d_H(u, v)$.

The following theorem states the main result of this paper:

**Theorem 2.2.5 (Convergence of weighted distance).** Consider the configuration model $CM_n(d)$ with empirical degree distribution satisfying Assumptions 2.2.1-2.2.3 and let $u$ and $v$ be uniformly chosen vertices from $[n]$. Suppose the edge weights are i.i.d. from distribution $L$ with distribution function $F_L(x)$ that satisfies

$$
\sum_{k=1}^{\infty} F_L^{(-1)}(e^{-x_k}) = \infty
$$

(2.2.5)

Then

$$
d_L(u, v) \sqrt{\frac{\log \log n}{\log(\tau - 2)}} \xrightarrow{p} 1
$$

(2.2.6)

i.e., the lhs fraction converges in probability.

**Remark 2.2.6 (Convergence of the hopcount).** From the proof of Theorem 2.2.5 it follows that

$$
d_H(u, v) \sqrt{\frac{\log \log n}{\log(\tau - 2)}} \xrightarrow{d} 1.
$$

(2.2.7)

We refrain ourselves from giving the proof here.
In addition we show the result for the shortest $L$-distance also holds in the erased configuration model, this is a simple graph constructed out of the original configuration model (i.e. there are no multiple-edges and self-loops in this graph anymore). After the configuration model is constructed and the edge weights are assigned, the self-loops are deleted and for all multiple-edges one of the edges is chosen uniformly at random independent of the edge weights and the other edges are deleted. The resulting graph is called the erased configuration model, shortly denoted by $ECM_n(d)$, let us denote the shortest distance in this graph by $d^*_L(u,v)$.

**Theorem 2.2.7** (Convergence of weighted distance in the erased graph). Consider the erased configuration model $CM_n(d)$ with empirical degree distribution satisfying Assumptions 2.2.1-2.2.3 and let $u$ and $v$ be uniformly chosen from $[n]$. Suppose the edge weights are i.i.d. from distribution $L$ with distribution function $F_L(x)$ that satisfies (2.2.5). Then

$$d^*_L(u,v) \leq \sqrt{2 \sum_{i=1}^{\left\lfloor \log \log n \right\rfloor} F_L^{(-1)} \left( e^{-\left( \frac{\tau}{\tau - 1} \right)^i} \right)} \xrightarrow{p} 1.$$  \hspace{1cm} (2.2.8)

**Remark 2.2.8.** Theorems 2.2.5 and 2.2.7 both use the assumption that the empirical degree distribution satisfies the bounds for a power-law random variable i.e., Assumption 2.2.1. It can be shown that Assumption 2.2.1 is satisfied whp when the degrees are i.i.d. coming from a cumulative distribution function $F(x)$ satisfying the bounds of a power-law i.e.,

$$\frac{1}{x^{\gamma - 1}} e^{-C(\log x)^\gamma} \leq 1 - F(x) \leq \frac{1}{x^{\gamma - 1}} e^{C(\log x)^\gamma}.$$  \hspace{1cm} (2.2.9)

To show that Assumption 2.2.1 is satisfied when the degrees are i.i.d., we write $1 - F_n(x) = 1/n \sum_{i=1}^n 1_{\{d_i > x\}}$ and $X_n = \text{Bin}(n, P(d > x))$, from there it follows from the concentration of a binomial random variable and a union bound over all $x \in [1, n^{(1-\omega)/\gamma - 1})$.

By setting the edge weights to be deterministic and equal to 1 in Theorem 2.2.5, we obtain the following corollary.

**Corollary 2.2.9** (Convergence of graph distance). Consider the configuration model $CM_n(d)$ with empirical degree distribution satisfying Assumptions 2.2.1-2.2.3 and let $u$ and $v$ be uniformly chosen vertices from $[n]$. Then

$$d_G(u,v) \leq \sqrt{2 \left\lfloor \log \log n \right\rfloor} \xrightarrow{p} 1.$$  \hspace{1cm} (2.2.10)

Next we discuss the results of this paper and show how they could be improved upon in the future.

### 2.3 Discussion and open problems

If we zoom in on the configuration model, we see that a vertex locally has a tree-like structure. The exploration around a vertex can therefore be coupled to a branching process. Since we additionally have weights this is modeled as an age-dependent branching process, in which individuals have an i.i.d. lifetime and give birth to their offspring upon death. The offspring distribution follows a power-law with exponent $\tau \in (1, 2)$ due to the size-biased effect and the lifetime of an individual is the respective edge through which the vertex is reached. The exploration of a branching process was a long-standing open question. In 2013 Amini et al [5] gave a condition on the lifetime distribution for which an age-dependent branching process, in which the offspring follows a power-law distribution with exponent $\tau \in (1, 2)$, is explosive. In an unpublished note [35] simplified this criterion to the sum in (2.2.5) being finite.

We give universality results for the configuration model with power-law degrees with exponent $\tau \in (2, 3)$ and edge weights satisfying (2.2.5). There are similar results regarding the distance metrics on random graphs for the same setting but with a different power-law exponent: infinite
mean degrees, \( \tau \in (1, 2) \), is investigated in \([15, 45]\). The results they obtained are that whp the graph distance is 2 or 3, the weighted distance converges to the sum of 2 random variables and the hopcount is 2 plus a random variable. In the configuration model with finite variance degrees, \( \tau > 3 \), that is studied in \([13, 14, 27, 44]\), the graph distance is of order \( \log n \), the weighted distance minus a constant times \( \log n \) converges to a random variable and the hopcount minus a constant times \( \log n \) and divided by this satisfies the central limit theorem. It is, however, unclear what happens when \( \tau = 2 \) or 3. By \([10]\) we know that there are two universality classes when \( \tau \) is between 2 and 3, depending on whether the sum in (2.2.5) is finite or not i.e., the describing age-dependent branching process is explosive. There are specific conditions for the age-dependent branching with power-law offspring distribution with \( \tau = 2 \) \([5]\), the setting in which \( \tau = 3 \) is however still unknown.

Theorem 2.2.5 does not give results about the fluctuations of the weighted distance around the sequence in the denominator on lhs in (2.2.6) and the same holds for the hopcount. This leads us to the following, open problem:

**Problem 2.3.1.** Consider the configuration model \( \text{CM}_n(d) \) with empirical degree distribution satisfying Assumptions 2.2.1-2.2.3 and let \( u \) and \( v \) be uniformly chosen vertices from \([n]\). Suppose the edge weights are i.i.d. from distribution \( L \) with distribution function \( F_L(x) \) that satisfies (2.2.5). Determine the conditions under which

\[
d_L(u, v) = 2 \sum_{i=1}^{[\log \log n]} F_L^{(-1)} \left( e^{-\left(\frac{1}{\tau} - 1\right) i} \right)
\]

is a tight sequence of random variables. Consequently

\[
d_H(u, v) = 2 \frac{\log \log n}{|\log(\tau - 2)|}
\]

is a tight sequence of random variables.

## 2.4 Overview of the proof

Next we give an overview of the main ideas used for the proof of Theorem 2.2.5. The proof consists of two parts, first we show the lower bound of the convergence in probability and then the upper bound.

### 2.4.0.1 Lower bound

For the proof of the lower bound we do an exploration in the neighborhoods of the vertices \( u \) and \( v \). This exploration is done in a breadth-first-search (BFS) manner. These explorations are coupled to two independent branching processes (BPs), which are Galton-Watson processes with the first generation having distribution function \( F_D \) and all further generations have distribution function \( F_B \) from Assumption 2.2.3. We explore around vertices \( u \) and \( v \) until the generation where whp the explorations can still be coupled to BPs. The lower bound for the shortest path between \( u \) and \( v \) is then the sum of the two shortest paths from \( u \) and \( v \) to their respective last generation. This path is always larger than or equal to the sum of the minimal edge weights in each generation. Then we show that this sum of minima is always larger than \((1 - \varepsilon)\) times the denominator of the lhs of (2.2.6) whp. Which completes the proof of the lower bound.

### 2.4.0.2 Upper bound

The upper bound requires more work to prove, for this part we additionally connect the two exploration processes. We start with explorations from \( u, v \) until we reach a vertex of degree at least \( K_n \), denoted as \( u_{K_n}, v_{K_n} \). We need to choose \( K_n \) carefully such that \( d_L(u, u_{K_n}) \) and \( d_L(v, v_{K_n}) \) are of negligible length compared to the denominator of the lhs of (2.2.6). Then we want to connect the vertices \( u_{K_n} \) and \( v_{K_n} \).
CHAPTER 2. MODEL & RESULTS

To obtain an upper bound for the weighted distance we perform degree dependent percolation on the edges of the graph. The idea for the percolation is that the higher the degree of a vertex the more edges we delete. Heuristically, if a vertex has more edges there are also more edges with small weights. We construct two paths starting from both vertices $u_{K_n}$ and $v_{K_n}$ to connect the vertices. To these paths we add vertices with increasing degrees, the paths then use edges with smaller and smaller edge weights. We go on until we reach vertices of sufficiently high degree such that they are connected whp. We delete edges depending on the degree of the vertices an edge is attached to and its weight. For this we extend the idea of [11] to this setting, they give an extension to the construction by Janson [33]. We define the percolation in terms of keeping an edge in stead of a half-edge. We construct a percolation function $p(d)$ in terms of the distribution function $\mathbb{P}(L \leq \xi_{d,d'})$, with $\xi_{d,d'}$ that depends on the degrees of the vertices and the weight distribution $L$. We keep an edge attached to vertices with degrees $d$ and $d'$ if the weight satisfies $L \leq \xi_{d,d'}$. If we start with a degree distribution satisfying (2.2.1) and weight distribution (2.2.5) we can choose $p(d)$ such that the new vertex distribution still follows a power law distribution satisfying (2.2.1).

To show this we alter the proof of [11, Proposition 2.1] such that it fits this setting. We thus obtain a graph with an empirical degree distribution satisfying (2.2.1).

Like we just noted, we want to construct a path starting from vertices $u_{K_n}$ and $v_{K_n}$ with increasing degrees along the path. Therefore we define a sequence $y_i$ with $y_0 = K_n$ and layers in the graphs, this is done in the same way as done in [11]. The layers $\Gamma_i$ are sets of vertices defined as $\Gamma_i := \{v \in [n] : d_v \geq y_i\}$, for $0 \leq i \leq b(n)$ with $b(n)$ an upper bound on the path length from $u_{K_n}$ to a vertex with degree larger than $n^{1/2}$. The layers are descending in size with the property that $\Gamma_{i+1} \subset \Gamma_i$. In addition we show that a vertex in $\Gamma_i$ is connected to a vertex in $\Gamma_{i+1}$ whp, moreover

$$\lim_{n \to \infty} \sum_{i=1}^{b(n)} \mathbb{P}(\exists u \in \Gamma_i, u \not\sim \Gamma_{i+1}) = 0.$$  

This gives an upper bound on the paths from $u_{K_n}$ and $v_{K_n}$ to vertices with degrees larger than $n^{1/2}$, this set of high degree vertices form a complete graph, which connects the two paths.

