Modelling real-world Networks

Generalised Random Graph and Configuration Model

Master Thesis in Mathematics (MAT491)

Angela Meier
10-693-695

Supervisor: Prof. Dr. Jean Bertoin

Zurich, October 2023

University of Zurich[UZH]

Abstract

In this master's thesis we study properties of the generalised random graph model and the configuration model, following Chapters 6 and 7 of the book "Random Graphs and Complex Networks" by Remco van der Hofstad [Hof18]. Over the past few decades, the interest in real-world networks increased tremendously. The development of computers and the internet enabled the empirical study of real-world networks in a way that was impossible before. Several empirical studies discovered that many naturally occurring networks share two properties: the scale-free property and the small-world property. Consequently, mathematicians and other scientists were and are still trying to create models that exhibit these characteristics. The generalised random graph model and the configuration model are two such models. The main focus of the present master's thesis is on studying the degree distribution of the generalised random graph and the probability of simplicity of the configuration model. It turns out, that under some constraints, the degree sequence of the generalised random graph has a power-law distribution. Furthermore, imposing regularity conditions on the configuration model yields a positive probability of simplicity. The last subchapter of this master's thesis is devoted to the discussion of the relation between the two mentioned models.

Acknowledgments

I would like to thank my supervisor Prof. Dr. Jean Bertoin for the great opportunity to write my thesis on the topic of random graphs and complex networks and for his flexibility concerning the deadline of my thesis.

I thank all my fellow students who helped me with mathematical and technical issues, especially mentioned Bettina, Caroline, and Rebekka.

During this thesis, there were many more people supporting me in various non-mathematical matters. I would like to say thank you to all of you, especially to my husband Joris, and my family.

Contents

1 Introduction **1**

1.1 Brief History of Network Science 1

1.2 Graphs and Their Degree Sequences 3

1.3 Scale-free Property 4

2 Preliminaries **9**

2.1 Probabilistic Bounds 9

2.2 Convergence of Random Variables 10

2.3 Coupling and Total Variation Distance 12

2.4 Stochastic Ordering 15

3 Generalised Random Graphs **17**

3.1 Introduction of the Model 17

3.2 Example of $\mathrm{GRG}_n(\boldsymbol{w})$ with Deterministic Weights 23

3.3 Degrees in the Generalised Random Graph 26

3.4 Degree Sequence of the Generalised Random Graph 33

3.5 Generalised Random Graph Conditioned on Its Degrees 39

3.6 Asymptotic Equivalence of Inhomogeneous Random Graphs 42

4 Configuration Model **50**

4.1 Introduction of the Model 50

4.2 Example of the Configuration Model with Fixed Degrees 56

4.3 Erased Configuration Model 58

4.4 Repeated Configuration Model: Simplicity Probability 65

4.5 Configuration Model, Uniform Simple Graphs and Generalised Random Graphs . . 71

5 Conclusion and Outlook **80**

1 Introduction

A network can broadly be described as a set of discrete elements, called vertices, and a set of connections, called edges, that link the elements. Various types of networks can be distinguished based on their structural, functional, and contextual characteristics such as social networks, transportation networks, communication networks, biological networks, information networks, power grids, citation networks, and many more. It is therefore not surprising that not only mathematicians are interested in the study of networks and their properties, but also many scientists from other fields of research.

The purpose of this first chapter is to provide motivation and inspiration for the exploration of different random graph models. We start by having a brief look at the history of the study of networks and the newly developed research field called network science. In the second part, we explain the scale-free property that many real-world networks have in common in more detail. We further present some different types of real-world networks sharing this property, which gave rise to the development of different random graph models, two of which are the generalised random graph model and the configuration model.

1.1 Brief History of Network Science

The history of network science dates back several centuries and has its roots in various disciplines such as mathematics, sociology, physics, and computer science. The study of networks has increased considerably in recent decades and has become a multidisciplinary field with broad applications in several fields. This subsection gives an overview of the key milestones in the history of network sciences. Those readers interested in more details about the history of the study of networks are referred to [NBW06], [Hof18] and [Fre04].

The mathematical study of networks can be traced back to the 18th century with the work of the Swiss mathematician Leonhard Euler. In 1736, Euler solved the famous "Seven Bridges of Königsberg" problem, which laid the foundation for graph theory. Euler's solution introduced the concept of a graph, which is a mathematical object consisting of points, also called vertices, and lines, also called edges. The land masses of Königsberg were depicted by four vertices, while the seven bridges connecting these land masses were represented by edges. Through the process of abstraction, graph theory can effectively capture the essential topological characteristics of a problem, providing a level of clarity that would be unattainable if every intricate detail were maintained (see Figure 1.1).

The study of networks in a sociological context began in the early 20th century. Pioneers like Jacob Moreno and his work on social networks and sociometry played a significant role in understanding social interactions and relationships. Moreno's techniques involved visualising social networks using

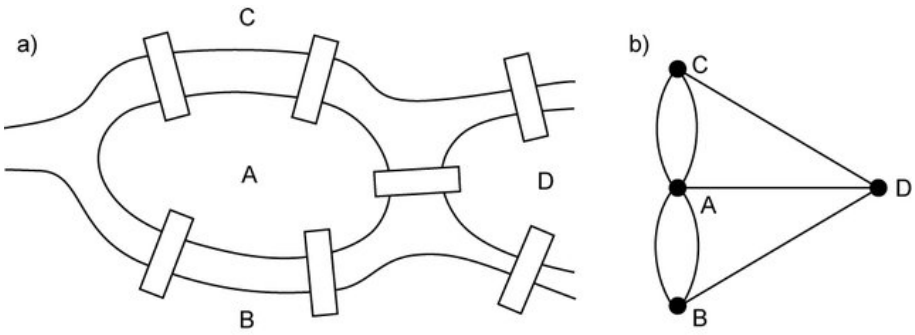

Figure 1.1: The Königsberg bridge problem
(a) land masses and seven bridges of Königsberg (b) graph representation
Source: Boguslawski, 2011, p.25.

sociograms. From the 1950s on, the mathematical language of graph theory, such as actor centrality, path lengths, and connected components, was borrowed or adapted by social scientists. Moreover, graph theory became an important tool to study the spread of information or disease.

At the end of the 1950s, the two Hungarian mathematicians Paul Erdős and Alfred Rényi initiated the field of random graphs. The Erdös-Rényi random graph, denoted by $\mathrm{ER}_n(p)$, is a graph generated by considering a fixed number of vertices n and introducing edges between pairs of vertices with a certain probability p. Each possible edge exists independently with the given probability, which means that for any two vertices $i, j \in [n]$ the probability that there is an edge linking them is p. This probabilistic approach results in a wide variety of graph structures depending on the values of n and p. Erdös and Rényi analysed the properties of random graphs and established important results related to the emergence of connected components and the existence of a giant component in such networks.

In the 1960s, American sociologist Stanley Milgram conducted the famous "small-world experiment", which showed that social networks often display a "small-world" property, where any two individuals in a large network are typically connected by a short chain of intermediate acquaintances. More specifically, Milgram's main conclusion was that the majority of individuals across the globe are linked through a series of at most six connections, a concept popularly known as "Six Degrees of Separation". This work laid the foundation for further research on the structure of real-world networks.

At the end of the 1990s, physicists Albert-László Barabási and Réka Albert made significant contributions to network science by introducing the concept of scale-free networks. They discovered that many real-world networks, including the World Wide Web and social networks, show a power-law degree distribution, where a few nodes have a disproportionately large number of connections.

From there on, network science experienced rapid growth and became an interdisciplinary field. Researchers from various disciplines, including physics, computer science, biology, neuroscience, economics, and more, started to apply network analysis techniques to their respective domains.

In recent years, network science has found applications in numerous fields. Among many others, it has been used to study the spread of diseases, analyse transportation and infrastructure systems, model information spreading in social media, understand brain connectivity in neuroscience, and analyse financial systems.

The evolution of network science has been greatly influenced by advancements in computational power, data collection techniques, and the increasing availability of large-scale network data sets. As the field continues to grow, it is expected to play an even more significant role in understanding complex systems and their interactions across different domains.

1.2 Graphs and Their Degree Sequences

Since this master's thesis is focused on random graphs, this subsection is devoted to presenting some basic definitions from graph theory following [Sud20], [Wes01] and [Hof18]. Let us start with the formal definition of a graph.

Definition 1.1 (Graph)**.** *([Sud20])*
A graph G is defined as a pair $G = (V; E)$ where $V = \{v_1, v_2, \ldots\}$ is a set of vertices and $E = \{e_1, e_2, \ldots\}$ is a set of unordered pairs of vertices $e = \{v_i, v_j\}$. The elements of E are called edges. We denote the set of vertices of a graph G by $V(G)$ and for the set of edges, we use the notation $E(G)$.

We distinguish between so-called simple graphs and multigraphs in the following way.

Definition 1.2 (Simple graphs and multigraphs)**.** *([Sud20])*
A self-loop is an edge $e = \{v, v\}$ whose endpoints are equal and $e = \{v_i, v_j\}$ is a multiple edge if it appears more than once in E. A graph that contains no self-loops or multiple edges is called a simple graph. Any graph that is not simple is called a multigraph.

Usually graphs are depicted by nodes and lines, where the nodes represent the vertices and the lines represent the edges. Figure 1.2 shows a few examples of graph representations.

In the course of this thesis, we will often be interested in the structural properties of the vertex degrees and the degree sequence of a random graph. Thus, let us introduce these concepts rigorously.

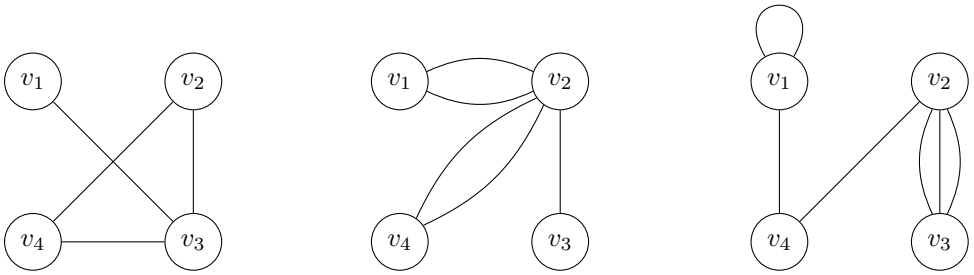

Figure 1.2: Example of a simple graph (a), a multigraph with multiple edges (b) and a multigraph with multiple edges and a self-loop (c).

Definition 1.3 (Vertex degree)**.** *([Wes01], [Hof18])*
Given a graph $G = (V; E)$ and a vertex $u \in V$, the degree of vertex u, written $d_G(u)$ or d_u, is the number of edges incident to u, except that each self-loop at u counts twice, i.e.,

$$d_u = \#\{v \in V : \{u, v\} \in E\}. \tag{1.2.1}$$

Definition 1.4 (Degree sequence of a graph)**.** *([Wes01], [Hof18])*
The degree sequence of a graph $G = (V; E)$ with n vertices $v_1, \dots, v_n$ is the list of vertex degrees $\boldsymbol{d} = (d_i)_{i\in[n]}$, where d_i is the degree of vertex $i \in [n] = \{1, \dots, n\}$.

Taking a look at the example in Figure 1.2, we observe that the simple graph (a) has the degree sequence $\boldsymbol{d} = (1, 2, 3, 2)$, whereas the two multigraphs (b)-(c) have the degree sequences $\boldsymbol{d} = (2, 5, 1, 2)$ and $\boldsymbol{d} = (3, 4, 3, 2)$, respectively.

Another important quantity, that we will use several times in the upcoming chapters, is the proportion of vertices with degree $k \in \mathbb{N}_0$ in G_n, where G_n is a graph with n vertices. $\left(P_k^{(n)}\right)_{k\geq 1}$ is often also called the empirical degree distribution of G_n.

Definition 1.5 (Proportion of vertices with degree k)**.** *([Hof18])*
The proportion of vertices with degree $k \in \mathbb{N}_0$ in G_n is given by

$$P_k^{(n)} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{d_i^{(n)}=k\}}, \tag{1.2.2}$$

where $d_i^{(n)}$ denotes the degree of vertex $i \in [n] = \{1, \dots, n\}$.

1.3 Scale-free Property

With the invention of computers and the internet, the storage and analysis of data changed tremendously. Millions of data can be analysed in no time which allows us to get a better understanding of real-world networks. Since the last few decades, many empirical studies have been conducted

which would not have been possible without the newly invented technologies. Different studies showed, that many real-world networks exhibit some common properties, such as the scale-free property, the small-world property, clustering, or a few others. In this subchapter, we focus on the description of the scale-free property, as we want to examine the two models presented in Chapters 3 and 4 with regard to this property. Those readers interested in other properties are referred to [Hof18], [New10] and [NBW06].

The scale-free property is a characteristic frequently observed in complex networks, where a small number of vertices have a significantly higher number of edges compared to the majority of vertices in the network. Frequently, the tail of the empirical degree distribution seems to decline as an inverse power of k. Degree sequences exhibiting such a distribution are called "power-law degree sequences", while the resultant graphs are referred to as "scale-free graphs". Mathematically speaking, if N_k is the number of vertices with degree k in a scale-free graph, then there exists some constant c_n and some exponent τ such that

$$N_k \approx c_n k^{-\tau}. \tag{1.3.1}$$

Moreover, taking the logarithm on both sides yields

$$\log N_k \approx \log c_n - \tau \log k. \tag{1.3.2}$$

Therefore, the plot $\log k \mapsto \log N_k$ is close to a straight line, which is the reason why degree sequences of networks are often depicted by a log-log plot. The power-law exponent τ can then be estimated by the slope of the line in the log-log plot.

Let us now take a look at two examples of real-world networks exhibiting a power-law degree distribution.

Example 1.6 (Internet)**.** *The infrastructure of the internet is one example of a huge real-world network, consisting of routers that are linked by cables. It can be modeled by a graph, where each router is represented by a vertex and an edge is present between two vertices when there is a physical cable linking them. In the case of the Faloutsos brothers (see [FFF03]) the vertices represent so-called Autonomous Systems (AS), a collection of IP networks, routers, etc. managed by a single administrative entity or organisation, whereas the edges represent links between those systems. They were the first to publish log-log plots of the degrees of the internet network (see Figure 1.3). While older studies represented the internet using a simple graph, multigraphs tend to be more common nowadays. Nevertheless, the log-log plot of a more recent study [NT18] still looks very similar as depicted in Figure 1.4. One can observe that in both studies the degree distribution is close to a straight line in the log-log plots and therefore the distribution of the degrees follows a power-law. Depending on the study, the power-law exponent is approximately* $\tau \approx 2.1 - 2.5$.

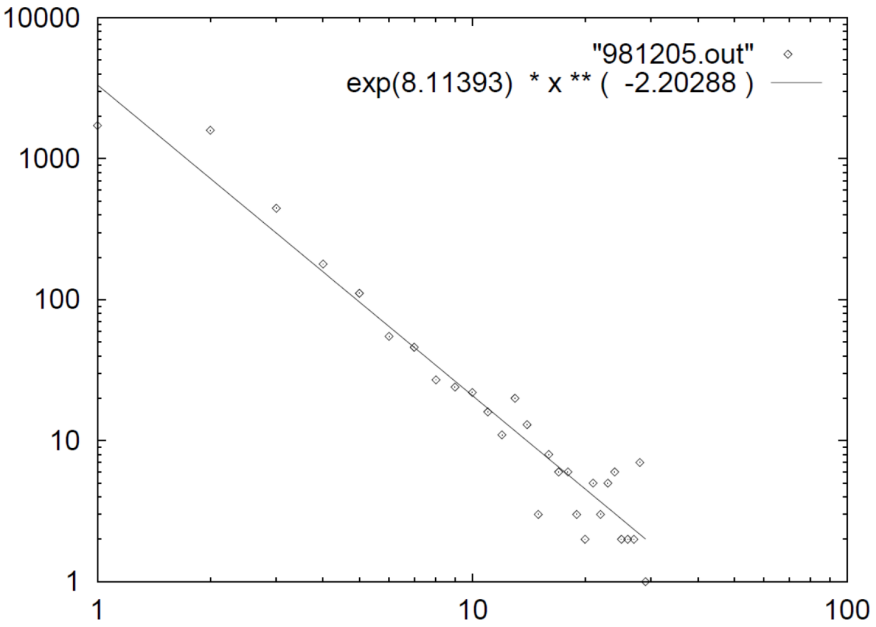

Figure 1.3: Degree sequence of Autonomous Systems (AS) on December 1998 on a log-log scale with exponent $\tau \approx 2.15 - 2.20$

Source: Faloutsos et al., 2003, p.6.

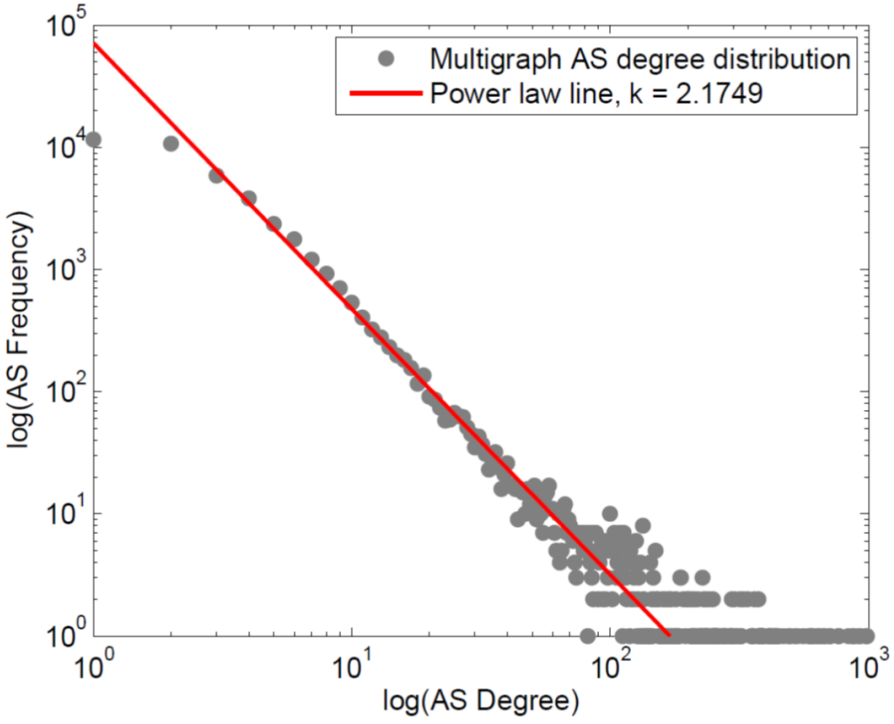

Figure 1.4: AS degree distribution on a log-log scale with exponent $\tau \approx 2.17$

Source: Nur and Tozal, 2018, p.63.

Example 1.7 (Facebook wall posts)**.** *When representing the social network of Facebook wall posts as a graph, vertices are used to symbolise individual Facebook users. An edge is established between two users when one of them posts a message onto the wall of the other. Consequently, given the possibility of contributing numerous posts on another user's wall, the network may feature multiple edges linking the same pair of vertices. Furthermore, since it is possible for a user to post something on his own wall, the network can contain self-loops. Hence, the resultant graph is a multigraph. The degree distribution of Facebook wall posts is presented in Figure 1.5, where it becomes evident that the tail of the degree distribution exhibits a power-law.*

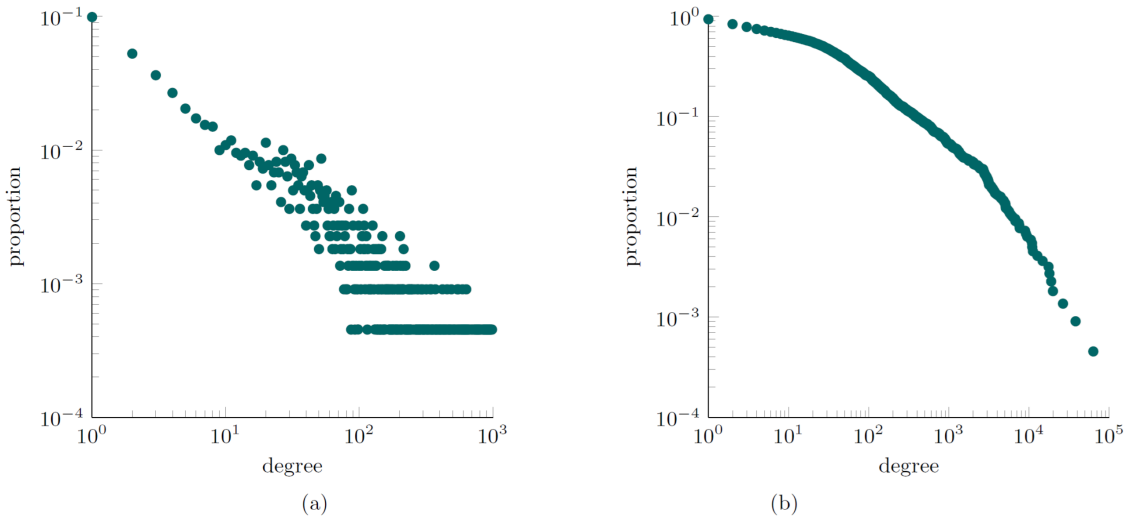

Figure 1.5: Degree distributions for networks of Facebook wall posts
(a) Log-log plot of the probability mass function. (b) Log-log plot of the complementary cumulative distribution function.
Source: Hofstad, 2018, p.214.

There are countless other real-world networks that have been studied in the last few decades that show power-law degree distributions. Moreover, many of them have power-law exponents in the interval $\tau \in (2, 3)$.

Up until now, our discussions have been centered on "scale-free" networks from an empirical perspective. In the next step, we explain what it means mathematically for a random graph model to satisfy this property. Given our interest in modeling real-world networks which are extremely large and grow in size with time, it is judicious to define the notion of scale-free random graphs as limiting statements when the size of the random graphs tends to infinity. To achieve this, we will work with graph sequences $(G_n)_{n\geq 1}$, where n denotes the number of vertices in G_n and we let $n \to \infty$.

Before defining scale-free graph sequences, we need to introduce the definition of sparse graph sequences.

Definition 1.8 (Sparse graph sequences)**.** *([Hof18], Definition 1.3)*
A graph sequences $(G_n)_{n\geq 1}$ is called sparse when

$$\lim_{n\to\infty} P_k^{(n)} = p_k, \quad k \geq 0, \tag{1.3.3}$$

where $P_k^{(n)}$ is the empirical degree distribution of G_n and $(p_k)_{k\geq 0}$ is some deterministic limiting probability distribution.

Since we deal with random graphs, i.e. with graphs whose edges are randomly generated, in Chapters 3 and 4, we will replace the convergence in (1.3.3) with convergence in probability. Moreover, since $(p_k)_{k\geq 0}$ sums up to 1, the degree of most vertices is bounded for large n, which explains the phrase "sparse" graphs. Now, we are ready to define the notion of scale-free graph sequences.

Definition 1.9 (Scale-free graph sequences)**.** *([Hof18], Definition 1.4)*
We call a graph sequence $(G_n)_{n\geq 1}$ scale-free with exponent τ when it is sparse and

$$\lim_{k\to\infty} \frac{\log[1 - F(k)]}{\log(\frac{1}{k})} = \tau - 1, \tag{1.3.4}$$

where $F(k) = \sum_{l\leq k} p_l$ denotes the cumulative distribution function corresponding to the probability mass function $(p_k)_{k\geq 0}$ defined in (1.3.3).

Therefore, a random graph process is called scale-free, when its degree distribution converges to a limiting probability distribution which has asymptotic power-law tails.

Note that there are many other possible definitions. Since in many empirical studies $\log P_k^{(n)}$ is plotted versus $\log(\frac{1}{k})$, it may be reasonable to replace (1.3.4) by

$$\lim_{k\to\infty} \frac{\log p_k}{\log(\frac{1}{k})} = \tau. \tag{1.3.5}$$

2 Preliminaries

In this chapter, we collect the preliminaries which are necessary to understand the content and proofs of the following chapters. Next to some definitions, we state a few important theorems, propositions, and lemmas without proof. Whenever possible, reference is given to books where the corresponding proofs can be found.

2.1 Probabilistic Bounds

Many proofs of the next two chapters rely on probabilistic bounds. First of all, we would like to consider the Markov inequality, also known as the first moment method.

Theorem 2.1 (Markov inequality)**.** *([Hof18], Theorem 2.17)*
Let X be a non-negative random variable with $\mathbb{E}[X] < \infty$. Then,

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}. \tag{2.1.1}$$

In particular, when X is non-negative and integer-valued with $\mathbb{E}[X] \leq m$, then

$$\mathbb{P}[X = 0] \geq 1 - m. \tag{2.1.2}$$

By (2.1.2), an integer-valued random variable with a small mean is equal to zero with high probability. This is a powerful tool.

Another important and useful probabilistic bound is given by the Chebychev inequality, also called the second moment method, which is often used to show that an integer-valued random variable is positive with high probability.

Theorem 2.2 (Chebychev inequality)**.** *([Hof18], Theorem 2.18)*
Assume that the random variable X has variance $Var(X) = \sigma^2$. Then,

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq a] \leq \frac{\sigma^2}{a^2}. \tag{2.1.3}$$

In particular, when X is non-negative and integer-valued with $\mathbb{E} \leq m$, then

$$\mathbb{P}[X = 0] \geq 1 - m. \tag{2.1.4}$$

The last inequality which we present in this chapter is called Boole's inequality (or the union bound). It states that for any at most countable collection of events, the probability that at least one of the events happens is at most the sum of the probabilities of the events in the collection.

Theorem 2.3 (Boole's inequality)**.** *([Fel68], p.23)*
Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $A_1, A_2, \ldots \in \mathcal{F}$ are events. Then,

$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] \leq \sum_{i=1}^{\infty} \mathbb{P}[A_i]. \tag{2.1.5}$$

2.2 Convergence of Random Variables

Just as there are different modes of convergence of functions on continuous spaces, we can distinguish between different notions for convergence of random variables. The following definition outlines three such notions. We refer readers who want to delve deeper into the different modes of convergence and their properties to the books [JP04] and [GS01].

Definition 2.4 (Convergence of random variables)**.** *([Hof18], Definition 2.1)*

(a) A sequence $(X_n)_{n\geq 1}$ of random variables converges in distribution to a limiting random variable X when

$$\lim_{n\to\infty} \mathbb{P}\left[X_n \leq x\right] = \mathbb{P}\left[X \leq x\right], \tag{2.2.1}$$

for every x for which $F(x) = \mathbb{P}[X \leq x]$ is continuous.
Notation: $X_n \xrightarrow{d} X$

(b) A sequence $(X_n)_{n\geq 1}$ of random variables converges in probability to a limiting random variable X when, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left[|X_n - X| > \varepsilon\right] = 0. \tag{2.2.2}$$

Notation: $X_n \xrightarrow{\mathbb{P}} X$

(c) A sequence $(X_n)_{n\geq 1}$ of random variables converges almost surely to a limiting random variable X when

$$\mathbb{P}\left[\lim_{n\to\infty} X_n = X\right] = 1. \tag{2.2.3}$$

Notation: $X_n \xrightarrow{a.s.} X$

Convergence in distribution is the weakest form of the three. In [JP04] one can find proof that almost sure convergence implies convergence in probability, which in turn implies convergence in distribution. However, in the coming chapters, we will mostly use convergence in probability and distribution.

The next theorem states that continuous functions preserve the modes of convergence, which will prove useful in later chapters. Those readers interested in a proof are referred to [Vaa98], p.8.

Theorem 2.5 (Continuous mapping theorem)**.** *([Vaa98], 2.3 Theorem (Continuous mapping))*
Let $g : \mathbb{R}^k \to \mathbb{R}^m$ be continuous at every point of a set C such that $\mathbb{P}[X \in C] = 1$.

(a) If $X_n \xrightarrow{d} X$, then $g(X_n) \xrightarrow{d} g(X)$.

(b) If $X_n \xrightarrow{\mathbb{P}} X$, then $g(X_n) \xrightarrow{\mathbb{P}} g(X)$.

(c) If $X_n \xrightarrow{a.s.} X$, then $g(X_n) \xrightarrow{a.s.} g(X)$.

There are a few other convenient criteria by which we can conclude that random variables converge in distribution. One of those criteria is stated in the following theorem and makes use of the moment-generating function of a random variable X, which is defined for $t \in \mathbb{R}$ by

$$M_X(t) = \mathbb{E}\left[e^{tX}\right]. \tag{2.2.4}$$

Those readers interested in other criteria which imply convergence in distribution as well, are referred to [Hof18], Theorem 2.3.

Theorem 2.6 (Criteria for convergence in distribution)**.** *([Hof18], Theorem 2.3)*
The sequence of random variables $(X_n)_{n\geq 1}$ converges in distribution to a random variable X when the moments $\mathbb{E}[X_n^r]$ converge to the moments $\mathbb{E}[X^r]$ of X for each $r = 1, 2, \ldots$, and $M_X(t)$, the moment-generating function of X, is finite for t in some neighbourhood of the origin.

One of the most well-known and important theorems in measure theory is Lebesgue's dominated convergence theorem (DCT), which is an essential tool in probability theory as well, as it gives sufficient conditions for the interchange of limit and expectation operations. We present here two slightly different versions, which both can be found in the appendix of [Hof18]. While almost sure convergence is one of the conditions in the first version, we only need convergence in distribution in the second one. Readers looking for proofs are referred to [JP04], p.55, and [Hof18], p.299.

Theorem 2.7 (Lebesgue's dominated convergence theorem (version 1))**.** *([Hof18], Theorem A.1) Let $(X_n)_{n\geq 1}$ and Y satisfy $\mathbb{E}[Y] < \infty$, $X_n \xrightarrow{a.s.} X$, and $|X_n| < Y$ almost surely. Then*

$$\mathbb{E}[X_n] \to \mathbb{E}[X] \tag{2.2.5}$$

and $\mathbb{E}[|X|] < \infty$.

Theorem 2.8 (Lebesgue's dominated convergence theorem (version 2))**.** *([Hof18], Theorem A.2) Let $(X_n)_{n\geq 1}$ and Y satisfy that $|X_n| < Y$ almost surely for all $n \geq 1$, where $\mathbb{E}[Y] < \infty$. Further, let $X_n \xrightarrow{d} X$. Then*

$$\mathbb{E}[X_n] \to \mathbb{E}[X] \tag{2.2.6}$$

and $\mathbb{E}[|X|] < \infty$.

In the further course of this section, we will state a few other theorems without giving their proofs which will be useful in Chapter 4 as well. We use the notation $(X)_r = X(X-1)\cdots(X-r+1)$, such that $\mathbb{E}[(X)_r]$ is the rth factorial moment of X. Some proofs or at least proof ideas can be found in [Hof18], Section 2.1, and [Bol], Section 1.4.

Theorem 2.9 (Convergence to a Poisson random variable)**.** *([Hof18], Theorem 2.4) A sequence of integer-valued random variables $(X_n)_{n\geq 1}$ converges in distribution to a Poisson random variable with parameter λ when, for all $r = 1, 2, ...,$*

$$\lim_{n\to\infty} \mathbb{E}\left[(X_n)_r\right] = \lambda^r. \tag{2.2.7}$$

Theorem 2.10 (Factorial moments of sums of indicators)**.** *([Hof18], Theorem 2.5)*
When $X = \sum_{i\in\mathcal{I}} I_i$ is a sum of indicators and $r \in \mathbb{N}$, then

$$\mathbb{E}\left[(X)_r\right] = \sum_{i_1,\dots,i_r\in\mathcal{I}}^{*} \mathbb{E}\left[\prod_{l=1}^{r} I_{i_l}\right] = \sum_{i_1,\dots,i_r\in\mathcal{I}}^{*} \mathbb{P}[I_{i_1} = \dots = I_{i_r} = 1], \tag{2.2.8}$$

where $\sum^{*}_{i_1,\dots,i_r\in\mathcal{I}}$ denotes a sum over distinct indices.

Theorem 2.11 (Convergence to independent Poisson random variables)**.** *([Hof18], Theorem 2.6)*
A vector of integer-valued random variables $((X_{1,n},\dots,X_{d,n}))_{n\geq 1}$ with $d \in \mathbb{N}$ converges in distribution to a vector of independent Poisson random variables with parameters $\lambda_1,\dots,\lambda_d$ when,

$$\lim_{n\to\infty} \mathbb{E}\left[(X_{1,n})_{r_1}\cdots(X_{d,n})_{r_d}\right] = \lambda_1^{r_1}\cdots\lambda_d^{r_d}, \tag{2.2.9}$$

for all $r_1,\dots r_d \in \mathbb{N}$.

Theorem 2.12 (Factorial moments of sums of indicators)**.** *([Hof18], Theorem 2.7)*
When $d \in \mathbb{N}$ and $X_{l,n} = \sum_{i\in\mathcal{I}_l} I_{i,l}$ for all $l = 1,\dots,d$ are sums of indicators, then

$$\begin{aligned}\mathbb{E}\left[(X_{1,n})_{r_1}\cdots(X_{d,n})_{r_d}\right] = \\ \sum_{i_1^{(1)},\dots,i_{r_1}^{(1)}\in\mathcal{I}_1}^{*} \cdots \sum_{i_1^{(d)},\dots,i_{r_d}^{(d)}\in\mathcal{I}_d}^{*} \mathbb{P}\left[I_{i_s}^{(l)} = 1 \text{ for all } l = 1,\dots,d,\ s = 1,\dots,r_l\right],\end{aligned} \tag{2.2.10}$$

where $\sum^{*}_{i_1^{(l)},\dots,i_{r_l}^{(l)}\in\mathcal{I}_l}$ denotes a sum over distinct indices for all $l = 1,\dots,d$.

