University of Bath

UNIVERSITY OF
BATH

PHD

Distances in Preferential Attachment Networks

Mönch, Christian

Award date:
2013

Awarding institution:
University of Bath

Link to publication

Distances in preferential attachment networks

submitted by

Christian Mönch

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

December 2013

Signature of Author . . . .

<div align="center">Christian Mönch</div>

Summary

Preferential attachment networks with power law degree sequence undergo a phase transition when the power law exponent τ changes. For $\tau > 3$ typical distances in the network are logarithmic in the size of the network and for $2 < \tau < 3$ they are doubly logarithmic. In this thesis, we identify the correct scaling constant for $\tau \in (2,3)$ and discover a surprising dichotomy between preferential attachment networks and networks without preferential attachment. This contradicts previous conjectures of universality. Moreover, using a model recently introduced by Dereich and Mörters, we study the critical behaviour at $\tau = 3$, and establish novel results for the scale of the typical distances under lower order perturbations of the attachment function.

Acknowledgements

Contents

Chapter 1

Introduction

1.1 Overview

The theory of random networks provides tools to model a multitude of phenomena in the physical sciences, computer science, social sciences and even linguistics. A network, in the most general sense, consists of two classes of objects: vertices and edges. For many applications, the number of vertices and edges is very large and changes over time according to non-deterministic rules. Of particular interest, especially in the social sciences and computer science is the *small world problem*[1], i.e. the observation that even in very large sparse[2] networks, most vertices can be connected by relatively short sequences of edges.

The goal of this work is to derive asymptotics for the *distance of typical vertices* in random networks as they grow large. One fundamental parameter in the modelling of large scale networks is the degree distribution. It has been observed, see e.g. [AB02] for a survey, that real world networks often display an approximate power law distribution. *Preferential attachment* (PA), which has attracted lots of attention in the mathematical community recently, see e.g. [Ath07; Bor+07; AGS08; BL12], is a simple modelling paradigm[3] for the network formation which explains the occurrence of power laws.

If the power law exponent τ of a random network satisfies $2 < \tau < 3$, then the network topology is especially rich. For instance in [NR08] the authors observe a *soft hierarchy* structure, which only appears for infinite variance degrees. In these networks it is possible to remove many high degree vertices without significantly changing the size of the largest con-

---

[1] See the famous study [Mil67], the phenomenon has been popularised e.g. in [Wat03] and [Bar02] and seems to capture the imagination of wide audiences, similarly to the famous *butterfly effect*.

[2] 'Sparse' means that the number of edges is of the same order as the number of vertices. This holds for all models considered in this work.

[3] PA schemes have been studied in other contexts much earlier, e.g. in [Yul25]. There is also combinatorial literature studying PA-like self-reinforcing processes, see e.g.[Szy87] and of course closely related *urn processes* have been studied extensively by probabilists. However, the idea to use PA to generate power law networks appeared only later, first in [BA99] as a model for the web graph.

nected component. More generally, it is assumed, see e.g. [HH08], that there are two universality classes: the class of *small worlds* in which $\tau > 3$ and typical distances are logarithmic in the total size of the network and the class of *ultrasmall worlds* in which $2 < \tau < 3$ and the scale is doubly logarithmic. The transition between these two classes occurs at $\tau = 3$ and we devote special attention to this case.

Our two main results and their proofs extend the understanding of the network topology for the class of preferential attachment models. They are presented and discussed in detail in Chapter 2. Both results feature the nonlinear preferential attachment model introduced in [DM09] and studied in [DM13]. A survey of known results for this model can be found in [DM11]. Theorem I covers the situation in which $\tau \in (2,3)$ and the arguments used in its proof are robust enough to establish a new and surprising general dichotomy between models with preferential attachment and models without preferential attachment. Theorem II addresses the case $\tau = 3$, in which the nonlinearity of the model comes into play and allows us to calculate distances in situations not previously covered, shedding some light on the structural properties of networks at this transition point.

In Chapter 3 we give a non-rigorous account of the proofs of our main results, explaining the most important ideas. Chapter 4 contains prerequisite results, most importantly a branching process approximation which was developed in [DM13] and is crucial for the proofs and our understanding of the model. The formal proofs of the novel results of Chapter 2 are given in Chapters 5 and 6.

In the remainder of this opening chapter we set out the notation used throughout this thesis. This is done in Section 1.2. Section 1.3 introduces the *sublinear preferential attachment model*, for which our main results are formulated. This section also contains a discussion of some of the models' features and surveys some results that are needed as background for subsequent chapters. Finally, in Section 1.4, we briefly discuss some of the literature regarding distances in other power law random graphs.

1.2 Notation

We start by introducing the basic notation which is used throughout this thesis. We use the notations $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ for the set of non-negative integers and $[n]$ for the set of positive integers less than or equal to n. The Gamma function is denoted by $(\Gamma(x))_{x \geq 0}$, i.e. for $n \in \mathbb{N}$ we have $\Gamma(n+1) = n!$. For a function $f(p,n)$ defined for integer values $n \in \mathbb{N}_0$ and parametrised by p we use the notation $\Delta f(p,n)$ for the increment $f(p,n+1) - f(p,n)$, $n \in \mathbb{N}_0$. If f is defined on a continuum we write $\Delta f(p,t) = f(p,t) - \lim_{s \uparrow t} f(p,s)$ if the latter exists.

Identical constants appearing in several proofs are indexed according to the place they

are first used, such that, say, $C_{1.2}$ is derived in Lemma 1.2 and $D_{(3.4)}$ is introduced in equation (3.4). γ_{EM} frequently denotes the Euler-Mascheroni constant $\lim_{n\to\infty} \int \left(\lfloor \frac{1}{x} \rfloor - \frac{1}{x}\right) \mathrm{d}x$.

Given a function $f : \mathbb{N} \longrightarrow \mathbb{R}$ we furthermore use the notations

$$f[m] := \sum_{i=1}^{m} f(i) \quad \text{for } m \in \mathbb{N}$$

and in general

$$f^j(A) = \sum_{i \in A} f(i)^j \quad \text{for } A \subset \mathbb{N}, j \in \mathbb{N},$$

where we suppress the superscript if $j = 1$ and $\#A = f^0(A)$ denotes the cardinality of A.

We are usually interested in statements about a graph $\mathscr{G}_N$ on $[N]$ which are very likely to be true if the system size N gets sufficiently large. We use the term *with high probability* or *w.h.p.* to describe this. Formally this means that if E_N is a property of graphs of size N, defined as a subset of all graphs on $[N]$, then

$$\lim_{N\to\infty} \mathbb{P}\{\mathscr{G}_N \in E_N\} = \lim_{N\to\infty} \mathbb{P}\{E_N\} = 1.$$

Other common phrases to denote this form of convergence in law are 'asymptotically almost surely' and 'almost every $\mathscr{G}_N$ satisfies E_N'.

We also frequently use the asymptotic o-notation, i.e. for non-negative real functions f, g, the statement $f(x) = o(g(x))$ means that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$ and $f(x) = O(g(x))$ means that $\limsup_{x\to\infty} \frac{f(x)}{g(x)} < \infty$. We also write $f \sim g$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.

We are now ready to introduce the main objects of our study: families of growing random graphs.