The percolated graph is equal in distribution to an induced subgraph of a new configuration model. Therefore anything we know about the configuration model, we can apply to the percolated graph. The just described layering structure is known to exist in $\text{CM}_{\alpha}(d)$ with power-law degrees, which means that it is also still present in the percolated graph.

First we do the exploration around the vertices $u$ and $v$ until we reach a vertex of a certain degree $K_n$, which whp has degree at least $K_n$ after percolation. Then we perform degree-dependent percolation on the graph. In the percolated graph we construct the path to connect the vertices $u_{K_n}$ and $v_{K_n}$. Thus we obtain,

$$d_L(u, v) \leq d_L(u, u_{K_n}) + d_L(v, v_{K_n}) + 2 \sum_{i=0}^{b(n)} \xi_{y_i, y_{i+1}}. \quad (2.4.1)$$

The first two terms on the rhs, coming from the branching processes, are negligible due to the choice of $K_n$. The last term, coming from the percolation threshold $\xi_{y_i, y_{i+1}}$, is an upper bound on the edge length between two consecutive layers $\Gamma_i$ and $\Gamma_{i+1}$, where the last term of the sum (for $i = b(n)$) represents the connection of the two paths in the highest layer.
Chapter 3

Exploration around two vertices

In this section we consider the configuration model $CM_n(d)$ with given degree sequence $d$ of which the empirical degree distribution satisfies (2.2.1). The goal of this section is to do an exploration around two vertices in the graph and couple these to two independent branching processes (BP). Then we first show that we can do these explorations until generation $\kappa_n \geq \log \log n/|\log(\tau - 2)|$ such that the coupling error still converges to zero. Then we give the maximum number of generations needed to reach a vertex with at least some degree $K_n$. At last, we give an upper- and lower bound for the length of the path from the source of the BP to its last generation. We refer to the source of a BP or exploration as the vertex from which we start the exploration. For this we consider two uniformly chosen vertices $u$ and $v$. We start with the coupling of breadth-first-search explorations in the neighborhood of vertices $u$ and $v$ which are coupled to branching process trees.

3.1 Coupling of the exploration to a branching process

First we explain how the exploration around the vertices $u$ and $v$ is coupled to branching processes. We do this in a similar way as done in [30]. Since the pairing of the half-edges is done uniformly at random, we are at any time allowed to pick any of the not-yet-paired half-edges and uniformly pair it to one of the remaining half-edges. Hence, we do the pairing such that we do an BFS exploration around the vertices $u$ and $v$. We start with pairing all the half-edges of $u$ and $v$ uniformly at random, then we pair all the half-edges of the neighbors of $u$ and $v$, and so on. The direct neighbors are at distance 1, the neighbors of the neighbors at distance 2 etc. By $U_t, V_t$ we denote the subgraphs consisting of vertices at a distance of at most $t$ from $u$ and $v$, respectively. [13, Proposition 4.7] (see also [9, Lemma 4.3] and [30, Lemma 2.2]) shows that that the number of vertices and their forward degrees $1$ in the exploration process can be coupled to i.i.d. degrees having distribution function $F_{B_n}$ from (2.2.2), as long as the total number of vertices in the exploration is not too large. This lemma is adjusted to our setting to obtain the following lemma.

Lemma 3.1.1 (Coupling error of the exploration process). Consider the configuration model $CM_n(d)$ with empirical degree distribution satisfying Assumptions 2.2.1-2.2.3. The forward degrees $(X_k^{(n)})_{k \leq s_n}$ of the first $s_n$ newly discovered vertices in the exploration processes started from $u$ and $v$, that are randomly chosen from $[n]$, can be coupled to an i.i.d. sequence $B_{k,n}$ with distribution $B_n$ and i.i.d. sequence $B_k$ from distribution $B$. So there is a coupling $(X_k^{(n)}, B_{k,n}, B_k)_{k \leq s_n}$ with the following error bound

$$\mathbb{P}(\exists k \leq s_n, B_{k,n} \neq X_k^{(n)} \text{ or } B_{n,k} \neq B_k) \leq Cs_n^{(2\tau - 2 - 2\epsilon)/(\tau - \epsilon)}n^{(2 - \tau + \epsilon)/(\tau - \epsilon)} + Cs_n^{2(\epsilon + \epsilon)/(\tau - \epsilon)}n^{-1} + s_n n^{-\kappa}. \quad (3.1.1)$$

This leads to the following corollary

---

Forward degree means the number of newly available half-edges when a new vertex is discovered, that will be paired to new vertices later on.

Weighted distances in the configuration model with conservative edge-weights
CHAPTER 3. EXPLORATION AROUND TWO VERTICES

Corollary 3.1.2 (Whp coupling of the exploration to BPs). In the configuration model satisfying Assumption 2.2.1 and 2.2.3, let \( t \) be such that

\[
\left| U \cup V \right| \leq \min\left\{ n^{1-(1+\varepsilon)/(\tau-1)-\delta}/2, n^{-(\tau-2-2\varepsilon)/(\tau-1-\varepsilon), n^{(2-2\varepsilon)/(\tau-\varepsilon), n^{\kappa-\delta}} \right\}
\]

for some \( \delta > 0 \). Then \( (U_t, V_t) \) can be whp coupled to two i.i.d. BPs with generation sizes \((Z_k^{(u)}, Z_k^{(v)})\) \( k \geq 0 \) with distribution \( F_U \) for the offspring in the second and further generations, and with distribution \( F \) for the offspring in the first generation.

Now we state the proof of Lemma 3.1.1.

Lemma 2.2. The goal of the proof is to couple \((X_k^{(u)})_{k \leq s_n} \) to an i.i.d. sample of size \( s_n \) from distribution \( B_n \) as in (2.2.2) and then couple these to \( B \) as well and consecutively find an upper bound on the coupling error. In order to achieve this we sample vertices with and without replacement as in [9].

To sample i.i.d. random variables \( B_{k,n} \) from distribution \( B_n \), we uniformly at random choose a half-edge \( h_k \) independently from the total set of half-edges then we set \( B_{k,n} \) to be \( d_n(h_k) - 1 \), where \( v(h) \) denotes the vertex that half-edge \( h \) is attached to. To achieve the coupling of \((X_k^{(u)})_{k \leq s_n} \) to \( B_{k,n} \), we first couple random variables \( B_{k,n} \) to another sequence of random variables. We keep a list of vertices that are incident to the uniformly chosen half-edges: \( S_k := \{ u, d, d(h_1), \ldots, d(h_k) \} \) and we define \( C_k^{(u)} := d_n(h_k) - 1 \) if \( h_k \) is not in \( S_{k-1} \). We do this is to make sure that the forward degrees are i.i.d. The coupling of \( B_{k,n} \) and \( C_k^{(u)} \) fails when \( h_k \) \( s_{k-1} \). In addition we have to guarantee that the exploration process is a tree, so none of the half-edges from \( C_k^{(u)} \) can be paired to vertices in \( S_{k-1} \). Otherwise loops are created in the graph and the coupling to a branching process fails. Hence, the forward degree \( X_k^{(u)} \) equals \( C_k^{(u)} \) minus the edges paired to \( S_{k-1} \). The coupling to a BP fails if \( X_k^{(u)} < C_k^{(u)} \) since this implies cycles are created. At first we couple the initial vertices \( u, v \) to two i.i.d. copies of \( D_n \). We choose two vertices uniform from \( [n] \), they are equal with error probability \( 1/n \). The probability that their half-edges create cycles is at most \( \mathbb{E}[B_n]/n^{1+\varepsilon}/(\tau-1) / \mathcal{H}_n \). These error probabilities are later merged into the total coupling error, which we derive now first.

We focus first on the coupling between \( B_{k,n} \) and \( C_k^{(u)} \), which, like noted, fails when \( h_k \) \( S_{k-1} \). In [9], Lemma 4.3] this probability is bounded by

\[
\begin{aligned}
\mathbb{P}\left( B_{k,n} \neq C_k^{(u)} \mid S_{k-1} \right) & \leq \frac{1}{\mathcal{H}_n} \left( d_u + d_v + \sum_{s=1}^{k-1} (C_k^{(u)} + 1) \right) \\
& \leq \frac{1}{\mathcal{H}_n} \left( d_u + d_v + \sum_{s=1}^{k-1} (C_k^{(u)} + 1) \right).
\end{aligned}
\]

Using that \( \mathcal{H}_n = \mathbb{E}[D_n]/n \) and thus \( \text{order of } n \) and the maximal degree in the graph is \( n^{1+\varepsilon}/(\tau-1-1) \), the rhs in (3.1.3) is at most \( Ckn^{1+\varepsilon}/(\tau-1) \). Using an union bound and summing up to \( s_n \) yields

\[
\mathbb{P}\left( \exists k \leq s_n : B_{k,n} \neq C_k^{(u)} \mid S_{k-1} \right) \leq C_s^2 n^{\frac{1+\varepsilon}{1-\varepsilon}} - 1.
\]

Next we focus on the second error term, that is when \( X_k^{(u)} < C_k^{(u)} \). A bound for the probability of creating cycles at step \( k \) is given by [9] as

\[
\mathbb{P}(X_k^{(u)} < C_k^{(u)} \mid S_{k-1}) \leq \frac{C_k^{(u)}}{\mathcal{H}_n - 1 - d_u - d_v - \sum_{s=1}^{k-1} (C_k^{(u)} + 1)} \left( d_u + d_v + \sum_{s=1}^{k} (C_k^{(u)} + 1) \right).
\]

Then, conditioning on the event that \( \{ C_k^{(u)} = B_{k,n} \}_{k \leq s_n} \), we can replace \( C_k^{(u)} \) by \( B_{k,n} \) in (3.1.5). Then we apply a truncation argument for the random variables \( B_{k,n} \) to cope with their infinite expectation. We choose some truncation value \( T_n \) later, but such that \( s_n T_n = o(n) \). By this choice there exists some constant \( c > 0 \) such that the denominator on the rhs in (3.1.5) is at least \( cn \).

\[
\begin{aligned}
\mathbb{P}(\exists k \leq s_n : X_k^{(u)} \neq C_k^{(u)} \mid \forall k \leq s_n : B_{k,n} = C_k^{(u)}) & \leq \mathbb{P}(\exists k \leq s_n : B_{k,n} > T_n) \\
& + \frac{1}{cn} \sum_{k=1}^{s_n} \sum_{s=1}^{k} \mathbb{E}[B_{k,n} B_{j,n} 1\{ B_{k,n}, B_{j,n} \leq T_n \}].
\end{aligned}
\]
CHAPTER 3. EXPLORATION AROUND TWO VERTICES