2.3 Coupling and Total Variation Distance

In this section, we introduce the concept of coupling random variables, show the connection to the total variation distance, and give a few useful results without proof. Readers looking for more detailed information on the coupling of random variables as well as proofs are referred to the books of [LPW+17]", Chapters 4 and 5, and [Hof18], Chapter 2.2.

Definition 2.13 (Coupling of random variables)**.** *([Hof18], Definition 2.8)*
The random variables $(\hat{X}_1,\dots,\hat{X}_n)$ are a coupling of the random variables $X_1,\dots,X_n$ when $(\hat{X}_1,\dots,\hat{X}_n)$ are defined on the same probability space, and are such that the marginal distribution of $\hat{X}_i$ is the same as that of X_i for all $i = 1,\dots,n$, i.e. for all measurable subsets $\mathcal{E} \subset \mathbb{R}$,

$$\mathbb{P}[\hat{X}_i \in \mathcal{E}] = \mathbb{P}[X_i \in \mathcal{E}]. \tag{2.3.1}$$

It is important to note that while the random variables $X_1,\dots,X_n$ need not necessarily be defined on one probability space, the coupled random variables $(\hat{X}_1,\dots,\hat{X}_n)$ are all defined on the same probability space. Since the marginal distributions of $(\hat{X}_1,\dots,\hat{X}_n)$ are equal to those of the original random variables $X_1,\dots,X_n$, the coupled random variables $(\hat{X}_1,\dots,\hat{X}_n)$ are related to the random

variables $X_1, \ldots, X_n$. For example, if the random variables $X_1, \ldots, X_n$ are dependent, we can take the coupling $(\hat{X_1}, \ldots, \hat{X_n})$ to be independent, but such that $\hat{X_i}$ has the same distribution as X_i.

However, there are many different ways how random variables can be coupled. So one might wonder if there is something like an optimal coupling, which maximises the probability that $\hat{X}$ is equal to $\hat{Y}$. The desired result is connected to the total variation distance, which is defined in the following way.

Definition 2.14 (Total variation distance)**.** *([Hof18], p.60)*
For two probability measures μ and ν, the total variation distance is given by

$$d_{TV}(\mu, \nu) = \sup_A |\mu(A) - \nu(A)|, \tag{2.3.2}$$

where $\mu(A)$ and $\nu(A)$ are the probabilities of the event A under the measures μ and ν, and the supremum is over all subsets A of the state space S.

Now let X and Y be two discrete random variables with

$$\mathbb{P}[X = x] = p_x, \quad \mathbb{P}[Y = y] = q_y, \quad x \in \mathcal{X}, y \in \mathcal{Y}, \tag{2.3.3}$$

where $(p_x)_{x\in\mathcal{X}}$ and $(q_y)_{y\in\mathcal{Y}}$ are two probability mass functions on two subsets $\mathcal{X}$ and $\mathcal{Y}$ of the same state space S.
Then, we can define for $A \subset \mathcal{X} \cup \mathcal{Y}$

$$\mu[A] = \sum_{a\in A} p_a, \quad \nu[A] = \sum_{a\in A} q_a, \tag{2.3.4}$$

where $p_a = 0$ whenever $a \notin \mathcal{X}$ and $q_a = 0$ whenever $a \notin \mathcal{Y}$.
Setting $d_{TV}(p, q) = d_{TV}(\mu, \nu)$, the total variation distance of two discrete probability distributions can be calculated in the following ways.

Proposition 2.15 (Total variation distance of probability mass functions)**.** *([Hof18], p.61)*
Let $(p_x)_{x\in\mathcal{X}}$ and $(q_x)_{x\in\mathcal{Y}}$ be two probability mass functions of the random variables X and Y respectively, then the total variation distance is given by

$$d_{TV}(p, q) = \frac{1}{2} \sum_x |p_x - q_x|. \tag{2.3.5}$$

Proposition 2.16 (Total variation distance of probability mass functions)**.** *([Hof18], p.61)*
Let $(p_x)_{x\in\mathcal{X}}$ and $(q_x)_{x\in\mathcal{Y}}$ be two probability mass functions of the random variables X and Y respectively, then the total variation distance is given by

$$d_{TV}(p, q) = \sum_x (p_x - \min\{p_x, q_x\}) = \sum_x (q_x - \min\{p_x, q_x\}). \tag{2.3.6}$$

Readers looking for proofs of Propositions 2.15 and 2.16 are referred to [LPW$^+$17] p.48.

From Proposition 2.16, we can easily deduce the following result.

Proposition 2.17 (Pointwise convergence implies total variation convergence)**.** *([Hof18], Exercise 2.16)*
Let $p^{(n)} = (p_x^{(n)})_{x\in\mathcal{X}}$ and $p = (p_x)_{x\in\mathcal{X}}$ be probability mass functions such that $\lim_{n\to\infty} p_x^{(n)} = p_x$ for every $x \in \mathcal{X}$, then also $\lim_{n\to\infty} d_{TV}(p^{(n)}, p) = 0$.

Proof. By Proposition 2.16, we can deduce that

$$d_{TV}(p^{(n)}, p) = \sum_x (p_x - \min\{p_x, p_x^{(n)}\}) \leq \sum_x p_x = 1. \tag{2.3.7}$$

Therefore, by Lebesgue's dominated convergence theorem, we can interchange limits and sums to get that

$$\lim_{n\to\infty} d_{TV}(p^{(n)}, p) = \lim_{n\to\infty} \sum_x (p_x - \min\{p_x, p_x^{(n)}\}) \tag{2.3.8}$$

$$= \sum_x \lim_{n\to\infty} (p_x - \min\{p_x, p_x^{(n)}\}) = \sum_x 0 = 0. \tag{2.3.9}$$

□

The main result linking a coupling of two discrete random variables with their total variation distance is given by the following theorem.

Theorem 2.18 (Maximal coupling)**.** *([Hof18], Theorem 2.9)*
For any two discrete random variables X and Y, there exists a coupling $(\hat{X}, \hat{Y})$ of X and Y such that

$$\mathbb{P}[\hat{X} \neq \hat{Y}] = d_{TV}(p, q), \tag{2.3.10}$$

while, for any coupling $(\hat{X}, \hat{Y})$ of X and Y,

$$\mathbb{P}[\hat{X} \neq \hat{Y}] \geq d_{TV}(p, q). \tag{2.3.11}$$

To achieve the maximal coupling, we need to define the random vector $(\hat{X}, \hat{Y})$ by

$$\mathbb{P}[\hat{X} = \hat{Y} = x] = (p_x \wedge q_x), \tag{2.3.12}$$

$$\mathbb{P}[\hat{X} = x, \hat{Y} = y] = \frac{(p_x - (p_x \wedge q_x))(q_y - (p_y \wedge q_y))}{\frac{1}{2}\sum_z |p_z - q_z|}, \quad x \neq y. \tag{2.3.13}$$

In [Hof18] on p.61-62 a proof can be found, that the above definition leads indeed to the maximal coupling, which maximizes the probability that $\hat{X}$ and $\hat{Y}$ are equal.

In the upcoming chapters, we often want to compare binomial random variables to Poisson random variables. The following theorem will prove useful for this purpose. Readers looking for a proof are again referred to [Hof18].

Theorem 2.19 (Poisson limit for binomial random variables)**.** *([Hof18], Theorem 2.10)*
Let $(I_i)_{i=1}^n$ be independent with $I_i \sim Be(p_i)$, and let $\lambda = \sum_{i=1}^n p_i$. Let $X = \sum_{i=1}^n I_i$ and let Y be a Poisson random variable with parameter λ. Then, there exists a coupling $(\hat{X}, \hat{Y})$ of X and Y such that

$$\mathbb{P}[\hat{X} \neq \hat{Y}] \leq \sum_{i=1}^{n} p_i^2. \tag{2.3.14}$$

Consequently, for every $\lambda \geq 0$ and $n \in \mathbb{N}$, there exists a coupling $(\hat{X}, \hat{Y})$ of X and Y, where $X \sim Bin(n, \lambda/n)$ and $Y \sim Poi(\lambda)$, such that

$$\mathbb{P}[\hat{X} \neq \hat{Y}] \leq \frac{\lambda^2}{n}. \tag{2.3.15}$$

2.4 Stochastic Ordering

In this section, we introduce the notion of stochastic ordering, which can be used to compare random variables. Moreover, we state a few results, which will prove useful in Chapter 4. Additionally, we present a third version of the Lebesgue's dominated convergence theorem which relies on stochastic domination. A few more examples and theorems can be found in [Hof18], Chapter 2.3. However, those readers who want to delve deeper into the topic of stochastic ordering are referred to [SS07].

Definition 2.20 (Stochastic domination)**.** *([Hof18], Definition 2.11)*
Let X and Y be two random variables, not necessarily living on the same probability space. The random variable X is stochastically smaller than the random variable Y when, for every $x \in \mathbb{R}$, the inequality

$$\mathbb{P}[X \leq x] \geq \mathbb{P}[Y \leq x] \tag{2.4.1}$$

holds. We denote this by $X \preceq Y$.

The following lemma shows how stochastic ordering is related to the coupling of random variables. The proof can be found in [Hof18], p.64.

Lemma 2.21 (Relation between stochastic domination and coupling)**.** *([Hof18], Lemma 2.12)*
The random variable X is stochastically smaller than the random variable Y if and only if there exists a coupling $(\hat{X}, \hat{Y})$ of X, Y such that

$$\mathbb{P}[\hat{X} \leq \hat{Y}] = 1. \tag{2.4.2}$$

The next theorem claims that stochastic ordering is preserved under monotone functions. The corresponding proof can be found in [Hof18], Chapter 2.3.3.

Theorem 2.22 (Preservation of ordering under monotone functions)**.** *([Hof18], Theorem 2.16)*
Suppose $X \preceq Y$, and $g : \mathbb{R} \to \mathbb{R}$ is non-decreasing. Then $g(X) \preceq g(Y)$.

Next we state another slightly different version of the Lebesgue's dominated convergence theorem (DCT), which includes stochastic domination. The proof of this theorem can be found in the appendix of [Hof18], p.300.

Theorem 2.23 (Lebesgue's dominated convergence theorem (version 3))**.** *([Hof18], Theorem A.4) Let $(X_n)_{n\geq 1}$ and Y satisfy that $|X_n| \preceq Y$ for all $n \geq 1$, where $\mathbb{E}[Y] < \infty$. Further, $X_n \xrightarrow{d} X$. Then*

$$\mathbb{E}[X_n] \to \mathbb{E}[X] \tag{2.4.3}$$

and $\mathbb{E}[|X|] < \infty$.

3 Generalised Random Graphs

Within this chapter, we will closely examine a particular model known as the generalised random graph model. This model is an adaptation of the Erdös-Renyi model $\mathrm{ER}_n(p)$ which was already mentioned in the first chapter. One major drawback of the Erdös-Renyi random graph is the fact that its degree distribution is asymptotically close to a Poisson distribution (see [Hof18], Chapter 4) and hence is thin-tailed. This means, that there are not many vertices that exhibit large degrees, and therefore, the Erdös-Renyi model is not suitable for modelling real-world networks.

To increase the amount of variability in the degrees, the equal edge probabilities of $\mathrm{ER}_n(p)$ need to be replaced by edge occupation statuses that are independent but not equally distributed. In the generalised random graph model the edge probabilities are moderated by certain vertex weights. Vertices with high weights are more likely to have many edges than vertices with small weights.

Our exploration of the generalised random graph model will include an assessment of the model's strengths and limitations, along with its suitability for effectively representing real-world networks, specifically the scale-free characteristic observed in such networks. If the weights satisfy certain conditions, it will become evident that this model generates graphs with a power-law degree distribution. Furthermore, we will compare the generalised random graph model to another inhomogeneous random graph model using the notion of asymptotic equivalence at the end of this chapter.

3.1 Introduction of the Model

In this section, we provide a formal introduction to the generalised random graph model. Additionally, we present certain regularity conditions which are frequently applied to the vertex weights in subsequent sections.

In the generalised random graph model, the edge probability of the edge between vertices i and j, where $i \neq j$, is defined as

$$p_{ij} = p_{ij}^{(\mathrm{GRG})} = \frac{w_i w_j}{\ell_n + w_i w_j}, \tag{3.1.1}$$

where $\boldsymbol{w} = (w_i)_{i\in[n]}$ are the vertex weights (i.e. w_i is the weight of vertex i for any $i \in [n]$) and ℓ_n is the total weight of all vertices given by

$$\ell_n = \sum_{i\in[n]} w_i. \tag{3.1.2}$$

The resulting graph is denoted by $\mathrm{GRG}_n(\boldsymbol{w})$. Without loss of generality, we assume that $w_i > 0$ as otherwise vertex $i \in [n]$ will be isolated and can be removed from the graph. As anticipated, the

structure of the generalised random graph is heavily influenced by the choice of the vertex weights $\boldsymbol{w} = (w_i)_{i\in[n]}$.

Remark 3.1. *When we take $w_i \equiv \frac{n\lambda}{n-\lambda}$ for some fixed $\lambda \in \mathbb{R}^+$ and for any $i \in [n]$, then $p_{ij} = \frac{\lambda}{n}$ for all $i, j \in [n]$ such that $i \neq j$. Thus, we recover the Erdös-Rényi random graph $ER_n(\frac{\lambda}{n})$.*

In order to describe the empirical properties of the weights, we need to define their empirical distribution function.

Definition 3.2 (Empirical distribution function of vertex weights)**.** *([Hof18], p.183)*
The empirical distribution function of the vertex weights $\boldsymbol{w} = (w_i)_{i\in[n]}$ is given by

$$F_n(x) = \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{w_i\leq x\}}, \quad x \geq 0. \tag{3.1.3}$$

The empirical distribution function F_n can be interpreted as the distribution of the weight of a uniformly chosen vertex. Let us denote the weight of a uniformly chosen vertex U in $[n]$ by $W_n = w_U$. In that case, W_n has the cumulative distribution function F_n.

Since we want to prove that under some constraints the degree distribution of $\mathrm{GRG}_n(\boldsymbol{w})$ obeys a power-law, we need to study the asymptotic degree distribution in the generalised random graph. In particular, we aim to prove that $P_k^{(n)}$, the proportion of vertices of degree k, approaches a limit when $n \to \infty$. As a matter of fact, the degree distribution can only converge when the vertex weights are sufficiently regular. Therefore, we often impose the following regularity conditions on the vertex weights.

Condition 3.3 (Regularity conditions for vertex weights)**.** *([Hof23], Condition 1.1)*
There exists a distribution function F such that, as $n \to \infty$ the following conditions hold:

(a) Weak convergence of vertex weight.
As $n \to \infty$,

$$W_n \xrightarrow{d} W, \tag{3.1.4}$$

where W_n and W have cumulative distribution functions F_n and F, respectively.
Equivalently, for any x for which $x \mapsto F(x)$ is continuous,

$$\lim_{n\to\infty} F_n(x) = F(x). \tag{3.1.5}$$

(b) Convergence of average vertex weight.
As $n \to \infty$,

$$\lim_{n\to\infty} \mathbb{E}[W_n] = \mathbb{E}[W] \in (0, \infty), \tag{3.1.6}$$

where W_n and W have cumulative distribution functions F_n and F from part (a), respectively.

(c) Convergence of second moment vertex weight.

As $n \to \infty$,

$$\lim_{n\to\infty} \mathbb{E}\left[W_n^2\right] = \mathbb{E}[W^2] < \infty, \tag{3.1.7}$$

where W_n and W have cumulative distribution functions F_n and F from part (a), respectively.

While the first condition implies that the weight of a random chosen vertex is close to a random variable W which is independent of n, the second condition guarantees that the average weight of the vertices in $\mathrm{GRG}_n(\boldsymbol{w})$ converges to the expectation of the limiting weight variable W. Additionally, the third condition ensures that the second moment of the vertex weights converges to the second moment of W.

In most upcoming theorems we need to assume Conditions 3.3 (a)-(b) to get the desired results. In some theorems, we will also make use of Condition 3.3 (c).

In the next proposition, we investigate bounds on the maximal weight $\max_{i\in[n]} w_i$ assuming that the above regularity conditions are satisfied.

Proposition 3.4 (Bound on the maximal weight assuming regularity conditions)**.** *([Hof18], Exercise 6.3)*
Let $(w_i)_{i\in[n]}$ be vertex weights fulfiling Conditions 3.3 (a)-(b). Then,

$$\max_{i\in[n]} w_i = o(n). \tag{3.1.8}$$

When 3.3 (c) holds as well, then

$$\max_{i\in[n]} w_i = o\big(\sqrt{n}\big). \tag{3.1.9}$$

Proof. Let us first assume that Conditions 3.3 (a)-(b) hold. Moreover, let $\varepsilon > 0$ be arbitrary and $M > 0$ large enough such that

$$\mathbb{E}[W] - \mathbb{E}[W \wedge M] < \varepsilon. \tag{3.1.10}$$

By the dominated convergence theorem and the assumed regularity conditions, we can deduce that

$$\begin{aligned} \mathbb{E}[W \wedge M] &= \int_0^M \mathbb{P}[W > t]dt = \int_0^M \lim_{n\to\infty} \mathbb{P}[W_n > t]dt \\ &= \lim_{n\to\infty} \int_0^M \mathbb{P}[W_n > t]dt = \lim_{n\to\infty} \mathbb{E}[W_n \wedge M]. \end{aligned} \tag{3.1.11}$$

Next, note that

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}[W_n - (W_n \wedge M)] &= \lim_{n\to\infty} \frac{1}{n} \sum_{i\in[n]:w_i>M} (w_i - M) \\ &= \mathbb{E}[W] - \mathbb{E}[W \wedge M] < \varepsilon. \end{aligned} \tag{3.1.12}$$

If there exists some $n \in \mathbb{N}$ such that $\max_{i\in[n]} w_i > 2\varepsilon n + M$, then

$$\sum_{i\in[n]:w_i>M} (w_i - M) \geq \max_{i\in[n]} w_i - M > 2\varepsilon n. \tag{3.1.13}$$

However, this can only happen for a finite number of n, since otherwise $\mathbb{E}[W] - \mathbb{E}[W \wedge M] \geq \varepsilon$, which contradicts equation (3.1.12). Therefore, we can deduce that

$$\limsup_{n\to\infty} \frac{1}{n} \max_{i\in[n]} w_i \leq 2\varepsilon. \tag{3.1.14}$$

As this holds for any $\varepsilon > 0$, we can conclude that

$$\limsup_{n\to\infty} \frac{1}{n} \max_{i\in[n]} w_i = 0, \tag{3.1.15}$$

and hence $\max_{i\in[n]} w_i = o(n)$. Using similar arguments, it can be shown that $\max_{i\in[n]} w_i = o\big(\sqrt{n}\big)$, when Condition 3.3 (c) holds as well. □

Another theorem that proves to be advantageous in subsequent chapters concerns the total number of edges in $\mathrm{GRG}_n(\boldsymbol{w})$.

Theorem 3.5 (Total number of edges in $\mathrm{GRG}_n(\boldsymbol{w})$)**.** *([Hof18], Theorem 6.6)*
Assume that the prescribed weights $\boldsymbol{w}$ of $GRG_n(\boldsymbol{w})$ satisfy Conditions 3.3 (a)-(b). Then,

$$\frac{1}{n} E(GRG_n(\boldsymbol{w})) \xrightarrow{\mathbb{P}} \frac{1}{2}\mathbb{E}[W], \tag{3.1.16}$$

where $E(GRG_n(\boldsymbol{w}))$ denotes the total number of edges in $GRG_n(\boldsymbol{w})$.

Proof. In a first step, we explore the first moment of the number of edges and prove that

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))] = \frac{1}{2}\mathbb{E}[W]. \tag{3.1.17}$$

Recall that p_{ij} denotes the probability that there is an edge between vertex i and j. Thus, we can observe that

$$\begin{aligned}
\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))] &= \frac{1}{2} \sum_{i,j\in[n]:i\neq j} p_{ij} \\
&= \frac{1}{2} \sum_{i,j\in[n]:i\neq j} \frac{w_i w_j}{\ell_n + w_i w_j} \\
&\leq \frac{1}{2} \sum_{i,j\in[n]} \frac{w_i w_j}{\ell_n} \\
&= \frac{\ell_n}{2},
\end{aligned} \tag{3.1.18}$$

which gives an upper bound. To find a lower bound, we have to use a truncation argument, as $w_i w_j$ can be rather large even though Conditions 3.3 (a)-(b) hold. Thus, we fix a real-valued sequence

$a_n \to \infty$. Using that $w_i \geq w_i \wedge a_n$ and the fact that $x \mapsto \frac{x}{\ell_n + x}$ is an increasing function, we can establish a lower bound on the expected number of edges by

$$\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))] \geq \frac{1}{2} \sum_{i,j \in [n]: i \neq j} \frac{(w_i \wedge a_n)(w_j \wedge a_n)}{\ell_n + (w_i \wedge a_n)(w_j \wedge a_n)}. \tag{3.1.19}$$

Let $\ell_n(a_n) = \sum_{i \in [n]} (w_i \wedge a_n)$. Using the derived lower bound, we get

$$\begin{aligned}
&\frac{\ell_n(a_n)^2}{\ell_n} - 2 \cdot \mathbb{E}\left[E(\mathrm{GRG}_n(\boldsymbol{w}))\right] \\
&= \frac{(\sum_{i \in [n]} (w_i \wedge a_n))^2}{\ell_n} - \sum_{i,j \in [n]: i \neq j} \frac{w_i w_j}{\ell_n + w_i w_j} \\
&\leq \frac{\sum_{i \in [n]} (w_i \wedge a_n)^2 + \sum_{i,j \in [n]: i \neq j} (w_i \wedge a_n)(w_j \wedge a_n)}{\ell_n} - \sum_{i,j \in [n]: i \neq j} \frac{(w_i \wedge a_n)(w_j \wedge a_n)}{\ell_n + (w_i \wedge a_n)(w_j \wedge a_n)} \\
&= \sum_{i \in [n]} \frac{(w_i \wedge a_n)^2}{\ell_n} + \sum_{i,j \in [n]: i \neq j} \frac{(w_i \wedge a_n)(w_j \wedge a_n)}{\ell_n} - \sum_{i,j \in [n]: i \neq j} \frac{(w_i \wedge a_n)(w_j \wedge a_n)}{\ell_n + (w_i \wedge a_n)(w_j \wedge a_n)} \\
&= \sum_{i \in [n]} \frac{(w_i \wedge a_n)^2}{\ell_n} + \sum_{i,j \in [n]} (w_i \wedge a_n)(w_j \wedge a_n) \left[\frac{1}{\ell_n} - \frac{1}{\ell_n + (w_i \wedge a_n)(w_j \wedge a_n)} \right] \\
&\qquad - \sum_{i \in [n]} \left[\frac{(w_i \wedge a_n)^2}{\ell_n} - \frac{(w_i \wedge a_n)^2}{\ell_n + (w_i \wedge a_n)^2} \right] \\
&= \sum_{i \in [n]} \frac{(w_i \wedge a_n)^2}{\ell_n + (w_i \wedge a_n)^2} + \sum_{i,j \in [n]} \frac{(w_i \wedge a_n)^2 (w_j \wedge a_n)^2}{\ell_n (\ell_n + (w_i \wedge a_n)(w_j \wedge a_n))} \\
&\leq \sum_{i \in [n]} \frac{(w_i \wedge a_n)^2}{\ell_n} \left(1 + \sum_{i \in [n]} \frac{(w_i \wedge a_n)^2}{\ell_n} \right) \\
&\leq a_n(1 + a_n),
\end{aligned} \tag{3.1.20}$$

where we used that

$$\sum_{i \in [n]} \frac{(w_i \wedge a_n)^2}{\ell_n} \leq \sum_{i \in [n]} \frac{w_i \cdot a_n}{\ell_n} = a_n. \tag{3.1.21}$$

Therefore, when we choose $a_n = o\big(\sqrt{n}\big)$, then $a_n(1 + a_n) = o(n)$ and

$$\lim_{n \to \infty} \frac{1}{n} \left(\frac{\ell_n(a_n)^2}{\ell_n} - 2\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))] \right) = 0. \tag{3.1.22}$$

Hence, to prove that $\lim_{n \to \infty} \frac{1}{n} \mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))] = \frac{1}{2}\mathbb{E}[W]$, it suffices to show that

$$\lim_{n \to \infty} \frac{\ell_n}{n} = \mathbb{E}[W], \tag{3.1.23}$$

$$\lim_{n \to \infty} \frac{\ell_n(a_n)^2}{n \ell_n} = \mathbb{E}[W]. \tag{3.1.24}$$

By assumption Condition 3.3 (b) holds, thus $\mathbb{E}[W_n] = \frac{\ell_n}{n} \to \mathbb{E}[W]$.

Next, we observe that we can bound $\frac{\ell_n(a_n)}{n} = \mathbb{E}[(W_n \wedge a_n)]$ from above by

$$\mathbb{E}[(W_n \wedge a_n)] \leq \mathbb{E}[W_n] \to \mathbb{E}[W]. \tag{3.1.25}$$

As $a_n \to \infty$, we get that, for every $K \geq 0$ and n sufficiently large, $\mathbb{E}[(W_n \wedge a_n)] \geq \mathbb{E}[(W_n \wedge K)]$. Using Condition 3.3 (a) and Lebesgue's dominated convergence theorem 2.8, we get that

$$\mathbb{E}[(W_n \wedge K)] \to \mathbb{E}[(W \wedge K)]. \tag{3.1.26}$$

By the monotone convergence theorem, $\mathbb{E}[(W \wedge K)] \to \mathbb{E}[W]$, as $K \to \infty$, from which we can establish the lower bound $\liminf_{n\to\infty} \mathbb{E}[(W_n \wedge a_n)] \geq \mathbb{E}[W]$. Therefore, we can deduce that

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))] = \frac{1}{2}\mathbb{E}[W], \tag{3.1.27}$$

as desired.

In the next step, we intend to bound $\mathrm{Var}(E(\mathrm{GRG}_n(\boldsymbol{w})))$. To do that, let us introduce the edge statuses $(X_{ij})_{1\leq i<j\leq n}$, where X_{ij} is the indicator that the edge ij is occupied. The random variables $(X_{ij})_{1\leq i<j\leq n}$ are independent Bernoulli variables with probability $\mathbb{P}[X_{ij} = 1] = p_{ij}$. Thus, we get

$$\begin{aligned}
\mathrm{Var}(E(\mathrm{GRG}_n(\boldsymbol{w}))) &= \frac{1}{2}\sum_{i,j\in[n]:i\neq j} \mathrm{Var}(X_{ij}) \\
&= \frac{1}{2}\sum_{i,j\in[n]:i\neq j} p_{ij}(1-p_{ij}) \\
&\leq \frac{1}{2}\sum_{i,j\in[n]:i\neq j} p_{ij} \\
&= \mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))].
\end{aligned} \tag{3.1.28}$$

By the facts that $\mathbb{E}[W] \geq 0$ by Conditions 3.3 (a)-(b) and $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w})] = \frac{1}{2}\mathbb{E}[W]$, we obtain that $\mathrm{Var}(E(\mathrm{GRG}_n(\boldsymbol{w}))) = o(\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))]^2)$. Using Chebychev's inequality 2.2, we can deduce that, for any $\varepsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left[\left|\frac{E(\mathrm{GRG}_n(\boldsymbol{w}))}{\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))]} - 1\right| > \varepsilon\right] \\
&= \mathbb{P}[|E(\mathrm{GRG}_n(\boldsymbol{w})) - \mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))]| > \varepsilon\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))]] \\
&\leq \frac{\mathrm{Var}(E(\mathrm{GRG}_n(\boldsymbol{w})))}{\varepsilon^2\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))]^2} \to 0.
\end{aligned} \tag{3.1.29}$$

That yields $E(\mathrm{GRG}_n(\boldsymbol{w})) \xrightarrow{\mathbb{P}} \mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))]$. As $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[E(\mathrm{GRG}_n(\boldsymbol{w}))] = \frac{1}{2}\mathbb{E}[W]$, we conclude that $\frac{1}{n}E(\mathrm{GRG}_n(\boldsymbol{w})) \xrightarrow{\mathbb{P}} \frac{1}{2}\mathbb{E}[W]$. □

3.2 Example of $\mathrm{GRG}_n(\boldsymbol{w})$ with Deterministic Weights

Within this subsection, we will present an approach for determining the vertex weights of the generalised random graph in a manner that ensures the fulfilment of 3.3 (a)-(c).

We start by choosing any cumulative distribution function F for which $F(0) = 0$ and fix

$$w_i = [1 - F]^{-1}\left(\frac{i}{n}\right), \tag{3.2.1}$$

where $[1 - F]^{-1}$ is the generalised inverse function of $1 - F$ defined by

$$[1 - F]^{-1}(u) = \inf\{x : [1 - F](x) \leq u\}, \tag{3.2.2}$$

for $u \in (0, 1)$.

Conventionally, we set $[1 - F]^{-1}(1) = 0$. The definition of $[1 - F]^{-1}$ is chosen in such a way that

$$[1 - F]^{-1}(1 - u) = F^{-1}(u) = \inf\{x : F(x) \geq u\}, \tag{3.2.3}$$

which implies that for the uniform random variable U on $(0, 1)$, $[1 - F]^{-1}(U)$ has distribution function F. For the choice of the weights $(w_i)_{i\in[n]}$ in (3.2.1) and due to the preceding explanations, we can now compute F_n explicitly as

$$\begin{aligned} F_n(x) &= \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{w_i\leq x\}} = \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{[1-F]^{-1}(\frac{i}{n})\leq x\}} = \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_{\{[1-F]^{-1}(1-\frac{j}{n})\leq x\}} \\ &= \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_{\{F^{-1}(\frac{j}{n})\leq x\}} = \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_{\{\frac{j}{n}\leq F(x)\}} = \frac{1}{n}(\lfloor nF(x)\rfloor + 1) \wedge 1. \end{aligned} \tag{3.2.4}$$

Let us now show, that the above choice of weights satisfies the regularity Conditions 3.3 (a)-(c), when F is the distribution function of a random variable W s.t. $\mathbb{E}[W] < \infty$ and $\mathbb{E}[W^2] < \infty$.

Proposition 3.6 (Regularity condition 3.3 (a))**.** *([Hof18], Exercise 6.4)*
Let F be a distribution function for which $F(0) = 0$ and $w_i = [1 - F]^{-1}\left(\frac{i}{n}\right)$, then the weights $(w_i)_{i\in[n]}$ satisfy Condition 3.3 (a).

Proof. We are going to show that $F_n(x)$ converges pointwise to $F(x)$ for any x for which $x \mapsto F(x)$ is continuous.
First we note that by (3.2.4)

$$F_n(x) = \frac{1}{n}\left(\lfloor nF(x)\rfloor + 1\right) \wedge 1 \leq \frac{1}{n}(nF(x) + 1) = F(x) + \frac{1}{n}. \tag{3.2.5}$$

Moreover, $\lfloor x\rfloor + 1 \geq x$ implies that

$$F_n(x) = \frac{1}{n}(\lfloor nF(x)\rfloor + 1) \wedge 1 \geq \frac{1}{n}(\lfloor nF(x)\rfloor) \wedge 1 \geq \frac{1}{n}(nF(x)) \wedge 1 = F(x), \tag{3.2.6}$$

since F is a distribution function and hence $F(x) \leq 1$ for any x.
We deduce that

$$|F_n(x) - F(x)| \leq \frac{1}{n}. \tag{3.2.7}$$

Therefore,

$$\lim_{n\to\infty} F_n(x) = F(x), \tag{3.2.8}$$

which proves the pointwise convergence of F_n to F, as desired. □

In order to show that the weights $(w_i)_{i\in[n]}$ fulfil Conditions 3.3 (b) and (c) as well, we are going to prove two more general results, from which we can deduce the desired conditions.

Theorem 3.7 (Boundaries of means using non-decreasing functions)**.** *([Hof18], Exercise 6.5) Let F be a distribution function for which $F(0) = 0$ and $w_i = [1 - F]^{-1}\left(\frac{i}{n}\right)$. Then, for every non-decreasing function $x \mapsto h(x)$,*

$$\frac{1}{n}\sum_{i\in[n]} h(w_i) \leq \mathbb{E}\left[h(W)\right], \tag{3.2.9}$$

where W is a random variable with cumulative distribution function F.

Proof. First, let us note that $x \mapsto F(x)$ is non-decreasing, since F is a cumulative distribution function, while the function $x \mapsto [1 - F](x)$ is non-increasing. Hence, $x \mapsto [1 - F]^{-1}(x)$ in non-increasing as well.