Definition 1.1 (Network model)**.** As a *(dynamic) network model* we denote a sequence of random graphs $(\mathscr{G}_N)_{N\in\mathbb{N}}$ with the set of vertices $V(\mathscr{G}_N)$ of $\mathscr{G}_N$ given by $[N] = \{1, 2, \ldots, N\}$ and the set of unoriented edges $E(\mathscr{G}_N)$ of $\mathscr{G}_N$ given by a random symmetric subset of $[N] \times [N]$. We write $v \leftrightarrow w$ if the vertices v, w are connected by an edge in the graph $\mathscr{G}_N$. We will also use the shorthands $v \in \mathscr{G}_N$ (instead of $v \in V(\mathscr{G}_N)$) for vertices and $e \in \mathscr{G}_N$ (instead of $e \in E(\mathscr{G}_N)$) for edges.

Some of the examples discussed in Section 2.2 include models where multiple edges between one pair of vertices are allowed, but this will play no role for our considerations and we treat them like one single edge.

In this work we provide techniques to find bounds on distances in $\mathscr{G}_N$, using the convention that two vertices have infinite distance if they cannot be connected by a sequence

of edges. For connected subgraphs $A \subset \mathscr{G}_N$, we also call the maximal distance between two vertices in A its diameter. This is formalised in the following definition.

Definition 1.2 (Distances). The graph distance in $\mathscr{G}_N$ is given by

$$d_N(v, w) := \min\left\{n : \exists v = v_0, v_1, \ldots, v_n = w \in \mathscr{G}_N \text{ such that } v_{i-1} \leftrightarrow v_i \; \forall i \in [n]\right\},$$

and the notation generalises to sets of vertices $A, B \subset [N]$ by setting

$$d_N(A, B) = \min\{d_N(a, b) : a \in A, b \in B\}.$$

The *typical distance* is the asymptotic graph distance of two vertices $U, V \in [N]$ which are chosen independently and uniformly at random in the graph $\mathscr{G}_N$. Furthermore, the *diameter* of $A \subset \mathscr{G}_N$ is defined by

$$\text{diam}_N(A) := \max_{\substack{m,n \in [A] \\ m,n \text{ connected}}} d_N(m, n).$$

Note that if $\mathscr{G}_N$ is not connected, which is often the case, then we restrict ourselves to typical distances in the largest connected component of $\mathscr{G}_N$. This component is usually denoted by $\mathscr{C}_N$. Typically, there will be only one component $\mathscr{C}_N \subset \mathscr{G}_N$ which comprises a positive fraction of all vertices in the limit $N \to \infty$. Therefore, picking two random vertices U, V from $\mathscr{C}_N$ has the same effect as conditioning U, V to be connected.

The canonical example of a random network is the *Erdős-Rényi random graph* $\mathscr{G}_N = \mathscr{G}(N, p)$, in which all edges in $\mathscr{G}_N$ are mutually independent and exist with probability $p \in [0, 1]$. However, for most applications this standard tool of combinatorial probability is not a satisfactory model. For instance, in view of our emphasis on growth dynamics, in $\mathscr{G}(N, p)$ there is no immediate rule how to obtain $\mathscr{G}(N + 1, p)$ from $\mathscr{G}(N, p)$, if p depends on N. The rich theory about $\mathscr{G}(N, p)$ nevertheless provides many ideas and techniques for the study of more complex networks. An account of the classical results can be found in [Bol01].

As outlined above, the starting point for our considerations is the modelling of random networks whose degree sequences obey an approximate power law.

Definition 1.3 (Degrees and power laws). Let $v \in \mathscr{G}_N$. Then

$$Z[v, N] = \sum_{i \in [N]} \mathbb{1}_{\{(v,i) \in E(\mathscr{G}_N)\}} \text{ denotes the } \textit{degree of vertex } v \textit{ in } \mathscr{G}_N.$$

We also call

$$\mathcal{Z}[v, N] = \sum_{i \in \{v+1, \ldots, N\}} \mathbb{1}_{\{(v,i) \in E(\mathscr{G}_N)\}} \text{ the } \textit{indegree} \text{ of } v.$$

The *empirical degree distribution* of $\mathscr{G}_N$ is denoted by $\mu(N) = (\mu_k(N))_{k=0}^{\infty}$, where

$$\mu_k(N) = \frac{1}{N} \sum_{v \in [N]} \mathbb{1}_{\{Z[v,N]=k\}}, \quad \text{for } k \in \mathbb{N}_0.$$

We speak of a *power law network* if there is a power law exponent $\tau > 0$ and a slowly varying function $L : \mathbb{N}_0 \longrightarrow [0,\infty)$ such that, for all $k \geq 0$,

$$\lim_{N\to\infty} \mu_k(N) = k^{-\tau} L(k), \quad \text{in probability.}$$

Remark 1.4. Many authors reserve the term power law for sequences $\mu(N)$ satisfying, for some $C > 0$ and $k \geq 1$,

$$\lim_{N\to\infty} \mu_k(N) = C k^{-\tau}, \quad \text{in probability.}$$

We call this a *strict power law*. The more general definition above is used, since we are particularly interested in the behaviour of the network under alteration of the lower order perturbation L.

As a consequence of the law of small numbers, the degrees in $\mathscr{G}(N, p)$ are asymptotically Poisson distributed with parameter $\lambda = \lim_{N\to\infty} N p(N)$ in the sparse case, i.e. if the latter expression is well defined and finite. Indeed, to generate power law networks we need to leave the framework of $\mathscr{G}(N, p)$.

1.3 The sublinear preferential attachment model

Preferential attachment is a dynamical modelling approach that was taken up to explain why power law degree sequences arise in growing real networks. The approach was pioneered in [BA99]. The principle of PA is that vertices which at a given point in time have a lot of connections accumulate more connections than other vertices subsequently as the network grows. This is an empirically and intuitively well grounded assumption in many applications – 'the rich get richer'[4]. Therefore, in general, one assumes in the PA-framework that the connectivity of a vertex, i.e. the probability that a new vertex connects to it, is an increasing function of its degree.

Indeed, if the connectivity of a vertex is an asymptotically linear (or affine) function of the total degree, then this type of model produces a power law graph. However, the degree sequence does not obey a power law as soon as this linear relation is broken up. This has been shown rigorously for a sublinear connectivity in [DM09] and for superlinear connectivities in [OS05]). The problem of justifying the modelling approach in applications has thus in some sense been shifted to the problem of justifying the use of a certain attachment

---

[4]In the social sciences the phenomenon is also known as the 'Matthew effect', after the line *"For unto every one that hath shall be given, and he shall have abundance; but from him that hath not shall be taken even that which he hath."* in the Gospel According to Matthew.

function.

We now describe the sublinear preferential attachment model studied in [DM09] and [DM13]. In the following subsections the model is introduced and some essential features and results which are important later in the text are stated.