For the first term on the rhs in (3.1.6) we have an upper bound using (2.2.2), and hence is at most $s_n T_n^{-2} e^{C_1^{(2)}(\log T_n)} \leq C s_n T_n^{-2+\varepsilon}$ for arbitrarily small $\varepsilon > 0$ and $T_n$ large enough. The second term is at most $C n^{-1} s_n T_n^2$ by the truncation argument. We set $T_n$ to minimize the total error:

$$T_n := \left( \frac{s_n}{n} \right)^{1/(\varepsilon - \tau)},$$

(3.1.7)

for which it holds that $T_n s_n = o(n)$ as long as $s_n = o(n)$. Using $T_n$ as in (3.1.7) for the total error term in (3.1.6) we obtain

$$\mathbb{P}\left( \exists k \leq s_n : X_k^{(n)} \neq C_k^{(n)} \mid \forall k \leq s_n B_{k,n} = C_k^{(n)} \right) \leq C s_n (2^{\varepsilon - 2 + \varepsilon}/(\tau - \varepsilon) n (2 - \varepsilon + \rho)/\tau).$$

(3.1.8)

As long as $s_n \leq n^{(2 - \varepsilon)/(2(1 - \varepsilon))}$ the rhs converges to zero. Recall that the coupling error for the first two vertices $u, v$ is of much smaller order and is hence merged into this coupling error. We extend the coupling by additionally coupling $B_k$ to $(X_k^{(n)}, B_{k,n})$. We use the optimal coupling that realizes the total variation distance between $B_{n}$ and $B$, conditioned on $X_k^{(n)} = B_{k,n} = j$.

$$P\left( B_k = j \mid X_k^{(n)} = B_{k,n} = j \right) := \frac{\min\{P(B = j), P(B_n = j)\}}{P(B_n = j)}.$$  

(3.1.9)

The coupling error then equals

$$P(B_k \neq B_{k,n}) = \sum_{j \geq 1} (1 - \min\{P(B = j), P(B_n = j)\}) = d_{TV}(F_B, F_{B_n}).$$  

(3.1.10)

Then the total coupling error is bounded by

$$\mathbb{P}\left( \exists k \leq s_n, B_{k,n} \neq B_k \right) \leq s_n d_{TV}(F_B, F_{B_n}) \leq s_n n^{-\kappa}.$$  

(3.1.11)

We can couple the two initial vertices $u, v$ to two i.i.d. copies of $D$ in a similar way with a coupling error of at most $d_{TV}(F, F_n)$. This finishes the proof. \qed

Now that we have this corollary and lemma we state some preliminaries we later use for the proof of the main theorem. The Lemma 3.1.1 shows that if we explore around the two initial vertices up to a restricted amount of vertices we can couple the two explorations to two BPs whp. From a theorem by Davies [22], that describes the growth rate of an infinite mean branching process, we know that we can write the size of the generations

$$Z_k^{(i)} = \exp \left\{ \left( \frac{1}{\tau - 2} \right)^k Y_k^{(i)} \right\},$$  

(3.1.12)

with $Y_k^{(i)}$ random variables with $Y_k^{(i)} \overset{a.s.}{\sim} Y^{(i)}$, where $i = u, v$. Using this theorem we can find the generation up to which we can explore such that we can couple the exploration to a BP whp.

**Lemma 3.1.3 (Last generation of the exploration).** Consider the configuration model $CM_n(d)$ with empirical degree distribution satisfying assumptions 2.2.1-2.2.3. Let $u$ and $v$ be two uniformly chosen vertices. Then we can couple the BFS-exploration around $u$ and $v$ to two BPs until generation

$$\kappa_n := \frac{\log \log n + \log (g\gamma / Y_n^{(i)})}{|\log (\tau - 2)|} - f_n,$$

(3.1.13)

where $f_n$ is the fractional part of the first term, such that the coupling error still converges to zero.

---

\[\text{Davies actually defines the random variables in terms on the generation sizes as } Y_n^{(i)} := (\tau - 2)^n \log(Z_n^{(i)}) \text{ and proves that this sequence of i.v. has a limit.}\]
CHAPTER 3. EXPLORATION AROUND TWO VERTICES

Proof. We denote the size of generation $k$ in the two i.i.d. BPs from Corollary 3.1.2 by $(Z_k^{(u)}, Z_k^{(v)})$. Both branching processes are Galton-Watson processes, coupled to a BFS exploration with sources $u$ and $v$. So the size of the first generation has distribution $F_L$ and of all the following generations have distribution $F_B$. In order to obtain the result of this lemma we use the coupling error of Lemma 3.1.1 to converge to zero. We define $\varrho := \min\{\kappa, (1 - \frac{1}{\tau})/2\}$ and $\varrho' := (\tau - 2)\varrho$. Now we let $n^\varrho$ be such that the coupling error converges to zero, since that requires $\varrho' < \varrho$. With this we can find the generation where we have to stop the exploration process. The goal is to find the last generation where the generation size is smaller than $n^\varrho$, so we define

$$t(n^\varrho, u) := \sup\{k : Z_k^{(u)} \leq n^\varrho\} \tag{3.1.14}$$

under the assumption that without loss of generality the BP with source $u$ reaches a size of $n^\varrho$ first. With the equation for the generation size (3.1.12), we find that

$$\kappa_n(u) := t(n^\varrho, u) = \frac{\log \log n + \log(\varrho'/Y_n^{(\varrho)})}{\log(\varrho' - 2)} - f_n \tag{3.1.15}$$

where $f_n$ is the fractional part of the fraction on the rhs. \hfill \square

For the proof of both the lower and upper bound we start with an exploration up to generation $k$, for some $k \geq 0$. As a result we give a lower and upper bound on the length of the path from the source up to generation $k$. First we give a lower bound for this length

**Lemma 3.1.4** (Lower bound on shortest path in exploration). Consider the configuration model $\text{CM}_n(d)$ with empirical degree distribution satisfying assumptions 2.2.1-2.2.3 with i.i.d. edge weights from distribution $L$ with distribution function $F_L$ satisfying (2.2.5). Let $u$ be a uniformly chosen vertex and $G_{\kappa_n}(u)$ be the set of vertices at distance $\kappa_n$ from $u$, with $\kappa_n$ as defined in (3.1.15). Then

$$\lim_{n \to \infty} \mathbb{P}\left(d_L(u, G_{\kappa_n}(u)) > (1 - \varepsilon) \sum_{i=1}^{\lceil \log \log n \rceil} F_L^{(-1)} \left( e^{-\left(\frac{1}{\tau - 2}\right)} \right)^i \right) = 1. \tag{3.1.16}$$

Proof. $d_L(u, G_{\kappa_n}(u))$ is the shortest path from the source $u$ to the last generation of the exploration $\kappa_n$, this path is certainly longer than the sum of the minimum per generation. I.e.

$$d_L(u, G_{\kappa_n}(u)) \geq \sum_{i=1}^{\kappa_n} \min\{L_{i,1}, \ldots, L_{i,Z_i^{(u)}}\}, \tag{3.1.17}$$

where $Z_i^{(u)}$ denotes the number of vertices at distance $i$ from source $u$. For the minimum we know that $\mathbb{P}(\min_{i \leq N} L_i > z(N)) = (1 - F_L(z(N)))^N \geq 1 - NF_L(z(N))$, so we set

$$z(N) = F_L^{(-1)}(1/N^{1+\xi}) \tag{3.1.18}$$

for some $\xi > 0$, which additionally gives us the error probability (the probability that the minimum is smaller than $z(N)$) as $e_N = 1/N^\xi$. We let

$$C_n := \max\{\sup_{1 \leq k \leq \kappa_n} Y_k^{(n)}, f(n)\} \tag{3.1.19}$$

with $f(n)$ a function defined later on. By (3.1.12) we have that $Z_i := \exp((\tau - 2)^{-1} Y_i)$ is the size of generation $i$ combined with (3.1.19) we obtain that

$$\sum_{i=1}^{\kappa_n} \min\{L_{i,1}, \ldots, L_{i,Z_i^{(u)}}\} \geq \sum_{i=1}^{\kappa_n} F_L^{(-1)}(e^{-\left(\frac{1}{\tau - 2}\right)^i C_n(1+\xi)}), \tag{3.1.20}$$

Weighted distances in the configuration model with conservative edge-weights
with error probability given by
\[ E(C_n) := \mathbb{P}(\exists i \leq \kappa_n \text{ s.t. } \min(L_{i,1}, \ldots, L_{i,N_i}) < z(Z_i)). \] (3.1.21)

By a union bound we obtain an upper bound
\[ E(C_n) \leq \sum_{i=1}^{\kappa_n} \mathbb{P}(\min(L_{i,1}, \ldots, L_{i,Z_i}) < z(Z_i)) \leq \sum_{i=1}^{\kappa_n} e^{-\left(\frac{1}{\tau+2}\right)^\epsilon C_n \epsilon} \leq C_1 e^{-\left(\frac{1}{\tau+2}\right)^\epsilon C_n \epsilon} \] (3.1.22)
for some constant $C_1 > 0$. Combining (3.1.17) and (3.1.20) yields that with error probability given in (3.1.21) it holds that
\[ d_L(u, G_{\kappa_n}(u)) \geq \sum_{i=1}^{\kappa_n} F_L^{(-1)}(1 + \epsilon) \] (3.1.23)

For the rhs in equation (3.1.23) we do a variable transformation $(1/(\tau+2))^\epsilon C_n (1 + \epsilon)$ to get rid of the $C_n (1 + \epsilon)$ term in the exponent. We use the following inequality for a lower bound twice, first for the sum and then the obtained integral
\[ \sum_{k=2}^{n} g(k) \leq \int_{1}^{n} g(x) dx \leq \sum_{k=1}^{n} g(k). \] (3.1.24)

First we transform the sum to an integral then we do a variable transformation to obtain the desired exponent and then we transform the integral back to a sum. The variable transformation shifts the boundaries over a value, for which we let $\tilde{C}_n = \log(C_n (1 + \epsilon)) / |\log(\tau - 2)|$ to obtain that
\[ \sum_{i=1}^{\kappa_n} F_L^{(-1)} \left( e^{-\left(\frac{1}{\tau+2}\right)^\epsilon C_n (1 + \epsilon)} \right) \geq \sum_{i=[\tilde{C}_n + 2]}^{\kappa_n + |\tilde{C}_n|} F_L^{(-1)} \left( e^{-\left(\frac{1}{\tau+2}\right)^\epsilon} \right) \] (3.1.25)

In order to make sure the error probability converges to zero, $C_n$ should tend to infinity as $n$ goes to infinity. By the definition of $C_n$, $f(n)$ should therefore tend to infinity with $n$ to let $C_n$ tend to infinity, since $\sup Y_k$ is bounded as a result of Davies’ Theorem. On the other hand $C_n$ cannot be too big, since we want the rhs of (3.1.25) to be larger than $(1 - \epsilon)$ times the denominator of the lhs of (2.2.6). For any fixed $\epsilon > 0$ we define
\[ R_n(\epsilon) := \max \left\{ \sum_{k=1}^{\kappa_n} F_L^{(-1)} \left( e^{-\left(\frac{1}{\tau+2}\right)^\epsilon} \right) \leq \epsilon \sum_{i=1}^{\kappa_n} F_L^{(-1)} \left( e^{-\left(\frac{1}{\tau+2}\right)^\epsilon} \right) \right\} \] (3.1.26)