Then, let U be a uniform random variable on $(0, 1)$ and W_n be a weight drawn uniformly at random from all weights, also called a "typical" weight. Note that

$$\frac{1}{n}\sum_{i=1}^{n} h(w_i) = \mathbb{E}[h(W_n)] = \mathbb{E}\left[h\left([1 - F]^{-1}(\lceil Un\rceil/n)\right)\right]. \tag{3.2.10}$$

Now, since $\lceil Un\rceil/n \geq U$ and $u \mapsto [1 - F]^{-1}(u)$ is non-increasing, we can deduce that

$$[1 - F]^{-1}(\lceil Un\rceil/n) \leq [1 - F]^{-1}(U). \tag{3.2.11}$$

Moreover, since $x \mapsto h(x)$ is non-decreasing by assumption,

$$h\left([1 - F]^{-1}(\lceil Un\rceil)/n)\right) \leq h\left([1 - F]^{-1}(U)\right). \tag{3.2.12}$$

Since $[1 - F]^{-1}(U)$ has distribution function F when U is uniform on $(0, 1)$, we can deduce that

$$\frac{1}{n}\sum_{i=1}^{n} h(w_i) = \mathbb{E}\left[h\left([1 - F]^{-1}\left(\lceil Un\rceil/n\right)\right)\right] \leq \mathbb{E}\left[h\left([1 - F]^{-1}(U)\right)\right] = \mathbb{E}[h(W)]. \tag{3.2.13}$$

□

Theorem 3.8 (Moments of $\boldsymbol{w}$ and F)**.** *([Hof18], Exercise 6.6)*
Let $\alpha > 0$ and W be a random variable with distribution function F. Assume that $\mathbb{E}[W^\alpha] < \infty$. Then

$$\frac{1}{n}\sum_{i\in[n]} w_i^\alpha \to \mathbb{E}[W^\alpha]. \tag{3.2.14}$$

Proof. Let U be a uniform random variable on $(0,1)$ and W_n be a weight drawn uniformly at random from all weights, then

$$\frac{1}{n}\sum_{i=1}^{n} w_i^\alpha = \mathbb{E}[W_n^\alpha] = \mathbb{E}\left[\left([1-F]^{-1}(\lceil Un\rceil/n)\right)^\alpha\right]. \tag{3.2.15}$$

By Theorem 3.7 and using that $x \mapsto x^\alpha$ is non-decreasing, while $\lceil Un\rceil/n \geq U$, we get that

$$\frac{1}{n}\left([1-F]^{-1}(\lceil Un\rceil/n)\right)^\alpha \leq \frac{1}{n}\left([1-F]^{-1}(U)\right)^\alpha. \tag{3.2.16}$$

Notice that $\lceil Un\rceil/n$ converges to U pointwise, which implies almost sure convergence. Therefore, by the continuous mapping theorem, we obtain that the left-hand side of the above inequality converges almost surely to the right-hand side. Using the assumption that $\mathbb{E}[W^\alpha] < \infty$ and the dominated convergence theorem, we get that

$$\frac{1}{n}\sum_{i\in[n]} w_i^\alpha \to \mathbb{E}[W^\alpha], \tag{3.2.17}$$

as desired. □

Corollary 3.9 (Regularity conditions 3.3 (b) and (c))**.** *([Hof18], Exercise 6.6)*
Let W be a random variable with cumulative distribution function F, where $F(0) = 0$ and such that $\mathbb{E}[W] < \infty$ and $\mathbb{E}[W^2] < \infty$. Set $w_i = [1-F]^{-1}\left(\frac{i}{n}\right)$, then the weights $(w_i)_{i\in[n]}$ satisfy Conditions 3.3 (b) and (c).

Proof. Let W_n be a vertex weight drawn uniformly at random from all weights $(w_i)_{i\in[n]}$. Choosing $\alpha = 1$, we get by Theorem 3.8 and by the assumption $\mathbb{E}[W] < \infty$ that, as $n \to \infty$,

$$\mathbb{E}[W_n] = \frac{1}{n}\sum_{i\in[n]} w_i \to \mathbb{E}[W]. \tag{3.2.18}$$

Similarly, when we set $\alpha = 2$, we get by Theorem 3.8 and the assumption $\mathbb{E}[W^2] < \infty$ that, as $n \to \infty$,

$$\mathbb{E}\left[W_n^2\right] = \frac{1}{n}\sum_{i\in[n]} w_i^2 \to \mathbb{E}[W^2]. \tag{3.2.19}$$

□

If we want that the degree distribution of the generalised random graph follows asymptotically a power-law distribution we can take F to have a power-law distribution.

More specifically, for some real-valued $a \geq 0$ and $\tau > 1$, let

$$F(x) = \begin{cases} 0 & \text{for } x \leq a, \\ 1 - (a/x)^{\tau-1} & \text{for } x > a. \end{cases} \tag{3.2.20}$$

Then

$$[1 - F]^{-1}(u) = au^{-1/(\tau-1)}, \tag{3.2.21}$$

so that

$$w_i = a\left(\frac{i}{n}\right)^{-1/(\tau-1)}. \tag{3.2.22}$$

3.3 Degrees in the Generalised Random Graph

In this section, we study the degree of vertices in $\mathrm{GRG}_n(\boldsymbol{w})$ as well as the degree of a uniformly chosen vertex in $\mathrm{GRG}_n(\boldsymbol{w})$. Before stating the main results, we need to introduce some notations and definitions.

Let us write D_i for the degree of vertex i in $\mathrm{GRG}_n(\boldsymbol{w})$. That is,

$$D_i = \sum_{j \in [n]} X_{ij}, \tag{3.3.1}$$

where X_{ij} is the indicator that the edge ij is occupied. Clearly, we need that $X_{ij} = X_{ji}$. The random variables $(X_{ij})_{1 \leq i < j \leq n}$ are independent Bernoulli random variables with $\mathbb{P}[X_{ij} = 1] = p_{ij}$. Using Theorem 2.19 we can deduce our first result.

Lemma 3.10 (Coupling of the degree D_i and a Poisson random variable)**.** *([Hof18], p.190)*
Let D_i be the degree of vertex i in $GRG_n(\boldsymbol{w})$ and Y_i a Poisson random variable with parameter $\lambda_i = \sum_{j \neq i} p_{ij} = \sum_{j \neq i} \frac{w_i w_j}{\ell_n + w_i w_j}$. Then, there exists a coupling $(\hat{D}_i, \hat{Y}_i)$ of D_i and Y_i such that

$$\mathbb{P}\left[\hat{D}_i \neq \hat{Y}_i\right] \leq \frac{w_i^2}{\ell_n} \frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]}. \tag{3.3.2}$$

Proof. Since $D_i = \sum_{j \in [n]: j \neq i} X_{ij}$ is a sum of independent Bernoulli random variables where $X_{ij} \sim Be(p_{ij})$, we can make use of Theorem 2.19 to couple D_i with a Poisson random variable Y_i with parameter $\lambda_i = \sum_{j \in [n]: j \neq i} p_{ij}$ in such a way that

$$\mathbb{P}\left[\hat{D}_i \neq \hat{Y}_i\right] \leq \sum_{j \in [n]: j \neq i} p_{ij}^2. \tag{3.3.3}$$

From that we can deduce that

$$\sum_{j \in [n]: j \neq i} p_{ij}^2 = \sum_{j \in [n]: j \neq i} \frac{w_i^2 w_j^2}{(\ell_n + w_i w_j)^2} \leq w_i^2 \sum_{j \in [n]} \frac{w_j^2}{\ell_n^2} = \frac{w_i^2}{\ell_n} \frac{n}{\ell_n} \sum_{j \in [n]} \frac{w_j^2}{n} = \frac{w_i^2}{\ell_n} \frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]}, \tag{3.3.4}$$

where we used that $\mathbb{E}[W_n] = \frac{\ell_n}{n}$ and $\mathbb{E}\left[W_n^2\right] = \frac{1}{n}\sum_{i\in[n]} w_i^2$.

Thus, we can conclude

$$\mathbb{P}\Big[\hat{D}_i \neq \hat{Y}_i\Big] \leq \frac{w_i^2}{\ell_n}\frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]}. \tag{3.3.5}$$

□

This lemma will come in handy to prove our first main result concerning vertex degrees. A noteworthy implication of the following theorem is, that the degree of vertex i is close to Poisson with parameter w_i if the weights $(w_i)_{i\in[n]}$ are sufficiently regular. We will delve deeper into this outcome after the statement and the proof of the theorem.

Theorem 3.11 (Degree of GRG with deterministic weights)**.** *([Hof18], Theorem 6.7)*

(a) There exists a coupling $(\hat{D}_i, \hat{Z}_i)$ of the degree D_i of vertex i and a Poisson random variable Z_i with parameter w_i, such that

$$\mathbb{P}\Big[\hat{D}_i \neq \hat{Z}_i\Big] \leq \frac{w_i^2}{l_n}\left(1 + 2\frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]}\right). \tag{3.3.6}$$

(b) Assume that Conditions 3.3 (a)-(b) hold. When $p_{ij} = \frac{w_i\cdot w_j}{l_n + w_i\cdot w_j}$ are all such that $\lim_{n\to\infty}\max_{1\leq i<j\leq m} p_{ij} = 0$, then the degrees $D_1, \ldots, D_m$ of vertices $1, \ldots, m$ are asymptotically independent.

Proof.

(a) As seen in Lemma 3.10, we can couple the degree of vertex i in $\mathrm{GRG}_n(\boldsymbol{w})$ with a Poisson random variable Y_i such that

$$\mathbb{P}\Big[\hat{D}_i \neq \hat{Y}_i\Big] \leq \frac{w_i^2}{\ell_n}\frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]}. \tag{3.3.7}$$

In a next step, we show how to couple $(\hat{D}_i, \hat{Y}_i)$ to $(\hat{D}_i, \hat{Z}_i)$ where $\hat{Z}_i$ is a Poisson random variable with parameter w_i, such that

$$\mathbb{P}\Big[\hat{Y}_i \neq \hat{Z}_i\Big] \leq \frac{w_i^2}{\ell_n}\left(1 + \frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]}\right). \tag{3.3.8}$$

To accomplish this, we first note that

$$\lambda_i = \sum_{j\neq 1}\frac{w_i w_j}{\ell_n + w_i w_j} \leq \sum_{j\neq i}\frac{w_i w_j}{\ell_n} \leq \frac{w_i}{\ell_n}\sum_{j\in[n]} w_j = \frac{w_i}{\ell_n}\cdot\ell_n = w_i. \tag{3.3.9}$$

Thus, we define $\varepsilon_i := w_i - \lambda_i$ and observe that $\varepsilon_i \geq 0$ by (3.3.9). Then, we let $\hat{V}_i \sim Poi(\varepsilon_i)$ be independent of $(\hat{D}_i, \hat{Y}_i)$, and set $\hat{Z}_i := \hat{Y}_i + \hat{V}_i$. As $\hat{Z}_i$ is the sum of two independent Poisson random variables, we get that $\hat{Z}_i \sim Poi(\lambda_i + \varepsilon_i) = Poi(w_i)$.

By Markov's inequality 2.1 we derive that

$$\mathbb{P}\Big[\hat{Y}_i \neq \hat{Z}_i\Big] = \mathbb{P}\Big[\hat{V}_i \neq 0\Big] = \mathbb{P}\Big[\hat{V}_i \geq 1\Big] \leq \mathbb{E}\Big[\hat{V}_i\Big] = \varepsilon_i, \tag{3.3.10}$$

where the last equation holds due to the fact that $\mathbb{E}[X] = \lambda$ for a Poisson random variable X with parameter λ.

We can bound ε_i by noting that

$$\begin{aligned}
\varepsilon_i = w_i - \lambda_i = w_i - \sum_{j \neq i} \frac{w_i w_j}{\ell_n + w_i w_j} &= \frac{w_i}{\ell_n} \cdot \ell_n - \sum_{j \neq i} \frac{w_i w_j}{\ell_n + w_i w_j} \\
&= \frac{w_i}{\ell_n} \sum_{j \in [n]} w_j - \sum_{j \in [n]} \frac{w_i w_j}{\ell_n + w_i w_j} + \frac{w_i^2}{\ell_n + w_i^2} \\
&= \sum_{j \in [n]} w_i w_j \left(\frac{1}{\ell_n} - \frac{1}{\ell_n + w_i w_j} \right) + \frac{w_i^2}{\ell_n + w_i^2} \\
&= \sum_{j \in [n]} \frac{w_i^2 w_j^2}{\ell_n (\ell_n + w_i w_j)} + \frac{w_i^2}{\ell_n + w_i^2} \leq \frac{w_i^2}{\ell_n} + \sum_{j \in [n]} \frac{w_i^2 w_j^2}{\ell_n^2} \\
&= \frac{w_i^2}{\ell_n} \left(1 + \frac{n}{\ell_n} \sum_{j \in [n]} \frac{w_j^2}{n} \right) = \frac{w_i^2}{\ell_n} \left(1 + \frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]} \right),
\end{aligned} \tag{3.3.11}$$

where we used again that $\mathbb{E}[W_n] = \frac{\ell_n}{n}$ and $\mathbb{E}\left[W_n^2\right] = \frac{1}{n} \sum_{i \in [n]} w_i^2$.

Thus, we have proved that we can couple $(\hat{D}_i, \hat{Y}_i)$ to $(\hat{D}_i, \hat{Z}_i)$ where $\hat{Z}_i$ is a Poisson random variable with parameter w_i in such a way that

$$\mathbb{P}\left[\hat{Y}_i \neq \hat{Z}_i\right] \leq \frac{w_i^2}{\ell_n} \left(1 + \frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]} \right). \tag{3.3.12}$$

We can now conclude Theorem 3.11 part (a) by noting that

$$\begin{aligned}
\mathbb{P}\left[\hat{D}_i \neq \hat{Z}_i\right] &\leq \mathbb{P}\left[\hat{D}_i \neq \hat{Y}_i\right] + \mathbb{P}\left[\hat{Y}_i \neq \hat{Z}_i\right] \\
&\leq \frac{w_i^2}{\ell_n} \frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]} + \frac{w_i^2}{\ell_n} \left(1 + \frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]} \right) \\
&= \frac{w_i^2}{\ell_n} \left(1 + 2 \frac{\mathbb{E}\left[W_n^2\right]}{\mathbb{E}[W_n]} \right).
\end{aligned} \tag{3.3.13}$$

(b) To prove Theorem 3.11 part (b), we use again a coupling argument. In particular, we prove that we can couple the degrees $(D_i)_{i \in [m]}$ to an independent random vector $(\hat{D}_i)_{i \in [m]}$ such that

$$\mathbb{P}\left[(D_i)_{i \in [m]} \neq (\hat{D}_i)_{i \in [m]} \right] = o(1). \tag{3.3.14}$$

The first step is to recall that X_{ij} denotes the indicator that the edge ij is occupied and that $X_{ij} = X_{ji}$. The random variables $(X_{ij})_{1 \leq i < j \leq n}$ are independent Bernoulli random variables with parameters $(p_{ij})_{1 \leq i < j \leq n}$.

Then, we let $(\hat{X}_{ij})_{1 \leq i < j \leq n}$ be an independent copy of $(X_{ij})_{1 \leq i < j \leq n}$ with $\hat{X}_{ij} = \hat{X}_{ji}$ and define, for $i \in [n]$,

$$\hat{D}_i := \sum_{j < i} \hat{X}_{ij} + \sum_{j = i+1}^{n} X_{ij}. \tag{3.3.15}$$

Due to this definition, $\hat{D}_i$ has the same distribution as D_i for every $i \in [n]$. Moreover, for fixed $i < j$ we observe that, while D_i and D_j are dependent because both contain $X_{ij} = X_{ji}$, $\hat{D}_i$ and $\hat{D}_j$ are independent since $\hat{D}_i$ contains X_{ij}, while $\hat{D}_j$ contains $\hat{X}_{ji} = \hat{X}_{ij}$, which is an independent copy of X_{ij}. Therefore, $(\hat{D}_i)_{i\in[m]}$ are sums of mutually independent Bernoulli random variables, which implies that $(\hat{D}_i)_{i\in[m]}$ are independent random variables.

Hence we can deduce that $(D_i)_{i\in[m]} \neq (\hat{D}_i)_{i\in[m]}$ if and only if there exists at least one edge ij with $i, j \in [m]$ such that $X_{ij} \neq \hat{X}_{ij}$. However, $X_{ij} \neq \hat{X}_{ij}$ when either $X_{ij} = 0$ while $\hat{X}_{ij} = 1$ or $X_{ij} = 1$ while $\hat{X}_{ij} = 0$.

For the final step, we make use of Boole's inequality 2.3 twice to deduce that

$$\begin{aligned}
\mathbb{P}\left[(D_i)_{i\in[m]} \neq (\hat{D}_i)_{i\in[m]}\right] &= \mathbb{P}\left[\bigcup_{i=1}^{m}(D_i \neq \hat{D}_i)\right] \leq \sum_{i=1}^{m}\mathbb{P}\left[D_i \neq \hat{D}_i\right] \\
&= \sum_{i=1}^{m}\mathbb{P}\left[\bigcup_{j=1}^{m}(X_{ij} \neq \hat{X}_{ij})\right] \leq \sum_{i=1}^{m}\sum_{j=1}^{m}\mathbb{P}\left[X_{ij} \neq \hat{X}_{ij}\right] \\
&= \sum_{i=1}^{m}\sum_{j=1}^{m}\left(\mathbb{P}\left[X_{ij} = 0 \wedge \hat{X}_{ij} = 1\right] + \mathbb{P}\left[X_{ij} = 1 \wedge \hat{X}_{ij} = 0\right]\right) \\
&\leq 2\sum_{i=1}^{m}\sum_{j=1}^{m}\mathbb{P}\left[X_{ij} = 1\right] = 2\sum_{i=1}^{m}\sum_{j=1}^{m}p_{ij}.
\end{aligned} \tag{3.3.16}$$

Since by assumption $\lim_{n\to\infty}\max_{1\leq i<j\leq m} p_{ij} = 0$, it follows for every fixed $m \geq 2$ that

$$\mathbb{P}\left[(D_i)_{i\in[m]} \neq (\hat{D}_i)_{i\in[m]}\right] \leq 2\sum_{i=1}^{m}\sum_{j=1}^{m}p_{ij} = o(1). \tag{3.3.17}$$

Thus, we conclude that the degrees $D_1, \ldots, D_m$ of vertices $1, \ldots, m$ are asymptotically independent, because $(D_i)_{i\in[m]}$ and the independent random vector $(\hat{D}_i)_{i\in[m]}$ are asymptotically equal. □

In order to prove similar results for uniformly chosen vertices, we need to introduce the definition of a mixed Poisson distribution first.

Definition 3.12 (Mixed Poisson distribution)**.** *([Hof18], Definition 6.8)*
A random variable X has a mixed Poisson distribution with mixing distribution F when, for every $k \in \mathbb{N}_0$,

$$\mathbb{P}[X = k] = \mathbb{E}\left[e^{-W}\frac{W^k}{k!}\right], \tag{3.3.18}$$

where W is a random variable with distribution function F.

The following lemma gives us some useful insights into the mean and variance of variables with a mixed Poisson distribution.

Lemma 3.13 (Mean and variance of mixed Poisson distribution)**.** *([Hof18], Exercise 6.11)*
Let X have a mixed Poisson distribution with mixing distribution F, where W is a random variable with distribution function F. Then,

$$\mathbb{E}[X] = \mathbb{E}[W] \tag{3.3.19}$$

and

$$Var(X) = Var(W) + \mathbb{E}[W]. \tag{3.3.20}$$

Proof. By the law of total expectation, we know that $\mathbb{E}[\mathbb{E}[X \mid W]] = \mathbb{E}[X]$. Taking a closer look at the left-hand side, we observe that

$$\begin{aligned}
\mathbb{E}[X \mid W = w] &= \sum_{k=0}^{\infty} k\mathbb{P}[X = k \mid W = w] \\
&= \sum_{k=0}^{\infty} k e^{-w} \frac{w^k}{k!} \\
&= 0 \cdot e^{-w} \frac{w^0}{0!} + \sum_{k=1}^{\infty} e^{-w} \frac{w^k}{(k-1)!} \\
&= we^{-w} \sum_{k=1}^{\infty} \frac{w^{k-1}}{(k-1)!} \\
&= we^{-w}e^{w} = w.
\end{aligned} \tag{3.3.21}$$

$\mathbb{E}[X \mid W = w] = w$ implies that $\mathbb{E}[X \mid W] = W$ and therefore $\mathbb{E}[\mathbb{E}[X \mid W]] = \mathbb{E}[W]$. Therefore, we can deduce that $\mathbb{E}[X] = \mathbb{E}[W]$.

To find the variance of X, we first consider $\mathbb{E}[\mathbb{E}[X(X-1) \mid W]] = \mathbb{E}[X(X-1)]$. Taking a closer look at the left-hand side, we get

$$\begin{aligned}
\mathbb{E}[X(X-1) \mid W = w] &= \sum_{k=0}^{\infty} k(k-1)\mathbb{P}[X(X-1) = k \mid W = w] \\
&= \sum_{k=0}^{\infty} k(k-1) e^{-w} \frac{w^k}{k!} \\
&= 0 \cdot (-1) \cdot e^{-w} \frac{w^0}{0!} + 1 \cdot 0 \cdot e^{-w} \frac{w^1}{1!} + \sum_{k=2}^{\infty} e^{-w} \frac{w^k}{(k-2)!} \\
&= w^2 e^{-w} \sum_{k=2}^{\infty} \frac{w^{k-2}}{(k-2)!} \\
&= w^2 e^{-w} e^{w} = w^2.
\end{aligned} \tag{3.3.22}$$

$\mathbb{E}[X(X-1) \mid W = w] = w^2$ yields $\mathbb{E}[X(X-1) \mid W] = W^2$. Taking expectations on both sides of the equation gives us $\mathbb{E}[X(X-1)] = \mathbb{E}[W^2]$, which implies that $\mathbb{E}[X^2] = \mathbb{E}[W^2] + \mathbb{E}[X] = \mathbb{E}[W^2] + \mathbb{E}[W]$.

Therefore, we can deduce that

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \mathbb{E}[W^2] + \mathbb{E}[W] - \mathbb{E}[W]^2 \\ &= \mathrm{Var}(W) + \mathbb{E}[W].\end{aligned} \tag{3.3.23}$$

□

According to Theorem 3.11, the degree of vertex i is close to Poisson with parameter w_i if we impose some additional regularity conditions on the weights $(w_i)_{i\in[n]}$. Thus, when we choose a vertex uniformly at random and denote the outcome by U, we expect that the degree of the chosen vertex is close to a Poisson distribution with the random parameter $w_U = W_n$. This leads to the following corollary.

Corollary 3.14 (Degree of uniformly chosen vertex in $\mathrm{GRG}_n(\boldsymbol{w})$)**.** *([Hof18], Corollary 6.9) Assume that Conditions 3.3 (a)-(b) hold. Then,*

(a) the degree of a uniformly chosen vertex converges in distribution to a mixed Poisson random variable with mixing distribution F.

(b) the degrees of m uniformly chosen vertices in [n] are asymptotically independent.

Proof.
(a) By assumption the regularity Conditions 3.3 (a)-(b) hold, which imply that $max_{i\in[n]} w_i = o(n)$ by Proposition 3.4. Thus, we start by fixing some $a_n \to \infty$ sufficiently slowly, but in such a way that the weight of a uniformly chosen vertex $W_n = w_U$ satisfies $w_U \leq a_n$ with high probability.

Let us assume that this is the case. Then, we can split

$$D_U = D_U^{(1)} + D_U^{(2)}, \tag{3.3.24}$$

where $D_U^{(1)}$ is equal to the number of edges of U to vertices of weight at most a_n, while $D_U^{(2)}$ counts the number of edges of U to vertices of weight larger than a_n.

Next, we observe that

$$\begin{aligned}\mathbb{E}\left[D_U^{(2)}\right] &= \mathbb{E}\left[\sum_{j\in[n]} X_{Uj}\cdot \mathbb{1}_{\{w_U\leq a_n<w_j\}}\right] \\ &= \sum_{j\in[n]} \mathbb{E}\left[X_{Uj}\cdot \mathbb{1}_{\{w_U\leq a_n<w_j\}}\right] \\ &= \sum_{j\in[n]} \mathbb{P}\left[X_{Uj}\cdot \mathbb{1}_{\{w_U\leq a_n<w_j\}} = 1\right] \\ &= \sum_{j\in[n]}\sum_{i\in[n]} \mathbb{P}\left[X_{Uj}\cdot \mathbb{1}_{\{w_U\leq a_n<w_j\}} = 1 | U=i\right]\cdot \mathbb{P}[U=i]\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n} \sum_{j\in[n]} \sum_{i\in[n]} \mathbb{P}\left[X_{ij} \cdot \mathbb{1}_{\{w_i \le a_n < w_j\}} = 1\right] \\
&\le \frac{1}{n} \sum_{j\in[n]} \sum_{i\in[n]} \mathbb{P}\left[X_{ij} = 1\right] \cdot \mathbb{1}_{\{w_i \le a_n < w_j\}} \\
&= \frac{1}{n} \sum_{j\in[n]} \sum_{i\in[n]} \frac{w_i w_j}{\ell_n + w_i w_j} \cdot \mathbb{1}_{\{w_i \le a_n < w_j\}} \\
&\le \frac{1}{n} \sum_{j\in[n]} w_j \cdot \mathbb{1}_{\{a_n < w_j\}} \sum_{i\in[n]} \frac{w_i}{\ell_n} \cdot \mathbb{1}_{\{w_i \le a_n\}} \\
&\le \frac{1}{n} \sum_{j\in[n]} w_j \cdot \mathbb{1}_{\{a_n < w_j\}} = \mathbb{E}\left[W_n \cdot \mathbb{1}_{\{a_n < W_n\}}\right] = o(1),
\end{aligned}
\tag{3.3.25}
$$

since we chose a_n in such a way that $\mathbb{P}[W_n \le a_n] \to 1$. Therefore, $\mathbb{P}\left[D_U^{(2)} \ge 1\right] = o(1)$ and hence, with high probability $D_U = D_U^{(1)}$.

Using similar arguments as in Theorem 3.11 (a), we get that $D_U^{(1)} \xrightarrow{d} D$, where $D \sim Poi(W)$, which finishes the proof of part (a).

(b) This proof is a minor adaptation of the proof of Theorem 3.11 (b). We start by letting $(U_i)_{i\in[m]}$ be independent vertices chosen uniformly at random from $[n]$.

By Boole's inequality 2.3,

$$
\begin{aligned}
\mathbb{P}\left[\bigcup_{i=1}^{m}\bigcup_{j=1}^{m} U_j = U_i\right] &\le \sum_{i=1}^{m}\sum_{j=1}^{m} \mathbb{P}[U_j = U_i] \\
&\le \sum_{i=1}^{m}\sum_{j=1}^{m} \frac{1}{n} \le \frac{m^2}{n},
\end{aligned}
\tag{3.3.26}
$$

and therefore, all vertices $(U_i)_{i\in[m]}$ are distinct with high probability.

When $(U_i)_{i\in[m]}$ are distinct, the dependence between the degrees of the vertices $(U_i)_{i\in[m]}$ arises only through the occupied edges between the vertices $(U_i)_{i\in[m]}$. Thus, in a next step, we want to bound the expected number of edges between $(U_i)_{i\in[m]}$, which we denote by N.

Note that,

$$
\begin{aligned}
\mathbb{E}[N|(U_i)_{i\in[m]}] &= \mathbb{E}\left[\sum_{i,j=1}^{m} X_{U_iU_j}|(U_i)_{i\in[m]}\right] \\
&= \sum_{i,j=1}^{m} \mathbb{P}\left[X_{U_iU_j} = 1|(U_i)_{i\in[m]}\right] \\
&\le \sum_{i,j=1}^{m} \frac{w_{U_i} w_{U_j}}{\ell_n + w_{U_i} w_{U_j}} \\
&\le \sum_{i,j=1}^{m} \frac{w_{U_i} w_{U_j}}{\ell_n} = \frac{1}{\ell_n}\left(\sum_{i=1}^{m} w_{U_i}\right)^2.
\end{aligned}
\tag{3.3.27}
$$

Since the random variables $(w_{U_i})_{i\in[m]}$ are i.i.d., the expected number N of occupied edges between m uniformly chosen vertices can be bounded from above by:

$$\begin{aligned}\mathbb{E}[N] &= \mathbb{E}[\mathbb{E}[N|(U_i)_{i\in[m]}]] \\ &\leq \frac{1}{\ell_n}\mathbb{E}\left[\left(\sum_{i=1}^{m} w_{U_i}\right)^2\right] \\ &= \frac{1}{\ell_n}\sum_{i=1}^{m}\mathbb{E}[w_{U_i}^2] + \frac{1}{\ell_n}\sum_{i=1}^{m}\sum_{j\neq i}\mathbb{E}[w_{U_i}w_{U_j}] \\ &= \frac{m}{\ell_n}\mathbb{E}[w_{U_1}^2] + \frac{m(m-1)}{\ell_n}\mathbb{E}[w_{U_1}]^2.\end{aligned} \tag{3.3.28}$$

Using Proposition 3.4, we get that

$$\mathbb{E}\left[w_{U_1}^2\right] \leq (\max_{i\in[n]} w_i)\mathbb{E}\left[w_{U_1}\right] = o(n)\mathbb{E}\left[w_{U_1}\right] = o(n). \tag{3.3.29}$$

Observing that

$$\frac{m}{\ell_n}\mathbb{E}\left[w_{U_1}^2\right] + \frac{m(m-1)}{\ell_n}\mathbb{E}\left[w_{U_1}\right]^2 = \frac{m}{\mathbb{E}[W_n]}\frac{\mathbb{E}\left[w_{U_1}^2\right]}{n} + \frac{m(m-1)}{\mathbb{E}[W_n]}\frac{\mathbb{E}\left[w_{U_1}\right]^2}{n} = o(1), \tag{3.3.30}$$

we can deduce that the expected number of edges between the vertices $(U_i)_{i\in[m]}$ is $o(1)$. Thus, with high probability there are no edges between m uniformly chosen vertices and therefore, the degrees of m uniformly chosen vertices in $[n]$ are asymptotically independent. □

3.4 Degree Sequence of the Generalised Random Graph

While we investigated the degree of a single vertex in the generalised random graph $\mathrm{GRG}_n(\boldsymbol{w})$ in the last section, we now extend the result to the degree distribution of $\mathrm{GRG}_n(\boldsymbol{w})$. In particular, we show that the degree distribution of the generalised random graph is close to a mixed Poisson distribution.

Before we state the main result of this section, we need to introduce some notation. For $k \geq 0$, we let

$$P_k^{(n)} = \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{D_i=k\}} \tag{3.4.1}$$

denote the empirical degree distribution of $\mathrm{GRG}_n(\boldsymbol{w})$.

Moreover, we denote the probability mass function of a mixed Poisson distribution for $k \geq 0$ by

$$p_k = \mathbb{E}\left[e^{-W}\frac{W^k}{k!}\right], \tag{3.4.2}$$

where W is a random variable having distribution function F from Conditions 3.3.

As one might expect due to the last section, we can prove that the degree distribution $\left(P_k^{(n)}\right)_{k\geq 0}$ is close to the mixed Poisson distribution with probability mass function $(p_k)_{k\geq 0}$.

Theorem 3.15 (Degree distribution of $\mathrm{GRG}_n(\boldsymbol{w})$)**.** *([Hof18], Theorem 6.10)*
Assume that Conditions 3.3 (a)-(b) hold. Then, for every $\varepsilon > 0$,

$$\mathbb{P}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \to 0, \tag{3.4.3}$$

where $(p_k)_{k\geq 0}$ is the probability mass function of the mixed Poisson distribution with mixing distribution F as in (3.4.2).

When we choose the mixing distribution F to have the Pareto distribution which has a power-law tail, then also the mixed Poisson distribution follows a power-law tail as stated in [VHHK19], p.15-16. Therefore, if we choose the weights in an appropriate way, the degree distribution exhibits a power-law tail.

Before proving Theorem 3.15, we state and prove two lemmas, which will facilitate the proof of the main theorem.

Lemma 3.16 (Maximal distance of probability mass functions)**.** *([Hof18], p.193)*
Let $P_k^{(n)} = \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{D_i=k\}}$ be the empirical degree distribution of $GRG_n(\boldsymbol{w})$ and $p_k = \mathbb{E}\left[e^{-W}\frac{W^k}{k!}\right]$ the probability mass function of a mixed Poisson distribution, where W is a random variable with distribution function F as in Conditions 3.3.
For every $\varepsilon > 0$,

$$\mathbb{P}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \to 0 \tag{3.4.4}$$

if and only if

$$\mathbb{P}\left[\max_{k\geq 0}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \to 0. \tag{3.4.5}$$

Proof.
($\Rightarrow$) Let us first assume that $\mathbb{P}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \to 0$. Then we note that for any n,

$$\max_{k\geq 0}\left|P_k^{(n)} - p_k\right| \leq \sum_{k=0}^{\infty}\left|P_k^{(n)} - p_k\right|, \tag{3.4.6}$$

and hence, for every $\varepsilon > 0$,

$$\mathbb{P}\left[\max_{k\geq 0}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \leq \mathbb{P}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \to 0. \tag{3.4.7}$$

($\Leftarrow$) Let us now assume that $\mathbb{P}\left[\max_{k\geq 0}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \to 0$, which implies that for any $k \geq 0$ $P_k^{(n)} \xrightarrow{\mathbb{P}} p_k$.
By Markov's inequality 2.1 we get that

$$\mathbb{P}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \leq \frac{\mathbb{E}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)} - p_k\right|\right]}{\varepsilon}. \tag{3.4.8}$$

Then, we observe that

$$\sum_{k=0}^{\infty}\left|P_k^{(n)}-p_k\right| = \sum_{k=0}^{\infty}\left(P_k^{(n)}-p_k\right)\mathbb{1}_{\{P_k^{(n)}\geq p_k\}} + \sum_{k=0}^{\infty}\left(p_k-P_k^{(n)}\right)\mathbb{1}_{\{P_k^{(n)}<p_k\}}, \tag{3.4.9}$$

while

$$0 = 1-1 = \sum_{k=0}^{\infty}\left(p_k-P_k^{(n)}\right) = \sum_{k=0}^{\infty}\left(p_k-P_k^{(n)}\right)\mathbb{1}_{\{P_k^{(n)}\geq p_k\}} + \sum_{k=0}^{\infty}\left(p_k-P_k^{(n)}\right)\mathbb{1}_{\{P_k^{(n)}<p_k\}}. \tag{3.4.10}$$

Adding the two equations above, we obtain that

$$\sum_{k=0}^{\infty}\left|P_k^{(n)}-p_k\right| = 2\sum_{k=0}^{\infty}\left(p_k-P_k^{(n)}\right)\mathbb{1}_{\{P_k^{(n)}<p_k\}}. \tag{3.4.11}$$

Moreover,

$$\sum_{k=0}^{\infty}\left(p_k-P_k^{(n)}\right)\mathbb{1}_{\{P_k^{(n)}<p_k\}} = \sum_{k=0}^{\infty}\left(p_k-\left(P_k^{(n)}\wedge p_k\right)\right), \tag{3.4.12}$$

where $P_k^{(n)}\wedge p_k = \min\left(P_k^{(n)},p_k\right)$.