1.3.1 Model definition

Definition 1.5 (Sublinear PA)**.** The *sublinear PA model* is a Markov chain $(\mathcal{G}_N)_{N\geq 1}$ on the set of all finite directed graphs with labels in $\mathbb{N}$ constructed using the following ingredients:

(I) a monotonically increasing *attachment rule* $f : \mathbb{N}_0 \longrightarrow (0,\infty)$;

(II) the initial state $\mathcal{G}_1$ represented by a single vertex labelled 1;

(III) the random attachment mechanism, i.e. given $\mathcal{G}_N$, the graph $\mathcal{G}_{N+1}$ is obtained by

- adding a new vertex labelled $N+1$,
- inserting directed edges from vertex $N+1$ to each existing vertex with label $m \in [N]$ with probability
$$\frac{f(\mathcal{Z}[m,N])}{N} \wedge 1$$
independently of everything else.

Remark 1.6. (i) We interpret the labels as time and also do not distinguish verbally between vertices and their labels.

(ii) The rules given above generate a directed graph, however the direction of the edges is determined solely by the order of the vertices and can be ignored.

(iii) Instead of using the connection probability $\frac{f(\mathcal{Z}[m,N])}{N} \wedge 1$, one also could demand $f(k) \leq k + 1$, for all $k \in \mathbb{N}_0$, which is the original definition in [DM09]. The results do not depend on this choice. For technical reasons it is often necessary to restrict the consideration to concave attachment rules. In some situations this restriction can be dropped, see e.g. [EM13], but in particular for the rigorous formulation of the branching process approximation, which is one main tool of our investigation, it is very useful.

Generating the network in the way described in Definition 1.5 has some straightforward consequences, which will be important in our considerations later.

Observation 1.7 (Stochastic domination I). (i) If $A, B \subset [N]$ and $N' > N$ then surely

$$d_{N'}(A, B) \leq d_N(A, B),$$

since edges are never removed.
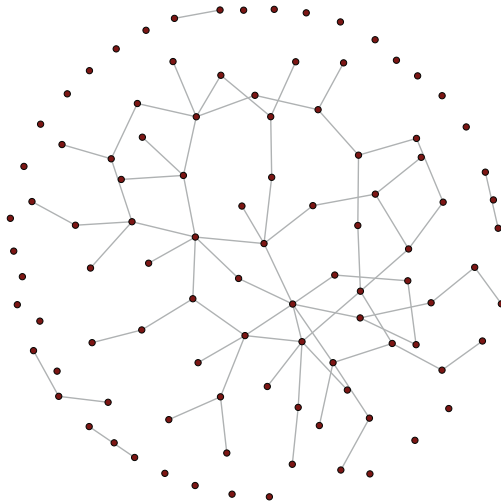
Figure 1-1: Realisation of $\mathscr{G}_N$ for $f(k) = \frac{1}{2}k + \frac{1}{2}$, $N = 100$

(ii) If $f \leq \bar{f}$ are attachment rules used to generate graph sequences $(\mathscr{G}_N)_{N\geq1}$ and $(\bar{\mathscr{G}}_N)_{N\geq1}$, then $\bar{\mathscr{G}}_N$ stochastically dominates $\mathscr{G}_N$, i.e. we can find a coupling $(\mathscr{G}^{(c)}, \bar{\mathscr{G}}^{(c)})$ of the sequences such that almost surely $\mathscr{G}_N^{(c)} \leq \bar{\mathscr{G}}_N^{(c)}$ for all $N \in \mathbb{N}$, in the sense that $v \leftrightarrow w$ in $\bar{\mathscr{G}}_N^{(c)}$ if $v \leftrightarrow w$ in $\mathscr{G}_N^{(c)}$ for all $v, w \in [N]$.

Induced parameters which are important in many of the calculations hereafter are

$$\gamma^- = \inf_k \Delta f(k), \quad \gamma^+ = \sup_k \Delta f(k), \quad \text{and } \gamma = \lim_{k\to\infty} \Delta f(k),$$

if f is such that the last expression is well defined. We shall often assume that f is *concave*, in which case $\gamma^+ = \Delta f(0)$ and $\gamma = \gamma^-$ exists. Some estimates also require that $\gamma^+ < 1$, however this is for simplicity only and could be replaced by $\limsup_{k\to\infty} \Delta f(k) < 1$.

Central to most of our arguments are the *degree evolutions* of the vertices in our network.

Definition 1.8 (Degree evolutions)**.** (i) Let $v \in \mathbb{N}$. We call the Markov chain $(\mathscr{Z}[v,n])_{n\geq v}$ counting the edges linking to v during the growth of $(\mathscr{G}_N)_{N=v}^\infty$ the *indegree evolution* of v or just its *degree evolution*.

(ii) The *idealised degree evolution* $(Z_t)_{t\geq0}$ is the continuous time pure jump Markov process which starts at $Z_0 = 0$ and jumps from state k into state $k+1$ at rate $f(k)$.

(iii) We call the independent families $\{(\mathcal{Z}^{(k)}[i,n])_{n=i}^{\infty};\ i,k \in \mathbb{N}\}$ and $\{(Z^{(k)})_{t\geq 0};\ k \in \mathbb{N}\}$ *generalised degree evolutions.* The Markov chain $(\mathcal{Z}^{(k)}[i,n])_{n=i}^{\infty}$ is started in $\mathcal{Z}^{(k)}[i,i] = k$, jumps by 1 at time $n > i$ with probability $\frac{f(\mathcal{Z}^{(k)}[i,n-1])}{n-1} \wedge 1$ and stays constant otherwise. The continuous time process $(Z_t^{(k)})_{t\geq 0}$ is started in $Z_0^{(k)} = k$ and jumps from state j into state $j+1$ at rate $f(j)$ for $j \geq k$.

A discrete-time generalised degree evolution behaves like the indegree of a vertex which is introduced with already k edges pointing towards it.The idealised degree evolutions are instrumental in constructing a limiting object for the sublinear PA graph, namely a typed tree. One main prerequisite for this thesis is the link between the existence of a *giant component* in the network and the survival of this tree.

The next two sections survey those results of [DM09] and [DM13] about the model which are the foundation of most of the work in this thesis.

1.3.2 The degree sequence

To put the sublinear PA model into the context of power law random networks, it is necessary to study the degree sequence. This (and much more) is the subject of [DM09]. The convergence of the empirical degree distribution is covered by the following two statements.

Proposition 1.9 (Convergence of degree distribution, [DM09, Theorem 1.1])**.** *Let f be any attachment rule.*

(I) Let $\mu^{(in)}(N)$ denote the empirical indegree distribution of $\mathcal{G}_N$, then

$$\lim_{N\to\infty} \mu^{(in)}(N) = \mu, \ \ almost\ surely\ in\ total\ variation\ norm,$$

where $\mu = (\mu_k)_{k\geq 0}$ is the probability measure on $\mathbb{N}_0$ with

$$\mu_k = \frac{1}{1+f(k)} \prod_{i=0}^{k-1} \frac{f(i)}{1+f(i)}, \ \ for\ k \in \mathbb{N}_0.$$

(II) If f satisfies $f(k) \leq \eta k + 1$ for some $\eta \in (0,1)$, then the conditional distribution of the outdegree of vertex $n+1$ given $\mathcal{G}_n$ converges almost surely in total variation norm to a Poisson distribution with parameter

$$\lambda = \sum_{k=0}^{\infty} \prod_{i=0}^{k} \frac{f(i)}{1+f(i)} = \sum_{k=0}^{\infty} \mu(k)f(k).$$