Since the sum on the rhs between the brackets tends to infinity with $n$, so will $R_n(\epsilon)$. Let the function $f(n)$ from before in the definition of $C_n$ (3.1.19) be defined as $f(n) := 1/(1 + \xi)(\tau - 2)^{-R_n(\epsilon)}$. For large $n$, $C_n = f(n)$ and $\tilde{C}_n = R_n(\epsilon)$ and since $R_n(\epsilon)$ tends to infinity, so will $\tilde{C}_n$ and thus $C_n$ tends to infinity as well. In addition we have a lower bound on the upper summation bound, so we use that
\[ \bar{C}_n = \frac{\log(\max(\sup_{1 \leq k \leq \kappa_n} Y_{k}^{(n)}, f(n)))}{|\log(\tau - 2)|} + \frac{\log(1 + \xi)}{|\log(\tau - 2)|} \geq \frac{\log Y_{\kappa_n}}{|\log(\tau - 2)|} \] (3.1.27)

By the definition of (3.1.26), the inequality (3.1.27) and using that $\kappa_n \geq \log \log n / |\log(\tau - 2)| - \log Y_{\kappa_n}^{(n)} / |\log(\tau - 2)|$, we obtain that for all $\epsilon > 0$
\[ \sum_{i=[\bar{C}_n + 2]}^{\kappa_n + |\bar{C}_n|} F_L^{(-1)} \left( e^{-\left(\frac{1}{\tau+2}\right)^\epsilon} \right) > (1 - \epsilon) \sum_{i=1}^{\kappa_n} F_L^{(-1)} \left( e^{-\left(\frac{1}{\tau+2}\right)^\epsilon} \right) \] (3.1.28)
CHAPTER 3. EXPLORATION AROUND TWO VERTICES

with error probability

\[ P \left( d_L(u, G_{K_n}(u)) \leq \sum_{i=1}^{\log \log n} \frac{F_L^{(-1)} \left( e^{-\left( \frac{\tau}{\tau + 2} \right)^{i}} \right)}{1 + \frac{\log \log n}{1 + \frac{\log \log n}{2}}} \right) \leq E(C_n) \]

Since \( C_n \) tends to infinity with \( n \) we know by (3.1.22) that the rhs tends to zero. This finishes the proof of the lower bound.

Next we do some preparations for the proof of the upper bound. First we investigate the number of generations we need to explore before, whp, we reach a vertex of degree \( K_n \).

Lemma 3.1.5 (Generations to reach degree \( K_n \)). Consider the configuration model \( CM_n(d) \) with empirical degree distribution satisfying Assumptions 2.2.1-2.2.3. Let \( u \) be a uniformly chosen vertex and \( (K_n)_{n \geq 1} = O(\log(n)) \) a sequence that tends to infinity with \( n \). Fix \( M \) with \( M|\log(\tau - 2)| > 1 \), then

\[ \lim_{n \to \infty} P \left( \max_{v \in G_{M_n}(u)} d_v \leq K_n \right) = 0 \]

with \( M_n = M \log \log K_n \) and \( G_{M_n}(u) \) the set of vertices at distance \( M_n \) from \( u \).

Proof. We write \( 1 + \delta = M|\log(\tau - 2)| \) and condition on whether \( Y_{M_n} \) is less than \( 1/(\log K_n)^{\delta/2} \) or not. Then

\[ P \left( \max_{v \in G_{M_n}(u)} d_v \leq K_n \right) \leq P \left( \max_{v \in G_{M_n}(u)} d_v \leq K_n \mid Y_{M_n} > 1/(\log K_n)^{\delta/2} \right) + P \left( Y_{M_n} < 1/(\log K_n)^{\delta/2} \right). \]

The exploration is coupled to a branching process as in Corollary 3.1.2. Recall that \( |G_{M_n}^{(\alpha)}| = Z_{M_n}^{(\alpha)} \), we write \( F_B(x) \) for the size-biased distribution function. By Assumption 2.2.2 the BP cannot go extinct and therefore \( P(Y_{M_n} = 0) = 0 \) and thus the second term on the rhs in (3.1.30) converges to zero. Since the degrees are i.i.d. in the BP, we obtain

\[ P \left( \max_{v \in G_{M_n}(u)} d_v \leq K_n \mid Z_{M_n}^{(\alpha)} \right) = (F_B(K_n))^{Z_{M_n}^{(\alpha)}}. \]

Using the lower bound on \( F_B, 1 - F_B(x) \geq L(x)/x^{\tau - 2} \) from (2.2.3), with \( L(x) = \exp(-c|\log x|) \)

we obtain

\[ (F_B(K_n))^{Z_{M_n}^{(\alpha)}} \leq \left( 1 - \frac{L(K_n)}{K_n} \right)^{Z_{M_n}^{(\alpha)}} \leq \exp \left( -L(K_n)Z_{M_n}^{(\alpha)}/K_n^{\tau - 2} \right). \]

Using that \( Z_{M_n}^{(\alpha)} = \exp((\tau - 2)^{-M_n} Y_{M_n}^{(\alpha)}) \) by (3.1.12) in (3.1.31), we obtain that

\[ P \left( \max_{v \in G_{M_n}(u)} d_v \leq K_n \mid Y_{M_n}^{(\alpha)} \right) \leq \exp \left\{ - \frac{\exp \left( -L(K_n) \gamma \right) \exp \left( (\log K_n)^{1+\delta/2} \right)}{K_n^{\tau - 2}} \right\} \]

\[ \leq \exp \left\{ - \frac{\exp \left( (\log K_n)^{1+\delta/2} \right)}{K_n^{\tau - 2}} \right\} \]

\[ \leq \exp \left\{ - \frac{1}{2} (\log K_n)^{1+\delta/2} \right\} \]

where we used that \( M_n = M \log \log K_n \), with \( 1 + \delta = M|\log(\tau - 2)| \), the last inequality holds for \( K_n \) large. Thus both probabilities on the rhs in (3.1.30) tend to zero as \( K_n \) tends to infinity, which completes the proof.

In the proof of the upper bound of the main theorem we run the exploration algorithm until we reach a vertex of degree \( K_n \). The path to this vertex is the sum of i.i.d. copies of edge weights. We show that whp this sum is less than some \( \varepsilon_1 > 0 \) times the denominator of the lhs of (2.2.6).
Lemma 3.1.6 (Upper bound on the length of the path in the exploration process). Let $L_i$ be i.i.d. random variables from distribution $F_i$, satisfying (2.2.5). Then for all $\varepsilon_1 > 0$ there exists a choice of $M_n$ such that $M_n$ tends to infinity with $n$ and

$$\lim_{n \to \infty} \mathbb{P} \left( \sum_{i=0}^{M_n} L_i < \varepsilon_1 \sum_{i=1}^{\log \log n} F_i^{-1} \left( e^{-\left( \frac{\varepsilon_1}{n} \right)^2} \right) \right) = 1. \tag{3.1.33}$$

Proof. Let us denote the sum at the rhs of the inequality within the probability sign by $a_n$ for readability here. We distinguish three cases. For the first we assume the expectation of $L_i$ is finite and in the other two cases the expectation is infinite but we have an additional assumption on the tail. We start with the case where the expectation is finite, where we let

$$M_n^{(1)} := \max \{ z \leq a_n^\delta \} \tag{3.1.34}$$

for some $\delta \in (0, 1)$. Then $M_n^{(1)}$ tends to infinity with $n$ since $a_n$ does. To bound (3.1.33) we use Markov’s inequality to obtain that

$$\mathbb{P} \left( \sum_{i=0}^{M_n^{(1)}} L_i \geq \varepsilon_1 a_n \right) \leq \frac{M_n^{(1)} \mathbb{E}[L]}{\varepsilon_1 a_n} \tag{3.1.35}$$

by the choice of $M_n^{(1)}$ in (3.1.34) this tends to zero with $n$. Now we assume that expectation is infinite and additionally $\mathbb{P}(L_i > x) \leq 1/g(x)$ for some monotonously increasing function $g(x)$ specified later. We distinguish between two cases for $g$, either the tail is bounded by a power-law or a slowly varying function. Any infinite mean random variable can be bounded by one of the two. We define $M_n^{(2)}$ to be such that the following holds

$$M_n^{(2)} := \left( M_n^{(2)} g^{-1}((M_n^{(2)})^{(1+\varepsilon(2))}) \right)^{1+\delta} = \varepsilon a_n, \tag{3.1.36}$$

for some $\varepsilon(2) > 0$. Since $g(x)$ is monotonously increasing and $a_n$ tends to infinity with $n$, $M_n^{(2)}$ tends to infinity as well. We condition (3.1.33) on the maximum of the $L_i$ being larger than $\hat{M}_n^{(2)} := g^{-1}((M_n^{(2)})^{(1+\varepsilon(2))})$ or not, which gives us the following upper bound

$$\mathbb{P} \left( \sum_{i=0}^{M_n^{(2)}} L_i \geq \varepsilon_1 a_n \right) \leq \mathbb{P} \left( \exists i \leq M_n^{(2)} : L_i \geq \hat{M}_n^{(2)} \right) + \mathbb{P} \left( \sum_{i=0}^{M_n^{(2)}} L_i \mathbb{1}_{\{ L_i \leq \hat{M}_n^{(2)} \}} \geq \hat{M}_n^{(2)} \right). \tag{3.1.37}$$

First we focus on the first term in (3.1.37)

$$\mathbb{P} \left( \exists i \leq M_n^{(2)} : L_i \geq \hat{M}_n^{(2)} \right) \leq M_n^{(2)} \mathbb{P} \left( L_i \geq \hat{M}_n^{(2)} \right) \leq M_n^{(2)} \frac{1}{(M_n^{(2)})^{1+\varepsilon(2)}} = (M_n^{(2)})^{-\varepsilon(2)}, \tag{3.1.38}$$

so this probability converges to zero as $n$ tends to infinity. Next we investigate the second term in (3.1.37) which we bound with Markov’s inequality,

$$\mathbb{P} \left( \sum_{i=0}^{M_n^{(2)}} L_i \mathbb{1}_{\{ L_i \leq \hat{M}_n^{(2)} \}} \geq \hat{M}_n^{(2)} \right) \leq \frac{M_n^{(2)} \mathbb{E}[L] \mathbb{1}_{\{ L_i \leq \hat{M}_n^{(2)} \}}}{\hat{M}_n^{(2)}}. \tag{3.1.39}$$

In order to calculate the expectation on the rhs in (3.1.39) we need to specify the function $g(x)$. We distinguish two cases, either the distribution of $L_i$ can be bounded from above by a power-law, so $g(x) = x^\alpha$ for some $\alpha \in (0, 1)$, which gives us

$$\mathbb{E}[L \mathbb{1}_{\{ L \leq y \}}] \leq \int_0^y \mathbb{P}(L_i > x) dx \leq \int_0^y \frac{c}{x^\alpha} dx = \frac{cy^{1-\alpha}}{1-\alpha}.$$
Using this result in (3.1.39) we obtain

\[
\frac{M_n^{(2)}E[L_i 1_{\{L_i \leq \hat{M}_n^{(2)}\}}]}{M_n^{(2)}} \leq \frac{c}{1-\alpha} \frac{M_n^{(2)}(\hat{M}_n^{(2)})^{1-\alpha}}{M_n^{(2)}} \leq \frac{c}{1-\alpha} M_n^{-\delta}.
\] (3.1.40)