Using the fact that $\sum_{k=0}^{\infty}\left|P_k^{(n)}-p_k\right| = 2\cdot\sum_{k=0}^{\infty}\left(p_k-\left(P_k^{(n)}\wedge p_k\right)\right) \leq 2\cdot\sum_{k=0}^{\infty}p_k = 2$ and Lebesgue's dominated convergence theorem 2.8, we deduce that

$$\begin{aligned}\mathbb{P}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)}-p_k\right|\geq\varepsilon\right] &\leq \frac{\mathbb{E}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)}-p_k\right|\right]}{\varepsilon}\\ &= \frac{\sum_{k=0}^{\infty}\mathbb{E}\left[\left|P_k^{(n)}-p_k\right|\right]}{\varepsilon}.\end{aligned} \tag{3.4.13}$$

Since $\left|P_k^{(n)}-p_k\right|\in[0,1]$ it holds that $\left|P_k^{(n)}-p_k\right| = \left|P_k^{(n)}-p_k\right|\wedge 1$. Moreover, it is known that $P_k^{(n)}\xrightarrow{\mathbb{P}}p_k$ if and only if $\mathbb{E}\left[\left|P_k^{(n)}-p_k\right|\wedge 1\right]\to 0$ as $n\to\infty$ (see [JP04], p.144).

Therefore,

$$\begin{aligned}\mathbb{P}\left[\sum_{k=0}^{\infty}\left|P_k^{(n)}-p_k\right|\geq\varepsilon\right] &\leq \frac{\sum_{k=0}^{\infty}\mathbb{E}\left[\left|P_k^{(n)}-p_k\right|\right]}{\varepsilon}\\ &= \frac{\sum_{k=0}^{\infty}\mathbb{E}\left[\left|P_k^{(n)}-p_k\right|\wedge 1\right]}{\varepsilon}\to 0,\end{aligned} \tag{3.4.14}$$

which proves the desired result. □

Lemma 3.17 (Uniform convergence of mean degree sequence)**.** *([Hof18], Ex. 6.17)*
Let $P_k^{(n)} = \frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{D_i=k\}}$ be the empirical degree distribution of $GRG_n(\boldsymbol{w})$ and $p_k = \mathbb{E}\left[e^{-W}\frac{W^k}{k!}\right]$ the probability mass function of a mixed Poisson distribution, where W is a random variable with distribution function F as in the regularity Conditions 3.3. Then, for every $\varepsilon>0$ and for n sufficiently large,

$$\max_{k\geq 0}\left|\mathbb{E}\left[P_k^{(n)}\right]-p_k\right| \leq \frac{\varepsilon}{2}. \tag{3.4.15}$$

Proof. First of all, we know by Corollary 3.14 part (a) that

$$\lim_{n\to\infty} \mathbb{E}\left[P_k^{(n)}\right] = \lim_{n\to\infty} \mathbb{P}\left[D_U = k\right] = p_k. \tag{3.4.16}$$

Now, let us fix some arbitrary $\varepsilon > 0$ and take $M \geq 1$ large enough such that

$$\sum_{k\geq M} p_k \leq \frac{\varepsilon}{4}. \tag{3.4.17}$$

Since M is finite, we can take n to be so large that for every $k < M$

$$\left|\mathbb{E}\left[P_k^{(n)}\right] - p_k\right| \leq \frac{\varepsilon}{4M}. \tag{3.4.18}$$

As we chose M to be at least 1, we can deduce that for every $k < M$

$$\left|\mathbb{E}\left[P_k^{(n)}\right] - p_k\right| \leq \frac{\varepsilon}{2}. \tag{3.4.19}$$

Using that $\sum_{k\geq 0} \mathbb{E}\left[P_k^{(n)}\right] = \sum_{k\geq 0} \mathbb{P}[D_U = k] = \sum_{k\geq 0} p_k = 1$, we get that

$$\begin{aligned}
\sum_{k\geq M} \mathbb{E}\left[P_k^{(n)}\right] &= 1 - \sum_{k<M} \mathbb{E}\left[P_k^{(n)}\right] \\
&\leq 1 - \sum_{k<M} \left(p_k - \frac{\varepsilon}{4M}\right) \\
&= 1 - \sum_{k<M} p_k + \sum_{k<M} \frac{\varepsilon}{4M} \\
&= \sum_{k\geq M} p_k + \frac{\varepsilon}{4} \leq \frac{\varepsilon}{2}.
\end{aligned} \tag{3.4.20}$$

Therefore, it must hold for every $k \geq M$ that $\mathbb{E}\left[P_k^{(n)}\right] \leq \frac{\varepsilon}{2}$ and $p_k \leq \frac{\varepsilon}{2}$, which implies for every $k \geq M$ that

$$\left|\mathbb{E}\left[P_k^{(n)}\right] - p_k\right| \leq \frac{\varepsilon}{2}. \tag{3.4.21}$$

Hence, we can conclude that for any $\varepsilon > 0$ and n sufficiently large

$$\max_{k\geq 0}\left|\mathbb{E}\left[P_k^{(n)}\right] - p_k\right| \leq \frac{\varepsilon}{2}. \tag{3.4.22}$$

□

Now we have discovered all necessary tools to prove Theorem 3.15, which declares that the degree distribution of the generalised random graph is close to a mixed Poisson distribution.

Proof of Theorem 3.15.

By Lemma 3.16 it suffices to show that, for every $\varepsilon > 0$,

$$\mathbb{P}\left[\max_{k\geq 0}\left|P_k^{(n)} - p_k\right| \geq \varepsilon\right] \to 0. \tag{3.4.23}$$

We use that

$$\mathbb{P}\left[\max_{k\geq 0}\left|P_k^{(n)}-p_k\right|\geq\varepsilon\right]\leq\sum_{k\geq 0}\mathbb{P}\left[\left|P_k^{(n)}-p_k\right|\geq\varepsilon\right]. \tag{3.4.24}$$

By Lemma 3.17 we know that, for every $\varepsilon>0$ and for n sufficiently large,

$$\max_{k\geq 0}\left|\mathbb{E}\left[P_k^{(n)}\right]-p_k\right|\leq\frac{\varepsilon}{2}. \tag{3.4.25}$$

Therefore, for any $\varepsilon>0$ and n sufficiently large,

$$\begin{aligned}
\mathbb{P}\left[\max_{k\geq 0}\left|P_k^{(n)}-p_k\right|\geq\varepsilon\right]&\leq\sum_{k\geq 0}\mathbb{P}\left[\left|P_k^{(n)}-p_k\right|\geq\varepsilon\right]\\
&=\sum_{k\geq 0}\mathbb{P}\left[\left|P_k^{(n)}-\mathbb{E}\left[P_k^{(n)}\right]+\mathbb{E}\left[P_k^{(n)}\right]-p_k\right|\geq\varepsilon\right]\\
&\leq\sum_{k\geq 0}\mathbb{P}\left[\left|P_k^{(n)}-\mathbb{E}\left[P_k^{(n)}\right]\right|+\left|\mathbb{E}\left[P_k^{(n)}\right]-p_k\right|\geq\varepsilon\right]\\
&\leq\sum_{k\geq 0}\mathbb{P}\left[\left|P_k^{(n)}-\mathbb{E}\left[P_k^{(n)}\right]\right|+\frac{\varepsilon}{2}\geq\varepsilon\right]\\
&=\sum_{k\geq 0}\mathbb{P}\left[\left|P_k^{(n)}-\mathbb{E}\left[P_k^{(n)}\right]\right|\geq\frac{\varepsilon}{2}\right].
\end{aligned} \tag{3.4.26}$$

By Chebychev's inequality 2.2,

$$\mathbb{P}\left[\left|P_k^{(n)}-\mathbb{E}\left[P_k^{(n)}\right]\right|\geq\frac{\varepsilon}{2}\right]\leq\frac{4}{\varepsilon^2}\mathrm{Var}\left(P_k^{(n)}\right). \tag{3.4.27}$$

$P_k^{(n)}=\frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{D_i=k\}}$ implies that $\left(P_k^{(n)}\right)^2=\frac{1}{n^2}\left(\sum_{i\in[n]}\mathbb{1}_{\{D_i=k\}}\right)^2=\frac{1}{n^2}\sum_{i,j\in[n]}\mathbb{1}_{\{D_i=D_j=k\}}$ and thus,

$$\begin{aligned}
\mathbb{E}\left[\left(P_k^{(n)}\right)^2\right]&=\frac{1}{n^2}\sum_{i,j\in[n]}\mathbb{P}[D_i=D_j=k]\\
&=\frac{1}{n^2}\sum_{i\in[n]}\mathbb{P}[D_i=k]+\frac{1}{n^2}\sum_{i,j\in[n]:i\neq j}\mathbb{P}[D_i=D_j=k].
\end{aligned} \tag{3.4.28}$$

Using that

$$\begin{aligned}
\mathbb{E}\left[P_k^{(n)}\right]^2&=\mathbb{P}[D_U=k]^2\\
&=\left(\frac{1}{n}\sum_{i\in[n]}\mathbb{P}[D_i=k]\right)^2\\
&=\frac{1}{n^2}\left(\sum_{i\in[n]}\mathbb{P}[D_i=k]^2+\sum_{i,j\in[n]:i\neq j}\mathbb{P}[D_i=k]\mathbb{P}[D_j=k]\right),
\end{aligned} \tag{3.4.29}$$

we can deduce that

$$\begin{aligned}\mathrm{Var}\left(P_k^{(n)}\right) &= \mathbb{E}\left[\left(P_k^{(n)}\right)^2\right] - \mathbb{E}\left[P_k^{(n)}\right]^2 \\ &= \frac{1}{n^2}\sum_{i\in[n]}\left[\mathbb{P}[D_i = k] - \mathbb{P}[D_i = k]^2\right] \\ &\quad + \frac{1}{n^2}\sum_{i,j\in[n]:i\neq j}\left[\mathbb{P}[D_i = D_j = k] - \mathbb{P}[D_i = k]\mathbb{P}[D_j = k]\right].\end{aligned} \tag{3.4.30}$$

Recall that $(X_{ij})_{1\leq 1<j\leq n}$ are independent Bernoulli random variables with success probabilities $\mathbb{P}[X_{ij} = 1] = p_{ij}$, where X_{ij} is the indicator that the edge ij is occupied. Then, we set

$$Y_{i;j} = \sum_{l\in[n]:l\neq i,j} X_{il}, \qquad Y_{j;i} = \sum_{l\in[n]:l\neq i,j} X_{jl}, \tag{3.4.31}$$

such that the law of (D_i, D_j) is equal to the one of $(Y_{i;j}+X_{ij}, Y_{j;i}+X_{ij})$. Further, let $\hat{X}_{ij}$ have the same distribution as X_{ij}, but be independent of it. Then $(Y_{i;j}+X_{ij}, Y_{j;i}+\hat{X}_{ij})$ are two independent random variables with the same marginal distribution as (D_i, D_j).

Hence,

$$\mathbb{P}[D_i = D_j = k] = \mathbb{P}\left[(Y_{i;j} + X_{ij}, Y_{j;i} + X_{ij}) = (k,k)\right], \tag{3.4.32}$$

$$\mathbb{P}[D_i = k]\mathbb{P}[D_j = k] = \mathbb{P}\left[(Y_{i;j} + X_{ij}, Y_{j;i} + \hat{X}_{ij}) = (k,k)\right]. \tag{3.4.33}$$

In a next step we use that $\mathbb{P}[A] - \mathbb{P}[B] \leq \mathbb{P}[A \cap B^c]$ for any events A, B. Therefore,

$$\begin{aligned}&\mathbb{P}[D_i = D_j = k] - \mathbb{P}[D_i = k]\mathbb{P}[D_j = k] \\ &\leq \mathbb{P}\left[(Y_{i;j} + X_{ij}, Y_{j;i} + X_{ij}) = (k,k), (Y_{i;j} + X_{ij}, Y_{j;i} + \hat{X}_{ij}) \neq (k,k)\right].\end{aligned} \tag{3.4.34}$$

When $(Y_{i;j} + X_{ij}, Y_{j;i} + X_{ij}) = (k,k)$, but $(Y_{i;j} + X_{ij}, Y_{j;i} + \hat{X}_{ij}) \neq (k,k)$, we must have that $X_{ij} \neq \hat{X}_{ij}$. Thus, either $X_{ij} = 1$ while $\hat{X}_{ij} = 0$, or vice versa. In the case that $X_{ij} = 1$, we have $\hat{X}_{ij} = 0$ and $Y_{j;i} + X_{ij} = k$. Otherwise, if $X_{ij} = 0$, it must hold that $\hat{X}_{ij} = 1$ and $Y_{i;j} + X_{ij} = k$. Note that the events $\hat{X}_{ij} = 0$ and $Y_{j;i} + X_{ij} = k$ are independent, as well as the two events $\hat{X}_{ij} = 1$ and $Y_{i;j} + X_{ij} = k$. Let us set $A \coloneqq \{(Y_{i;j}+X_{ij}, Y_{j;i}+X_{ij}) = (k,k), (Y_{i;j}+X_{ij}, Y_{j;i}+\hat{X}_{ij}) \neq (k,k)\}$, then we can deduce by above observations that

$$\begin{aligned}\mathbb{P}[A] &= \mathbb{P}[A|X_{ij} = 1]\mathbb{P}[X_{ij} = 1] + \mathbb{P}\left[A|\hat{X}_{ij} = 1\right]\mathbb{P}\left[\hat{X}_{ij} = 1\right] \\ &= p_{ij}\left(\mathbb{P}[A|X_{ij} = 1] + \mathbb{P}[A|\hat{X}_{ij} = 1]\right) \\ &\leq p_{ij}\left(\mathbb{P}[D_i = k] + \mathbb{P}[D_j = k]\right).\end{aligned} \tag{3.4.35}$$

Therefore, by Tonelli's theorem,

$$\begin{aligned}\sum_{k\geq 0}\mathrm{Var}\left(P_k^{(n)}\right) &\leq \frac{1}{n^2}\sum_{k\geq 0}\Bigg[\sum_{i\in[n]}\left[\mathbb{P}[D_i=k]-\mathbb{P}[D_i=k]^2\right]\\ &\quad+\sum_{i,j\in[n]:i\neq j}\left[\mathbb{P}[D_i=D_j=k]-\mathbb{P}[D_i=k]\mathbb{P}[D_j=k]\right]\Bigg]\\ &\leq \frac{1}{n^2}\sum_{i\in[n]}\sum_{k\geq 0}\mathbb{P}[D_i=k]+\frac{1}{n^2}\sum_{i,j\in[n]:i\neq j}p_{ij}\sum_{k\geq 0}\mathbb{P}[D_i=k]\\ &\quad+\frac{1}{n^2}\sum_{i,j\in[n]:i\neq j}p_{ij}\sum_{k\geq 0}\mathbb{P}[D_j=k]\\ &=\frac{1}{n}+\frac{2}{n^2}\sum_{i,j\in[n]:i\neq j}p_{ij}\end{aligned} \tag{3.4.36}$$

Clearly,

$$\begin{aligned}\sum_{i,j\in[n]:i\neq j}p_{ij}&=\sum_{i,j\in[n]:i\neq j}\frac{w_iw_j}{\ell_n+w_iw_j}\\ &\leq\sum_{i,j\in[n]:i\neq j}\frac{w_iw_j}{\ell_n}\\ &\leq\frac{1}{\ell_n}\sum_{i\in n}w_i\sum_{j\in[n]}w_j\\ &=\ell_n=n\mathbb{E}[W_n]=O(n),\end{aligned} \tag{3.4.37}$$

by the regularity Condition 3.3 (b).

Hence, we conclude that for any $\varepsilon>0$ and $n\to\infty$,

$$\begin{aligned}\mathbb{P}\left[\max_{k\geq 0}\left|P_k^{(n)}-p_k\right|\geq\varepsilon\right]&\leq\frac{4}{\varepsilon^2}\sum_{k\geq 0}\mathrm{Var}\left(P_k^{(n)}\right)\\ &\leq\frac{4}{\varepsilon^2}\left(\frac{1}{n}+\frac{2}{n^2}\sum_{i,j\in[n]:i\neq j}p_{ij}\right)\to 0.\end{aligned} \tag{3.4.38}$$

□

Remark 3.18. *Theorem 3.15 implies that* $P_0^{(n)}\xrightarrow{\mathbb{P}}p_0$ *and since* $p_0=\mathbb{E}\left[e^{-W}\right]>0$, *we can deduce that* $GRG_n(\boldsymbol{w})$ *has a positive proportion of vertices of degree* 0 *and thus is not connected when Conditions 3.3 (a)-(b) hold.*

3.5 Generalised Random Graph Conditioned on Its Degrees

In this section, we will have a closer look at the distribution of $\mathrm{GRG}_n(\boldsymbol{w})$. The main goal of this section is to prove that the generalised random graph conditioned on its degree sequence is a uniform simple random graph with that degree sequence. Expressed slightly differently, using the generalised random graph model, the likelihood to generate one specific random graph is equal for

any random graph with the same degree sequence. Before we state and prove the main result, we need to introduce some notation which will be useful in the rest of this section.

Let $X = (X_{ij})_{1\le i<j\le n}$, where X_{ij} are independent Bernoulli random variables with

$$\mathbb{P}[X_{ij} = 1] = 1 - \mathbb{P}[X_{ij} = 0] = p_{ij}, \tag{3.5.1}$$

where $p_{ij} = \frac{w_i w_j}{\ell_n + w_i w_j}$. Let us assume that $p_{ij} < 1$ for every $i, j \in [n]$. Then, with $q_{ij} = 1 - p_{ij}$ and for $x = (x_{ij})_{1\le i<j\le n}$, where $x_{ij} \in \{0, 1\}$ for any $i, j \in [n]$,

$$\mathbb{P}[X = x] = \prod_{1\le i<j\le n} p_{ij}^{x_{ij}} q_{ij}^{1-x_{ij}}. \tag{3.5.2}$$

In a first step, we define the *odds-ratios* $(r_{ij})_{1\le i<j\le n}$ by

$$r_{ij} \coloneqq \frac{p_{ij}}{q_{ij}}, \tag{3.5.3}$$

which is well-defined, since $q_{ij} = 1 - p_{ij} > 0$ by the assumption $p_{ij} < 1$.

Since

$$\frac{r_{ij}}{1 + r_{ij}} = \frac{\frac{p_{ij}}{q_{ij}}}{1 + \frac{p_{ij}}{q_{ij}}} = \frac{\frac{p_{ij}}{q_{ij}}}{\frac{q_{ij}+p_{ij}}{q_{ij}}} = p_{ij}, \tag{3.5.4}$$

whereas

$$\frac{1}{1 + r_{ij}} = \frac{1}{1 + \frac{p_{ij}}{q_{ij}}} = \frac{1}{\frac{q_{ij}+p_{ij}}{q_{ij}}} = q_{ij}, \tag{3.5.5}$$

we can deduce that

$$\begin{aligned}
\mathbb{P}[X = x] &= \prod_{1\le i<j\le n} p_{ij}^{x_{ij}} q_{ij}^{1-x_{ij}} \\
&= \prod_{1\le i<j\le n} \left(\frac{r_{ij}}{1 + r_{ij}}\right)^{x_{ij}} \left(\frac{1}{1 + r_{ij}}\right)^{1-x_{ij}} \\
&= \prod_{1\le i<j\le n} \frac{1}{1 + r_{ij}} \prod_{1\le i<j\le n} r_{ij}^{x_{ij}}.
\end{aligned} \tag{3.5.6}$$

In a next step, we adapt to the setting of the generalised random graph by choosing

$$r_{ij} = u_i u_j, \tag{3.5.7}$$

for some weights $(u_i)_{i\in[n]}$.

For example, we can choose

$$u_i \coloneqq \frac{w_i}{\sqrt{\ell_n}}, \tag{3.5.8}$$

in which case we return to

$$p_{ij} = \frac{r_{ij}}{1 + r_{ij}} = \frac{u_i u_j}{1 + u_i u_j} = \frac{w_i w_j}{\ell_n + w_i w_j}. \tag{3.5.9}$$

Setting

$$G(u) := \prod_{1\le i<j\le n} (1+u_iu_j), \tag{3.5.10}$$

and using that $\frac{1}{1+r_{ij}} = \frac{1}{1+u_iu_j}$ we obtain,

$$\begin{aligned}\mathbb{P}[X=x] &= \prod_{1\le i<j\le n} \frac{1}{1+r_{ij}} \prod_{1\le i<j\le n} r_{ij}^{x_{ij}} \\ &= G(u)^{-1} \prod_{1\le i<j\le n} (u_iu_j)^{x_{ij}}.\end{aligned} \tag{3.5.11}$$

Let us take a closer look at

$$\prod_{1\le i<j\le n} (u_iu_j)^{x_{ij}}. \tag{3.5.12}$$

For some fixed $k \in [n]$, the above product contains either $(u_ku_j)^{x_{kj}}$ when $k < j$, or $(u_iu_k)^{x_{ik}}$ when $k > i$. Therefore, for any $m \in [n]\backslash\{k\}$, either $u_k^{x_{km}}$ or $u_k^{x_{mk}}$ is a factor of the product in (3.5.12). Since the generalised random graph shall not contain any self-loops, we can assume that $x_{ii} = 0$ for any $i \in [n]$. Moreover, it must hold that $x_{ij} = x_{ji}$ for any $i, j \in [n]$. Hence, if we set

$$d_i(x) = \sum_{j\in[n]} x_{ij}, \tag{3.5.13}$$

we get that

$$\prod_{1\le i<j\le n} (u_iu_j)^{x_{ij}} = \prod_{i\in n} u_i^{d_i(x)}. \tag{3.5.14}$$

Note that $d_i(x)$ is the degree of vertex i in the generalised random graph with configuration $x = (x_{ij})_{1\le i<j\le n}$.

Based on all these manipulations, we can deduce that

$$\begin{aligned}\mathbb{P}[X=x] &= G(u)^{-1} \prod_{1\le i<j\le n} (u_iu_j)^{x_{ij}} \\ &= G(u)^{-1} \prod_{i\in n} u_i^{d_i(x)}.\end{aligned} \tag{3.5.15}$$

Since $\sum_x \mathbb{P}[X=x] = 1$, it follows that

$$G(u) = \prod_{1\le i<j\le n} (1+u_iu_j) = \sum_x \prod_{i\in n} u_i^{d_i(x)}. \tag{3.5.16}$$

Moreover, we can now prove that the distribution of X conditioned on $d_i(X) = d_i \ \forall 1 \le i \le n$ is uniform. This suggests that graphs having an identical degree sequence share an equal probability of being generated. This interesting result is formulated in the following main theorem of this section.

Theorem 3.19 ($\mathrm{GRG}_n(\boldsymbol{w})$ conditioned on degrees has uniform law)**.** *([Hof18], Theorem 6.15) The generalised random graph $GRG_n(\boldsymbol{w})$ with edge probabilities $(p_{ij})_{1\le i<j\le n}$ given by*

$$p_{ij} = \frac{u_i u_j}{1 + u_i u_j}, \tag{3.5.17}$$

conditioned on $d_i(X) = d_i \ \forall i \in [n]$, where $X = (X_{ij})_{1\le i<j\le n}$ with $X_{ij} \sim Be(p_{ij})$, is uniform over all graphs with degree sequence $(d_i)_{i\in[n]}$.

Proof. Let x be such that $d_i(x) = d_i$ for all $i \in [n]$. Then, using the law of total probability, we get that

$$\begin{aligned} \mathbb{P}\left[X = x | d_i(X) = d_i \ \forall i \in [n]\right] &= \frac{\mathbb{P}[X = x]}{\mathbb{P}[d_i(X) = d_i \ \forall i \in [n]]} \\ &= \sum_{y : d_i(y) = d_i \ \forall i \in [n]} \mathbb{P}[X = y]. \end{aligned} \tag{3.5.18}$$

By (3.5.15), we can deduce that

$$\begin{aligned} \mathbb{P}[X = x | d_i(X) = d_i \ \forall i \in [n]] &= \frac{\prod_{i\in n} u_i^{d_i(x)}}{\sum_{y : d_i(y) = d_i \ \forall i \in [n]} \prod_{i\in n} u_i^{d_i(y)}} \\ &= \frac{\prod_{i\in n} u_i^{d_i}}{\sum_{y : d_i(y) = d_i \ \forall i \in [n]} \prod_{i\in n} u_i^{d_i}} \\ &= \frac{1}{\sum_{y : d_i(y) = d_i \ \forall i \in [n]} 1} \\ &= \frac{1}{\#\{y : d_i(y) = d_i \ \forall i \in [n]\}}. \end{aligned} \tag{3.5.19}$$

Therefore, the distribution of the generalised random graph conditioned on its degrees is uniform over all graphs with the prescribed degree sequence. □

3.6 Asymptotic Equivalence of Inhomogeneous Random Graphs

The generalised random graph model is just one of many different inhomogeneous random graph models. There are many papers introducing similar models having independent (or at least conditionally independent) edges as well. In this section, we define what it means for two inhomogeneous random graphs to be asymptotically equivalent. Moreover, we investigate some sufficient criterion for two inhomogeneous random graphs to be asymptotically equivalent. At the end of this chapter, we will introduce the Chung-Lu model which is related to the generalised random graph, and investigate under which circumstances they are asymptotically equivalent.

Let us start by introducing the notion of asymptotic equivalence for general sequences of probability measures.

Definition 3.20 (Asymptotic equivalence of sequences of probability measures)**.** *([Hof18], Definition 6.17)*
Let $(\mathcal{X}_n, \mathcal{F}_n)$ be a sequence of measurable spaces. Let $\mathbb{P}_n$ and $\mathbb{Q}_n$ be two probability measures on $(\mathcal{X}_n, \mathcal{F}_n)$. Then, we say that the sequences $(\mathbb{P}_n)_{n\geq 1}$ and $(\mathbb{Q}_n)_{n\geq 1}$ are asymptotically equivalent if, for every sequence $\mathcal{E}_n \in \mathcal{F}_n$ of events,

$$\lim_{n\to\infty} \mathbb{P}_n(\mathcal{E}_n) - \mathbb{Q}_n(\mathcal{E}_n) = 0. \tag{3.6.1}$$

Hence, $(\mathbb{P}_n)_{n\geq 1}$ and $(\mathbb{Q}_n)_{n\geq 1}$ are asymptotically equivalent when they have asymptotically equal probabilities.

The following theorem is the main result of this section and gives a sharp criterion on when two inhomogeneous random graph sequences are asymptotically equivalent.

Theorem 3.21 (Asymptotic equivalence of inhomogeneous random graphs)**.** *([Hof18], Theorem 6.18)*
Let $\mathrm{IRG}_n(\boldsymbol{p})$ and $\mathrm{IRG}_n(\boldsymbol{q})$ be two inhomogeneous random graphs with edge probabilities $\boldsymbol{p} = (p_{ij})_{1\leq i<j\leq n}$ and $\boldsymbol{q} = (q_{ij})_{1\leq i<j\leq n}$ respectively. Assume there exists $\varepsilon > 0$ such that $\max_{1\leq i<j\leq n} p_{ij} \leq 1-\varepsilon$. Then $\mathrm{IRG}_n(\boldsymbol{p})$ and $\mathrm{IRG}_n(\boldsymbol{q})$ are asymptotically equivalent when

$$\lim_{n\to\infty} \sum_{1\leq i<j\leq n} \frac{(p_{ij}-q_{ij})^2}{p_{ij}} = 0. \tag{3.6.2}$$

The following four lemmas are the keys to prove Theorem 3.21. Thus, we need to consider those first.

Lemma 3.22 (Asymptotic equivalence and total variation distance)**.** *([Hof18], Exercise 6.30)*
Let $p^{(n)} = (p_x^{(n)})_{x\in\mathcal{X}}$ and $q^{(n)} = (q_x^{(n)})_{x\in\mathcal{X}}$ be two sequences of probability mass functions. Then, they are asymptotically equivalent if and only if $\lim_{n\to\infty} d_{TV}\left(p^{(n)}, q^{(n)}\right) = 0$.

Proof.
($\Rightarrow$) Let us first assume that $p^{(n)}$ and $q^{(n)}$ are asymptotically equivalent. Thus, for every sequence $(\mathcal{E}_n)_{n\geq 1}$ of events

$$\lim_{n\to\infty} \mu^{(n)}(\mathcal{E}_n) - \nu^{(n)}(\mathcal{E}_n) = 0, \tag{3.6.3}$$

where $\mu^{(n)}(\mathcal{E}_n) = \sum_{x\in\mathcal{E}_n} p_x^{(n)}$ and $\nu^{(n)}(\mathcal{E}_n) = \sum_{x\in\mathcal{E}_n} q_x^{(n)}$ are the laws corresponding to the probability mass functions.

In the next step, recall that according to Corollary 2.15

$$d_{TV}\left(p^{(n)}, q^{(n)}\right) = \sum_{x\in\mathcal{X}} \left(p_x^{(n)} - \min\{p_x^{(n)}, q_x^{(n)}\}\right), \tag{3.6.4}$$

which we can rewrite as

$$d_{TV}\left(p^{(n)}, q^{(n)}\right) = \sum_{x\in\mathcal{X}} \left(p_x^{(n)} - q_x^{(n)}\right) \cdot \mathbb{1}_{\{p_x^{(n)} > q_x^{(n)}\}}. \tag{3.6.5}$$

Note that, when we set $\mathcal{E}_n = \{x \in \mathcal{X} : p_x^{(n)} > q_x^{(n)}\}$ for any $n \geq 1$, we get that

$$d_{TV}\left(p^{(n)}, q^{(n)}\right) = \sum_{x \in \mathcal{E}_n} \left(p_x^{(n)} - q_x^{(n)}\right) = \sum_{x \in \mathcal{E}_n} p_x^{(n)} - \sum_{x \in \mathcal{E}_n} q_x^{(n)} = \mu^{(n)}(\mathcal{E}_n) - \nu^{(n)}(\mathcal{E}_n). \tag{3.6.6}$$

Since by assumption $\lim_{n\to\infty} \mu^{(n)}(\mathcal{E}_n) - \nu^{(n)}(\mathcal{E}_n) = 0$ for any sequence $(\mathcal{E}_n)_{n\geq 1}$, we can deduce that $\lim_{n\to\infty} d_{TV}\left(p^{(n)}, q^{(n)}\right) = 0$.

($\Leftarrow$) Let us now assume that $\lim_{n\to\infty} d_{TV}\left(p^{(n)}, q^{(n)}\right) = 0$ and let again $\mu^{(n)}$ and $\nu^{(n)}$ be the laws corresponding to the probability mass functions $p^{(n)}$ and $q^{(n)}$ respectively. Then, recall that according to the Proposition 2.15

$$d_{TV}\left(p^{(n)}, q^{(n)}\right) = \frac{1}{2} \sum_{x \in \mathcal{X}} \left|p_x^{(n)} - q_x^{(n)}\right|. \tag{3.6.7}$$

Clearly, it holds for any sequence of events $(\mathcal{E}_n)_{n\geq 1}$ that

$$0 \leq \left|\mu^{(n)}(\mathcal{E}_n) - \nu^{(n)}(\mathcal{E}_n)\right| \leq \sum_{x \in \mathcal{E}_n} \left|p_x^{(n)} - q_x^{(n)}\right| \leq \sum_{x \in \mathcal{X}} \left|p_x^{(n)} - q_x^{(n)}\right| = 2 \cdot d_{TV}\left(p^{(n)}, q^{(n)}\right). \tag{3.6.8}$$

Thus, when $\lim_{n\to\infty} d_{TV}\left(p^{(n)}, q^{(n)}\right) = 0$, we get that

$$\lim_{n\to\infty} \sum_{x \in \mathcal{E}_n} p_x^{(n)} - \sum_{x \in \mathcal{E}_n} q_x^{(n)} = \lim_{n\to\infty} \mu^{(n)}(\mathcal{E}_n) - \nu^{(n)}(\mathcal{E}_n) = 0, \tag{3.6.9}$$

and therefore, $p^{(n)} = \left(p_x^{(n)}\right)_{x\in\mathcal{X}}$ and $q^{(n)} = \left(q_x^{(n)}\right)_{x\in\mathcal{X}}$ are asymptotically equivalent. □

Lemma 3.22 shows that the notion of asymptotic equivalence of random variables is strongly related to the concept of coupling. Two sequences of random variables are asymptotically equivalent precisely when they can be coupled, such that they are equal with high probability. Therefore, in particular, $\mathrm{IRG}_n(\boldsymbol{p})$ and $\mathrm{IRG}_n(\boldsymbol{q})$ are asymptotically equivalent if and only if they can be coupled in such a way that $\mathbb{P}\left[\mathrm{IRG}_n(\boldsymbol{p}) \neq \mathrm{IRG}_n(\boldsymbol{q})\right] = o(1)$.