If f is chosen to be affine, i.e. $f(k) = f(0)+\gamma k$, with $f(0),\gamma \in (0,1]$, then Proposition 1.9 (I)

yields, see [DM09, Example 1.3],

$$\mu_k = \frac{1}{\gamma} \frac{\Gamma\left(k + \frac{f(0)}{\gamma}\right)\Gamma\left(\frac{f(0)+1}{\gamma}\right)}{\Gamma\left(k + \frac{f(0)+1}{\gamma}\right)\Gamma\left(\frac{f(0)}{\gamma}\right)}, \quad \text{for all } k \in \mathbb{N}_0$$

and Stirling's formula yields

$$\mu_k \sim \frac{\Gamma\left(\frac{f(0)+1}{\gamma}\right)}{\Gamma\left(\frac{f(0)}{\gamma}\right)} k^{-(1+\frac{1}{\gamma})}, \quad \text{as } k \to \infty,$$

thus the indegree distribution converges to a power law. More generally, $\mathcal{G}_N$ is a power law network with power law exponent $\tau = 1 + \frac{1}{\gamma}$, if $\lim_{k\to\infty} \Delta f(k) = \gamma \in (0,1)$. In this case the outdegree distribution has no influence, since the tails of the Poisson distribution decay exponentially and do not distort the power law asymptotics.

Although we are concerned with the more important regime of asymptotically linear functions, it is worthwhile mentioning at this point, that if $\lim_{k\to\infty} \Delta f(k) = 0$, then the model does not produce a power law network. This is especially clear for constant attachment rules, in which case the attachment is uniform and not preferential and the model yields a version of *Dubins' model*, which is considered in e.g. in [She89]. For a suitable choice of p, this network is asymptotically equivalent to $\mathcal{G}(N, p)$, as $N \to \infty$, and thus provides a dynamical version of $\mathcal{G}(N, p)$ which is not a power law network.

Between the constant and asymptotically linear regime, the limiting degree distribution depends on f in a more intricate way. For instance, the choice $f(k) \sim \beta k^\alpha$, for $\beta > 0$ and $0 < \alpha < 1$, yields stretched exponential tails

$$\log \mu_k \sim \frac{1}{\beta} \frac{1}{1-\alpha} k^{1-\alpha}, \quad \text{as } k \to \infty,$$

according to [DM09, Example 1.4].

1.3.3 The giant component

It is easily seen from the attachment mechanism, that $\mathcal{G}_N$ will in general be a disconnected graph. If we restrict ourselves to distances between vertices in the largest connected component $\mathcal{C}_N \subset \mathcal{G}_N$, then, to make a meaningful statement about *typical* distances in $\mathcal{G}_N$, we need to make sure that the relative size $\frac{\#\mathcal{C}_N}{N}$ of this component does not vanish in the limit $N \to \infty$. If $\frac{\#\mathcal{C}_N}{N}$ converges in probability to a positive value, then we say that a *giant component* exists. The existence and uniqueness of a giant component is one of the main results of [DM13], which is reproduced below as Proposition 1.12.

Before we can state the result, we need to introduce an important tool, which is also of central significance for the later chapters. This is the *idealised branching random walk* (IBRW). It is a typed random walk on $\mathbb{R}$ with type space $\mathscr{S} = [0,\infty] \cup \{\ell\}$, where ℓ is an arbitrarily chosen non-numerical symbol. This random walk will help us to approximate the neighbourhood of a typical vertex $V \in \mathscr{G}_N$ for very large N, i.e. a vertex chosen randomly from the uniform distribution on $[N]$. The types are necessary to keep track of the dependencies between edges which are induced by the preferential attachment.

Definition 1.10 (Idealised branching random walk). Recall that $(Z_t \colon t \geq 0)$ is the pure birth process started in zero with generator G given by

$$Gg(k) = f(k)\Delta g(k), \quad g \colon \mathbb{N}_0 \longrightarrow \mathbb{R}, k \in \mathbb{N}_0.$$

For $\sigma \in [0,\infty]$, we denote by $(Z_t^\sigma)_{t\geq 0}$ a version of $(Z_t)_{t\geq 0}$ conditioned to jump at $\sigma \geq 0$. The *idealised branching random walk* X consists of a collection of *particles* $x = (s(x), \tau(x))$, where $s(x) \in \mathbb{R}$ is the *position* of x and $\tau(x) \in \mathscr{S}$ its *type*, endowed with a genealogical structure[5]. It is constructed as follows:

(I) The walk is *initialised* with one particle $x_\emptyset$ of type ℓ in random position $s(x_\emptyset)$, where $-s(x_\emptyset)$ is exponentially distributed with parameter 1.

(II) Every particle of the walk *produces offspring* according to the following rules depending on its type and position:

- an ℓ-type particle at $s \in \mathbb{R}$ produces offspring particles
  - to the right in positions relative to s given by jumps of $(Z_t)_{t\geq 0}$ and of type ℓ,
  - to the left in position relative to s distributed like a Poisson point process with intensity measure $e^t \mathbb{E}f[(Z_{-t})]\,\mathrm{d}t$, typed by their distance to s;
- a τ-type particle at $s \in \mathbb{R}$ produces offspring particles
  - to the right in position relative to s given by jumps of $(Z_t^\tau - \mathbb{1}_{[\tau,\infty)}(t))_{t\geq 0}$ and of type ℓ,
  - to the left like an ℓ-type particle.

Having constructed the infinite random walk X, we need to relate X to the sequence $(\mathscr{G}_N)_{N\geq 1}$ of finite networks. This is done by removing all particles on the positive half line.

Definition 1.11 (Idealised neighbourhood tree). Let X denote the IBRW, then the *idealised neighbourhood tree* (INT) $\mathfrak{T}$ is the typed tree obtained from X by removing all particles $x \in X$

---

[5]Formally, every $x \in X$ also has a unique *label* $l(x) \in \bigcup_{n\in\mathbb{N}} \mathbb{N}^n \cup \{\emptyset\}$, where $\emptyset$ is the label of the root particle and the set of labels $\{l(x), x \in X\}$ obeys the obvious order rules. However, we do not need to refer to the labels explicitly anywhere in our argumentation.

with $s(x) > 0$ and their descendants. We denote by $\#\mathfrak{T}$ the number of particles in $\mathfrak{T}$. If $\#\mathfrak{T} = \infty$, then the INT *survives*, and we denote

$$p(f) = \mathbb{P}\{\#\mathfrak{T} = \infty\}.$$

Furthermore, we denote by π_N the projection of $(-\infty]$ to $[N]$ which maps a point $s \in (-\infty, 0]$ onto the smallest $m \in [N]$ with $s \leq -\sum_{i=1}^{N-1} \frac{1}{i} + \sum_{i=1}^{m-1} \frac{1}{i}$.

Informally, the inverse of π_N can be thought of mapping the network $\mathscr{G}_N$ to the INT. The vertex N is mapped to a particle with position 0, the vertex 1 to a particle at $-\log N$ and the order of the vertices is preserved, therefore the whole network is projected to the negative half line.

The survival of the INT reflects the existence of a giant component in the network. Moreover, the survival probability corresponds to the relative size of the giant component.