In the other case \(g(x) = \ell(x)\), with \(\ell(x)\) a monotonously increasing slowly varying function. A function \(\ell(x)\) is called a slowly varying at infinity if \(\lim_{x \to \infty} \ell(cx)/\ell(x) = 1\) for all \(c \in \mathbb{R}\). By a theorem by Potter [17, Theorem 1.5.6.] we know that

\[
x^\delta \ell(x) \geq \inf_y \{y^\delta \ell(y)\} =: m_\ell < \infty
\]

Now we can calculate the expectation on the rhs of (3.1.39) and bound this using Potter’s Theorem

\[
E[L_i 1_{\{L_i \leq y\}}] \leq \int_0^y \frac{c}{\ell(x)} \, dx = \int_0^y x^\delta \frac{c}{x^\delta \ell(x)} \, dx \leq \int_0^y x^\delta \frac{c}{m_\ell} = \frac{c}{m_\ell} y^{1+\delta}
\] (3.1.41)

We use this result in (3.1.39) then we obtain

\[
\frac{M_n^{(2)}E[L_i 1_{\{L_i \leq \hat{M}_n^{(2)}\}}]}{M_n^{(2)}} \leq \frac{c}{m_\ell(1+\delta)} \frac{M_n^{(2)}(\hat{M}_n^{(2)})^{1+\delta}}{M_n^{(2)}} = \frac{c}{m_\ell(1+\delta)} (M_n^{(2)})^{-\delta}
\] (3.1.42)

By (3.1.38), (3.1.40) and (3.1.42) we know that (3.1.37) tends to zero as \(n\) tends to infinity in either case, since \(\hat{M}_n^{(2)}\) tends to infinity. Let \(M_n^{(\ast)}\) as defined in (3.1.34) and \(M_n^{(2)}\) such that it satisfies (3.1.36), now we define \(M_n := \min\{M_n^{(\ast)}, M_n^{(2)}\}\) and the proof is completed. \(\square\)
Chapter 4

Percolation of the graph

In this section we define the degree dependent-percolation we use and then show that the degree distribution after percolation still satisfies the power-law conditions. At first, we define the degree-dependent percolation, this method was first given in [33] and later adjusted for the configuration model in [11]. For this we define a function that returns a probability for every half-edge, depending on the degree of the vertex that this half-edge is attached to. The resulting probability is the probability with which we keep this half-edge. Let

\[ p(d) : \mathbb{N} \rightarrow [0, 1] \]  

be a monotone decreasing function of \( s \). For a half edge \( s \) we write the percolation probability shortly as \( p_s := p(d_{(s)}) \) with \( v(s) \) the vertex that half edge \( s \) is attached to and \( d_{(s)} \) the vertex’ degree. Now we define two different ways to percolate the graph, but first we give the definition of an induced subgraph as it is used in the first percolation.

**Definition 4.0.1 (Induced subgraph).** Let \( S \) be a set of vertices. The induced subgraph of \( G \) on vertex set \( S \) is the largest subgraph of \( G \) with edges that have both endpoints in \( S \). We denote the induced graph of a graph \( G \) restricted to the vertices in a set \( S \) by \( G_{|S} \).

Next we define two different ways to percolate the configuration model and after that we show equality in distribution for the two different percolated graphs.

**Definition 4.0.2 (Half-edge percolation).** Given a degree sequence \( d = (d_1, \ldots, d_n) \) and a half-edge \( s \), we keep a half-edge with probability \( p_s \) independently. If we do not keep it, then we delete this half-edge and we create a new vertex with one half-edge corresponding to the deleted half-edge. The total number of artificial edges is denoted by \( A \). After this procedure is carried out for all \( s \in [H_n] \) we pair the half-edges. At last we take the induced subgraph on the \( n \) original vertices. We denote the resulting graph by \( CM_p(d)_{n}(d) \).

**Definition 4.0.3 (Edge percolation).** Given a configuration model \( CM_n(d) \) that is already paired. We delete any edge in the graph independently of all other edges with probability \( p(s)p(t) \), where \( s \) and \( t \) are the degrees of the vertices the edge is attached to. We denote the resulting graph by \( CM_{n}(d) \).

TO DO: cite the paper [11], but after the changes such that we can use their Lemma 3.3 about the probability that the two matchings result in the same graph.

As a result of this lemma we can use the following corollary about the two graphs after percolation.

**Corollary 4.0.4 (Equality in distribution of two percolated graphs).** Consider a function \( p(d) \) satisfying (4.0.1), a graph \( CM_p(d)_{n}(d) \) percolated as described in Definition 4.0.2 and a graph \( CM_{n}(d) \) the result of edge percolation as in Definition 4.0.3. Then \( CM_p(d)_{n}(d) = CM_{n}(d) \).
CHAPTER 4. PERCOLATION OF THE GRAPH

With Corollary 4.0.4 we can understand the connectivity properties of the graph after percolation CM<sub>n</sub>(d) by studying the configuration model CM<sub>n+A</sub>(d', 1(A)) restricted to the first n vertices by changing the order of percolation and pairing. We can do this since the resulting graphs are equal in distribution. So from now on we study CM<sub>n+A</sub>(d', 1(A))|[n].

Next we investigate the degree sequence after half-edge percolation restricted to the original vertices, which we denote by d<sup>'</sup> := {d'_1, ..., d''_n}, with empirical degree distribution F<sub>n</sub>'(x) as defined in Section (2.2). The goal is to show that this degree sequence still satisfies the conditions of (2.2.1). For this we start with a configuration model with degree sequence d of which the empirical distribution satisfies (2.2.1).

**Lemma 4.0.5** (Empirical degree distribution after percolation). Let CM<sub>n</sub>(d) be a configuration model with given degree sequence satisfying (2.2.1). We perform half-edge percolation as described in Definition 4.0.2 on this graph with percolation function p(d) as in (4.0.1) satisfying

\[
p(d) > b \exp\{-C(\log(d))^\gamma\}
\]

for some constants b, C > 0 and γ ∈ (0, 1). Then there exists a θ such that for all x ∈ [θ, n(1−ɛ)/(τ−1)] the empirical degree distribution F<sub>n</sub>'(x) of the degrees after percolation still satisfies the power-law conditions of (2.2.1) whp, with some small ɛ > 0.

**Proof.** We have a configuration model CM<sub>n</sub>(d) with degree sequence d and empirical degree distribution F<sub>n</sub> satisfying (2.2.1). The degree sequence after percolation described in Definition 4.0.2 is d<sup>'</sup> with empirical degree distribution F<sub>n</sub>'(x). For the sake of the proof we write α = (1−ɛ)/(τ−1). Note that d''<sub>i</sub> = Bin(d<sub>i</sub>, p(d<sub>i</sub>)), by Bin(n, p) we denote a binomial random variable with parameters n and p. As a result the random variables (d''<sub>i</sub>)<sub>i≤n</sub> are independent given the initial degrees (d<sub>i</sub>). First we show the upper bound

\[
1 - F_n'(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{d''_i > x\}} \leq \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{d_i > x\}} = 1 - F_n(x). \tag{4.0.3}
\]

We know that the F<sub>n</sub> satisfies the upper bound and hence F<sub>n</sub>' as well. Next we show the lower bound, first we define for all x < n<sup>α</sup>

\[
S(x) := \{v : d_v ≥ y(x)\} \tag{4.0.4}
\]

where y(x) is a function of x that is defined later. We use S(x) to bound the empirical degree distribution after percolation from below. For this we use that

\[
1 - F_n'(x) ≥ \frac{1}{n} \sum_{i \in S(x)} \mathbb{1}_{\{d''_i > x\}} \tag{4.0.5}
\]

We choose the value of y(x) such that the probability that the indicators within the sum are 1 with high enough probability. We can use the concentration of binomial random variables [28, Theorem 2.21] to get an upper bound on the probability that this indicator function is 1. Thus we choose y(x) such that the expectation of the binomial, d<sub>i</sub>p(d<sub>i</sub>), is higher than 2x for all vertices in S(x). Let

\[
y(x) = \frac{2x}{b} e^{2c(\log(2x/b))^{γ}} \tag{4.0.6}
\]

then we find that

\[
y(x)p(y(x)) ≥ 2xe^{2c(\log(2x/b))^{γ}} e^{-c(\log((2x/b)e^{2c(\log(2x/b))^{γ}})})^{γ} = 2xe^{2c(\log(2x/b))^{γ}} e^{-c(\log(2x/b)(1+2c(\log(2x/b))^{γ−1}))^{γ}}
\]

Then for all x ≥ b<sup>γ</sup> exp {−4e<sup>1−γ</sup>} := b<sup>γ</sup> we find that for all d<sub>i</sub> > y(x)

\[
d_i p(d_i) ≥ y(x)p(y(x)) ≥ 2xe^{c/2(\log 2x/b)^{γ}} ≥ 2x, \tag{4.0.7}
\]

xxiv
Weighted distances in the configuration model with conservative edge-weights
which gives for all \( d_i > y(x) \)
\[
\mathbb{P}(\text{Bin}(d_i, p(d_i)) > x) \leq \mathbb{P}\left( \frac{\text{Bin}(d_i, p(d_i)) > \frac{y(x)p(y(x))}{2}}{\frac{y(x)p(y(x))}{2}} \right) \leq e^{-y(x)p(y(x))/8} \leq e^{-x/4} < \frac{1}{8},
\]
whenever \( x > 4 \log 8 \). Using this we get for all \( x \geq \max \left\{ \frac{1}{2} \tilde{\theta}, 4 \log 8 \right\} := \theta \)
\[
\mathbb{P}\left( n(1 - F_n^\gamma(x)) \leq \frac{\left| S(x) \right|}{4} \right) \leq \mathbb{P}\left( \sum_{i \in S(x)} \mathbb{1}_{(\text{Bin}(d_i, p(d_i)) > x)} \leq \frac{\left| S(x) \right|}{4} \right)
\leq \mathbb{P}\left( \text{Bin}(\left| S(x) \right|, 7/8) \leq \frac{\left| S(x) \right|}{4} \right) \leq e^{-|S(x)|/8}.
\]
Combining this estimate with a union bound,
\[
\mathbb{P}\left( \exists x \in [\theta, n^\alpha] : n(1 - F_n^\gamma(x)) \leq \frac{|S(x)|}{4} \right) \leq \sum_{x = \theta}^{n^\alpha} e^{-|S(x)|/8} \leq n^\alpha e^{-|S(n^\alpha)|/8}, \tag{4.0.8}
\]
since \( |S(x)| \) decreases as \( x \) increases. Using (2.2.1) \( |S(x)| \) can be bounded from below as follows
\[
|S(x)| = n(1 - F_n(y(x))) \geq n \frac{1}{y(x)^\tau - 1} e^{-c(\log y(x))^\gamma}.
\]
Using this bound for \( |S(x)| \) within the probability sign in (4.0.8) and for \( |S(n^\alpha)| \) on the rhs of (4.0.8) we arrive at:
\[
\mathbb{P}\left( \exists x \in [\theta, n^\alpha] : 1 - F_n^\gamma(x) \leq \frac{1}{y(x)^\gamma} e^{-c(\log y(x))^\gamma} \right) \leq e^{\alpha \log n} e^{-\gamma n^\alpha e^{-c(\log n)^\gamma} \rightarrow \infty} 0. \tag{4.0.9}
\]

**Lemma 4.0.6** (Degree after percolation). Consider the configuration model \( \text{CM}_n(d) \), a percolation function \( p(d) \) as in (4.0.1) satisfying (4.0.2), \( \tilde{K}_n = O(\log n) \) an arbitrary sequence that tends to infinity with \( n \). We define
\[
K_n := \sup \left\{ m : 2m \leq \tilde{K}_n b e^{-c(\log \tilde{K}_n)^\gamma} \right\}. \tag{4.0.10}
\]

We use degree-dependent percolation as described in Definition 4.0.2. Then a vertex \( v \) of degree \( \tilde{K}_n \) in \( \text{CM}_n(d) \) has degree \( K_n \) in \( \text{CM}_n^{(d)}(d) \) whp.