To prove Theorem 3.21, rather than using the total variation distance between two measures, it is more convenient to work with the so-called Hellinger distance, which is defined, for discrete measures $p = (p_x)_{x\in\mathcal{X}}$ and $q = (q_x)_{x\in\mathcal{X}}$ by

$$d_{\mathrm{H}}(p, q) = \sqrt{\frac{1}{2} \sum_{x \in \mathcal{X}} (\sqrt{p_x} - \sqrt{q_x})^2}. \tag{3.6.10}$$

However, there is an intimate relation between the total variation distance and the Hellinger distance which is stated in the following lemma.

Lemma 3.23 (Total variation and Hellinger distance)**.** *([Hof18], Exercise 6.32)*
Let $p = (p_x)_{x\in\mathcal{X}}$ and $q = (q_x)_{x\in\mathcal{X}}$ be two discrete probability measures. Then,

$$d_{\mathrm{H}}(p, q)^2 \leq d_{\mathrm{TV}}(p, q) \leq \sqrt{2} \cdot d_{\mathrm{H}}(p, q). \tag{3.6.11}$$

Proof. In a first step, we recall Proposition 2.15 and rewrite it as

$$\begin{aligned} d_{\mathrm{TV}}(p,q) &= \frac{1}{2}\sum_{x\in\mathcal{X}}|p_x - q_x| = \frac{1}{2}\sum_{x\in\mathcal{X}}(\sqrt{p_x}+\sqrt{q_x})|\sqrt{p_x}-\sqrt{q_x}| \\ &= \frac{1}{2}\sum_{x\in\mathcal{X}}\sqrt{p_x}|\sqrt{p_x}-\sqrt{q_x}| + \frac{1}{2}\sum_{x\in\mathcal{X}}\sqrt{q_x}|\sqrt{p_x}-\sqrt{q_x}|. \end{aligned} \tag{3.6.12}$$

In a next step, we recall the Cauchy-Schwarz inequality, which states that

$$\left(\sum_{i=1}^{n} u_i v_i\right)^2 \leq \left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right). \tag{3.6.13}$$

We use the Cauchy-Schwarz inequality to obtain that

$$\sum_{x\in\mathcal{X}}\sqrt{p_x}|\sqrt{p_x}-\sqrt{q_x}| \leq \sqrt{\sum_{x\in\mathcal{X}}p_x}\sqrt{\sum_{x\in\mathcal{X}}(\sqrt{p_x}-\sqrt{q_x})^2} = \sqrt{2}\cdot d_{\mathrm{H}}(p,q) \tag{3.6.14}$$

$$\sum_{x\in\mathcal{X}}\sqrt{q_x}|\sqrt{p_x}-\sqrt{q_x}| \leq \sqrt{\sum_{x\in\mathcal{X}}q_x}\sqrt{\sum_{x\in\mathcal{X}}(\sqrt{p_x}-\sqrt{q_x})^2} = \sqrt{2}\cdot d_{\mathrm{H}}(p,q). \tag{3.6.15}$$

Therefore,

$$d_{\mathrm{TV}}(p,q) \leq \frac{1}{2}\sqrt{2}\cdot d_{\mathrm{H}}(p,q) + \frac{1}{2}\sqrt{2}\cdot d_{\mathrm{H}}(p,q) = \sqrt{2}\cdot d_{\mathrm{H}}(p,q). \tag{3.6.16}$$

To get the lower bound, we observe that

$$d_{\mathrm{H}}(p,q)^2 = \frac{1}{2}\sum_{x\in\mathcal{X}}(\sqrt{p_x}-\sqrt{q_x})^2 \leq \frac{1}{2}\sum_{x\in\mathcal{X}}(\sqrt{p_x}+\sqrt{q_x})|\sqrt{p_x}-\sqrt{q_x}| = d_{\mathrm{TV}}(p,q). \tag{3.6.17}$$

Hence, we conclude that

$$d_{\mathrm{H}}(p,q)^2 \leq d_{\mathrm{TV}}(p,q) \leq \sqrt{2}\cdot d_{\mathrm{H}}(p,q). \tag{3.6.18}$$

□

Combining the results of Lemma 3.22 and Lemma 3.23, we get the following result.

Lemma 3.24 (Asymptotic equivalence and Hellinger distance)**.** *([Hof18], Exercise 6.33) Let $p^{(n)} = \left(p_x^{(n)}\right)_{x\in\mathcal{X}}$ and $q^{(n)} = \left(q_x^{(n)}\right)_{x\in\mathcal{X}}$ be two sequences of probability mass functions. Then, they are asymptotically equivalent if and only if $\lim_{n\to\infty} d_{\mathrm{H}}\left(p^{(n)}, q^{(n)}\right) = 0$.*

Proof.
($\Rightarrow$) Let us first assume that $p^{(n)} = \left(p_x^{(n)}\right)_{x\in\mathcal{X}}$ and $q^{(n)} = \left(q_x^{(n)}\right)_{x\in\mathcal{X}}$ are asymptotically equivalent. This implies that $\lim_{n\to\infty} d_{\mathrm{TV}}\left(p^{(n)}, q^{(n)}\right) = 0$ by Lemma 3.22. From Lemma 3.23, we know that

$$d_{\mathrm{H}}(p,q)^2 \leq d_{\mathrm{TV}}(p,q), \tag{3.6.19}$$

from which we can deduce that

$$d_{\mathrm{H}}(p,q) \leq \sqrt{d_{\mathrm{TV}}(p,q)}. \tag{3.6.20}$$

Thus, $\lim_{n\to\infty} d_{\mathrm{TV}}\left(p^{(n)}, q^{(n)}\right) = 0$ implies in turn that $\lim_{n\to\infty} d_{\mathrm{H}}\left(p^{(n)}, q^{(n)}\right) = 0$.

($\Leftarrow$) Let us now assume that $\lim_{n\to\infty} d_{\mathrm{H}}\left(p^{(n)}, q^{(n)}\right) = 0$. By Lemma 3.23, we know that

$$d_{\mathrm{H}}(p,q)^2 \leq d_{\mathrm{TV}}(p,q) \leq \sqrt{2} \cdot d_{\mathrm{H}}(p,q). \tag{3.6.21}$$

Since $\lim_{n\to\infty} d_{\mathrm{H}}\left(p^{(n)}, q^{(n)}\right) = 0$, both the left- and right-hand side of those inequalities converge to zero, which implies by the sandwich theorem that $\lim_{n\to\infty} d_{\mathrm{TV}}\left(p^{(n)}, q^{(n)}\right) = 0$. Using again Lemma 3.22, we can conclude that $p^{(n)} = (p_x^{(n)})_{x\in\mathcal{X}}$ and $q^{(n)} = (q_x^{(n)})_{x\in\mathcal{X}}$ are asymptotically equivalent. □

Before stating the fourth and last lemma of this section, we need to introduce another definition. For p and q corresponding to Bernoulli distributions, we define

$$\rho(p,q) \coloneqq 2d_{\mathrm{H}}(Be(p), Be(q))^2 = \left(\sqrt{p} - \sqrt{q}\right)^2 + \left(\sqrt{1-p} - \sqrt{1-q}\right)^2. \tag{3.6.22}$$

The next lemma gives us an upper bound for $\rho(p,q)$.

Lemma 3.25 (Bound on Hellinger distance for Bernoulli variables)**.** *([Hof18], Exercise 6.35) For p and q corresponding to Bernoulli distributions, it holds that*

$$\rho(p,q) \leq (p-q)^2 \left(p^{-1} + (1-p)^{-1}\right). \tag{3.6.23}$$

Moreover, when $p \leq 1-\varepsilon$ for some $\varepsilon > 0$, then there exists $C = C(\varepsilon) > 0$ such that

$$\rho(p,q) \leq C\frac{(p-q)^2}{p}. \tag{3.6.24}$$

Proof. By definition we have that

$$\rho(p,q) = \left(\sqrt{p} - \sqrt{q}\right)^2 + \left(\sqrt{1-p} - \sqrt{1-q}\right)^2. \tag{3.6.25}$$

Using that

$$\left(\sqrt{p} - \sqrt{q}\right)^2 = \frac{\left(\sqrt{p} - \sqrt{q}\right)^2\left(\sqrt{p} + \sqrt{q}\right)^2}{\left(\sqrt{p} + \sqrt{q}\right)^2} = \frac{(p-q)^2}{\left(\sqrt{p} + \sqrt{q}\right)^2}, \tag{3.6.26}$$

$$\begin{aligned}\left(\sqrt{1-p} - \sqrt{1-q}\right)^2 &= \frac{\left(\sqrt{1-p} - \sqrt{1-q}\right)^2\left(\sqrt{1-p} + \sqrt{1-q}\right)^2}{\left(\sqrt{1-p} + \sqrt{1-q}\right)^2} \\ &= \frac{(q-p)^2}{\left(\sqrt{1-p} + \sqrt{1-q}\right)^2},\end{aligned} \tag{3.6.27}$$

we obtain that

$$\begin{aligned}\rho(p,q) &= (p-q)^2\left(\left(\sqrt{p} + \sqrt{q}\right)^{-2} + \left(\sqrt{1-p} + \sqrt{1-q}\right)^{-2}\right) \\ &\leq (p-q)^2\left(\left(\sqrt{p}\right)^{-2} + \left(\sqrt{1-p}\right)^{-2}\right) \\ &= (p-q)^2\left(p^{-1} + (1-p)^{-1}\right).\end{aligned} \tag{3.6.28}$$

Now let us assume that there exists some $\varepsilon > 0$ such that $p \leq 1 - \varepsilon$. Then, we can observe that

$$\rho(p,q) \leq (p-q)^2 \left(p^{-1} + \varepsilon^{-1}\right) = (p-q)^2 \left(\frac{\varepsilon + p}{p\varepsilon}\right) \leq (p-q)^2 \left(\frac{1}{p\varepsilon}\right). \tag{3.6.29}$$

Therefore, when we choose $C = C(\varepsilon) \coloneqq \varepsilon^{-1}$, we can conclude that

$$\rho(p,q) \leq C \frac{(p-q)^2}{p}. \tag{3.6.30}$$

□

Now we have laid all the necessary foundations to complete the proof of Theorem 3.21, which provides us with a sufficient criterion for the asymptotic equivalence of two inhomogeneous random graphs.

Proof of Theorem 3.21.
Let $\mathrm{IRG}_n(\boldsymbol{p})$ and $\mathrm{IRG}_n(\boldsymbol{q})$ with $\boldsymbol{p} = (p_{ij})_{1 \leq i < j \leq n}$ and $\boldsymbol{q} = (q_{ij})_{1 \leq i < j \leq n}$ be two inhomogeneous random graphs. We say that $\mathrm{IRG}_n(\boldsymbol{p})$ and $\mathrm{IRG}_n(\boldsymbol{q})$ are asymptotically equivalent, if their edge variables $X_{ij} \sim Be(p_{ij})$ and $Y_{ij} \sim Be(q_{ij})$ are asymptotically equivalent. Lemma 3.24 states that the edge variables are asymptotically equivalent precisely when $\lim_{n\to\infty} d_{\mathrm{H}}(\boldsymbol{p}, \boldsymbol{q}) = 0$, which is what we are going to prove.

For two discrete probability measures $p = (p_x)_{x \in \mathcal{X}}$ and $q = (q_x)_{x \in \mathcal{X}}$, we denote

$$H(p,q) = 1 - d_{\mathrm{H}}(p,q)^2 = 1 - \frac{1}{2}\rho(p,q). \tag{3.6.31}$$

For $\mathrm{IRG}_n(\boldsymbol{p})$ and $\mathrm{IRG}_n(\boldsymbol{q})$ with $\boldsymbol{p} = (p_{ij})_{1 \leq i < j \leq n}$ and $\boldsymbol{q} = (q_{ij})_{1 \leq i < j \leq n}$, the edges are independent, such that

$$H(\boldsymbol{p}, \boldsymbol{q}) = \prod_{1 \leq i < j \leq n} \left(1 - \frac{1}{2}\rho(p_{ij}, q_{ij})\right). \tag{3.6.32}$$

Since

$$d_{\mathrm{H}}(\boldsymbol{p}, \boldsymbol{q}) = \sqrt{1 - H(\boldsymbol{p}, \boldsymbol{q})}, \tag{3.6.33}$$

$d_{\mathrm{H}}(\boldsymbol{p}, \boldsymbol{q}) = o(1)$ precisely when $H(\boldsymbol{p}, \boldsymbol{q}) = 1 - o(1)$.

Using that $(1-x)(1-y) \geq 1 - x - y$ and $1 - x \leq \exp(-x)$ for any $x, y \geq 0$, we can bound $H(\boldsymbol{p}, \boldsymbol{q})$ from below and above by

$$1 - \frac{1}{2} \sum_{1 \leq i < j \leq n} \rho(p_{ij}, q_{ij}) \leq H(\boldsymbol{p}, \boldsymbol{q}) \leq \exp\left(-\frac{1}{2} \sum_{1 \leq i < j \leq n} \rho(p_{ij}, q_{ij})\right). \tag{3.6.34}$$

Hence, $H(\boldsymbol{p}, \boldsymbol{q}) = 1 - o(1)$ if and only if $\sum_{1 \leq i < j \leq n} \rho(p_{ij}, q_{ij}) = o(1)$.

By assumption there exists $\varepsilon > 0$ such that $\max_{1\le i<j\le n} p_{ij} \le 1-\varepsilon$. Thus, by Lemma 3.25 and by the theorem's assumptions, we obtain that

$$\sum_{1\le i<j\le n} \rho(p_{ij}, q_{ij}) \le C \sum_{1\le i<j\le n} \frac{(p_{ij}-qij)^2}{p_{ij}} = o(1), \tag{3.6.35}$$

which completes the proof. □

We now introduce the Chung-Lu model, which is another inhomogeneous random graph and, moreover, related to the generalised random graph. We can show that Conditions 3.3 (a)-(c) ensure the asymptotic equivalence of the Chung-Lu model and the generalised random graph. Let $\boldsymbol{w} = (w_i)_{i\in[n]}$ be deterministic vertex weights. The edge probabilities in the Chung-Lu random graph are then given by

$$p_{ij}^{(CL)} = \frac{w_i w_j}{\ell_n} \wedge 1, \tag{3.6.36}$$

where $\ell_n = \sum_{i\in[n]} w_i$ as in the generalised random graph. We denote the resulting graph by $\mathrm{CL}_n(\boldsymbol{w})$.

Note that, when $\max_{i\in[n]} w_i^2 \le \ell_n$ we can neglect the minimum with 1 in (3.6.36). Moreover, when $\frac{w_i}{\sqrt{\ell_n}}$ is small, there is hardly any difference between the edge weights $p_{ij} = \frac{w_i w_j}{\ell_n + w_i w_j}$ and $p_{ij}^{(CL)} = \frac{w_i w_j}{\ell_n}$. That is why one might expect that $\mathrm{CL}_n(\boldsymbol{w})$ and $\mathrm{GRG}_n(\boldsymbol{w})$ behave similarly. The following theorem gives a sufficient condition under which the two inhomogeneous random graphs are asymptotically equivalent.

Theorem 3.26 (Asymptotic equivalence of $\mathrm{CL}_n(\boldsymbol{w})$ and $\mathrm{GRG}_n(\boldsymbol{w})$ with deterministic weights)**.** *([Hof18], Theorem 6.19)*

The random graphs $CL_n(\boldsymbol{w})$ and $GRG_n(\boldsymbol{w})$ with deterministic weights $\boldsymbol{w} = (w_i)_{i\in[n]}$ are asymptotically equivalent when

$$\sum_{i\in[n]} w_i^3 = o\left(n^{\frac{3}{2}}\right). \tag{3.6.37}$$

Proof. We want to make use of Theorem 3.21. Thus, we have to show that

$$\lim_{n\to\infty} \sum_{1\le i<j\le n} \frac{\left(p_{ij} - p_{ij}^{(CL)}\right)^2}{p_{ij}} = 0. \tag{3.6.38}$$

To do that, we start by bounding $p_{ij}^{(CL)} - p_{ij}$ from above. Using the fact that $1 - \frac{1}{1+x} \le x$, we obtain

$$p_{ij}^{(CL)} - p_{ij} \le \frac{w_i w_j}{\ell_n} - \frac{w_i w_j}{\ell_n + w_i w_j} = \frac{w_i w_j}{\ell_n}\left[1 - \frac{1}{1+\frac{w_i w_j}{\ell_n}}\right] \le \frac{w_i^2 w_j^2}{\ell_n^2}. \tag{3.6.39}$$

Moreover, $\max_{i\in[n]} w_i^3 \le \sum_{i\in[n]} w_i^3 = o\left(n^{\frac{3}{2}}\right)$ implies that $\max_{i\in[n]} w_i^2 = o(n)$. Therefore, when we choose n sufficiently large, we get

$$p_{ij} = \frac{w_i w_j}{\ell_n + w_i w_j} \ge \frac{w_i w_j}{\ell_n + \max_{i\in[n]} w_i^2} \ge \frac{w_i w_j}{2\ell_n}. \tag{3.6.40}$$

Therefore, it follows that

$$\sum_{1\leq i<j\leq n} \frac{\left(p_{ij}-p_{ij}^{(CL)}\right)^2}{p_{ij}} \leq \frac{2}{\ell_n^3} \sum_{1\leq 1<j\leq n} w_i^3 w_j^3 \leq \frac{1}{\ell_n^3}\left(\sum_{i\in[n]} w_i^3\right)^2 = o(1), \tag{3.6.41}$$

by the assumption that $\sum_{i\in[n]} w_i^3 = o\left(n^{\frac{3}{2}}\right)$. Thus, by Theorem 3.21 we can conclude that $\mathrm{CL}_n(\boldsymbol{w})$ and $\mathrm{GRG}_n(\boldsymbol{w})$ are indeed asymptotically equivalent under the given assumptions. □

The next corollary is a direct consequence of Theorem 3.26.

Corollary 3.27 (Asymptotic equivalence of $\mathrm{CL}_n(\boldsymbol{w})$ and $\mathrm{GRG}_n(\boldsymbol{w})$ under regularity conditions)**.** *([Hof18], Corollary 6.20)*
Assume that the deterministic weights $\boldsymbol{w} = (w_i)_{i\in[n]}$ satisfy Conditions 3.3 (a)-(c). Then, the random graphs $CL_n(\boldsymbol{w})$ and $GRG_n(\boldsymbol{w})$ are asymptotically equivalent.

Proof. When Conditions 3.3 (a)-(c) hold, we know by Proposition 3.4 that $\max_{i\in[n]} w_i = o\left(\sqrt{n}\right)$. Thus,

$$\sum_{i\in[n]} w_i^3 \leq \max_{i\in[n]} w_i \sum_{i\in[n]} w_i^2 = n^2 o\left(\sqrt{n}\right)\mathbb{E}\left[W_n^2\right] = o\left(n^{\frac{3}{2}}\right). \tag{3.6.42}$$

By Theorem 3.26 we conclude that $\mathrm{CL}_n(\boldsymbol{w})$ and $\mathrm{GRG}_n(\boldsymbol{w})$ are asymptotically equivalent. □

4 Configuration Model

In this chapter we will have a closer look at the configuration model. In comparison to the generalised random graph, we choose the degree sequence beforehand and generate a random graph with precisely the same degrees. As one might expect, this will often lead to graphs with self-loops as well as multiple edges. Therefore, the configuration model is particularly useful if we want to model real-world networks such as Facebook wall posts which contain such structures.

However, there are many real-world networks like for example the internet infrastructure or transportation networks which we often want to model by a simple graph without any self-loops or multiple edges. Thus, we will show that under some conditions the configuration model leads asymptotically to simple graphs. Additionally, we will prove that when conditioned on simplicity, the generated graph is uniform over all simple graphs.

We will finish this chapter by investigating connections between the generalised random graph and the configuration model.

4.1 Introduction of the Model

The configuration model can be used to generate a random graph with prescribed degrees. We first fix the number of vertices n and choose a degree sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$. We choose $d_j \geq 1$ for all $j \in [n]$, since when $d_j = 0$, vertex j is isolated and can be removed from the graph. To construct the multigraph where vertex j has degree d_j for every $j \in [n]$, we attach d_j half-edges to vertex j (see Figure 4.1) and pair all the half-edges in a uniform way. As two half-edges together form one edge of the graph, we need to choose the degree sequence in such a way that the sum of all degrees $\ell_n = \sum_{j\in[n]} d_j$, also called the total degree, is even. By connecting every half-edge to another half-edge, we generate the graph. To do that, the half-edges are numbered in an arbitrary way from 1 to ℓ_n. Then, we connect the first half-edge randomly with one of the $\ell_n - 1$ remaining half-edges. Hence, these two half-edges form a single edge of the multigraph and are removed from the list of half-edges that need to be connected. This procedure is continued until all half-edges are paired. The resulting graph is called the configuration model with degree sequence $\boldsymbol{d}$, abbreviated as $\mathrm{CM}_n(\boldsymbol{d})$.

Figure 4.1 shows an example of a realisation of the pairing process of half-edges given the degree sequence $\boldsymbol{d} = (3, 2, 2, 1)$, whereas Figure 4.2 displays all possible resulting graphs having the prescribed degrees $\boldsymbol{d} = (3, 2, 2, 1)$.

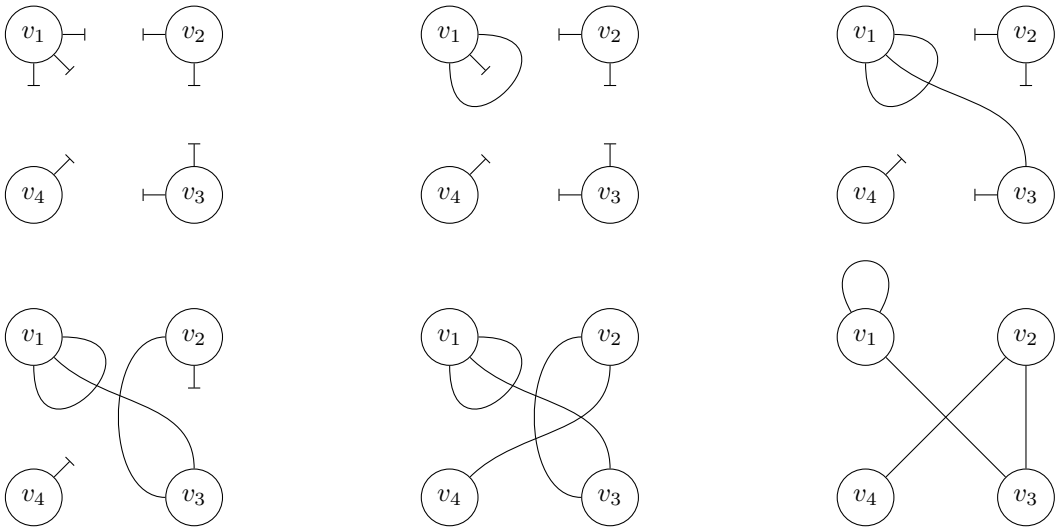

Figure 4.1: The pairing of half-edges in the configuration model with $n = 6$ and $\boldsymbol{d} = (3, 2, 2, 1)$. In each step a new edge is formed connecting two half-edges. The last graph shows the result of the pairing process in a more recognisable way.

As seen in Figures 4.1 and 4.2, some vertices in the configuration model may have self-loops as well as multiple edges. Clearly, when the total degree ℓ_n is even, then the above procedure does produce a multigraph with the right degree sequence. In the degree sequence of a multigraph, a self-loop contributes two to the degree of the vertex incident to it, while each of the multiple edges contributes one to the degree of each of the two vertices incident to it.

Although Figure 4.2 might give the impression that simple graphs are quite rare, we will prove that multiple edges and self-loops are relatively scarce when $n \to \infty$ and the degree sequence is subject to some regularity conditions.

As some may worry about the influence of the arbitrary ordering of the half-edges in the pairing procedure, we will use the following lemma to justify that the process of pairing half-edges is exchangeable, meaning that the order does not matter for the distribution of the final outcome.

Lemma 4.1 (The number of configurations)**.** *([Hof18], Exercise 7.2)*
Let $\boldsymbol{d} = (d_i)_{i\in[n]}$ be a deterministic degree sequence with total degree $\ell_n = \sum_{j\in[n]} d_j$ and let d_i be the number of half-edges attached to vertex i. Then, there are $(\ell_n - 1)!! = (\ell_n - 1)(\ell_n - 3)\cdots 3 \cdot 1$ different ways of pairing the ℓ_n half-edges.

Proof. For ℓ_n half-edges $\frac{\ell_n}{2}$ pairing steps are needed to pair all half-edges. In each step two half-edges are paired with each other. Let h_1 be any of the ℓ_n half-edges. There are $\ell_n - 1$ different possibilities to pair h_1 to any of the other half-edges. For step $i + 1$, we have already paired $2i$ half-edges. Choosing one of the remaining half-edges randomly and denoting it by h_{i+1}, we can observe that it can be paired with $\ell_n - 2i - 1$ other possible half-edges. For the last step $\frac{\ell_n}{2}$, we have already paired $2(\frac{\ell_n}{2} - 1) = \ell_n - 2$ half-edges. This gives us for all pairing steps the total

amount of possibilities to be

$$(\ell_n - 1)(\ell_n - 3) \cdots (\ell_n - (\ell_n - 2) - 1) = (\ell_n - 1)!!. \tag{4.1.1}$$

Therefore, we can conclude that there are $(\ell_n - 1)$ different possible configurations for a given degree sequence with total degree ℓ_n. □

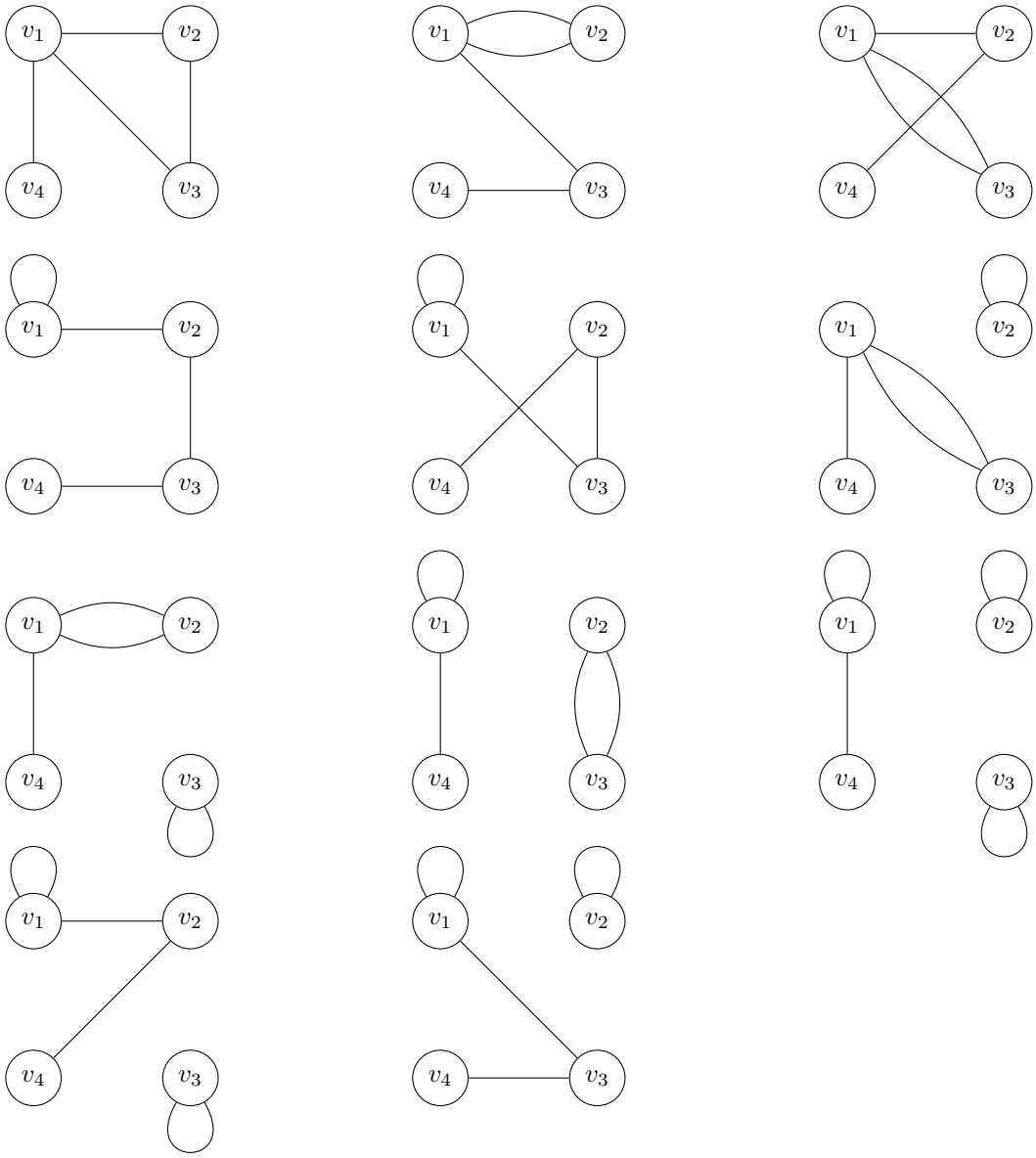

Figure 4.2: All possible resulting (multi)graphs with the prescribed degree sequence $\boldsymbol{d} = (3, 2, 2, 1)$.

According to Lemma 4.1, the probability to get one specific configuration is $\frac{1}{(\ell_n-1)!!}$ Moreover, it gives us an easy way to see that the order of pairings does not matter for the final outcome. As long as each configuration has the same probability, the precise order in which the pairings take place is irrelevant.

However, it is important to note that not all resulting multigraphs have the same probability to occur since some configurations lead to the same resulting (multi)graph while others don't. Let us look at an example to make this better understandable. If we consider the degree sequence $\boldsymbol{d} = (2, 2)$ with $\ell_n = 4$, then there are three possible configurations, which are displayed in Figure 4.3. But the second and third configuration result in the same multigraph. Therefore, given the degree sequence $\boldsymbol{d} = (2, 2)$, it is twice as likely to get a multigraph with multiple edges as to get a multigraph with self-loops using the configuration model. Indeed, we will see that there is a weight $\frac{1}{j!}$ for every edge of multiplicity j and a factor $\frac{1}{2}$ for every self-loop. Hence, there is not a uniform measure on all multigraphs with the prescribed degree sequence.

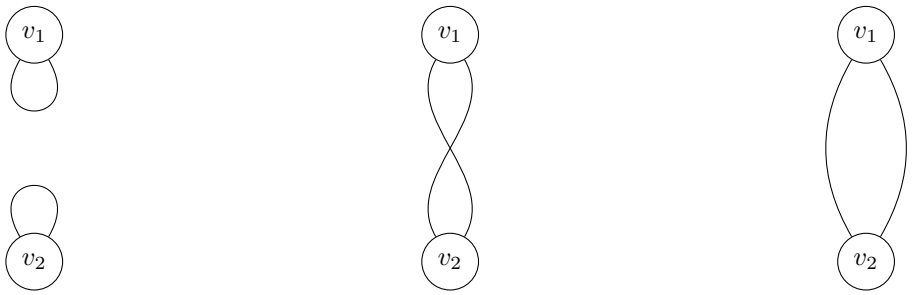

Figure 4.3: All possible configurations with the prescribed degree sequence $\boldsymbol{d} = (2, 2)$.

Before we turn to the next proposition, let us first introduce some notation. Let $G = (x_{ij})_{i,j\in[n]}$ denote a multigraph on the vertices [n], where x_{ij} is the number of edges between vertices $i, j \in [n]$ and x_{ii} is the number of self-loops of vertex $i \in [n]$. Then, for all $i \in [n]$, $d_i = x_{ii} + \sum_{j\in[n]} x_{ij}$. As one can see, the number of self-loops of vertex i appears twice, so that a self-loop contributes two to the degree. By convention, we again set $x_{ij} = x_{ji}$. Let us now find the probability of getting a specific multigraph given a prescribed degree sequence using the configuration model.

Proposition 4.2 (The law of $\mathrm{CM}_n(\boldsymbol{d})$)**.** *([Hof18], Proposition 7.7)*
Let $\boldsymbol{d} = (d_i)_{i\in[n]}$ be a prescribed degree sequence and $G = (x_{ij})_{i,j\in[n]}$ be a multigraph on the vertices [n] which is such that

$$d_i = x_{ii} + \sum_{j\in[n]} x_{ij}. \tag{4.1.2}$$

Then,

$$\mathbb{P}[CM_n(\boldsymbol{d}) = G] = \frac{1}{(\ell_n - 1)!!} \frac{\prod_{i\in[n]} d_i!}{\prod_{i\in[n]} 2^{x_{ii}} \prod_{1\le i\le j\le n} x_{ij}!}. \tag{4.1.3}$$

Proof. By Lemma 4.1, the number of possible configurations with degree sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$ is equal to $(\ell_n - 1)!!$, where ℓ_n is the sum of all degrees. As each configuration has the same probability

$$\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G] = \frac{1}{(\ell_n - 1)!!} N(G), \tag{4.1.4}$$

where $N(G)$ is the number of configurations that give rise to the multigraph G.