Proposition 1.12 (The giant component, [DM13, Theorem 1.8])**.** *Let f be a concave attachment rule and let $p(f)$ be the survival probability of the INT. If $\mathscr{C}_N^{(1)}$ and $\mathscr{C}_N^{(2)}$ denote the largest and second largest connected component of $\mathscr{G}_N$, then*

$$\frac{\#\mathscr{C}_N^{(1)}}{N} \to p(f) \ \text{and} \ \frac{\#\mathscr{C}_N^{(2)}}{N} \to 0, \ \text{in probability.}$$

In particular, there exists a unique giant component if and only if $p(f) > 0$.

Remark 1.13. One can analyse the growth of the IBRW by studying the family of compact linear operators $(A_\theta)_{0<\theta<1}$ on the Banach space $\mathfrak{C}_b$ of continuous bounded real valued functions on $\mathscr{S} = [0,\infty] \cup \{\ell\}$, which is given by

$$A_\theta g(\tau) = \int_0^\infty g(t) e^{\theta t} \, \mathrm{d}M(t) + \int_0^\infty g(\ell) e^{-\theta t} \, \mathrm{d}M^\tau(t), \ \ g \in \mathfrak{C}_b$$

where increasing functions M, resp. M^τ, are given by

$$M(t) = \int_0^t e^{-s} \mathbb{E}\big[f(Z_s)\big] \, \mathrm{d}s, \quad M^\ell(t) = \mathbb{E}\big[Z_t\big], \ \text{for } t \geq 0,$$

$$\text{and } M^\tau(t) = \mathbb{E}\big[Z_t | \Delta Z_\tau = 1\big] - \mathbb{1}_{[\tau,\infty)}(t) \ \ \text{for } t \geq 0, \tau \in [0,\infty].$$

Note that the functions M featuring in the definition of the operators are derived from the IBRW as follows: $M(t)$ is the expected number of particles within distance t to the left of any given particle, and $M^\tau(t)$ is the expected number of particles within distance t to the right of a given particle of type τ. Indeed, $(A_\theta)_{\theta \in (0,1)}$ plays the role of the Laplace transform of the offspring measure in classical branching random walk results, see e.g. [Big77]. Consequently, the 'Malthusian parameter' for the branching random walk can be derived from $(A_\theta)_{\theta \in (0,1)}$ via the spectral radii $\varrho(A_\theta), 0 < \theta < 1$, and $p(f) > 0$ if and only if there is $\theta \in (0,1)$ such that $\varrho(A_\theta) > 1$. This is the content of Theorem 1.1 of [DM13].

In particular, if $\gamma \geq \frac{1}{2}$, then $p(f) > 0$ and for affine attachment rules, the abstract criterion of survival of $\mathfrak{T}$ can be reformulated explicitly.

Proposition 1.14 (Conditions for the existence of a giant component,[DM13, Example 1.2, Proposition 1.3])**.** *If $\frac{1}{2} \leq \liminf_{k \to \infty} \Delta f(k)$, then a giant component exists. Moreover, if $f(k) = f(0) + \gamma k$, $k \in \mathbb{N}_0$, for some $0 < \gamma < 1$ and $f(0) > 0$, then a giant component exists if and only if*

$$\gamma \geq \frac{1}{2} \quad or \quad f(0) \geq \frac{\left(\frac{1}{2} - \gamma\right)^2}{1 - \gamma}.$$

The above results give some further justification to focus on situations in which $\gamma \geq \frac{1}{2}$, i.e. in which $\mathcal{G}_N$ is a power law network with power law exponent $\tau = 1 + \frac{1}{\gamma} \in (2, 3]$, and explain why it is reasonable to call the case $\tau = 3$ 'critical' in this context. If no giant component exists, then the problem of 'typical' distances needs to be reformulated. If a giant component exists, yet $\gamma < \frac{1}{2}$, then typical distances are of logarithmic scale and completely determined by the tree structure of the INT. To calculate the exact asymptotics for this case, i.e. to derive the correct scaling constant from the attachment function f is beyond the scope of this thesis. However, the problem does not appear to be very difficult and lies within the range of the techniques already developed for the sublinear PA model.

1.3.4 Relation to other preferential attachment models

Apart from the sublinear PA model, there are many ways of formalising the preferential attachment approach. The first rigorous mathematical formulation of the mechanism suggested in [BA99] was given in [Bol+01]. The model defined there yields a power law network with exponent $\tau = 3$. Later the model was extended to cover the range $\tau \in (2, \infty)$, see [Hof13, Chapters 8 and 11] for an extensive survey of the results. We will refer to this model as the *classical* or *affine* PA model and revisit it in more detail in Section 2. Recall that in the sublinear model $\tau = 1 + \frac{1}{\gamma}$, hence $\tau = 3$ corresponds to $\gamma = \frac{1}{2}$. Remarkably, the distance results about the difficult critical case for this model pre-date the results for other values and even rigorous results for simpler models without preferential attachment at $\tau = 3$, see in particular [BR04].[6]

We now survey results on the diameter and typical distances for the classical model, which have been obtained in [BR04] and [DHH10]. The key feature of this model – or in fact this family of models, since there are several non-equivalent formulations of it – is that the graph $\mathcal{G}_N = \mathcal{G}_{N+1}^{\mathfrak{m}, \delta}$ is obtained from the graph $\mathcal{G}_N^{\mathfrak{m}, \delta}$ by adding one new vertex with a fixed number of $\mathfrak{m}$ outgoing edges which connect to each one of the already existing vertices with a probability proportional to its total degree plus a constant δ. The parameter $\delta \in (-\mathfrak{m}, \infty)$ is allowing us to construct power law networks for any exponent $\tau > 2$. Although the exact

---

[6]Although not published until 2004 this article had already been written by 2001.

attachment mechanism differs slightly from model to model within the family, the asymptotic behaviour is the same. We will revisit the model in Example 2.7.

Certain model variants allow self-loops which affects the connectedness of the graph, since there is a tiny probability that an incoming vertex connects only to itself and thus stays isolated for some time. However, these events are so rare that they do not influence the distance results. If $\mathfrak{m} = 1$, then the resulting graph is a tree or a forest and diameter as well as average distances grow logarithmically. The behaviour of these trees is studied in [Pit94]. We exclude this case here, as we are only interested in models which allow for cycles, i.e. 'true' network models.

The results for the order of the diameter of these networks (or of its largest connected component, respectively) are given in the following table, note that the power law exponent is given in terms of the parameters $\mathfrak{m}$ and δ, we have $\tau = 3 + \frac{\delta}{\mathfrak{m}}$. For convenience we have also included the corresponding parameter γ of the sublinear model.