**Proof.** First we investigate the expected degree of a vertex after percolation. Consider a vertex \( v \) with degree \( d_v \), as before \( d_v^p \) denotes the degree after percolation.
\[
\mathbb{E}[d_v^p] \geq \mathbb{E}\left[ \text{Bin}(d_v, b e^{-c(\log d_v)^\gamma}) \right] = d_v b e^{-c(\log d_v)^\gamma} = b \log d_v (1 - e^{-c(\log d_v)^\gamma})^{-1}.
\]
We can see now that the value on the rhs tends to infinity with \( d \). Therefore \( K_n \) tends to infinity when \( \tilde{K}_n \) does. The degree after percolation is the random variable \( \text{Bin}\left( \tilde{K}_n, p(\tilde{K}_n) \right) \). By definition (4.0.10), we have that the expectation of the binomial is larger than \( 2K_n \). Knowing that, we can use the concentration of binomial random variables [28, Theorem 2.21] to obtain a bound on the probability that the binomial is smaller than \( K_n \), i.e.
\[
\mathbb{P}\left( \text{Bin}(\tilde{K}_n, p(\tilde{K}_n)) < K_n \right) \leq \exp \left\{ -K_n / 4 \right\},
\]
since \( K_n \) tends to infinity this finishes the proof. \( \square \)
Chapter 5
Proof of the main Theorem

In this section we give the proofs of Theorems 2.2.5 and 2.2.7. We start with the main result as stated in Theorem 2.2.5, after that we give the proof of Theorem 2.2.7. The first proof consist of two parts, namely a lower- and an upper bound. We start with the lower bound as stated in the following lemma

Lemma 5.0.1 (Lower bound on the weighted graph distance). Consider the configuration model \( CM_n(d) \) with empirical degree distribution satisfying Assumptions 2.2.1-2.2.3 and let \( u \) and \( v \) be uniformly chosen from \([n]\). Suppose the edge weights are i.i.d. from distribution \( L \) with distribution function \( F_L(x) \) with inverse \( F_L^{-1}(x) \) that satisfies (2.2.5). Then for all \( \varepsilon > 0 \)

\[
\lim_{n \to \infty} \mathbb{P}\left( d_L(u, v) > (1 - \varepsilon)2 \sum_{i=1}^{\left\lfloor \frac{\log \log n}{\log(\tau - 2)} \right\rfloor} F_L^{-1}\left( e^{-\left( \tau - 2 \right)} \right) \right) = 1 \quad (5.0.1)
\]

Proof. We consider two uniformly chosen vertices \( u \) and \( v \). We do a BFS-exploration on both sides and by Lemma 3.1.3, we can couple these explorations whp to two independent BPs until generation \( \kappa_n(u) \), \( \kappa_n(v) \) respectively. We write \( G_{\kappa_n}(i) \) for the set of vertices in generation \( \kappa_n(i) \) of the exploration started from vertex \( i = u, v \), respectively. By the coupling, these explorations are disjoint whp. Since the sum of the two shortest paths from \( u, v \) to the last generations of their explorations is certainly smaller than the entire shortest path between \( u \) and \( v \) we know that

\[
d_L(u, v) \geq d_L(u, G_{\kappa_n}(u)) + d_L(v, G_{\kappa_n}(v)) \quad (5.0.2)
\]

By Lemma 3.1.4 we have a lower bound for the path length from the vertex to the last generation and thus

\[
\lim_{n \to \infty} \mathbb{P}\left( d_L(u, G_{\kappa_n}(u)) > (1 - \varepsilon)2 \sum_{i=1}^{\left\lfloor \frac{\log \log n}{\log(\tau - 2)} \right\rfloor} F_L^{-1}\left( e^{-\left( \tau - 2 \right)} \right) \right) = 1 \quad (5.0.3)
\]

Since the two terms at the rhs of (5.0.2) are equal in distribution and using the inequality of (5.0.2) the result of the lemma follows. \( \square \)

Next we show the upper bound for which we first need a proposition, similar to [11, Proposition 2.1]. This proposition gives an upper bound on the path length between two vertices of a fixed degree of at least \( K \). In this setting these vertices have a degree of at least \( K_n \) with \( K_n \) tending to infinity with \( n \). In addition we provide the adjusted proof since the adjustments are non-trivial.

Proposition 5.0.2. Consider the configuration model \( CM_n(d) \) with empirical degree distribution satisfying (2.2.1) for all \( x \in [\theta, \theta^{1-\varepsilon}/(\tau - 1)] \) for some given \( \theta \in \mathbb{R} \) and some small \( \varepsilon > 0 \). Let \( u_{K_n} \) be a vertex with degree at least \( K_n \). Then, whp, there exists a path from \( u_{K_n} \) to a vertex \( u^* \)
with degree at least \( n^{(\tau - 2)(1 + \zeta)/(\tau - 1)} \) for \( \zeta > 0 \) such that the degree \( y_i \) of the \( i \)th vertex on the path satisfies

\[
y_i \geq (K_n^{1 - \kappa_n})^{(\tau - 2)},
\]

with \( \kappa_n = D(\log K_n)^{-1} \). In particular, whp the length of this path is at most

\[
\log \log n \left| \frac{\log(\tau - 2)}{\log(\tau - 2)} \right| - \log \log K_n.
\]

**Proof.** We denote the path length from \( u_{K_n} \) to \( u^* \) by \( b(n) \) and we define sets of vertices

\[
\Gamma_i := \{ v \in [n] : d_v \geq y_i \}
\]

for some sequence \( y_i \) to be determined shortly. These sets of vertices can be seen as layers of the graph. Our goal is to prove that there exists a sequence \( y_i \) such that the following holds:

\[
\lim_{n \to \infty} \sum_{i=0}^{\infty} \mathbb{P}(u_i \in \Gamma_i, u_i \Rightarrow \Gamma_{i+1} | d_{u_0} \geq K_n) = 0,
\]

where \( u_i \) is a vertex chosen from \( \Gamma_i \) according to the size-biased distribution (In other words, the vertex \( u_i \) is the vertex of a uniformly chosen half-edge from \( \Gamma_i \)). If this holds then a vertex in the layer \( \Gamma_i \) is whp connected to the consecutive layer \( \Gamma_{i+1} \) for all \( i = 1, 2, \ldots b(n) - 1 \). Conditioning on the total number of half edges \( H_n \) and in the set \( \Gamma_i \) denoted by \( H_{y_i} \) we have the upper bound

\[
\mathbb{P}(v_i \in \Gamma_i, v_i \Rightarrow \Gamma_{i+1} | H_{y_i}, H_n) \leq \left( 1 - \frac{H_n}{H_n} \right)^{y_i/2}.
\]

Since any vertex in \( \Gamma_i \) has degree larger than \( y_i \), and \( |\Gamma_i| = n(1 - F_n(y_i)) \), we have the following lower bound amount of half-edges \( H_{y_i} \) in \( \Gamma_i \):

\[
H_{y_i} \geq y_i n (1 - F_n(y_i)).
\]

Assuming that \( H_n \leq \varphi n \) for some \( \varphi \in \mathbb{R} \) we have by (5.0.8)

\[
\mathbb{P}(v_i \in \Gamma_i, v_i \Rightarrow \Gamma_{i+1}) \leq \exp \left\{ - \frac{y_i y_{i+1} (1 - F_n(y_{i+1}))}{2\varphi} \right\}.
\]

For now we focus on the term in the exponent. By (2.2.1) we have a lower bound on this

\[
\frac{y_i y_{i+1} (1 - F_n(y_{i+1}))}{2\varphi} \geq \bar{c} y_{i+1}^{2-\tau-C(\log y_{i+1})^\gamma} \geq \bar{c} y_{i+1}^{2-\tau-C(\log y_{i+1})^{-1}},
\]

with \( C \) defined in (2.2.1) and \( \bar{c} \) some positive constant. Now we would like to choose the sequence \( y_i \) such that (5.0.10) converges to zero in particular that (5.0.7) holds. We claim that this holds when \( y_i \) is given by the following recursion

\[
y_0 = K_n, \quad y_{i+1} = y_i^{(\tau - 2 + D(\log y_i)^{-1})^{-1}}
\]

with \( D > 0 \) defined later. Note that by the choice of \( y_0 \), for sufficiently large \( K_n \), since \( \gamma < 1 \),

\[
\tau - 2 + D(\log y_0)^{-1} < 1
\]

Now let \( \kappa_n = D(\log K_n)^{-1} \) then

\[
y_{i+1} \geq y_i^{(2 - \tau + \kappa_n)^{-1}} \geq \ldots \geq K_n^{(\tau - 2 + \kappa_n)^{-1}}.
\]

We use the recursion relation of (5.0.12) in (5.0.11)

\[
\bar{c} y_{i+1}^{2-\tau-C(\log y_{i+1})^{-1}} = \bar{c} y_i^{2-\tau-C(\log y_{i+1})^{-1}} \geq \bar{c} y_i^{D(\log y_i)^{-1} - C(\log y_{i+1})^{-1}}
\]

Weighted distances in the configuration model with conservative edge-weights xxvii
Now we choose $D \geq 2C$ and use that the sequence $y_i$ is increasing and the lower bound in (5.0.14)

\[
\tilde{c}y_i \geq \tilde{c} \exp \{ C(\log y_i)^\gamma \} \geq \tilde{c} \exp \left\{ \frac{C(\log K_n)^\gamma}{(\tau - 2 + x_n)^\gamma} \right\}. \tag{5.0.16}
\]