Notice that if we permute the half-edges incident to a specific vertex, the resulting multigraph remains unchanged while the configuration is in many cases different. In total there are $\prod_{i\in[n]} d_i!$ ways to permute the half-edges incident to all vertices. However, some of these mentioned permutations give rise to the same configuration.

Let us assume that the multigraph G contains a self-loop and let h_1 and h_2 be the half-edges that were paired to form the self-loop. In that case, the permutation of h_1 and h_2 does not change the configuration. Since this is true for every self-loop of G, we need to divide the total amount of permutations by a factor 2 for every self-loop contained in G. The total compensation needed due to self-loops can be written as $\prod_{i\in[n]} 2^{x_{ii}}$.

Additionally, we might need to add compensation for the case where the multigraph G contains multiple edges. The number x_{ij} denotes the number of multiple edges between the vertices $i, j \in [n]$. Thus, there are $x_{ij}!$ possibilities to permute these multiple edges between vertex i and j (and the half-edges incident to them) which leave the configuration unchanged. Therefore, the total compensation needed on account of multiple edges is $\prod_{1\le i\le j\le n} x_{ij}!$.

Thus, the number of configurations $N(G)$ that give rise to the multigraph G can be computed as

$$N(G) = \frac{\prod_{i\in[n]} d_i!}{\prod_{i\in[n]} 2^{x_{ii}} \prod_{1\le i\le j\le n} x_{ij}!}, \tag{4.1.5}$$

and hence we can conclude that,

$$\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G] = \frac{1}{(\ell_n - 1)!!} \frac{\prod_{i\in[n]} d_i!}{\prod_{i\in[n]} 2^{x_{ii}} \prod_{1\le i\le j\le n} x_{ij}!}. \tag{4.1.6}$$

□

Remark 4.3. *A similar argumentation as in the proof of Proposition 4.2 can be found in [JK93], Section 1.*

Compared to the flexibility in choosing the vertex weights $\boldsymbol{w}$ as in Chapter 3, the choice of a degree sequence allows a much more direct control over the topology of the generated graph. In many applications, such flexibility is rather convenient, as it allows us for example to generate a graph with precisely the same degrees as a real-world network. This gives us the opportunity to compare the structure of a real-world network with the generated random graph and to investigate whether the real-world network is similar to it.

As mentioned beforehand, it is necessary to impose some regularity conditions on the degree sequence $\boldsymbol{d}$ if we want to generate random graphs that are simple with high probability. Before we state these assumptions, we introduce some notation. We denote the degree of a uniformly chosen vertex U in $[n]$ by $D_n = d_U$. The random variable D_n has distribution function F_n given by

$$F_n(x) = \frac{1}{n} \sum_{j\in[n]} \mathbb{1}_{\{d_j \le x\}}, \tag{4.1.7}$$

which is the empirical distribution of the degrees.

Condition 4.4 (Regularity conditions for vertex degrees)**.** *([Hof23], Condition 1.7)*

(a) Weak convergence of vertex degrees.

There exists a distribution function F such that, as $n \to \infty$,

$$D_n \xrightarrow{d} D, \tag{4.1.8}$$

where D_n and D have cumulative distribution functions F_n and F, respectively.

Equivalently, for any $x \in \mathbb{R}$,

$$\lim_{n\to\infty} F_n(x) = F(x). \tag{4.1.9}$$

Further, we assume that $F(0) = 0$, i.e., $\mathbb{P}[D \geq 1] = 1$.

(b) Convergence of average vertex degrees.

As $n \to \infty$,

$$\lim_{n\to\infty} \mathbb{E}[D_n] = \mathbb{E}[D] < \infty, \tag{4.1.10}$$

where D_n and D have cumulative distribution functions F_n and F from part (a), respectively.

(c) Convergence of second moment vertex degrees.

As $n \to \infty$,

$$\lim_{n\to\infty} \mathbb{E}[D_n^2] = \mathbb{E}[D^2] \in (0, \infty), \tag{4.1.11}$$

where D_n and D have cumulative distribution functions F_n and F from part (a), respectively.

Remark 4.5. *At the very end of this chapter, we will discuss the relation between the generalised random graph $GRG_n(\boldsymbol{w})$ conditioned on the degrees and the configuration model $CM_n(\boldsymbol{d})$ conditioned on simplicity. As the degrees obtained from the degree sequence of the generalised random graph $GRG_n(\boldsymbol{w})$ are random themselves, we have to adapt the regularity Conditions 4.4. Thus, in the case where the degrees of the vertices are random themselves, we replace Condition 4.4 (a) by the condition that, for every $k \in \mathbb{N}$,*

$$\mathbb{P}_n[D_n = k] \xrightarrow{\mathbb{P}} \mathbb{P}[D = k], \tag{4.1.12}$$

where $\mathbb{P}_n$ denotes the conditional probability given the (random) degrees $(d_i)_{i\in[n]}$.
Similarly, Conditions 4.4 (b) and (c) can be replaced by

$$\mathbb{E}_n[D_n] \xrightarrow{\mathbb{P}} \mathbb{E}[D], \tag{4.1.13}$$

$$\mathbb{E}_n[D_n^2] \xrightarrow{\mathbb{P}} \mathbb{E}[D^2], \tag{4.1.14}$$

where $\mathbb{E}_n$ denotes the expectation with respect to $\mathbb{P}_n$.

Similarly, as in Chapter 3, we almost always assume that Conditions 4.4 (a)-(b) hold, while Condition 4.4 (c) is used only occasionally. Since the degrees d_i are restricted to integer values, so does

D_n, and henceforth so must the limiting random variable D. Consequently, the limiting distribution function F is constant between integers but has a jump at $x \in \mathbb{N}$ of the size $\mathbb{P}[D = x]$. This implies that the distribution function F does have discontinuity points, and the weak convergence in (4.1.8) usually implies (4.1.9) only at continuity points. However, since F_n is constant between integers, statements (4.1.8) and (4.1.9) are indeed equivalent.

Notice that, instead of defining $\mathrm{CM}_n(\boldsymbol{d})$ in terms of the degrees, we could have defined it in terms of the number of vertices with a fixed degree. More specifically, for any $k \in \mathbb{N}$, we could have denoted

$$n_k = \sum_{i \in [n]} \mathbb{1}_{\{d_i = k\}}, \tag{4.1.15}$$

as the number of vertices with degree k. Then the degree sequence $\boldsymbol{d}$ is uniquely determined by the sequence $(n_k)_{k \geq 0}$, apart from the vertex labels. In that case, Condition 4.4 (a) is equivalent to the statement that $\lim_{n \to \infty} \frac{n_k}{n} = \mathbb{P}[D = k]$, while Condition 4.4 (b) can be equivalently stated as $\lim_{n \to \infty} \sum_{k \geq 0} \frac{k n_k}{n} = \mathbb{E}[D]$. In some cases this slightly different definition will prove useful.

4.2 Example of the Configuration Model with Fixed Degrees

In this subchapter, we present two canonical ways of choosing a degree sequence $\boldsymbol{d}$ such that the regularity Conditions 4.4 are satisfied.

Similarly, as we did for the vertex weights of the generalised random graph $\mathrm{GRG}_n(\boldsymbol{w})$, we start by fixing a distribution function F of an integer-valued random variable D.

A first way to determine the degrees would be to take $d_i = [1 - F]^{-1}\left(\frac{i}{n}\right)$, which is analogous to the way in which we chose the vertex weights in Subchapter 3.2. In that case, the degrees are integer-valued as F is a distribution function of an integer-valued random variable by choice. By adapting the arguments in Subchapter 3.2, it can be demonstrated that such a degree sequence fulfils all regularity conditions.

Let us now focus on a slightly different construction. Let the number of vertices with degree k be equal to

$$n_k = \lceil nF(k) \rceil - \lceil nF(k-1) \rceil, \tag{4.2.1}$$

and take the degree sequence $\boldsymbol{d} = (d_i)_{i \in [n]}$ to be the unique ordered degree sequence compatible with $(n_k)_{k \geq 0}$.

Next, we will prove that the above choice of degrees satisfies the regularity Conditions 4.4 (a)-(c), whenever F is the distribution function of an integer-valued random variable D such that $\mathbb{E}[D] < \infty$ and $\mathbb{E}[D^2] < \infty$.

Proposition 4.6 (Regularity Condition 4.4 (a)). *([Hof18], Exercise 7.6)*
Let $\mathrm{CM}_n(\boldsymbol{d})$ *be such that there are precisely* $n_k = \lceil nF(k)\rceil - \lceil nF(k-1)\rceil$ *vertices with degree* k*, where* F *is the distribution function of an integer-valued random variable* D *for which* $F(0) = 0$*. Then, the corresponding degree sequence* $\boldsymbol{d} = (d_i)_{i\in[n]}$ *satisfies Condition 4.4(a).*

Proof. First note that $\mathbb{P}[D_n = k] = \frac{n_k}{n}$ and thus, $\mathbb{P}[D_n \le k] = \frac{1}{n}\sum_{i=1}^{k} n_i$. By construction $n_k = \lceil nF(k)\rceil - \lceil nF(k-1)\rceil$. Thus, we can deduce that

$$\mathbb{P}[D_n \le k] = \frac{1}{n}\sum_{i=1}^{k}(\lceil nF(i)\rceil - \lceil nF(i-1)\rceil) = \frac{1}{n}(\lceil nF(k)\rceil - \lceil nF(0)\rceil) = \frac{1}{n}\lceil nF(k)\rceil, \tag{4.2.2}$$

where we used the property of telescopic sums as well as the assumption that $F(0) = 0$.

Taking into account that,

$$nF(k) \le \lceil nF(k)\rceil \le nF(k) + 1, \tag{4.2.3}$$

and dividing the inequalities by n, we obtain

$$F(k) \le \frac{\lceil nF(k)\rceil}{n} \le F(k) + \frac{1}{n}. \tag{4.2.4}$$

Taking limits on all sides, we get that

$$\lim_{n\to\infty} \mathbb{P}[D_n \le k] = F(k). \tag{4.2.5}$$

Since $F(k) = \mathbb{P}[D \le k]$, we conclude that Condition 4.4 (a) holds. □

Proposition 4.7 (Regularity Condition 4.4 (b)). *([Hof18], Exercise 7.7)*
Let $\mathrm{CM}_n(\boldsymbol{d})$ *be such that there are precisely* $n_k = \lceil nF(k)\rceil - \lceil nF(k-1)\rceil$ *vertices with degree* k*, where* F *is the distribution function of an integer-valued random variable* D *for which* $F(0) = 0$*. Assume that* $\mathbb{E}[D] < \infty$*. Then, the corresponding degree sequence* $\boldsymbol{d}$ *satisfies Condition 4.4(b).*

Proof. First of all, we notice that, as

$$\mathbb{P}[D_n \le k] = \frac{1}{n}\lceil nF(k)\rceil \ge \frac{1}{n}nF(k) = \mathbb{P}[D \le k], \tag{4.2.6}$$

D_n is stochastically smaller than D (see Definition 2.20).
By assumption $\mathbb{E}[D] \le \infty$ and $D_n \xrightarrow{d} D$ by Proposition 4.6. Thus, we can use Lebesgue's dominated convergence theorem 2.23 to conclude that

$$\lim_{n\to\infty} \mathbb{E}[D_n] = \mathbb{E}[D] < \infty. \tag{4.2.7}$$

□

Proposition 4.8 (Regularity Condition 4.4 (c)).
Let $\mathrm{CM}_n(\boldsymbol{d})$ *be such that there are precisely* $n_k = \lceil nF(k)\rceil - \lceil nF(k-1)\rceil$ *vertices with degree* k*, where* F *is the distribution function of an integer-valued random variable* D *for which* $F(0) = 0$*. Assume that* $\mathbb{E}[D^2] < \infty$*. Then, the corresponding degree sequence* $\boldsymbol{d}$ *satisfies Condition 4.4(c).*

Proof. By Proposition 4.6 $D_n \xrightarrow{d} D$ under the given assumptions. Since $g : \mathbb{R} \to \mathbb{R}; x \mapsto x^2$ is a continuous function, we can deduce by the Continuous mapping theorem 2.5 that $D_n^2 \xrightarrow{d} D^2$.
Next, we use that

$$\mathbb{P}[D_n \leq k] = \frac{1}{n}\lceil nF(k)\rceil \geq \frac{1}{n}nF(k) = \mathbb{P}[D \leq k], \tag{4.2.8}$$

implies $D_n^2 \preceq D^2$ due to Theorem 2.22, as the function g is non-decreasing on $\mathbb{R}_0^+$.

By Lebesgue's dominated convergence theorem 2.23 we conclude that

$$\lim_{n\to\infty} \mathbb{E}[D_n^2] = \mathbb{E}[D^2] < \infty. \tag{4.2.9}$$

□

4.3 Erased Configuration Model

In this subsection, we define and discuss the erased configuration model. As mentioned earlier, one of the major drawbacks of the configuration model is the fact that the generated random graph might have self-loops as well as multiple edges. One way to get a simple graph out of it, is just to erase all self-loops and to merge all multiple edges, which we then call the erased configuration model.

Clearly, this might drastically change the degree sequence that we chose at the beginning, as can be seen in Figure 4.4. While the prescribed degree sequence in that illustrated example is $\boldsymbol{d} = (4, 4, 5, 5, 1, 3)$, the degree sequence of the erased configuration model is $\boldsymbol{D}^{(er)} = (3, 2, 3, 1, 1, 2)$. Thus, the total degree got reduced from 22 to 12.

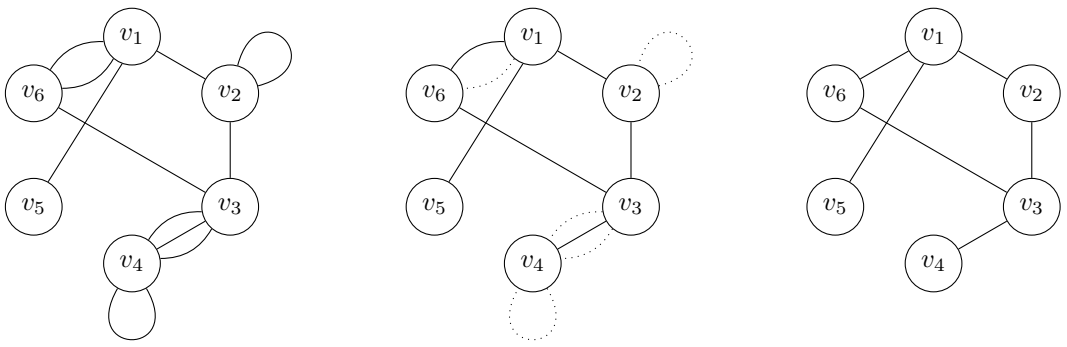

Figure 4.4: Visual depiction of the erasure process. Self-loops and multiple edges of the configuration model $\mathrm{CM}_n(\boldsymbol{d})$ with $\boldsymbol{d} = (4, 4, 5, 5, 1, 3)$ are erased to get the erased configuration model with $\boldsymbol{D}^{(er)} = (3, 2, 3, 1, 1, 2)$.

However, we can prove that the expected number of self-loops and removed multiple edges can be bounded from above. Furthermore, under the regularity Conditions 4.4 and when the number of vertices n is sufficiently large, the degree sequence of the erased configuration model will be close to the prescribed degree sequence.

Let us denote the degrees in the erased configuration model by $\boldsymbol{D}^{(er)} = (D_i^{(er)})_{i\in[n]}$. Then,

$$D_i^{(er)} = d_i - 2s_i - m_i, \tag{4.3.1}$$

where $(d_i)_{i\in[n]}$ are the degrees in $\mathrm{CM}_n(\boldsymbol{d})$, $s_i = x_{ii}$ is the number of self-loops of vertex i in $\mathrm{CM}_n(\boldsymbol{d})$, and

$$m_i = \sum_{j\neq i}(x_{ij}-1)\mathbb{1}_{\{x_{ij}\geq 2\}}, \tag{4.3.2}$$

is the number of multiple edges that got removed from vertex i in $\mathrm{CM}_n(\boldsymbol{d})$.

Denote the empirical distribution of the degree sequence $(p_k^{(n)})_{k\geq 1}$ in $\mathrm{CM}_n(\boldsymbol{d})$ by

$$p_k^{(n)} = \mathbb{P}[D_n = k] = \frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{d_i=k\}}, \tag{4.3.3}$$

and denote the related distribution of the degree sequence in the erased configuration model $(P_k^{(er)})_{k\geq 1}$ by

$$P_k^{(er)} = \frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{D_i^{(er)}=k\}}. \tag{4.3.4}$$

As we work with a deterministic degree sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$, also $\left(p_k^{(n)}\right)_{k\geq 1}$ is deterministic. On the contrary, $\left(P_k^{(er)}\right)_{k\geq 1}$ is a random sequence.

Now we are ready to prove the first two results concerning the expected number of self-loops and removed multiple edges.

Proposition 4.9 (Bound on the expected number of self-loops)**.** *([Hof18], Proposition 7.11) The expected number of self-loops S_n in the configuration model $\mathrm{CM}_n(\boldsymbol{d})$ satisfies*

$$\mathbb{E}[S_n] = \frac{1}{2}\sum_{i\in[n]}\frac{d_i(d_i-1)}{\ell_n - 1} \leq \frac{1}{2}\sum_{i\in[n]}\frac{d_i^2}{\ell_n}. \tag{4.3.5}$$

Proof. For every vertex $i \in [n]$ we number the half-edges incident to it in an arbitrary way and label them as (i, s), where i is the vertex to which half-edge s is incident and $s \in [d_i]$. For $1 \leq s < t \leq d_i$, let $I_{st,i}$ be the indicator of the event that half-edge (i, s) is paired to half-edge (i, t). Then the number of self-loops S_n can be computed as

$$S_n = \sum_{i\in[n]}\sum_{1\leq s<t\leq d_i} I_{st,i}. \tag{4.3.6}$$

Therefore,

$$\mathbb{E}[S_n] = \sum_{i\in[n]}\sum_{1\leq s<t\leq d_i}\mathbb{E}[I_{st,i}] = \sum_{i\in[n]}\sum_{1\leq s<t\leq d_i}\mathbb{P}[(i,s) \text{ is paired to } (i,t)]. \tag{4.3.7}$$

Since the probability of producing a self-loop by pairing the half-edges (i, s) and (i, t) does not depend on s and t, we get that $\mathbb{P}[(i, s) \text{ is paired to } (i, t)] = \frac{1}{\ell_n - 1}$. Thus,

$$\mathbb{E}[S_n] = \sum_{i\in[n]}\sum_{1\leq s<t\leq d_i}\frac{1}{\ell_n-1} = \sum_{i\in[n]}\frac{d_i(d_i-1)}{2}\frac{1}{\ell_n-1} = \frac{1}{2}\sum_{i\in[n]}\frac{d_i(d_i-1)}{\ell_n-1}, \tag{4.3.8}$$

as there are $\frac{d_i(d_i-1)}{2}$ possibilities to choose s and t from the set $[d_i]$ such that $s < t$.

Next, observe that

$$\frac{d_i - 1}{\ell_n - 1} = \frac{d_i(1 - \frac{1}{d_i})}{\ell_n(1 - \frac{1}{\ell_n})} \leq \frac{d_i(1 - \frac{1}{\ell_n})}{\ell_n(1 - \frac{1}{\ell_n})} = \frac{d_i}{\ell_n}, \tag{4.3.9}$$

because $d_i \leq \ell_n$ implies that $\frac{1}{d_i} \geq \frac{1}{\ell_n}$.

Therefore, we conclude that the expected number of self-loops can be bounded from above by

$$\mathbb{E}[S_n] = \frac{1}{2} \sum_{i \in [n]} \frac{d_i(d_i - 1)}{\ell_n - 1} \leq \frac{1}{2} \sum_{i \in [n]} \frac{d_i^2}{\ell_n}. \tag{4.3.10}$$

□

Proposition 4.10 (Bound on the expected number of removed multiple edges)**.** *([Hof18], Proposition 7.11) The expected number of removed multiple edges M_n in the configuration model* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfies*

$$\mathbb{E}[M_n] \leq \frac{\ell_n^2}{(\ell_n - 1)(\ell_n - 3)} \left(\sum_{i \in [n]} \frac{d_i(d_i - 1)}{2\ell_n} \right)^2 \leq 2 \left(\sum_{i \in [n]} \frac{d_i(d_i - 1)}{\ell_n} \right)^2. \tag{4.3.11}$$

Proof. Note that when $\ell_n = 3$, there cannot be any multiple edges, thus we assume that $\ell_n \geq 4$. For any vertices i and j and for $1 \leq s_1 < s_2 \leq d_i$ and $1 \leq t_1 \neq t_2 \leq d_j$, we define $I_{s_1 t_1, s_2 t_2, ij}$ to be the indicator function of the event that half-edge (i, s_1) is paired to half-edge (j, t_1) and half-edge (i, s_2) is paired to half-edge (j, t_2). Whenever there are multiple edges between vertices i and j, then $I_{s_1 t_1, s_2 t_2, ij} = 1$ for some s_1, t_1 and s_2, t_2.

In a next step, we can bound the number of removed multiple edges of vertex i from above by

$$\begin{aligned} m_i &= \sum_{j \in [n]: j \neq i} (x_{ij} - 1) \mathbb{1}_{\{x_{ij} \geq 2\}} \\ &\leq \sum_{j \in [n]: j \neq i} \left(\binom{x_{ij}}{2} \right) \mathbb{1}_{\{x_{ij} \geq 2\}} \\ &= \sum_{j \in [n]: j \neq i} \sum_{1 \leq s_1 < s_2 \leq d_i} \sum_{1 \leq t_1 \neq t_2 \leq d_j} I_{s_1 t_1, s_2 t_2, ij}. \end{aligned} \tag{4.3.12}$$

Since $M_n = \frac{1}{2} \sum_{i \in [n]} m_i$, it follows that

$$M_n \leq \frac{1}{2} \sum_{1 \leq i \neq j \leq n} \sum_{1 \leq s_1 < s_2 \leq d_i} \sum_{1 \leq t_1 \neq t_2 \leq d_j} I_{s_1 t_1, s_2 t_2, ij}. \tag{4.3.13}$$

Therefore,

$$\begin{aligned} \mathbb{E}[M_n] &\leq \frac{1}{2} \sum_{1 \leq i \neq j \leq n} \sum_{1 \leq s_1 < s_2 \leq d_i} \sum_{1 \leq t_1 \neq t_2 \leq d_j} \mathbb{E}[I_{s_1 t_1, s_2 t_2, ij}] \\ &= \frac{1}{2} \sum_{1 \leq i \neq j \leq n} \sum_{1 \leq s_1 < s_2 \leq d_i} \sum_{1 \leq t_1 \neq t_2 \leq d_j} \mathbb{E}[I_{11,22,ij}], \end{aligned} \tag{4.3.14}$$

as the probability of producing multiple edges between the half-edges (i, s_1) and (j, t_1) as well as between (i, s_2) and (j, t_2) is independent of s_1, s_2, t_1, t_2.

Since $\sum_{1\leq s_1<s_2\leq d_i} 1 = \frac{d_i(d_i-1)}{2}$ and $\sum_{1\leq t_1\neq t_2\leq d_j} 1 = d_j(d_j-1)$, we get that

$$\mathbb{E}[M_n] \leq \frac{1}{4} \sum_{1\leq i\neq j\leq n} d_i(d_i-1)d_j(d_j-1)\mathbb{E}[I_{11,22,ij}]. \tag{4.3.15}$$

As $I_{11,22,ij}$ is an indicator function, $\mathbb{E}[I_{11,22,ij}] = \mathbb{P}[I_{11,22,ij} = 1]$. The probability that half-edge 1 of vertex i is paired to half-edge 1 of vertex j is $\frac{1}{\ell_n-1}$. When we remove the two paired half-edges from the list of all half-edges, there are only $\ell_n - 2$ half-edges left. Thus, the probability to pair half-edge 2 of vertex i with half-edge 2 of vertex j is then $\frac{1}{\ell_n-3}$.

Hence,

$$\mathbb{E}[I_{11,22,ij}] = \frac{1}{(\ell_n-1)(\ell_n-3)}. \tag{4.3.16}$$

Thus,

$$\mathbb{E}[M_n] \leq \sum_{i,j\in[n]} \frac{d_i(d_i-1)d_j(d_j-1)}{4(\ell_n-1)(\ell_n-3)} = \frac{\ell_n^2}{(\ell_n-1)(\ell_n-3)} \left(\sum_{i\in[n]} \frac{d_i(d_i-1)}{2\ell_n}\right)^2. \tag{4.3.17}$$

In a next step, we elucidate that $8(\ell_n-1)(\ell_n-3) \geq \ell_n^2$ whenever $\ell_n \geq 4$. In order to do so, let us define two functions $f(x) = 8(x-1)(x-3)$ and $g(x) = x^2$. First, note that $f(x) = g(x)$ for $x = \frac{16\pm 2\sqrt{22}}{7}$, where $x = \frac{16+2\sqrt{22}}{7} \approx 3.6$. Moreover, $f'(x) = 16x - 32$ and $g'(x) = 2x$. Since $f'(x) \geq g'(x)$ if and only if $x \geq \frac{16}{7} \approx 2.3$, we can deduce that the function f grows faster than g for any $x \geq \frac{16}{7}$. Thus, as $f(x) = g(x)$ for $x \approx 3.6$ and $f'(x) \geq g'(x)$ for any $x \geq 3.6$, it follows that clearly $f(x) \geq g(x)$ for any $x \geq 4$.

From the fact that $8(\ell_n-1)(\ell_n-3) \geq \ell_n^2$ whenever $\ell_n \geq 4$, it follows that

$$\mathbb{E}[M_n] \leq 2\left(\frac{\sum_{i\in[n]} d_i(d_i-1))}{\ell_n}\right)^2, \tag{4.3.18}$$

which concludes the proof. □

At last, we can state the main result of this subsection showing that under some assumptions the distribution of the degree sequence of the erased configuration model is close to the distribution of the prescribed degree sequence when n is sufficiently large. This shows us, that we can use the erased configuration model for huge graphs to get a simple graph with a degree sequence that is quite similar to the one we want to get.

Theorem 4.11 (Distribution of the degree sequence of the erased configuration model with fixed degrees)**.** *([Hof18], Theorem 7.10)*
For a fixed degree sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$ *satisfying Conditions 4.4(a)-(b), the distribution of the degree sequence of the erased configuration model* $(P_k^{(er)})_{k\geq 1}$ *converges in probability to* $(p_k)_{k\geq 1}$*. More precisely, for every* $\varepsilon > 0$*,*

$$\mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k\right| \geq \varepsilon\right] \to 0, \tag{4.3.19}$$

where $p_k = \mathbb{P}[D = k]$ *as in Condition 4.4(a).*

Proof. Let $p_k^{(n)} = \mathbb{P}[D_n = k]$. By Condition 4.4(a), we get that $\lim_{n\to\infty} p_k^{(n)} = p_k$ for any $k \in [n]$. By the fact that pointwise convergence of a probability mass function implies convergence in total variation distance (recall Proposition 2.17), we get that

$$\lim_{n\to\infty}\sum_{k=1}^{\infty}\left|p_k^{(n)} - p_k\right| = 0. \tag{4.3.20}$$

Therefore, for any $\varepsilon > 0$, we can take n so large that

$$\sum_{k=1}^{\infty}\left|p_k^{(n)} - p_k\right| \leq \frac{\varepsilon}{2}. \tag{4.3.21}$$

Our goal is to show that

$$\mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| < \frac{\varepsilon}{2}\right] \to 0. \tag{4.3.22}$$

We start by proving the result under the extra assumption that

$$\max_{i\in[n]} d_i = o\big(\sqrt{n}\big). \tag{4.3.23}$$

To do so, we have to bound $\mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| \geq \frac{\varepsilon}{2}\right]$. Note that

$$\begin{aligned}\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| &= \sum_{k=1}^{\infty}\left|\frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{D_i^{(er)}=k\}} - \frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{d_i=k\}}\right| \\ &\leq \frac{1}{n}\sum_{k=1}^{\infty}\sum_{i\in[n]}\left|\mathbb{1}_{\{D_i^{(er)}=k\}} - \mathbb{1}_{\{d_i=k\}}\right|.\end{aligned} \tag{4.3.24}$$

Since $D_i^{(er)} = d_i - 2s_i - m_i$ implies that $D_i^{(er)} \neq d_i$ if and only if $s_i + m_1 > 0$, it follows that

$$\mathbb{1}_{\{D_i^{(er)}=k\}} - \mathbb{1}_{\{d_i=k\}} = \mathbb{1}_{\{s_i+m_i>0\}}\left(\mathbb{1}_{\{D_i^{(er)}=k\}} - \mathbb{1}_{\{d_i=k\}}\right), \tag{4.3.25}$$

Thus,

$$\left|\mathbb{1}_{\{D_i^{(er)}=k\}} - \mathbb{1}_{\{d_i=k\}}\right| \leq \mathbb{1}_{\{s_i+m_i>0\}}\left(\mathbb{1}_{\{D_i^{(er)}=k\}} + \mathbb{1}_{\{d_i=k\}}\right), \tag{4.3.26}$$

so that

$$\begin{aligned}\sum_{k=1}^{\infty}\left|P_k^{(er)}-p_k^{(n)}\right| &\le \frac{1}{n}\sum_{k=1}^{\infty}\sum_{i\in[n]}\left|\mathbb{1}_{\{D_i^{(er)}=k\}}-\mathbb{1}_{\{d_i=k\}}\right| \\ &\le \frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{s_i+m_i>0\}}\sum_{k=1}^{\infty}\left(\mathbb{1}_{\{D_i^{(er)}=k\}}+\mathbb{1}_{\{d_i=k\}}\right) \\ &= \frac{2}{n}\sum_{i\in[n]}\mathbb{1}_{\{s_i+m_i>0\}} \\ &\le \frac{2}{n}\sum_{i\in[n]}(s_i+m_i) \\ &= \frac{2}{n}S_n+\frac{4}{n}M_n,\end{aligned} \tag{4.3.27}$$

where S_n is the number of self-loops and M_n is the number of removed multiple edges of $\mathrm{CM}_n(\boldsymbol{d})$.

Then, we obtain that

$$\mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)}-p_k^{(n)}\right|\ge\frac{\varepsilon}{2}\right]\le\mathbb{P}\left[2S_n+4M_n\ge\frac{\varepsilon n}{2}\right], \tag{4.3.28}$$

so that Theorem 4.11 follows if

$$\mathbb{P}\left[2S_n+4M_n\ge\frac{\varepsilon n}{2}\right]\to 0. \tag{4.3.29}$$

By Markov's inequality 2.1, we get

$$\mathbb{P}\left[2S_n+4M_n\ge\frac{\varepsilon n}{2}\right]\le\frac{4}{\varepsilon n}\left(\mathbb{E}[S_n]+2\mathbb{E}[M_n]\right). \tag{4.3.30}$$

By Proposition 4.9 and under the assumption that $\max_{i\in[n]}d_i=o\big(\sqrt{n}\big)$, we can deduce that

$$\mathbb{E}[S_n]\le\frac{1}{2}\sum_{i\in[n]}\frac{d_i^2}{\ell_n}\le\frac{\max_{i\in[n]}d_i\sum_{i\in[n]}d_i}{\ell_n}=\max_{i\in[n]}d_i=o\big(\sqrt{n}\big). \tag{4.3.31}$$

Similarly, by Proposition 4.10 and again under the assumption that $\max_{i\in[n]}d_i=o\big(\sqrt{n}\big)$, we obtain that

$$\mathbb{E}[M_n]\le 2\left(\sum_{i\in[n]}\frac{d_i(d_i-1)}{\ell_n}\right)^2\le 2\left(\frac{\max_{i\in[n]}d_i\sum_{i\in[n]}d_i}{\ell_n}\right)^2=o(n). \tag{4.3.32}$$

Hence, under the assumption that $\max_{i\in[n]}d_i=o\big(\sqrt{n}\big)$, we can summarise that

$$\begin{aligned}\mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)}-p_k^{(n)}\right|\ge\frac{\varepsilon}{2}\right] &\le \mathbb{P}\left[2S_n+4M_n\ge\frac{\varepsilon n}{2}\right] \\ &\le \frac{4}{\varepsilon n}\left(\mathbb{E}[S_n]+2\mathbb{E}[M_n]\right) \\ &= \frac{4}{\varepsilon n}o(n)=o(1).\end{aligned} \tag{4.3.33}$$

Therefore, we can conclude that for every $\varepsilon > 0$

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k\right| \geq \varepsilon\right] &\leq \lim_{n\to\infty} \mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| + \sum_{k=1}^{\infty}\left|p_k^{(n)} - p_k\right| \geq \varepsilon\right] \\ &\leq \lim_{n\to\infty} \mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| + \frac{\varepsilon}{2} \geq \varepsilon\right] \\ &= \lim_{n\to\infty} \mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| \geq \frac{\varepsilon}{2}\right] = 0.\end{aligned} \tag{4.3.34}$$

To prove the result assuming only Conditions 4.4 (a)-(b), we start by noting that these conditions imply that $\max_{i\in[n]} d_i = o(n)$. The proof of this statement is similar to the proof of Proposition 3.4. Moreover, notice that $\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| \geq \frac{\varepsilon}{2}$ if and only if $\sum_{k=1}^{\infty}\left|\sum_{i\in[n]} \mathbb{1}_{\{D_i^{(er)}=k\}} - \sum_{i\in[n]} \mathbb{1}_{\{d_i=k\}}\right| \geq \frac{\varepsilon n}{2}$, which implies that the degrees of at least $\frac{\varepsilon n}{2}$ vertices are changed by the erasure process.