Asymptotic bounds for $\mathrm{diam}_N(\mathscr{G}_N^{\mathfrak{m},\delta})$ for $\mathfrak{m} \geq 2$			
range of δ	$(-\mathfrak{m}, 0)$	$\{0\}$	$(0, \infty)$
range of τ	$(2, 3)$	$\{3\}$	$(3, \infty)$
range of γ	$\left(1, \frac{1}{2}\right)$	$\{\frac{1}{2}\}$	$\left(0, \frac{1}{2}\right)$
upper bound	$\left(\frac{4}{|\log(\tau-2)|} + \frac{4\sigma}{\mathfrak{m}}\right)\log\log N$ for any $\sigma > \frac{1}{3-\tau}$	$\frac{\log N}{\log\log N}$	$O(\log N)$
lower bound	$O(\log\log N)$	$\frac{\log N}{\log\log N}$	$O(\log N)$
published in	[DHH10]	[BR04]	[DHH10]

For the average distances between two uniformly chosen vertices in the largest connected component of $\mathscr{G}_N^{\mathfrak{m},\delta}$, the picture is less complete. Obviously, upper bounds for typical distances are implicit in the corresponding bounds on the diameter. Lower bounds for typical distances often follow directly from the proof of the lower bound statement for the diameter, since a first moment method is usually used to show that in the network there are with high probability no vertices at all which lie within a certain distance of each other.

Asymptotic bounds for $d_N(U, V)$ for $\mathfrak{m} \geq 2$, $U, V \in \mathscr{G}_N^{\mathfrak{m},\delta}$ uniform			
range of δ	$(-\mathfrak{m}, 0)$	$\{0\}$	$(0, \infty)$
range of τ	$(2, 3)$	$\{3\}$	$(3, \infty)$
range of γ	$\left(\frac{1}{2}, 1\right)$	$\{\frac{1}{2}\}$	$\left(0, \frac{1}{2}\right)$
upper bound	$\frac{4}{|\log(\tau-2)|}\log\log N$	$\frac{\log N}{\log\log N}$	$O(\log N)$
lower bound	$\frac{4}{|\log(\tau-2)|}\log\log N$	$\frac{\log N}{\log\log N}$	$O(\log N)$
published in	[DMM12]	[BR04]	[DHH10]

The problem of bounding the diameter is often somewhat more challenging than finding the typical distance. However, the main reason why we focus in our main results solely on the typical distances is that the typical distance has much more practical relevance as

a statistics of real world networks. Whereas the diameter can be influenced by one single extremely long path, the typical distances are much more robust under alteration of small subsets of the network.

Besides the classical and the sublinear model, there are many more ways of modelling preferential attachment, e.g. the model suggested in [CL06, Chapter 5]. Some authors work with very general models, which also include other effects like edge duplication or depletion, e.g. in [CF03]. From the viewpoint of applications, it is of course desirable to include as many effects as possible. From a mathematical point of view however, the models also need to be tractable. To obtain rigorous results beyond the asymptotics of the degree sequence, many of those models are too complicated as they have usually been designed with computer simulations of real world networks in mind.

Nevertheless, there has also been recent progress in deriving results for some extensions of pure preferential attachment models which include other formation principles. This success is partly due to a better understanding of preferential attachment on its own. Some mathematically interesting effects, like condensation, see [BB01] for an early treatment, require a combination of preferential attachment and other effects. One such way in which this can be realised is assigning to each vertex a fitness value, understood as a position in some abstract geometric space. This fitness then has an influence on the connectivity. Rigorous results about such models are obtained, e.g. in [Jor13; DO13; JM12]. These 'geometric' preferential attachment models lie outside the framework of this thesis and rigorous distance results are not available at present.

1.4   Other power law random networks

To provide an idea of the wider context of our results, we briefly discuss distance estimates for some well studied network models and give a short account of some related results.

Inhomogeneous random graphs are a class of (generally non-dynamic) models in which edges are conditionally independent. If $\mathbb{P}\{u \leftrightarrow v \text{ in } \mathcal{G}_N\} = \Phi_N(u)\Phi_N(v), \; u, v \in [N]$ for some function Φ_N on the set of vertices, then this is called the *rank 1* case. Bollobás, Janson and Riordan in [BJR07] provide diameter and distance results for a very general formulation of the model which covers power law graphs with $\tau > 3$. For the regime $\tau \in (2,3]$ there are several models suggested in the literature. The monograph [Hof13] provides a comprehensive overview including detailed proofs of distance results. Two inhomogeneous random graph models, the *expected degree random graph* of Chung and Lu (see e.g. [CL06; CL02]) and the *Poissonian random graph* of Norros and Reittu discussed in [NR06] appear as examples in Section 2.2, thus we give an account of some known results below.

The exact formulation of the Chung-Lu model is given in Example 2.10. The main parameter is a sequence of weights $w = (w_k)_{k \in [N]}$ which, if chosen to obey a power law with exponent $\tau > 2$, produces a power law network.

Proposition 1.15 (Average distances in the Chung-Lu model,[CL06, Theorem 7.7])**.** *Let* $\mathcal{H} = ([N], E(\mathcal{H}))$ *be a random graph with given expected degrees* $w = (w_1, \ldots, w_N)$*, where w satisfies*

  (i) for some $\varepsilon > 0$*, independent of N, we have*

$$\frac{\sum_{k \in [N]} w_k}{N} > 1 + \varepsilon,$$

  (ii) for $M(N) = \max_{i \in N} w_i$ *we have*

$$\frac{\log N}{\log \log N} = o(M(N)),$$

  (iii) w follows a power law with exponent $\tau \in (2, 3)$*.*

Then for U, V chosen uniformly and independently from $[N]$

$$d_N(U, V) \le (2 + o(1)) \frac{\log \log N}{\log \frac{1}{\tau - 2}} \quad \text{with high probability as } N \to \infty.$$

This result is the first rigorous mathematical result about distances in a power law random network and appeared in [CL02]. The proof in [CL06] includes a comparison of the set $\{v \in [N] : Z[v, N] \ge N^{(\log \log N)^{-1}}\}$ to an Erdős-Rényi graph, which combined with the technique of [DHH10] provided a model for the arguments of Section 4.6.

We are especially interested in the case where $w_i \sim c\sqrt{\frac{N}{i}}$ for some $c > 0$, corresponding to a strict power law with $\tau = 3$, in which the distances are of order $O\left(\frac{\log N}{\log \log N}\right)$. Before we state an exact result for this case in a more general framework, we leave the setting of the Chung-Lu model and briefly mention another closely related model, the Poissonian random graph. The Poissonian random graph allows a convenient representation as a *dynamic* network. The precise formulation is given in Example 2.12. The main difference to the Chung-Lu model is that the vertex weights $w = (w_k)_{k \in [N]}$ are replaced by random capacities $W_k, 1 \le k \le N$, and the number of edges between $v, w \in [N]$ is Poisson with parameter $\frac{W_v W_w}{\sum_{i \in [N]} W_i}$, so the model produces a graph with multiple edges.

In the monograph [Hof13] a host of previous results are summarised into a general framework for inhomogeneous random graph models specified by random or deterministic weight sequences. The distance result [Hof13, Theorem 9.4] includes the Norros-Reittu model and Proposition 1.15 as special cases. Consequently, the distances in the Norros-

Reittu model agree with the distances in the Chung-Lu model for $\tau \in (2,3)$.