Using this lower bound in (5.0.10), since $\tau - 2 + x_n < 1$, the rhs of (5.0.10) is summable in $i$. Summing the lhs of (5.0.10) over $i$ and then use the above bound we obtain

\[
\sum_{i=0}^\infty P(u_i \in \Gamma_1, u_i \not\rightarrow \Gamma_{i+1} \mid d_{u_0} \geq K_n) \leq \tilde{C} \exp \left\{ -\tilde{c} \exp \left\{ \frac{C(\log K_n)^\gamma}{(\tau - 2 + x_n)^\gamma} \right\} \right\}, \tag{5.0.17}
\]

which tends to zero with $n$ as $K_n$ tends to infinity with $n$. This result yields the statement of (5.0.7). Using the result of [11, Lemma 2.6], the lower bound in (5.0.14) can be improved to

\[
y_i \geq (y_0^{1-x_n})(\tau-2)^{-i}. \tag{5.0.18}
\]

The path in the statement of the lemma is then constructed as follows, starting from the first vertex $u_0 = u_{K_n}$. By the first term in the sum in (5.0.7), $u_0$ is whp connected to at least one vertex in $\Gamma_1$. By the fact that the pairs of the half edges of $u_0$ are chosen uniformly, $u_1$ is a vertex attached to a uniformly chosen half-edge in $\Gamma_1$. As a result, $u_1$ is chosen according to the size-biased distribution within $\Gamma_1$. Then we iterate this procedure to obtain $u_2, u_3, \ldots$ in $\Gamma_2, \Gamma_3, \ldots$ until we reach a vertex of degree at least $n^\alpha$ for $\alpha = (\tau - 2)/(1 + \zeta)/\gamma$. The constructed path uses at most all layers so the number of layers is an upper bound on the length of the path from $u_{K_n}$ to $u^*$. The layer containing no vertices with degrees lower than $n^\alpha$ is $\Gamma_{[1^r]}$ with $r^*$ the solution

\[
K_n^{(\tau-2)^{-r^*}} = n^\alpha. \tag{5.0.19}
\]

The following upper bound then holds

\[
b(n) \leq \frac{\log \log n - \log K_n + (\log \alpha)}{|\log(\tau - 2)|} \leq \frac{\log \log n - \log K_n}{|\log(\tau - 2)|}, \tag{5.0.20}
\]

which finishes the proof.

\[\square\]

**Lemma 5.0.3** (Upper bound on the weighted graph distance). Consider the configuration model $CM_n(d)$ with empirical degree distribution satisfying Assumptions 2.2.1-2.2.3 and $u, v$ two uniformly chosen vertices. Suppose the edge weights are i.i.d. from distribution $L$ with distribution function $F_L(x)$ with inverse $F^{-1}(x)$ that satisfies (2.2.5). Then for all $\epsilon > 0$

\[
\lim_{n \to \infty} P \left( d_L(u, v) < (1 + \epsilon)2 \sum_{i=1}^{\lfloor \frac{\log \log n}{\log(\tau - 2)} \rfloor} F_L^{-1}(\epsilon^{-1/2}) \right) = 1. \tag{5.0.21}
\]

**Proof.** The proof consists of two steps. At first we do an exploration around the two initial vertices $u$ and $v$ until we reach a vertex of degree at least $K_n$. We denote these vertices by respectively $u_{K_n}$ and $v_{K_n}$. Then we percolate the graph, after percolation a vertex with this degree still has degree $K_n$ whp by Lemma 4.0.6. Then we construct a path in the percolated graph that connects the vertices $u_{K_n}$ and $v_{K_n}$.

For the first part we give a bound on the path length from the two initial vertices $u, v$ until a vertex of high degree $K_n$. We define $M_n := \min\{M_n^{(1)}, M_n^{(2)}\}$ with $M_n^{(1)}$ as defined in (3.1.34) and $M_n^{(2)}$ such that it satisfies (3.1.36), $\tilde{K}_n = \exp \left\{ \exp(M_n/M) \right\}$ and $M$ such that $M[\log(\tau - 2)] > 1$. As a result of Lemma 3.1.5 we know that we need to explore up to at most generation $M_n$ until we

\[\footnote{This has been done before in [11], but then up to a fixed degree $K$. Since we want to show convergence in probability, we need to improve upon this result. For us the shortest path from $u$ to the first vertex found with degree $K_n$ is not bounded anymore since we let $K_n$ tend to infinity with $n$.}
reach a vertex of degree at least $\tilde{K}_n$ whp. The path from a vertex $q$ up to $q_{K_n}$, for $q \in \{u,v\}$ is the sum of at most $M_n$ i.i.d. copies of edge weights and by the result of Lemma 3.1.6 we have that

$$\lim_{n \to \infty} P\left( d_L(q, q_{K_n}) \geq \varepsilon_1 \sum_{i=1}^{\lfloor \log \log n \rfloor} F^{(i-1)}_L \left( e^{-\left(\frac{q}{\tilde{K}_n}\right)^\varepsilon} \right) \right) = 0. \quad (5.0.22)$$

Next we construct a connecting path between the two vertices $u_{K_n}$ and $v_{K_n}$ in the percolated graph such that we have an upper bound for $d_L(u_{K_n}, v_{K_n})$. For the degree-dependent percolation we define the following percolation function

$$p(d) = \exp\{-c(\log d)^\alpha\}, \quad (5.0.23)$$

with the parameter $d$ being the degree of the vertex a half-edge is attached to. The idea for the degree-dependent percolation is that we want to keep an edge, connecting two vertices with degrees $d, d'$, only if its edge-weight is at most $\xi_{d,d'}$ a sequence depending on the degrees. The degree-dependent percolation per definition keeps an edge with probability $p(d)p(d')$, where $d$ and $d'$ are the degrees of the vertices on the two ends of the edge. We can realize this percolation by using the edge weights to determine which edges to keep, by setting $\mathbb{I}\{e \text{ is kept}\} = \mathbb{I}\{L_e \leq \xi_{d,d'}\}$. By doing so, in the percolated graph, the edge-weights between vertices with degrees $d, d'$ have a deterministic upper bound $\xi_{d,d'}$. To determine $\xi_{d,d'}$ the following relation should hold:

$$\mathbb{P}(L_e \leq \xi_{d,d'}) = p(d)p(d'). \quad (5.0.24)$$

By Corollary 4.0.4 we know that the resulting graphs, independently of which percolation (Definition 4.0.2 and 4.0.3) we use, are equal in distribution. The percolation function (5.0.23) satisfies the conditions (4.0.1) and (4.0.2) so by Lemma 4.0.5 we know that the empirical degree distribution of the degree sequence after percolation $d^*$ still satisfies the power-law conditions in (2.2.1). Then by Lemma 4.0.6 the vertex with degree $\tilde{K}_n$ whp has degree $K_n$ after percolation. From $q_{K_n}$ we use the constructed path as described in Proposition 5.0.2 to reach a vertex $q^*$ with degree $n^\alpha$ in the percolated graph, with $\alpha = (\tau - 2)(1 + \zeta)/(\tau - 1)$. A lower bound on the degree of the $i$th vertex on this path is $y_i$. Thus, the edge-lengths on the constructed path are at most $\xi_{y_i, y_{i+1}}$ for $i = 0, 1, \ldots, b(n) - 1$, so we have to upper bound on the weighted path-length:

$$d_L(q_{K_n}, q^*) \leq \sum_{i=0}^{b(n)-1} \xi_{y_i, y_{i+1}}. \quad (5.0.25)$$

The sum is up to $b(n) - 1$ because it takes this many steps to get to the $b(n)$th layer. What remains is to connect the two high-degree vertices $u^*$ and $v^*$. Let us denote $w^*$ as the vertex in the graph with the highest degree, which is at least $n^{(1-\varepsilon)/(\tau-1)}$. Then for $q \in \{u,v\}$

$$\mathbb{P}(q^* \to w^*) \leq \left(1 - \frac{n^{(1-\varepsilon)/(\tau-1)}}{H_n}\right)^{n^\alpha/2} \leq \exp \left\{-\frac{n^{(1-\varepsilon)/(\tau-1)n^\alpha/2}}{n}\right\}$$

The exponent on the rhs can be rewritten to $n^{(\zeta(\tau-2)-\varepsilon)/(\tau-1)}$ which tends to infinity if $\varepsilon$ is chosen small enough. Thus, whp, the vertices $u^*$ and $v^*$ are connected to $w^*$ and

$$d_L(u^*, v^*) \leq 2\xi_{w^*, y^*(1-\varepsilon)/(\tau-1)}. \quad (5.0.26)$$

In what follows we determine $\xi_{y_i, y_{i+1}}$. We can re-write the percolation function in terms of the edges between two consecutive layers. Using the definition (5.0.12) of $y_i$

$$p(y_i) = \exp\{-c(\log y_i)^\alpha\} \leq \exp\left\{-\frac{c}{\alpha^\gamma} f(K_n)\right\}, \quad (5.0.27)$$

Weighted distances in the configuration model with conservative edge-weights xxix
with \( f(K_n) = (\log K_n)^\alpha \), \( \alpha = (\tau - 2)^{-\eta} \) and \( \epsilon = \alpha / (1 + \alpha) \). Using this representation for (5.0.24) we obtain
\[
\Pr(L_e \leq \xi_{y_i, y_i+1}) = p(y_i) p(y_{i+1}) \leq \exp \left\{ -\frac{c}{\alpha^i} (\log K_n)^\eta \right\} \exp \left\{ -\frac{c}{\alpha^{i+1}} (\log K_n)^\eta \right\} \\
= \exp \left\{ -\frac{c}{\alpha^i} (\log K_n)^\eta (1 + 1/\alpha) \right\} = \exp \left\{ -\frac{1}{\alpha^i} (\log K_n)^\eta \right\}
\]
(5.0.28)

Combining (5.0.24) and (5.0.28) we find that maximum edge length between two consecutive layers \( \Gamma_i \) and \( \Gamma_{i+1} \) is
\[
\xi_{y_i, y_{i+1}} \leq F_L^{(-1)}(e^{-(\log K_n)^\eta/\alpha^i} , \tau).
\]
(5.0.29)

By Proposition 5.0.2 we have an upper bound on \( b(n) \), combining this with (5.0.25) and (5.0.26) yields
\[
d_L(u_{K_n}, v_{K_n}) \leq 2 \sum_{i=0}^{\left\lfloor \log (\log n / \log K_n) \right\rfloor} F_L^{(-1)}(e^{-(\log K_n)^\eta/\alpha^i} , \tau).
\]
(5.0.30)

where \( \xi_{\eta^{1/2+\delta}, \eta^{1/2+\delta}} \) is merged into the sum by increasing the summation boundary by 1. Similar to the proof of Lemma 3.1.4, we need to transform the argument in the inverse distribution to the desired result in (5.0.21). Therefore we rewrite the sum to an integral, in the integral we can do the variable transformation \((1/\alpha)^\eta (\log K_n)^\eta = (1/(\tau - 2))^\eta\). This results in a shift of the summation boundaries of \( \eta \log \log K_n / \log(\tau - 2) \) and a multiplication of the integration variable with \( \eta \). Therefore we obtain
\[
\sum_{i=0}^{\left\lfloor \log (\log n / \log K_n) \right\rfloor} F_L^{(-1)}(e^{-(\log K_n)^\eta/\alpha^i} , \tau) \leq \frac{1}{\eta} \sum_{i=\left\lfloor \log (\log n / \log K_n) \right\rfloor}^{\left\lfloor \log (\log n / \log K_n) \right\rfloor} F_L^{(-1)}(e^{-1/(\tau - 2)^\eta} , \tau) \\
\leq \frac{1}{\eta} \sum_{i=1}^{\left\lfloor \log (\log n / \log K_n) \right\rfloor} F_L^{(-1)}(e^{-1/(\tau - 2)^\eta} , \tau)
\]
(5.0.31)

We can choose \( \eta \) in (5.0.23) such that \( 1/\eta < (1 + \varepsilon_2) \) so we obtain that
\[
d_L(u_{K_n}, v_{K_n}) \leq 2(1 + \varepsilon_2) \sum_{i=1}^{\left\lfloor \log (\log n / \log K_n) \right\rfloor} F_L^{(-1)}(e^{-1/(\tau - 2)^\eta} , \tau)
\]
(5.0.32)

For the shortest distance we have the upper bound
\[
d_L(u, v) \leq d_L(u, u_{K_n}) + d_L(u_{K_n}, v_{K_n}) + d_L(v, v_{K_n}),
\]
(5.0.33)

We can choose \( \varepsilon_1 + \varepsilon_2 < \varepsilon \) so by (5.0.22) and (5.0.32) this bound gives the result of the lemma, this finishes the proof.