Let us now introduce a real-valued sequence $a_n \to \infty$, which we determine more precisely later on. Then, as $\max_{i\in[n]} d_i = o(n)$, it is possible to choose n large enough such that there are at most $\frac{\varepsilon n}{4}$ vertices $i \in [n]$ of degree $d_i \geq a_n$. In that case, $\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| \geq \frac{\varepsilon}{2}$ implies that there must be at least $\frac{\varepsilon n}{4}$ vertices of degree at most a_n, whose degrees are changed by the erasure procedure.

Next, let us set

$$S_n(a_n) = \sum_{i\in[n]} s_i \mathbb{1}_{\{d_i\leq a_n\}}, \tag{4.3.35}$$

$$M_n(a_n) = \frac{1}{2}\sum_{i\in[n]} m_i \mathbb{1}_{\{d_i\leq a_n\}}, \tag{4.3.36}$$

to be the number of self-loops and removed multiple edges incident to those vertices with degree at most a_n. Hence, by the above reasoning and the use of Markov's inequality 2.1, we obtain

$$\mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| \geq \frac{\varepsilon}{2}\right] \leq \mathbb{P}\left[2S_n(a_n) + 4M_n(a_n) \geq \frac{\varepsilon n}{4}\right] \leq \frac{8}{\varepsilon n}\left(\mathbb{E}[S_n(a_n)] + 2\mathbb{E}[M_n(a_n)]\right). \tag{4.3.37}$$

Small adaptations of Propositions 4.9 and 4.10 yield that

$$\mathbb{E}[S_n(a_n)] \leq \frac{1}{2}\sum_{i\in[n]} \frac{d_i^2 \mathbb{1}_{\{d_i\leq a_n\}}}{\ell_n}, \tag{4.3.38}$$

$$\mathbb{E}[M_n(a_n)] \leq 2\sum_{i\in[n]} \frac{d_i^2 \mathbb{1}_{\{d_i\leq a_n\}}}{\ell_n} \sum_{j\in[n]} \frac{d_j^2}{\ell_n}. \tag{4.3.39}$$

Then, we get that

$$\mathbb{E}[S_n(a_n)] \leq a_n \sum_{i\in[n]} \frac{d_i}{\ell_n} = a_n, \tag{4.3.40}$$

and

$$\mathbb{E}[M_n(a_n)] \leq 2a_n \sum_{i\in[n]} \frac{d_i}{\ell_n} \cdot \max_{j\in[n]} d_j \sum_{j\in[n]} \frac{d_j}{\ell_n} = a_n \cdot \max_{j\in[n]} d_j. \tag{4.3.41}$$

Since $\max_{j\in[n]} d_j = o(n)$ under the Conditions 3.3(a)-(b), we can take $a_n \to \infty$ so slowly that $a_n \cdot \max_{j\in[n]} d_j = o(n)$.

Therefore,

$$\mathbb{P}\left[\sum_{k=1}^{\infty}\left|P_k^{(er)} - p_k^{(n)}\right| \geq \frac{\varepsilon}{2}\right] \leq \frac{8}{\varepsilon n}\left(\mathbb{E}[S_n(a_n)] + 2\mathbb{E}[M_n(a_n)]\right) = \frac{8}{\varepsilon n}o(n) = o(1), \tag{4.3.42}$$

which implies the desired result (similar as under the assumption that $\max_{i\in[n]} d_i = o\big(\sqrt{n}\big)$). □

4.4 Repeated Configuration Model: Simplicity Probability

Another way to get a simple graph using the configuration model is to repeat the generation process several times until you get a simple graph. This we call the repeated configuration model. One might wonder, however, how likely it is to get a simple graph. Therefore, we investigate the probability that the configuration model yields a simple graph in this section.

To get to the main result, we first need to show that the number of self-loops S_n and the number of additional multiple edges M_n converge to two independent Poisson random variables with means $\frac{\nu}{2}$ and $\frac{\nu^2}{4}$, respectively, where $\nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}$. To prove that statement, it turns out to be simpler to work with

$$\widetilde{M_n} = \sum_{1\leq i\neq j\leq n}\sum_{1\leq s_1<s_2\leq d_i}\sum_{1\leq t_1\neq t_2\leq d_j} I_{s_1t_1,s_2t_2,ij}, \tag{4.4.1}$$

such that $M_n \leq \widetilde{M_n}$ as seen in (4.4). As a result, $\widetilde{M_n} = 0$ implies that $M_n = 0$. It turns out that $M_n = \widetilde{M_n}$ occurs with high probability when Conditions 4.4 (a)-(c) hold, which is declared in the following lemma.

Lemma 4.12 ($M_n = \widetilde{M_n}$ occurs with high probability)**.** *([Hof18], Exercise 7.11)*
Assume that $\boldsymbol{d} = (d_i)_{i\in[n]}$ *satisfies Conditions 4.4 (a)-(c), then* $\mathbb{P}[M_n < \widetilde{M_n}] = o(1)$ *as* $n \to \infty$.

Proof. First of all, we note that $M_n \neq \widetilde{M_n}$ if and only if there exist at least two vertices i and j with three multiple edges or more.

In the next step, we try to bound the probability that there are at least three edges between vertices i and j from above. The probability that the first half-edge of vertex i pairs with any of the d_j half-edges of vertex j is $\frac{d_j}{\ell_n-1}$, where ℓ_n is the sum of all degrees. The probability that the second half-edge of vertex i pairs with any of the remaining half-edges of vertex j is then $\frac{d_j-1}{\ell_n-3}$ and the probability that the third half-edge of vertex i pairs with any of the remaining half-edges of vertex j is then $\frac{d_j-2}{\ell_n-5}$. However, it does not have to be specifically the initial three half-edges of vertex i that are matched with the half-edges of vertex j, rather, it could be any three half-edges of vertex i. There are $\binom{d_i}{3}$ possibilities to choose 3 of the d_i half-edges incident to vertex i. Thus, the probability that there are at least three edges between i and j can be bounded from above by

$$\mathbb{P}[x_{ij} \geq 3] = \binom{d_i}{3}\frac{d_j(d_j-1)(d_j-2)}{(\ell_n-1)(\ell_n-3)(\ell_n-5)} \leq \frac{d_i(d_i-1)(d_i-2)d_j(d_j-1)(d_j-2)}{(\ell_n-1)(\ell_n-3)(\ell_n-5)}. \tag{4.4.2}$$

By Boole's inequality 2.3, the probability that there exist vertices $i \neq j$ such that there are at least three edges between i and j is then bounded from above by

$$\mathbb{P}[\exists i,j \in [n] : x_{ij} \geq 3] \leq \sum_{i,j\in[n]} \frac{d_i(d_i-1)(d_i-2)d_j(d_j-1)(d_j-2)}{(\ell_n-1)(\ell_n-3)(\ell_n-5)}. \tag{4.4.3}$$

Applying Proposition 3.4 to the weights $\boldsymbol{w} = \boldsymbol{d}$ under the Conditions 4.4 (a)-(c), we get that $d_i = o\big(\sqrt{n}\big)$ as $n \to \infty$. Since $\ell_n \geq n$, as all vertices with degree 0 can be removed from the graph, we can conclude that

$$\mathbb{P}[M_n < \widetilde{M_n}] = \mathbb{P}[\exists i,j \in [n] : x_{ij} \geq 3] \leq \sum_{i,j\in[n]} \frac{d_i(d_i-1)(d_i-2)d_j(d_j-1)(d_j-2)}{(\ell_n-1)(\ell_n-3)(\ell_n-5)} = o(1), \tag{4.4.4}$$

as $n \to \infty$. □

To prove that S_n and M_n converge to two independent Poisson random variables, we also need the following result.

Lemma 4.13 (Characterisation moments of independent Poisson random variables)**.** *([Hof18], Exercise 7.14)*

The joint moments of (X,Y), where X and Y are independent Poisson random variables with parameters μ_X and μ_Y, are identified by the following two relations:

for $r \geq 1$

$$\mathbb{E}[X^r] = \mu_X \mathbb{E}[(X+1)^{r-1}], \tag{4.4.5}$$

and, for $r,s \geq 1$

$$\mathbb{E}[X^r Y^s] = \mu_Y \mathbb{E}[X^r (Y+1)^{s-1}]. \tag{4.4.6}$$

Proof. We start by observing that, for any $r \geq 1$,

$$\begin{aligned}
\mu_X \mathbb{E}\left[(X+1)^{r-1}\right] &= \mu_X \sum_{k=0}^{\infty} (k+1)^{r-1} \frac{e^{-\mu_X}\mu_X^k}{k!} \\
&= \sum_{k=0}^{\infty} (k+1)^r \frac{e^{-\mu_X}\mu_X^{k+1}}{(k+1)!} \\
&= \sum_{n=1}^{\infty} n^r \frac{e^{-\mu_X}\mu_X^n}{n!} + 0^r \frac{e^{-\mu_X}\mu_X^0}{0!} \\
&= \sum_{m=0}^{\infty} m^r \frac{e^{-\mu_X}\mu_X^m}{m!} \\
&= \mathbb{E}[X^r].
\end{aligned} \tag{4.4.7}$$

Using the independence of X and Y and above result, we obtain for any $r, s \geq 1$

$$\begin{aligned}\mathbb{E}[X^r Y^s] &= \mathbb{E}[X^r] \cdot \mathbb{E}[Y^s] \\ &= \mathbb{E}[X^r] \cdot \mu_Y \mathbb{E}\left[(Y+1)^{s-1}\right] \\ &= \mu_Y \mathbb{E}\left[X^r (Y+1)^{s-1}\right].\end{aligned} \tag{4.4.8}$$

□

Now, we have all the necessary tools to prove the Poisson limit of self-loops and multiple edges.

Proposition 4.14 (Poisson limit of self-loops and multiple edges)**.** *([Hof18], Proposition 7.13) Assume that $\boldsymbol{d} = (d_i)_{i\in[n]}$ satisfies Conditions 4.4 (a)-(c). Then (S_n, M_n) converges in distribution to (S, M), where S and M are two independent Poisson random variables with means $\frac{\nu}{2}$ and $\frac{\nu^2}{4}$, respectively, where $\nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}$.*

Proof. First of all, we note that, as S and M are by assumption independent Poisson random variables, both their moment-generating functions $M_S(t)$ and $M_M(t)$ are finite for t in some neighbourhood of the origin. Therefore, we can use a two-dimensional extension of Theorem 2.6, stating that when the joint moments $\mathbb{E}\left[S_n^r M_n^s\right]$ converge to the moments $\mathbb{E}[S^r M^s]$ for each $r, s = 0, 1, 2, \ldots$, and when the moment-generating functions are finite in some neighbourhood of the origin, then (S_n, M_n) converges in distribution to (S, M). To prove that the joint moments $\mathbb{E}[S_n^r M_n^s]$ converge to the moments $\mathbb{E}[S^r M^s]$, we make use of induction in both r and s. Therefore, we need to prove that

$$\lim_{n\to\infty} \mathbb{E}[S_n^r] = \mathbb{E}[S^r] = \mu_S \mathbb{E}\left[(S+1)^{r-1}\right], \tag{4.4.9}$$

and

$$\lim_{n\to\infty} \mathbb{E}[S_n^r M_n^s] = \mathbb{E}[S^r M^s] = \mu_M \mathbb{E}\left[S^r (M+1)^{s-1}\right], \tag{4.4.10}$$

where the second equality in (4.4.9) and (4.4.10) follow from Lemma 4.13.

Recall that

$$S_n = \sum_{i=1}^{n} \sum_{1\leq a<b\leq d_i} I_{ab,i}, \tag{4.4.11}$$

where $I_{ab,i}$ is the indicator function of the event that half-edge (i, a) is paired to the half-edge (i, b) to form a self-loop. Let us prove (4.4.9) by induction on r first. For $r = 0$, the statement is trivial, which initialises the induction hypothesis. The induction hypothesis is that (4.4.9) is true for any $k \leq r-1$, for $\mathrm{CM}_n(\boldsymbol{d})$ when $n \to \infty$ and for all $(d_i)_{i\in[n]}$ satisfying Conditions 4.4 (a)-(c).

To advance the induction hypothesis, we observe that

$$\begin{aligned}\mathbb{E}[S_n^r] &= \mathbb{E}\left[\sum_{i=1}^{n}\sum_{1\le a<b\le d_i} I_{ab,i}S_n^{r-1}\right]\\ &= \sum_{i=1}^{n}\sum_{1\le a<b\le d_i}\mathbb{E}\left[I_{ab,i}S_n^{r-1}\right]\\ &= \sum_{i=1}^{n}\sum_{1\le a<b\le d_i}\mathbb{P}[I_{ab,i}=1]\mathbb{E}\left[S_n^{r-1}|I_{ab,i}=1\right],\end{aligned} \tag{4.4.12}$$

by the linearity of the expected value and the law of total probability.

When $I_{ab,i}=1$, then the remaining half-edges of vertex i need to be paired in a uniform manner. The number of self-loops in the total graph in this pairing has the same distribution as $1+S_n'$, where S_n' is the number of self-loops in the configuration model with degrees $(d_i')_{i\in[n]}$, where $d_i'=d_i-2$ and $d_j'=d_j$ for all $j\neq i$. By construction, the degrees $(d_i')_{i\in[n]}$ still satisfy Conditions 4.4 (a)-(c), which means that by the induction hypothesis it holds for any $k\le r-1$

$$\lim_{n\to\infty}\mathbb{E}\left[(S_n')^k\right]=\mathbb{E}\left[S^k\right]. \tag{4.4.13}$$

Therefore, we can deduce that

$$\lim_{n\to\infty}\mathbb{E}\left[S_n^{r-1}|I_{ab,i}=1\right]=\lim_{n\to\infty}\mathbb{E}\left[(1+S_n')^{r-1}\right]=\mathbb{E}\left[(1+S)^{r-1}\right]. \tag{4.4.14}$$

As the limit does not depend on i, we get that

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\left[S_n^r\right] &= \mathbb{E}\left[(1+S)^{r-1}\right]\lim_{n\to\infty}\sum_{i=1}^{n}\sum_{1\le a<b\le d_i}\mathbb{P}[I_{ab,i}=1]\\ &= \mathbb{E}\left[(1+S)^{r-1}\right]\lim_{n\to\infty}\sum_{i=1}^{n}\sum_{1\le a<b\le d_i}\frac{1}{\ell_n-1}\\ &= \mathbb{E}\left[(1+S)^{r-1}\right]\lim_{n\to\infty}\sum_{i=1}^{n}\frac{d_i(d_i-1)}{2}\frac{1}{\ell_n-1}\\ &= \frac{1}{2}\mathbb{E}\left[(1+S)^{r-1}\right]\lim_{n\to\infty}\frac{\frac{\sum_{i=1}^{n}d_i(d_i-1)}{n}}{\frac{\ell_n-1}{n}}\\ &= \frac{1}{2}\mathbb{E}\left[(1+S)^{r-1}\right]\frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}\\ &= \frac{\nu}{2}\mathbb{E}\left[(1+S)^{r-1}\right]\\ &= \mathbb{E}\left[S^r\right].\end{aligned} \tag{4.4.15}$$

This advances the induction hypothesis and completes the proof of (4.4.9).

To prove (4.4.10), we perform a similar induction scheme.

Recall that

$$M_n=\sum_{1\le i<j\le n}(x_{ij}-1)\mathbb{1}_{\{x_{ij}\ge 2\}}, \tag{4.4.16}$$

while

$$\widetilde{M_n} = \sum_{1\le i<j\le n} \sum_{1\le s_1<s_2\le d_i} \sum_{1\le t_1\neq t_2\le d_j} I_{s_1t_1,s_2t_2,ij}. \tag{4.4.17}$$

By Lemma 4.12, it is enough to show that, for all $r \geq 0$, $\mathbb{E}[S_n^r \tilde{M}_n^s]$ converges to $\mathbb{E}[S^r M^s]$ using induction on s. The base case $s = 0$ follows from (4.4.9), which initialises the induction hypothesis, claiming that (4.4.10) is true for any $k \leq s-1$, for $\mathrm{CM}_n(\boldsymbol{d})$ when $n \to \infty$ and for all $(d_i)_{i\in[n]}$ satisfying Conditions 4.4 (a)-(c). To advance the induction hypothesis, we proceed similarly as for S_n above.

First, we note that

$$\begin{aligned}
\mathbb{E}\left[S_n^r \widetilde{M}_n^s\right] &= \mathbb{E}\left[S_n^r \sum_{1\le i<j\le n} \sum_{1\le s_1<s_2\le d_i} \sum_{1\le t_1\neq t_2\le d_j} I_{s_1t_1,s_2t_2,ij} \widetilde{M}_n^{s-1}\right] \\
&= \sum_{1\le i<j\le n} \sum_{1\le s_1<s_2\le d_i} \sum_{1\le t_1\neq t_2\le d_j} \mathbb{E}\left[I_{s_1t_1,s_2t_2,ij} S_n^r \widetilde{M}_n^{s-1}\right] \\
&= \sum_{1\le i<j\le n} \sum_{1\le s_1<s_2\le d_i} \sum_{1\le t_1\neq t_2\le d_j} \mathbb{P}[I_{s_1t_1,s_2t_2,ij} = 1]\mathbb{E}\left[S_n^r \widetilde{M}_n^{s-1} | I_{s_1t_1,s_2t_2,ij} = 1\right].
\end{aligned} \tag{4.4.18}$$

When $I_{s_1t_1,s_2t_2,ij} = 1$, we know that two half-edges of vertex i are paired to two half-edges of vertex j to form a multiple edge. Removing these two edges leaves us with a graph that is close to the configuration model with degrees $(d_i')_{i\in[n]}$, where $d_i' = d_i - 2$, $d_j' = d_j - 2$ and $d_t' = d_t$ for all $t \neq i, j$. The only difference is that when a half-edge connected to i is attached to a half-edge connected to j, then this increases the value of $\widetilde{M}_n$ by more than just one. Thus, if we denote $\widetilde{M}_n' = \sum_{1\le i<j\le n} \sum_{1\le s_1<s_2\le d_i'} \sum_{1\le t_1\neq t_2\le d_j'} I_{s_1t_1,s_2t_2,ij}$, it must hold that $\widetilde{M_n} \geq \widetilde{M}_n' + 1$.

However, as can be seen in the proof of Lemma 4.12,

$$\mathbb{P}[x_{ij} \geq 3] = o(1), \tag{4.4.19}$$

as $n \to \infty$, which implies that $\widetilde{M_n} = \widetilde{M}_n' + 1$ with high probability.

Hence, we can deduce that

$$\lim_{n\to\infty} \mathbb{E}\left[S_n^r \widetilde{M}_n^{s-1} | I_{s_1t_1,s_2t_2,ij} = 1\right] = \lim_{n\to\infty} \mathbb{E}\left[S_n^r (1 + \widetilde{M}_n')^{s-1}\right] = \mathbb{E}\left[S^r (1+M)^{s-1}\right], \tag{4.4.20}$$

where we used the induction hypothesis that $\lim_{n\to\infty} \mathbb{E}\left[S_n^r (\widetilde{M}_n')^k\right] = \mathbb{E}\left[S^r M^k\right]$ when $k \leq s-1$.

In the next step, let us set

$$\Sigma := \sum_{1\le i<j\le n} \sum_{1\le s_1<s_2\le d_i} \sum_{1\le t_1\neq t_2\le d_j} \mathbb{P}\left[I_{s_1t_1,s_2t_2,ij} = 1\right], \tag{4.4.21}$$

and observe that

$$\begin{aligned}
\Sigma &= \frac{1}{2} \sum_{1\le i\neq j\le n} \sum_{1\le s_1<s_2\le d_i} \sum_{1\le t_1\neq t_2\le d_j} \mathbb{P}[I_{s_1t_1,s_2t_2,ij}=1] \\
&= \frac{1}{2} \sum_{1\le i\neq j\le n} \sum_{1\le s_1<s_2\le d_i} \sum_{1\le t_1\neq t_2\le d_j} \frac{1}{(\ell_n-1)(\ell_n-3)} \\
&= \frac{1}{2(\ell_n-1)(\ell_n-3)} \sum_{1\le i\neq j\le n} \sum_{1\le s_1<s_2\le d_i} d_j(d_j-1) \\
&= \frac{1}{2(\ell_n-1)(\ell_n-3)} \sum_{1\le i\neq j\le n} \frac{d_i(d_i-1)}{2} d_j(d_j-1) \\
&= \frac{1}{4(\ell_n-1)(\ell_n-3)} \left(\sum_{1\le i,j\le n} d_i(d_i-1)d_j(d_j-1) - \sum_{1\le i\le n} d_i^2(d_i-1)^2 \right) \\
&= \frac{1}{4\frac{\ell_n-1}{n}\frac{\ell_n-3}{n}} \left(\frac{1}{n^2} \left(\sum_{1\le i\le n} d_i(d_i-1) \right)^2 - \frac{1}{n^2} \sum_{1\le i\le n} d_i^2(d_i-1)^2 \right) \\
&= \frac{1}{4(\mathbb{E}[D_n]-\frac{1}{n})(\mathbb{E}[D_n]-\frac{3}{n})} \left(\mathbb{E}[D_n(D_n-1)]^2 - \frac{1}{n^2} \sum_{1\le i\le n} d_i^2(d_i-1)^2 \right).
\end{aligned} \tag{4.4.22}$$

As under Conditions 4.4 (a)-(c) $d_i = o(\sqrt{n})$ (using Proposition 3.4 with $\boldsymbol{w} = \boldsymbol{d}$), we can deduce that $\frac{1}{n^2}\sum_{1\le i\le n} d_i^2(d_i-1)^2 = o(1)$. Therefore,

$$\lim_{n\to\infty} \sum_{1\le i<j\le n} \sum_{1\le s_1<s_2\le d_i} \sum_{1\le t_1\neq t_2\le d_j} \mathbb{P}[I_{s_1t_1,s_2t_2,ij}=1] = \frac{1}{4}\frac{\mathbb{E}[D(D-1)]^2}{\mathbb{E}[D]^2} = \frac{\nu^2}{4}. \tag{4.4.23}$$

Putting all the pieces together, we get that

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[S_n^r \widetilde{M}_n^s\right] &= \lim_{n\to\infty} \sum_{i<j} \sum_{s_1<s_2} \sum_{t_1\neq t_2} \mathbb{P}[I_{s_1t_1,s_2t_2,ij}=1] \mathbb{E}\left[S_n^r \widetilde{M}_n^{s-1} | I_{s_1t_1,s_2t_2,ij}=1\right] \\
&= \frac{\nu^2}{4} \mathbb{E}\left[S^r (1+M)^{s-1}\right] \\
&= \mathbb{E}\left[S^r M^s\right],
\end{aligned} \tag{4.4.24}$$

and hence it follows also by Lemma 4.12 that $\lim_{n\to\infty} \mathbb{E}\left[S_n^r M_n^s\right] = \mathbb{E}\left[S^r M^s\right]$. Therefore, we can conclude that (S_n, M_n) converges in distribution to (S, M), as desired. □

Finally, we can turn to the main theorem. Due to it, we know that when the prescribed degree sequence satisfies Conditions 4.4 (a)-(c), the probability of getting a simple graph is asymptotically positive.

Theorem 4.15 (Probability of simplicity of $\mathrm{CM}_n(\boldsymbol{d})$)**.** *([Hof18], Theorem 7.12)*
Assume that $\boldsymbol{d} = (d_i)_{i\in n}$ satisfies Conditions 4.4 (a)-(c). Then, the probability that $\mathrm{CM}_n(\boldsymbol{d})$ is a simple graph is asymptotically equal to

$$e^{-\frac{\nu}{2}-\frac{\nu^2}{4}}, \tag{4.4.25}$$

where

$$\nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}. \tag{4.4.26}$$

Proof. First, we have to realise that $\mathrm{CM}_n(\boldsymbol{d})$ is simple if and only if both $S_n = 0$ and $M_n = 0$, thus $\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}] = \mathbb{P}[S_n = M_n = 0]$. By the weak convergence result of Proposition 4.14 and the independence of S and M, where S and M are Poisson random variables with means $\frac{\nu}{2}$ and $\frac{\nu^2}{4}$, respectively, we get that

$$\lim_{n\to\infty} \mathbb{P}[S_n = M_n = 0] = \mathbb{P}[S = M = 0] = \mathbb{P}[S = 0]\mathbb{P}[M = 0] = e^{-\frac{\nu}{2}} e^{-\frac{\nu^2}{4}} = e^{-\frac{\nu}{2}-\frac{\nu^2}{4}}. \tag{4.4.27}$$

□

According to Theorem 4.15, we can use the configuration model repeatedly until we get a simple random graph with the prescribed degree sequence. However, it is essential that all three Conditions 4.4 (a)-(c) are met. It can be shown by an adaptation of the proof of Proposition 4.14 that when only Conditions 4.4 (a)-(b) are satisfied and when $\mathbb{E}[D^2] = \infty$, the probability of simplicity converges to 0. Those readers interested in the proof of that result are referred to [Hof18], Proposition 7.14.

4.5 Configuration Model, Uniform Simple Graphs and Generalised Random Graphs

In this last section of Chapter 4, we delve into exploring the connections among the configuration model, uniform simple random graphs having prescribed degrees, and the generalised random graph with given weights. These findings are well-known within the random graph community and allow us to use the configuration model to prove results for other models. We rely on the fact that $\mathrm{CM}_n(\boldsymbol{d})$ is simple with asymptotically positive probability whenever the chosen degree sequence $(d_i)_{i\in[n]}$ satisfies Conditions 4.4 (a)-(c), as seen in the subchapter before.

The first proposition deals with the relationships between the configuration model and uniform simple random graphs with a certain degree sequence. According to it, the configuration model conditioned on being simple can be used to generate uniform simple random graphs with a given degree sequence.

Proposition 4.16 (Uniform graphs with given degree sequence)**.** *([Hof18], Proposition 7.15)*
For any degree sequence $(d_i)_{i\in[n]}$, and conditionally on the event $\{\mathrm{CM}_n(\boldsymbol{d})$ is a simple graph$\}$, $\mathrm{CM}_n(\boldsymbol{d})$ is a uniform simple graph with degree sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$.

Proof. By Proposition 4.2, we know that for any multigraph $G = (x_{ij})_{i,j\in[n]}$ it holds that

$$\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G] = \frac{1}{(\ell_n - 1)!!} \frac{\prod_{i\in[n]} d_i!}{\prod_{i\in[n]} 2^{x_{ii}} \prod_{1\le i\le j\le n} x_{ij}!}. \tag{4.5.1}$$

When G is simple, it holds that for any $i, j \in [n]$ $x_{ii} = 0$ and $x_{ij} = 0$.

Therefore, for any simple graph G we obtain

$$\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G] = \frac{\prod_{i\in[n]} d_i!}{(\ell_n - 1)!!}. \tag{4.5.2}$$

Thus, for any simple graph G with a prescribed degree sequence $(d_i)_{i\in[n]}$, $\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G]$ is the same. This implies that also $\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G|\mathrm{CM}_n(\boldsymbol{d})$ is simple] is identical for every simple random graph G with the given degree sequence. □

An important consequence of Proposition 4.16 and Theorem 4.15 is that they enable us to compute the asymptotic number of simple graphs with a fixed degree sequence.

Corollary 4.17 (Number of graphs with a given degree sequence)**.** *([Hof18], Corollary 7.16) Assume that the degree sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$ satisfies Conditions 4.4 (a)-(c), and that $\ell_n = \sum_{i\in[n]} d_i$ is even. Then, the number of simple graphs with degree sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$ is asymptotically equal to*

$$e^{-\frac{\nu}{2}-\frac{\nu^2}{4}} \frac{(\ell_n - 1)!!}{\prod_{i\in[n]} d_i!}. \tag{4.5.3}$$

Proof. By Proposition 4.2, we obtain for any simple graph G that

$$\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G] = \frac{\prod_{i\in[n]} d_i!}{(\ell_n - 1)!!}. \tag{4.5.4}$$

Let us denote the set of all simple graphs with the given degree sequence by $\textsc{Simp}(\boldsymbol{d})$ and the number of simple graphs with the given degree sequence $\boldsymbol{d}$ by $N_G(\boldsymbol{d})$.

Then,

$$\begin{aligned} \mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}] &= \sum_{G\in\textsc{Simp}(\boldsymbol{d})} \mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G] \\ &= N_G(\boldsymbol{d}) \frac{\prod_{i\in[n]} d_i!}{(\ell_n - 1)!!}. \end{aligned} \tag{4.5.5}$$

When we multiply above equations by $(\ell_n - 1)!!$ and divide by $\prod_{i\in[n]} d_i!$, we get that

$$N_G(\boldsymbol{d}) = \mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}] \frac{(\ell_n - 1)!!}{\prod_{i\in[n]} d_i!}. \tag{4.5.6}$$

By Theorem 4.15, we know that, as $n \to \infty$, $\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}] \to e^{-\frac{\nu}{2}-\frac{\nu^2}{4}}$, hence we can conclude that $N_G(\boldsymbol{d})$ is asymptotically equal to $e^{-\frac{\nu}{2}-\frac{\nu^2}{4}} \frac{(\ell_n-1)!!}{\prod_{i\in[n]} d_i!}$. □

A special case of the configuration model arises when all degrees are equal to the same $r \in \mathbb{N}$. In that case, when we condition on the fact that the resulting graph of the configuration model has to be simple, we obtain a uniform random regular graph. In particular, we get by Corollary 4.17, that when the product nr is even, the number of regular graphs of degree r is asymptotically equal to

$$e^{-\frac{r-1}{2}-\frac{(r-1)^2}{4}}\frac{(rn-1)!!}{(r!)^n}. \tag{4.5.7}$$

Another interesting consequence of Theorem 4.15 is, that it allows us to prove a property for uniform simple random graphs with a given degree sequence by proving it for the configuration model with that same degree sequence.

Corollary 4.18 (Uniform graphs with given degree sequence and $\mathrm{CM}_n(\boldsymbol{d})$)**.** *([Hof18], Corollary 7.17)*

Assume that the degree sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$ satisfies Conditions 4.4 (a)-(c), and that $\ell_n = \sum_{i\in[n]} d_i$ is even. Then, an event $\mathcal{E}_n$ occurs with high probability for a uniform simple random graph with degrees $(d_i)_{i\in[n]}$ when it occurs with high probability for $\mathrm{CM}_n(\boldsymbol{d})$.

Proof. Let $\mathrm{UG}_n(\boldsymbol{d})$ denote a uniform simple random graph with degree sequence $\boldsymbol{d}$. Moreover, let $\mathcal{E}_n$ be a subset of all simple graphs with degree sequence $\boldsymbol{d}$. For example, $\mathcal{E}_n$ could contain all those simple graphs with degree sequence $\boldsymbol{d}$ which fulfil a specific property. We want to show that if $\lim_{n\to\infty}\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d})\in\mathcal{E}_n]=1$, then also $\lim_{n\to\infty}\mathbb{P}[\mathrm{UG}_n(\boldsymbol{d})\in\mathcal{E}_n]=1$. Alternatively, using that $\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d})\in\mathcal{E}_n]=1-\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d})\in\mathcal{E}_n^c]$, we can show that if $\lim_{n\to\infty}\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d})\in\mathcal{E}_n^c]=0$, then also $\lim_{n\to\infty}\mathbb{P}[\mathrm{UG}_n(\boldsymbol{d})\in\mathcal{E}_n^c]=0$.

By Proposition 4.16, we obtain that

$$\begin{aligned}\mathbb{P}[\mathrm{UG}_n(\boldsymbol{d})\in\mathcal{E}_n^c] &= \mathbb{P}[\mathrm{CM}_n(\boldsymbol{d})\in\mathcal{E}_n^c \mid \mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}]\\ &= \frac{\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d})\in\mathcal{E}_n^c), \mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}]}{\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}]}\\ &\le \frac{\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d})\in\mathcal{E}_n^c)]}{\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}]}.\end{aligned} \tag{4.5.8}$$

As $\boldsymbol{d}$ satisfies Conditions 4.4 (a)-(c), we can deduce by Theorem 4.15 that $\liminf_{n\to\infty}\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}] > 0$. Therefore, $\lim_{n\to\infty}\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d})\in\mathcal{E}_n^c]=0$ yields that $\lim_{n\to\infty}\mathbb{P}[\mathrm{UG}_n(\boldsymbol{d})\in\mathcal{E}_n^c]=0$, as required. □

Due to Corollary 4.18, we can derive properties of uniform simple random graphs with a prescribed degree sequence from the properties of the configuration model with the same degree sequence. Whenever we can prove that a statement holds with high probability for $\mathrm{CM}_n(\boldsymbol{d})$, then it follows with high probability for a uniform simple random graph with the same degrees. As $\mathrm{CM}_n(\boldsymbol{d})$ can be constructed in a rather simple manner, this makes it easier to prove properties for $\mathrm{CM}_n(\boldsymbol{d})$ rather than for a uniform simple random graph. However, the above strategy does not work when

$\mathbb{E}[D^2] = \infty$. This is for example the case when the degree distribution satisfies a power-law with exponent $\tau \in (2,3)$ which corresponds approximately to the distribution of the degrees in many real-world networks (as explained in Chapter 1). Therefore, it is still a big open problem how to effectively generate uniform random graphs with degree sequences having infinite variance.