For $\tau = 3$ the situation is similar, i.e. both models obey the same scaling of typical distances and the following result holds:

Proposition 1.16 (Inhomogeneous random graphs for $\tau = 3$, [Hof13, Theorem 9.22]). *Let $w = (w_i)_{i \in \mathbb{N}}$ be a sequence of non-negative weights specified by its distribution function*

$$F_N(x) = \frac{1}{N} \sum_{i \in [N]} \mathbb{1}_{\{w_i \le x\}}, \quad x \ge 0.$$

Let also $(W_N)_{N \ge 1}$, be random variables with distribution functions F_N, satisfying

(i) *there is a random variable W with distribution function F such that $\lim_{N \to \infty} F_N(x) = F(x)$ for all continuity points x of F,*

(ii) $\lim_{N \to \infty} \mathbb{E} W_N = \mathbb{E} W > 0$ *for some random variable W with distribution function F.*

Additionally, let the distribution of W follow a strict power law with exponent $\tau = 3$, i.e. there exist $C > c > 0$ and $a > 0$ such that

(a) *for all $x \le N^a$, we have $(1 - F_N)(x) \ge \frac{c}{x^2}$,*

(b) *for all $x \ge 0$, we have $(1 - F_N)(x) \le \frac{C}{x^2}$.*

Then, for uniformly and independently chosen $U, V \in [N]$, conditionally on being connected,

$$d_N(U, V) \sim \frac{\log N}{\log \log N} \quad \text{in probability, as } N \to \infty.$$

It should be mentioned at this point that the corresponding situation for preferential attachment treated in Theorem II is more general than the setup here, since we do not restrict ourselves to the strict definition of a power law adopted in Proposition 1.16.

The *configuration model* of Example 2.15 has been studied by several authors, starting with [BC78]. For the exact specifications of the configuration model, see Example 2.15. The defining parameter of the configuration model $(\mathcal{G}_N(D)_{N \ge 1})$ in the set up we wish to address here is the i.i.d. sequence $D = (D_i)_{i \in \mathbb{N}}$ of degrees in the network. Remarkably detailed bounds on typical distances are derived in [HHZ07], see below. In general, the distances in a given power law configuration model have mutas mutandis the same scaling as the distances in the inhomogeneous random graph. However, in the configuration model even the distribution of the fluctuation around the mean distance is known.

Proposition 1.17 (Fluctuations of typical distances, [Hof13, Theorem 10.46]). *Let $\tau \in (2,3)$ and let $D = D_1, D_2 \dots$ i.i.d. with distribution function F such that there exist $\eta \in [0,1), C < \infty$ with*

$$x^{-\tau + 1 - C(\log x)^{\eta - 1}} \le 1 - F(x) \le x^{-\tau + 1 + C(\log x)^{\eta - 1}}, \quad \text{for all sufficiently large } x.$$

Then, $\mathscr{G}_N(D)$ satisfies that there exist random variables $(R_a)_{a \in (-1,0]}$ such that, for $l \in \mathbb{N}_0$,

$$\mathbb{P}\left\{ d_N(U,V) = 2\left\lfloor \frac{\log\log N}{\log\frac{1}{\tau-2}} \right\rfloor + l \,\middle|\, d_N(U,V) < \infty \right\} = \mathbb{P}\{R_{a_N} = l\} + o(1), \quad as \ N \to \infty,$$

where

$$a_N = \left\lfloor \frac{\log\log N}{\log\frac{1}{\tau-2}} \right\rfloor - \frac{\log\log N}{\log\frac{1}{\tau-2}} \in (-1,0].$$

Moreover, the random variables $(R_a)_{a \in (1,0]}$ have an explicit interpretation in terms of the height of a tree associated to the network.

The results stated so far suggest that if the power law exponent τ is in the range $(2,3)$, then typical distances grow doubly logarithmic and that the constants $\frac{2}{|\log(\tau-2)|}$ in front of the double logarithm agree for all models discussed in this section. Even for the classical preferential attachment model, this has been the conjectured asymptotic. However, it turns out, see Chapter 2, that for $\tau \in (2,3)$ preferential attachment models have typical distances which are twice as long as the shortest paths in models without preferential attachment.

For the critical case of $\tau = 3$, it is reasonable to assume that exact asymptotics depend sensitively on the model specification and that the $\frac{\log N}{\log\log N}$ scaling breaks down, as soon as one moves away from the 'strict' power law setting. However, this case had not been treated for any of the models described above and only very recently a result based on *path counting* techniques has been established for general inhomogeneous random graphs following a strict power law, see [Hof13, Theorem 9.22]. The proof in [Hof13] seems to allow for a straightforward adaptation to the configuration model and also for an extension away from the strict power law setting in the spirit of our Theorem II. The path counting approach is different from the arguments presented in Chapter 6 to prove Theorem II, which are novel.

Chapter 2

Results

We now present our two main results for the sublinear preferential attachment model, Theorems I and II in Section 2.1. In Section 2.2 we demonstrate the consequences of our model independent approach to prove Theorem I by discussing the universality of the typical distances by means of several examples of complex network models.

2.1 Main theorems and discussion

In the doubly logarithmic regime, the typical distances only depend on the value of $\gamma \in \left(\frac{1}{2}, 1\right)$. The exact asymptotics are described in our first theorem.

Theorem I. *Let f be a concave attachment rule with $\gamma \in \left(\frac{1}{2}, 1\right)$, then, for $U, V \in \mathscr{C}_N$ chosen independently and uniformly,*

$$d_N(U, V) \le 4 \frac{\log\log N}{\log\left(\frac{\gamma}{1-\gamma}\right)} + O(1), \quad \textit{with high probability as } N \to \infty \tag{2.1}$$

and

$$d_N(U, V) \ge (4 + o(1)) \frac{\log\log N}{\log\left(\frac{\gamma}{1-\gamma}\right)}, \quad \textit{with high probability as } N \to \infty.$$

Moreover, if f is affine, then equality holds in (2.1).

Remark 2.1. Note that if f is affine, then the result implies that the distributions $(\nu_N)_{N \ge 1}$,

$$\nu_N = \mathscr{L}\left(d_N(U, V) - 4 \frac{\log\log N}{\log\left(\frac{\gamma}{1-\gamma}\right)}\right)$$

are tight. This paves the way for the investigation of accumulation points in the spirit of Proposition 1.17.

The lower bounds in Theorem I are special a case of a more general result for all network models which satisfy the following Assumption 2.2. They are in the same universality class as the sublinear preferential attachment model and are thus called *PA-type models*.

Assumption 2.2. For some $0 < \gamma < 1$, there exists $\kappa < \infty$, such that, for all N and pairwise distinct vertices $v_0, \dots, v_\ell \in [N]$,

$$\mathbb{P}\{v_0 \leftrightarrow v_1 \leftrightarrow v_2 \leftrightarrow \cdots \leftrightarrow v_\ell\} \le \prod_{k=1}^{\ell} \kappa\, (v_{k-1} \wedge v_k)^{-\gamma} (v_{k-1} \vee v_k)^{\gamma-1}.$$

In preferential attachment models with power law exponent τ, Assumption 2.2 is typically satisfied for all $\gamma > \frac{1}{\tau-1}$. Hence we expect these networks to be ultrasmall if and only if $\frac{1}{2} < \gamma < 1$. Theorem 2.3, one of the main results used to prove Theorem I, gives a lower bound on the typical distance under this assumption.