Now that we have Lemmas 5.0.1 and 5.0.3 we are ready to state the proof of Theorem 2.2.5.

Proof of Theorem 2.2.5. Lemma 5.0.1 states the proof of the lower bound and Lemma 5.0.3 the proof of the upper bound. These combined prove the statement of the theorem.

5.1 Erased configuration model

Now that we have completed the proof of Theorem 2.2.5, we can give the proof of Theorem 2.2.7, for this we adjust the proof of the lower- and upper bound where necessary. We alter the proof where results are not valid anymore due to the erasure to obtain a simple graph.
Proof. The proof of Lemma 5.0.1 consists of a BFS exploration around the two vertices \( u \) and \( v \). These explorations can whip be coupled to two BP and therefor we know that there are no multiple-edges or self-loops in the exploration. So this lemma remains valid after erasure and thus

\[
\lim_{n \to \infty} \mathbb{P} \left( d_L^i(u,v) > (1 - \varepsilon) \sum_{i=1}^{\lfloor \log \log \log n \rfloor} F_L^{(i-1)} \left( e^{-\left( \frac{1}{v_i^2} \right)^{10}} \right) \right) = 1. \tag{5.1.1}
\]

The proof of the upper bound, as noted before, consists of two steps. First we explore around the initial vertices \( u \) and \( v \) until we reach a vertex of degree at least \( K_n \). The path up to the first vertex with this degree reached is the sum of i.i.d. copies of edge-weights. Because we do an exploration and this can be coupled to a BP we know that there is only one path to this vertex, so by Lemmas 3.1.5 and 3.1.6 we have that

\[
\lim_{n \to \infty} \mathbb{P} \left( d_L^i(u,v_{K_n}) \geq \varepsilon \sum_{i=1}^{\lfloor \log \log n \rfloor} F_L^{(i-1)} \left( e^{-\left( \frac{1}{v_i^2} \right)^{10}} \right) \right) = 0. \tag{5.1.2}
\]

Next we investigate the constructed path after percolation. The erasure happens before the percolation, so edges of the path as constructed in Proposition 5.0.2 could be deleted in the erased configuration model after percolation. We need any edge on the path to be unique i.e. not part of a multiple edge, such that it is certainly not deleted.

**Lemma 5.1.1 (Unique edges on the path).** Let \( v_i \) and \( v_{i+1} \) be two consecutive vertices on the constructed path as in Proposition 5.0.2 and \( i = 0, \ldots, i^* - 2 \), then

\[
\lim_{n \to \infty} \mathbb{P} \left( \geq 2 \text{ edges connecting } v_i \leftrightarrow v_{i+1} \right) \geq 1 \text{ edge connecting } v_i \leftrightarrow v_{i+1} \right) = 0. \tag{5.1.3}
\]

**Proof.** We can rewrite the probability in (5.1.3) as

\[
\frac{1 - \mathbb{P}(1 \text{ edge } v_i \leftrightarrow v_{i+1}) - \mathbb{P}(\text{no edge } v_i \leftrightarrow v_{i+1})}{1 - \mathbb{P}(\text{no edge } v_i \leftrightarrow v_{i+1})}. \tag{5.1.4}
\]

We investigate the probabilities in (5.1.4) separately starting with the probability that there is exactly one edge between those two vertices. We denote the degrees of \( v_i \) and \( v_{i+1} \) by \( d_{v_i} \) and \( d_{v_{i+1}} \), respectively.

\[
\mathbb{P}(1 \text{ edge connecting } v_i \leftrightarrow v_{i+1}) = \sum_{j=1}^{d_{v_i}} \mathbb{P}(\text{connect only the } j\text{th half-edge of } v_i \text{ to } v_{i+1})
\]

\[
= \sum_{j=1}^{d_{v_i}} \left( \prod_{k=1}^{j-1} \frac{\mathcal{H}_n - d_{v_{i+1}} - 2(k - 1)}{\mathcal{H}_n - 2(k - 1) - 1} \right) \cdot \frac{d_{v_{i+1}}}{\mathcal{H}_n - 2(j - 1) - 1} \cdot \left( \prod_{k=j+1}^{d_{v_i}} \frac{\mathcal{H}_n - d_{v_{i+1}} - 2k}{\mathcal{H}_n - 2(k - 1) - 1} \right)
\]

In the equation above, the terms represent the probability of respectively connecting the first \( j - 1 \) half edges to any vertex but \( v_{i+1} \), then the probability of connecting the \( j \)th half edge to \( v_{i+1} \), and the probability of connecting the last half-edges to any vertex but \( v_{i+1} \). We need an upper bound for (5.1.3) and thus a lower bound for the probability for exactly one edge so we obtain

\[
\mathbb{P}(1 \text{ edge connecting } v_i \leftrightarrow v_{i+1}) \geq \sum_{j=1}^{d_{v_i}} \frac{d_{v_{i+1}}}{\mathcal{H}_n - 2d_{v_i}} \prod_{k=1}^{d_{v_i}} \left( 1 - \frac{d_{v_{i+1}}}{\mathcal{H}_n} \right)
\]

\[
\geq \frac{d_{v_{i+1}} d_{v_i}}{\mathcal{H}_n - 2d_{v_i}} \left( 1 - \frac{d_{v_{i+1}}}{\mathcal{H}_n} \right)^{d_{v_i}}. \tag{5.1.5}
\]
Next we want to find the probability that there is no edge between two vertices. Since this probability appears in the denominator and nominator in (5.1.4) we need a lower- and upper bound.

\[ P(\text{0 edges connecting } v_i \leftrightarrow v_{i+1}) \leq \prod_{k=1}^{\lfloor d_{v_i}/2 \rfloor} \left( 1 - \frac{d_{v_{i+1}}}{\mathcal{H}_n - 2d_{v_i}} \right) = \left( 1 - \frac{d_{v_{i+1}}}{\mathcal{H}_n} \right)^{d_{v_i}/2}, \quad (5.1.6) \]

\[ P(\text{0 edges connecting } v_i \leftrightarrow v_{i+1}) \geq \prod_{k=1}^{d_{v_i}} \left( 1 - \frac{d_{v_{i+1}}}{\mathcal{H}_n} \right) = \left( 1 - \frac{d_{v_{i+1}}}{\mathcal{H}_n} \right)^{d_{v_i}}. \quad (5.1.7) \]

We can now use series expansion for equations (5.1.5), (5.1.6) and (5.1.7) to obtain an upper bound for (5.1.4)

\[ 1 - P(\text{1 edge } v_i \leftrightarrow v_{i+1}) - P(\text{no edge } v_i \leftrightarrow v_{i+1}) \leq \frac{2d_{v_{i+1}}d_{v_i}}{\mathcal{H}_n} / \left( 1 - \frac{d_{v_{i+1}}d_{v_i}}{4(\mathcal{H}_n - d_{v_i})} \right). \quad (5.1.8) \]

Now we use that \( v_i \) and \( v_{i+1} \) are in the layers on the path as defined in (5.0.6). Suppose that \( d_{v_i} \leq y_{i+1} \), if this does not hold then the path through the layers is shorter by one edge and we shift the indices such that \( v_i \equiv r_{i+1} \) and in addition \( d_{v_{i+1}} \leq n^{1/(\tau - 1)} \), the maximum degree in the graph. Let us set the maximum value for \( i \) at \( i^* - 2 \) such that we find that \( d_{v_{i^* - 2}} \leq n^{(\tau - 2)/\tau - 1} \) and thus

\[ \frac{d_{v_{i+1}}d_{v_i}}{\mathcal{H}_n} \leq \frac{y_{i+1}n^{1/(\tau - 1)}}{n} \leq \frac{n^{(\tau - 2)/\tau - 1}}{n} = n^{-1/(\tau - 1)}. \]

This converges to zero as \( n \) tends to infinity and so we see that the right-hand side of (5.1.8) tends to zero as \( n \) tends to infinity and since the left-hand side is the same probability as (5.1.3) this completes the proof of Lemma 5.1.1.

For the last edges we do not know with certainty that they are still present after erasure so we add 4 i.i.d. copies from \( L \). As a result we know that

\[ d_L^e(u_{K_n}, v_{K_n}) \leq 2 + \sum_{i=0}^{b(n)-2} \xi_{y_i, y_{i+1}} + \sum_{i=0}^{4} L_i \quad (5.1.9) \]

For all \( \varepsilon > 0 \)

\[ \lim_{n \to \infty} \mathbb{P} \left( \sum_{i=0}^{4} L_i \leq \varepsilon \sum_{i=1}^{\left\lfloor \log \log n \right\rfloor} F_L^{(-1)} \left( e^{-\frac{n}{\varepsilon^2}} \right) \right) = 0. \]

Then we apply this to (5.1.9) and use the results of the upper bound to obtain that for all \( \varepsilon_2 > 0 \)

\[ \lim_{n \to \infty} \mathbb{P} \left( d_L^e(u_{K_n}, v_{K_n}) < (1 + \varepsilon_2) \sum_{i=1}^{\left\lfloor \log \log n \right\rfloor} F_L^{(-1)} \left( e^{-\frac{n}{\varepsilon_2^2}} \right) \right) = 0. \quad (5.1.10) \]

Combining (5.1.2) and (5.1.10) with

\[ d_L^e(u, v) \leq d_L^e(u, u_{K_n}) + d_L^e(u_{K_n}, v_{K_n}) + d_L^e(v_{K_n}, v), \quad (5.1.11) \]

proofs the upper bound for the erased configuration, together with the lower bound this also finishes the proof of Theorem (2.2.7).
Bibliography


Weighted distances in the configuration model with conservative edge-weights xxxiii


Weighted distances in the configuration model with conservative edge-weights


---

Weighted distances in the configuration model with conservative edge-weights xxxv