Comparing Proposition 4.16 with Theorem 3.19, we observe that $\mathrm{GRG}_n(\boldsymbol{w})$ conditioned on the degrees $\boldsymbol{d}$ and $\mathrm{CM}_n(\boldsymbol{d})$ conditioned on simplicity are equal in distribution. This enables us to prove results for $\mathrm{GRG}_n(\boldsymbol{w})$, and sometimes even for the Erdös-Rényi random graph, by proving the result for the configuration model instead. This finding is stated in the following theorem and partially explains the popularity of the configuration model.

Theorem 4.19 (Relation between $\mathrm{GRG}_n(\boldsymbol{w})$ and $\mathrm{CM}_n(\boldsymbol{d})$)**.** *([Hof18], Theorem 7.18)*
Let G be a simple graph with prescribed degrees $\boldsymbol{d} = (d_i)_{i\in[n]}$. Let $D_i = \sum_{j\in[n]} X_{ij}$ be the degree of vertex i in $GRG_n(\boldsymbol{w})$ and let $\boldsymbol{D} = (D_i)_{i\in[n]}$. Then,

$$\mathbb{P}[GRG_n(\boldsymbol{w}) = G \mid \boldsymbol{D} = \boldsymbol{d}] = \mathbb{P}[CM_n(\boldsymbol{d}) = G \mid CM_n(\boldsymbol{d}) \text{ is simple}]. \tag{4.5.9}$$

Let $\mathcal{E}_n$ be a subset of simple graphs for which $\mathbb{P}[CM_n(\boldsymbol{d}) \in \mathcal{E}_n] \xrightarrow{\mathbb{P}} 1$ when $\boldsymbol{d}$ satisfies the regularity Conditions 4.4 (a)-(c). Assume that the degree sequence $\boldsymbol{D} = (D_i)_{i\in[n]}$ of $GRG_n(\boldsymbol{w})$ satisfies Conditions 4.4 (a)-(c) in probability as explained in Remark 4.5. Then, also $\mathbb{P}[GRG_n(\boldsymbol{w}) \in \mathcal{E}_n] \to 1$.

Proof. First of all let us recall that by Theorem 3.19 the law of $\mathrm{GRG}_n(\boldsymbol{w})$ conditioned on its degree sequence is uniform over all simple graphs with the same degree sequence. On the other hand, we know by Proposition 4.16 that the law of $\mathrm{CM}_n(\boldsymbol{d})$ conditioned on simplicity is uniform over all simple graphs with degree sequence $\boldsymbol{d}$ as well. Therefore, it must hold that

$$\mathbb{P}[\mathrm{GRG}_n(\boldsymbol{w}) = G \mid \boldsymbol{D} = \boldsymbol{d}] = \mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) = G \mid \mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}], \tag{4.5.10}$$

which proves the first part of Theorem 4.19.

By above equation, we obtain, for every event $\mathcal{E}_n$,

$$\mathbb{P}[\mathrm{GRG}_n(\boldsymbol{w}) \in \mathcal{E}_n \mid \boldsymbol{D} = \boldsymbol{d}] = \mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \in \mathcal{E}_n \mid \mathrm{CM}_n(\boldsymbol{d}) \text{ is simple}]. \tag{4.5.11}$$

Using the law of total expectation (i.e. $\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[X]$) and $\mathbb{E}[\mathbb{1}_{X\in A} \mid B] = \mathbb{P}[X \in A \mid B]$, we can modify

$$\begin{aligned}
\mathbb{P}[\mathrm{GRG}_n(\boldsymbol{w}) \in \mathcal{E}_n^c] &= \mathbb{E}\left[\mathbb{P}[\mathrm{GRG}_n(\boldsymbol{w}) \in \mathcal{E}_n^c] \mid \boldsymbol{D}\right] \\
&= \mathbb{E}\left[\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \in \mathcal{E}_n^c \mid \mathrm{CM}_n(\boldsymbol{D}) \text{ is simple}]\right] \\
&= \mathbb{E}\left[\frac{\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \in \mathcal{E}_n^c, \mathrm{CM}_n(\boldsymbol{D}) \text{ is simple}]}{\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \text{ is simple}]}\right] \\
&\leq \mathbb{E}\left[\left(\frac{\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \in \mathcal{E}_n^c]}{\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \text{ is simple}]}\right) \wedge 1\right],
\end{aligned} \tag{4.5.12}$$

where we write $\mathbb{P}_n$ for the conditional law given the degrees $\boldsymbol{D}$ (see also Remark 4.5). Further, note that $\mathbb{E}$ is the expectation with respect to $\boldsymbol{D}$. By the assumptions that the degree sequence $\boldsymbol{D}$

satisfies Conditions 4.4 (a)-(c) in probability and that $\mathbb{P}[\mathrm{CM}_n(\boldsymbol{d}) \in \mathcal{E}_n] \xrightarrow{\mathbb{P}} 1$, we can deduce that $\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \in \mathcal{E}_n^c] \xrightarrow{\mathbb{P}} 0$. Moreover, by the same assumptions and using Theorem 4.15, we obtain

$$\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \text{ is simple}] \xrightarrow{\mathbb{P}} e^{-\frac{\nu}{2}-\frac{\nu^2}{4}} \geq 0. \tag{4.5.13}$$

Hence, we can use Lebesgue's dominated convergence theorem 2.8 to get

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\frac{\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \in \mathcal{E}_n^c]}{\mathbb{P}_n[\mathrm{CM}_n(\boldsymbol{D}) \text{ is simple}]}\right) \wedge 1\right] = 0, \tag{4.5.14}$$

from which it follows that $\lim_{n\to\infty} \mathbb{P}[\mathrm{GRG}_n(\boldsymbol{w}) \in \mathcal{E}_n^c] = 0$. Therefore, we can conclude that $\lim_{n\to\infty} \mathbb{P}[\mathrm{GRG}_n(\boldsymbol{w}) \in \mathcal{E}_n] = 1$, as desired. □

Theorem 4.19 is valuable in the sense that it is often easier to verify that Conditions 4.4 (a)-(c) are fulfilled than to prove directly that the event $\mathcal{E}_n$ occurs with high probability for $\mathrm{GRG}_n(\boldsymbol{w})$. In the following theorem, it is shown that when the weight sequence $\boldsymbol{w}$ satisfies the regularity Conditions 3.3 (a)-(c), than the degree sequence $\boldsymbol{D}$ of $\mathrm{GRG}_n(\boldsymbol{w})$ satisfies Conditions 4.4 (a)-(c) in probability, as we would hope for. Note that from now on, we write d_i instead of D_i for the degree of vertex i in $\mathrm{GRG}_n(\boldsymbol{w})$ and that d_i is a random variable. The reason is that the regularity Conditions 4.4 are phrased in terms of D_n, which is the degree of a uniformly chosen vertex and we do not want to confuse it with the degree of vertex n, which we therefore call d_n.

Theorem 4.20 (Regularity conditions of weights and degrees)**.** *([Hof18], Theorem 7.19)*
Let d_i be the degree of vertex i in $GRG_n(\boldsymbol{w})$ defined by $d_i = \sum_{j\in[n]} X_{ij}$, and let $\boldsymbol{d} = (d_i)_{i\in[n]}$. Then, $\boldsymbol{d}$ satisfies Conditions 4.4 (a)-(c) in probability when $\boldsymbol{w}$ satisfies Conditions 3.3 (a)-(c), where

$$\mathbb{P}[D = k] = \mathbb{E}\left[\frac{W^k}{k!}e^{-W}\right], \tag{4.5.15}$$

denotes the mixed Poisson distribution with mixing distribution W having distribution function F as in Condition 3.3 (a).

Proof. Let $P_k^{(n)} = \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{d_i=k\}}$ be the empirical distribution of the degrees in $\mathrm{GRG}_n(\boldsymbol{w})$. Since the weights of $\mathrm{GRG}_n(\boldsymbol{w})$ satisfy Conditions 3.3 (a)-(b) by assumption, we can use Theorem 3.15 to deduce that $P_k^{(n)} \xrightarrow{\mathbb{P}} \mathbb{P}[D = k]$. Therefore, Condition 4.4 (a) holds in probability.

Recalling Theorem 3.5, we know that $\frac{1}{n}E(\mathrm{GRG}_n(\boldsymbol{w})) \xrightarrow{\mathbb{P}} \frac{1}{2}\mathbb{E}[W]$. As $\sum_{i\in[n]} d_i = 2\cdot E(\mathrm{GRG}_n(\boldsymbol{w}))$, we obtain that

$$\mathbb{E}_n[D_n] = \frac{1}{n}\sum_{i\in[n]} d_i = \frac{2}{n}E(\mathrm{GRG}_n(\boldsymbol{w})) \xrightarrow{\mathbb{P}} \mathbb{E}[W]. \tag{4.5.16}$$

By Lemma 3.13 we know that $\mathbb{E}[D] = \mathbb{E}[W]$ and thus Condition 4.4 (b) holds in probability when Conditions 3.3 (a)-(b) hold.

To prove that Condition 4.4 (c) holds in probability as well, we use the fact that the generalised random graph $\mathrm{GRG}_n(\boldsymbol{w})$ is asymptotically equivalent to the Chung-Lu random graph $\mathrm{CL}_n(\boldsymbol{w})$,

when Conditions 3.3 (a)-(c) hold (see Corollary 3.27). By Proposition 3.4 $\max_{i\in[n]} w_i = o(\sqrt{n})$ and thus, for n sufficiently large, $p_{ij}^{(CL)} = \left(\frac{w_i w_j}{\ell_n} \wedge 1\right) = \frac{w_i w_j}{\ell_n}$, which is slightly simpler than $p_{ij}^{(GRG)} = \frac{w_i w_j}{\ell_n + w_i w_j}$. Therefore, it turns out to be easier to show $\frac{1}{n}\sum_{i\in[n]} d_i^2 \xrightarrow{\mathbb{P}} \mathbb{E}[D^2]$ for $\mathrm{CL}_n(\boldsymbol{w})$, than for $\mathrm{GRG}_n(\boldsymbol{w})$. Throughout the rest of this proof we will therefore use the edge probabilities $p_{ij} = \frac{w_i w_j}{\ell_n}$.

Note that it is sufficient to show that Conditions 3.3 (a)-(c) imply

$$\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1) \xrightarrow{\mathbb{P}} \mathbb{E}[D(D-1)] = \mathbb{E}[W^2], \tag{4.5.17}$$

where the last equality follows from the fact that D is a mixed Poisson random variable with mixing distribution W (see Lemma 3.13). Indeed, when $\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1) \xrightarrow{\mathbb{P}} \mathbb{E}[D(D-1)]$, then $\frac{1}{n}\sum_{i\in[n]} d_i^2 - \frac{1}{n}\sum_{i\in[n]} d_i \xrightarrow{\mathbb{P}} \mathbb{E}[D^2] - \mathbb{E}[D]$. Since we know that Condition 4.4 (b) holds in probability, $\frac{1}{n}\sum_{i\in[n]} d_i^2 - \frac{1}{n}\sum_{i\in[n]} d_i \xrightarrow{\mathbb{P}} \mathbb{E}[D^2] - \mathbb{E}[D]$ yields that $\frac{1}{n}\sum_{i\in[n]} d_i^2 \xrightarrow{\mathbb{P}} \mathbb{E}[D^2]$. To prove that $\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1) \xrightarrow{\mathbb{P}} \mathbb{E}[D(D-1)] = \mathbb{E}[W^2]$, we rewrite

$$\begin{aligned}
\sum_{i\in[n]} d_i(d_i-1) &= \sum_{i\in[n]}\sum_{j\in[n]} X_{ij}\left(\sum_{k\in[n]} X_{ik} - 1\right) \\
&= \sum_{i,j,k\in[n]:k\neq j} X_{ij}X_{ik} \\
&= 2\sum_{i,j,k\in[n]:j<k} X_{ij}X_{ik},
\end{aligned} \tag{4.5.18}$$

where $(X_{ij})_{1\le i<j\le n}$ are independent Bernoulli variables with success probability $\mathbb{P}[X_{ij}=1] = \frac{w_i w_j}{\ell_n}$ for any $j \neq i$ and $X_{ii} = 0$ for any $i \in [n]$. To be able to use the second moment method, we need to start by computing the first moment.

Using the independence of $(X_{ij})_{1\le i<j\le n}$, we obtain that

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right] &= \frac{1}{n}\sum_{i,j,k\in[n]:k\neq j} \mathbb{E}[X_{ij}X_{ik}] \\
&= \frac{1}{n}\sum_{i,j,k\in[n]:k\neq j} \mathbb{P}[X_{ij}=1, X_{ik}=1] \\
&= \frac{1}{n}\sum_{i,j,k\in[n]:j\neq i;k\neq i,j} \frac{w_i w_j}{\ell_n}\frac{w_i w_k}{\ell_n} \\
&= \frac{1}{n}\sum_{i\in[n]} w_i^2 \sum_{j\in[n]:j\neq i} \frac{w_j}{\ell_n} \sum_{k\in[n]:k\neq i,j} \frac{w_k}{\ell_n} \\
&= \frac{1}{n}\sum_{i\in[n]} w_i^2 \left(\sum_{j\in[n]} \frac{w_j}{\ell_n}\sum_{k\in[n]:k\neq i,j} \frac{w_k}{\ell_n} - \frac{w_i}{\ell_n}\sum_{k\in[n]:k\neq i,j}\frac{w_k}{\ell_n}\right) \\
&= \frac{1}{n}\sum_{i\in[n]} w_i^2 \left(\sum_{j\in[n]} \frac{w_j}{\ell_n}\left(1 - \frac{w_i+w_j}{\ell_n}\right) - \frac{w_i}{\ell_n}\sum_{k\in[n]:k\neq i,j}\frac{w_k}{\ell_n}\right)
\end{aligned} \tag{4.5.19}$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i\in[n]} w_i^2 \left(\sum_{j\in[n]} \frac{w_j}{\ell_n} - \sum_{j\in[n]} \frac{w_j}{\ell_n}\frac{w_i+w_j}{\ell_n} - \frac{w_i}{\ell_n} \sum_{k\in[n]:k\neq i,j} \frac{w_k}{\ell_n} \right) \\
&= \frac{1}{n}\sum_{i\in[n]} w_i^2 \left(1 - \sum_{j\in[n]} \frac{w_j}{\ell_n}\frac{w_i+w_j}{\ell_n} - \frac{w_i}{\ell_n} \sum_{k\in[n]:k\neq i,j} \frac{w_k}{\ell_n} \right) \\
&= \frac{1}{n}\sum_{i\in[n]} w_i^2 - \frac{1}{n}\sum_{i\in[n]} \frac{w_i^3}{\ell_n} \sum_{k\in[n]:k\neq i,j} \frac{w_k}{\ell_n} - \frac{1}{n}\sum_{i,j\in[n]} \frac{w_i^2 w_j(w_i+w_j)}{\ell_n^2} \\
&= \mathbb{E}\left[W_n^2\right] - \frac{1}{n}\sum_{i\in[n]} \frac{w_i^3}{\ell_n} \sum_{k\in[n]:k\neq i,j} \frac{w_k}{\ell_n} - \frac{1}{n}\sum_{i,j\in[n]} \frac{w_i^2 w_j(w_i+w_j)}{\ell_n^2}.
\end{aligned}
$$

Using that $\max_{i\in[n]} w_i = o\big(\sqrt{n}\big)$ by Proposition 3.4, we get that $\frac{1}{n}\sum_{i\in[n]} \frac{w_i^3}{\ell_n} \sum_{k\in[n]:k\neq i,j} \frac{w_k}{\ell_n} = o(1)$ and $\frac{1}{n}\sum_{i,j\in[n]} \frac{w_i^2 w_j(w_i+w_j)}{\ell_n^2} = o(1)$. Therefore, taking the limit for $n\to\infty$ on both sides yields $\mathbb{E}\left[\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right] \to \mathbb{E}[W^2]$, by Condition 3.3 (c).

In a next step we aim to show that $\mathrm{Var}(\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)) = o(1)$. To do that, we observe

$$
\begin{aligned}
\mathrm{Var}\left(\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right) &= \mathrm{Cov}\left(\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1), \frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right) \\
&= \frac{1}{n^2} \sum_{(i,j,k),(i',j',k')} \mathrm{Cov}(X_{ij}X_{ik}, X_{i'j'}X_{i'k'}),
\end{aligned}
\tag{4.5.20}
$$

where $\mathrm{Cov}(X,Y)$ denotes the covariance between two random variables X and Y, and where the sum is over (i,j,k) and (i',j',k') such that $i,j,k\in[n]$ with $j\neq k$ and $i',j',k'\in[n]$ with $j'\neq k'$.

When $i,j,k,i',j'k'$ are all distinct, we get by using the independence of the Bernoulli random variables, that

$$
\begin{aligned}
\mathrm{Cov}(X_{ij}X_{ik}, X_{i'j'}X_{i'k'}) &= \mathbb{E}[X_{ij}X_{ik}X_{i'j'}X_{i'k'}] - \mathbb{E}[X_{ij}X_{ik}] - \mathbb{E}[X_{i'j'}X_{i'k'}] \\
&= \mathbb{E}[X_{ij}]\mathbb{E}[X_{ik}]\mathbb{E}[X_{i'j'}]\mathbb{E}[X_{i'k'}] - \mathbb{E}[X_{ij}]\mathbb{E}[X_{ik}]\mathbb{E}[X_{i'j'}]\mathbb{E}[X_{i'k'}] \\
&= 0.
\end{aligned}
\tag{4.5.21}
$$

It turns out that the covariance is always 0, except when at least one edge out of $\{i,j\},\{i,k\}$ agrees with one edge out of $\{i',j'\},\{i',k'\}$. In the case where exactly one edge out of $\{i,j\},\{i,k\}$ is equal to one edge out of $\{i',j'\},\{i',k'\}$, the three different edges could be either incident to the same vertex, they could form a 3-cycle or they could form a path of length 3. In the case where both edges $\{i,j\},\{i,k\}$ agree with the edges $\{i',j'\},\{i',k'\}$, the two edges must form a path of length 2. So let us take a closer look at the four described shapes. It turns out that there are exactly four possibilities, how we can end up with three edges incident to one vertex. These possibilities are:

- $i=i', j=j', k\neq k'$,
- $i=i', j\neq j', k=k'$,
- $i=i', j=k', k\neq j'$,
- $i=i', j\neq k', k=j'$.

If we rename the three vertices which are not incident to all three edges by j, k, l, we get that

$$\mathrm{Cov}(X_{ij}X_{ik}, X_{i'j'}X_{i'k'}) \leq p_{ij}p_{ik}p_{il}. \tag{4.5.22}$$

Similarly, we can find out that there are exactly four ways to get a 3-cycle and four ways to get a path of length 3. For the 3-cycles we can deduce that

$$\mathrm{Cov}(X_{ij}X_{ik}, X_{i'j'}X_{i'k'}) \leq p_{ij}p_{ik}p_{jk}, \tag{4.5.23}$$

while for the paths of length 3 we get

$$\mathrm{Cov}(X_{ij}X_{ik}, X_{i'j'}X_{i'k'}) \leq p_{ij}p_{jk}p_{kl}. \tag{4.5.24}$$

To get a path of length 2, however, there are only two possibilities:

- $i = i', j = j', k = k'$,
- $i = i', j = k', k = j'$.

For each of them, we obtain

$$\mathrm{Cov}(X_{ij}X_{ik}, X_{i'j'}X_{i'k'}) \leq \mathbb{E}[X_{ij}^2 X_{ik}^2] = \mathbb{E}[X_{ij}X_{ik}]. \tag{4.5.25}$$

Using that $\sum_{i,j,k\in[n]:k\neq j} \mathbb{E}[X_{ij}X_{ik}] = \mathbb{E}\left[\sum_{i\in[n]} d_i(d_i-1)\right]$, we can bound the variance from above by

$$\begin{aligned}
\mathrm{Var}\left(\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right) &\leq \frac{4}{n^2}\sum_{i,j,k,l} p_{ij}p_{ik}p_{il} + \frac{4}{n^2}\sum_{i,j,k} p_{ij}p_{ik}p_{jk} \\
&\quad + \frac{4}{n^2}\sum_{i,j,k,l} p_{ij}p_{jk}p_{kl} + \frac{2}{n^2}\mathbb{E}\left[\sum_{i\in[n]} d_i(d_i-1)\right] \\
&= \frac{4}{n^2}\sum_{i,j,k,l} \frac{w_i^3 w_j w_k w_l}{\ell_n^3} + \frac{4}{n^2}\sum_{i,j,k} \frac{w_i^2 w_j^2 w_k^2}{\ell_n^3} \\
&\quad + \frac{4}{n^2}\sum_{i,j,k,l} \frac{w_i w_j^2 w_k^2 w_l}{\ell_n^3} + \frac{2}{n^2}\mathbb{E}\left[\sum_{i\in[n]} d_i(d_i-1)\right].
\end{aligned} \tag{4.5.26}$$

Using the fact that Conditions 3.3 (a)-(c) imply that $\max_{i\in[n]} w_i = o\big(\sqrt{n}\big)$, we can deduce that, as $n \to \infty$,

$$\frac{4}{n^2}\sum_{i,j,k,l} \frac{w_i^3 w_j w_k w_l}{\ell_n^3} \leq \frac{4\max_{i\in[n]} w_i^2}{n}\mathbb{E}[W_n] = o(1), \tag{4.5.27}$$

$$\frac{4}{n^2}\sum_{i,j,k} \frac{w_i^2 w_j^2 w_k^2}{\ell_n^3} \leq \frac{4\max_{i\in[n]} w_i^3}{n^2} = o(1), \tag{4.5.28}$$

$$\frac{4}{n^2}\sum_{i,j,k,l} \frac{w_i w_j^2 w_k^2 w_l}{\ell_n^3} \leq \frac{4\max_{i\in[n]} w_i^2}{n}\mathbb{E}[W_n] = o(1). \tag{4.5.29}$$

Recalling that we already discovered that $\mathbb{E}\left[\sum_{i\in[n]} d_i(d_i-1)\right] = \mathbb{E}\left[W_n^2\right] + o(1)$ and $\mathbb{E}[W] < \infty$, we get that

$$\mathrm{Var}\left(\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right) = o(1). \tag{4.5.30}$$

Let us take n large enough such that $\left|\mathbb{E}\left[\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right] - \mathbb{E}[W^2]\right| \leq \frac{\varepsilon}{2}$ and let us make use of Chebychev's inequality 2.2 to obtain that

$$\begin{aligned}
&\mathbb{P}\left[\left|\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1) - \mathbb{E}[W^2]\right| > \varepsilon\right] \\
&= \mathbb{P}\left[\left|\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1) - \mathbb{E}\left[\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right] + \mathbb{E}\left[\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right] - \mathbb{E}[W^2]\right| > \varepsilon\right] \\
&\leq \mathbb{P}\left[\left|\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1) - \mathbb{E}\left[\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right]\right| > \frac{\varepsilon}{2}\right] \\
&\leq \frac{4}{\varepsilon^2}\mathrm{Var}\left(\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1)\right) = o(1).
\end{aligned} \tag{4.5.31}$$

By the fact that this holds for any $\varepsilon > 0$, it follows that $\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1) \xrightarrow{\mathbb{P}} \mathbb{E}[D(D-1)] = \mathbb{E}[W^2]$ which yields that $\frac{1}{n}\sum_{i\in[n]} d_i^2 \xrightarrow{\mathbb{P}} \mathbb{E}[D^2]$. Therefore, we can conclude that under the assumption that Conditions 3.3 (a)-(c) hold, Condition 4.4 (c) holds in probability, as required. □

5 Conclusion and Outlook

In this master's thesis, we analysed two static models namely the generalised random graph model and the configuration model, primarily guided by the insights presented in the book by Remco van der Hofstad [Hof18]. As explained in the first chapter, real-world networks share some common properties. Among these is the scale-free property, indicating that numerous real-world networks exhibit power-laws or at least power-law tails. Consequently, for the generalised random graph model, we examined its degree sequence in more detail. A notable advantage of the configuration model lies in the flexibility to choose the degree sequence beforehand, thereby enabling us to specify the degrees such that they have a power-law distribution. However, it turns out that the resulting graphs of the configuration model frequently have multiple edges or self-loops. As a result, the main objective behind investigating the configuration model was to ascertain its suitability to model networks which are usually depicted by simple graphs. This chapter provides a concise overview of the key findings from both Chapters 3 and 4. In addition, at the end of this chapter one can find an outlook on further results that might be of interest to the reader.

One of the primary results of Chapter 3 reveals that the degree of a fixed vertex i with given vertex weight w_i is close to a Poisson random variable with parameter w_i. Building upon this, we proceeded to demonstrate that the degree of a uniformly chosen vertex converges in distribution to a mixed Poisson random variable with mixing distribution F, when the prescribed vertex weights are sufficiently regular. The most important result of that chapter is stated in Theorem 3.15. Thanks to this, it can be concluded that the generalised random graph model is suitable for the representation of networks featuring power-law distributions when we choose the vertex weights appropriately.

In Chapter 4 we described two ways of turning the resultant graph of the configuration model into a simple graph. The first way included erasing self-loops and merging multiple edges, which is known as the erased configuration model. We proved that, under some regularity assumptions, the distribution of the degrees of the erased configuration model is close to the prescribed degree distribution of the original configuration model when the number of vertices n is sufficiently large (see Theorem 4.11). By Theorem 4.15 we know that the probability of getting a simple graph when using the configuration model is positive. Thus, instead of erasing self-loops and merging multiple edges, we can just repeatedly use the configuration model until it yields a simple graph. This procedure is then called the repeated configuration model.

In the last part of Chapter 4 we investigated the relation between the generalised random graph model and the configuration model. It turns out that the generalised random graph conditioned on the degrees is equal in distribution to the configuration model conditioned on simplicity. Consequently, we can use the configuration model, whose construction is rather simple, to prove results for the generalised random graph model (see Theorem 4.19). This is why the configuration model

is very popular. In conclusion, both models have certain advantages and disadvantages. Still, both can be used to model scale-free networks.

Our focus in this thesis has lain in investigating the two models concerning deterministic weights and degrees. In the book of van der Hofstad [Hof18] additional results can be found where the weights of the $\mathrm{GRG}_n(\boldsymbol{w})$ and the degrees of the $\mathrm{CM}_n(\boldsymbol{d})$ are random themselves. When the weights of the generalised random graph are independent and identically distributed and the distribution function F of the weights has finite mean or the limiting variable W is bounded, then the results agree with the ones we showed for deterministic weights. However, when the weights have infinite mean, then the resulting degree distribution changes. Similarly, the results concerning the configuration model depend on the finiteness of the mean of the independent and identically distributed vertex degrees (see [Hof18] Chapter 7.6).

As elaborated in Chapter 3 the generalised random graph model introduced by van der Hofstad [Hof18] represents an adaptation of the Erdös-Rényi random graph model. Yet, there exist also other models categorised as inhomogeneous random graph models, including the Chung-Lu model studied by Chung and Lu [CL02] and the Norros-Reittu model introduced by [NR06]. We have already dealt with the Chung-Lu model at the end of Chapter 3 and shown that it is asymptotically equivalent to the generalised random graph model. While both the generalised random graph model and the Chung-Lu model yield simple graphs, the Norros-Reittu model generates multigraphs. Moreover, the graph is defined as a graph process, where at each time t, a new vertex is born and older edges are erased with a given probability. Nevertheless, the Norros-Reittu model is asymptotically equivalent to the generalised random graph as well.

As mentioned in the first chapter, many real-world networks exhibit not only the scale-free property, but also other properties such as clustering. Unfortunately, it turns out that the configuration model has low clustering. To overcome that problem each vertex in $\mathrm{CM}_n(\boldsymbol{d})$ can be replaced by a small graph to obtain a graph with two levels of hierarchy. This model is called the hierarchical configuration model investigated by van der Hofstad, van Leeuwaarden and Stegehuis [HLS15].

A second book by van der Hofstad is expected to be published in the summer or autumn of 2023. A preliminary version of volume 2 is already available online [Hof23]. In this book, among other things, the generalised random graph model and the configuration model will be examined with respect to further characteristics that many real-world networks have in common. Notably, the investigation will assess whether these models are suitable for generating graphs with a giant connected component or graphs fulfiling the small-world property.

Over the past few decades, a significant surge of interest has been witnessed in the realm of modeling real-world networks. Numerous different models have been introduced and investigated with regard to their applicability. It is foreseeable that many additional models will be developed in the years to come.

References

[BAJ00] A.-L. Barabási, R. Albert, and H. Jeong. Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: Statistical Mechanics and its Applications*, 281(1-4):69–77, 2000.

[Bar03] A.-L. Barabási. *Linked: the new science of networks*. Perseus Publishing, 2003.

[Bil95] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Statistics. Wiley, 3rd edition, 1995.

[Bog11] P. Boguslawski. Modelling and analysing 3d building interiors with the dual half-edge data structure. 2011.

[Bol] B. Bollobás. *Random graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2nd edition.

[CL02] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25):15879–15882, 2002.

[EESH23] A. Elshahhat, E.-S. A. El-Sherpieny, and A. S. Hassan. The Pareto–Poisson Distribution: Characteristics, Estimations and Engineering Applications. *Sankhya A: The Indian Journal of Statistics*, 85(1):1058–1099, 2023.

[Fel68] W. Feller. *An introduction to probability theory and its applications*, volume I. John Wiley and Sons, 3rd edition, 1968.

[FFF03] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *ACM SIGCOMM Computer Communication Review*, 29(4):251–262, 2003.

[Fre04] L. C. Freeman. *The development of social network analysis: a study in the sociology of science*. Empirical Press, 2004.

[GS01] G. R. Grimmett and D. Stirzaker. *Probability and random processes*. Oxford University Press, 3rd edition, 2001.

[HFC12] B. Huffaker, M. Fomenkov, and K. Claffy. Internet topology data comparison. Technical report, Cooperative Association for Internet Data Analysis (CAIDA), (May) 2012.

[HLS15] R. van der Hofstad, J. Leeuwaarden, and C. Stegehuis. Hierarchical configuration model. *Internet Mathematics*, 2015.

[Hof18] R. van der Hofstad. *Random graphs and complex networks*, volume I. Cambridge University Press, 2018.

[Hof23] R. van der Hofstad. Random graphs and complex networks, volume II. `https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf`, 13.02.2023.

[JK93] S. Janson, D. E. Knuth, T. Łuczak, and B. G. Pittel. The birth of the giant component. *Random Structures Algorithms*, 4(3):233–359, 1993.

[JP04] J. Jacod and P. Protter. *Probability essentials*. Springer Verlag, 2nd edition, 2004.

[KX05] D. Karlis and E. Xekalaki. Mixed poisson distributions. *International Statistical Review*, 73(1):35–58, 2005.

[LPW+17] D. A. Levin, Y. Peres, E. L. Wilmer, J. G. Propp, and D. B. Wilson. *Markov chains and mixing times*. American Mathematical Society, 2nd edition, 2017.

[Mil67] S. Milgram. The small-world problem. *Psychology Today*, 1(1):61–67, 1967.

[MN20] R. Molontay and M. Nagy. Two decades of network science: as seen through the co-authorship network of network scientists. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, page 578–583. Association for Computing Machinery, 2020.

[NBW06] M. E. J. Newman, A.-L. Barabási, and D. J. Watts. *The structure and dynamics of networks*. Princeton Studies in Complexity. Princeton University Press, 2006.

[New10] M. Newman. *Networks: an introduction*. OUP Oxford, 2010.

[NR06] I. Norros and H. Reittu. On a conditionally poissonian graph process. *Advances in Applied Probability*, 38(1):59–75, 2006.

[NT18] A. Y. Nur and M. E. Tozal. Cross-as (x-as) internet topology mapping. *Computer Networks*, 132:53–67, 2018.

[SS07] M. Shaked and J. G. Shanthikumar. *Stochastic orders*. Springer Series in Statistics. Springer New York, 2007.

[Sud20] B. Sudakov. Graph theory. pages 1–6, 2020.

[Vaa98] A. W. van der Vaart. *Asymptotic statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

[vdHL14] R. van der Hofstad and N. Litvak. Degree-degree dependencies in random graphs with heavy-tailed degrees. *Internet mathematics*, 10(3-4):287–334, 2014.

[VHHK19] I. Voitalov, P. van der Hoorn, R. van der Hofstad, and D. Krioukov. Scale-free networks well done. *Physical Review Research*, 1:033034, 2019.

[Wes01] D. B. West. *Introduction to graph theory*. Pearson Education, 2nd edition, 2001.

[YJB02] S.-H. Yook, H. Jeong, and A.-L. Barabási. Modeling the internet's large-scale topology. *Proceedings of the National Academy of Sciences of the United States of America*, 99(21):13382–13386, 11 2002.

[ZMS20] B. Zhou, X. Meng, and H. E. Stanley. Power-law distribution of degree-degree distance: A better representation of the scale-free property of complex networks. *Proceedings of the National Academy of Sciences*, 117(26):14812–14818, 2020.