Theorem 2.3 ([DMM12, Theorem 2]). *Let $(\mathcal{G}_N)_{N \in \mathbb{N}}$ be a dynamic network model that satisfies Assumption 2.2 for some $\frac{1}{2} < \gamma < 1$. Then, for random vertices U and V chosen independently and uniformly from $[N]$, we have*

$$d_N(U, V) \ge 4 \frac{\log\log N}{\log\left(\frac{\gamma}{1-\gamma}\right)} + O(1), \quad \text{with high probability as } N \to \infty.$$

Not only can Theorem 2.3 be applied to the sublinear model of Section 1.3, but also to the affine PA graph with fixed outdegree discussed in Section 1.4, see Example 2.7. In all variations of the affine PA model, Assumption 2.2 is satisfied for all $\gamma > (\tau - 1)^{-1}$, and thus Theorem 2.3 implies that

$$d_N(U, V) \ge \left(4 + o(1)\right) \frac{\log\log N}{-\log(\tau - 2)}, \quad \text{with high probability as } N \to \infty.$$

The difference in the error bound is due to the fact that only the sublinear PA model admits a rigorous branching process approximation for local neighbourhoods, see Sections 3.2 and 4.3. The error term of constant order in Theorems I and 2.3 corresponds to the height of a random tree, as becomes apparent in the proofs of Theorem 2.6 and Proposition 2.8. Those proofs also reveal that matching upper bounds usually require a more model specific approach.

The proof of Theorem 2.3 and a comparison of Assumption 2.2 to the path probabilities in inhomogeneous random graph models of rank one (see also [BJR07] for the general framework, which covers much more than the rank one case, but no power laws with $\tau \ge 3$) suggest a version of the lower bound result for models with a different connectivity structure, like the inhomogeneous random graph models of Section 1.4 and the configuration model. We call them *fitness type models* and require only that they satisfy the following assumption, which is strictly weaker[1] than Assumption 2.2.

Assumption 2.4. For some $\frac{1}{2} < \gamma < 1$, there exists $\kappa < \infty$, such that, for all N and pairwise

---

[1] This follows from $n^{\gamma-1} m^{-\gamma} \le n^{-\gamma} m^{-\gamma} N^{2\gamma-1}$ for all $n \ge m \in [N]$.

distinct vertices $v_0, \ldots, v_\ell \in [N]$,

$$\mathbb{P}\{v_0 \leftrightarrow v_1 \leftrightarrow v_2 \leftrightarrow \cdots \leftrightarrow v_\ell\} \leq \prod_{k=1}^{\ell} \kappa \, v_{k-1}^{-\gamma} \, v_k^{-\gamma} \, N^{2\gamma-1}.$$

Under this assumption, the following theorem holds.

Theorem 2.5 ([DMM12, Theorem 1]). *Let $(\mathscr{G}_N)_{N \in \mathbb{N}}$ be a dynamic network model that satisfies Assumption 2.4 for some $\frac{1}{2} < \gamma < 1$. Then, for random vertices U and V chosen independently and uniformly from $[N]$, we have*

$$d_N(U, V) \geq 2 \frac{\log \log N}{\log\left(\frac{\gamma}{1-\gamma}\right)} + O(1), \quad \text{with high probability as } N \to \infty.$$

Together with the matching upper bounds, Theorems 2.3 and 2.5 indicate that PA- and fitness type models are in different universality classes. Examples of network models in which Theorem 2.5 can be applied, are mentioned in the introduction chapter and will be discussed in some more detail in Section 2.2. Examples 2.15, 2.10 and 2.12 in that section refer to this class. For instance, in the configuration model, the fitness of a vertex equals its degree and in the Chung-Lu model it is described by the weight sequence w. In all cases, if the degree distribution has power law exponent $\tau \in (2, 3)$, Assumption 2.4 is satisfied for all $\gamma > \frac{1}{\tau-1}$, and the theorem implies that

$$d_N(U, V) \geq \left(2 + o(1)\right) \frac{\log \log N}{-\log(\tau - 2)}, \quad \text{with high probability as } N \to \infty.$$

For all listed examples of fitness type models, matching upper bounds are known from the literature. Generally, these are models in which there is no dynamical build up of strong preference for certain vertices. Instead, every vertex receives an a priori fitness value which determines its likelihood to form future edges. In the PA-type models, the need to build up high connectivity by first accumulating connections leads to a *dissociative* network structure: high degree vertices are not very likely to be directly connected. Instead, they tend to be connected by a path of length 2 via one low degree vertex. On the other hand, in fitness type models one usually finds edges between high degree vertices, i.e. they are *associative*. This explains the extra factor of 2 in the distances for PA-type models.

We now focus on the upper bound in Theorem I, which only holds in this form for the sublinear PA model. In Section 5.4 it is shown to be a straightforward consequence of the following similar result for the affine case.

Theorem 2.6. *Let $f(k) = \gamma k + \beta$, where $\beta \in (0, 1)$ and $\gamma \in \left(\frac{1}{2}, 1\right)$. Then, for $U, V \in \mathscr{C}_N$ chosen independently and uniformly,*

$$d_N(U, V) \leq 4 \frac{\log \log N}{\log\left(\frac{\gamma}{1-\gamma}\right)} + O(1), \quad \text{with high probability as } N \to \infty.$$

The proof of Theorem 2.6 in Section 5.3 may be model specific, but is following in broad strokes the approach of [DHH10], which is in turn a PA version of the arguments presented in [NR06] and [HHZ07]. There are also similarities to the upper bound of Theorem II, which is formulated below. In the affine PA model given in Example 2.7 we cannot obtain an error term of constant order, since the branching process approximating local neighbourhoods does not allow for an accessible representation.

The second main result of this work is concerned with $\gamma = \frac{1}{2}$, in which case the resulting graph has a degree distribution which is a power law with exponent $\tau = 3$. We focus on a special choice of attachment rule to demonstrate that in the critical case the asymptotic distance is not universal and instead dependent on model parameters beyond the power law exponent. The scale of the distances depends crucially on the features of f.

Theorem II. *Let f be a concave attachment rule satisfying, for some $\alpha \geq 0$, $\eta > 0$,*

$$f(k) = \frac{1}{2}k + \frac{\alpha}{2}\frac{k}{\log k} + o\left(\frac{k}{(\log k)^{1+\eta}}\right), \quad k \geq 2.$$

Then, for U, V chosen independently and uniformly from $\mathscr{C}_N$,

$$d_N(U, V) \sim \frac{1}{1+\alpha}\frac{\log N}{\log\log N}, \quad \text{in probability as } N \to \infty.$$

The attachment rule f is chosen to create distances at the specific scale of the theorem, i.e. to influence the scaling constant in front of the $\frac{\log N}{\log\log N}$ term. Other choices result in different scales interpolating between the $\frac{\log N}{\log\log N}$ and the $\log\log N$ regime. However, in those cases the dependence of the scaling constant on the asymptotic order of the perturbation $g(k) = f(k) - \frac{1}{2}k$ is not as transparent.

In the case of affine attachment, i.e. $g(k) \equiv f(0)$, the typical distance between two randomly chosen vertices of the giant component is asymptotically $\frac{\log N}{\log\log N}$, in accordance with the results for the LCD model in [BR04].

For $\alpha > 0$, the result concerns a regime that has not been considered so far and the model shows a behaviour which can not be obtained with an affine attachment rule. Therefore, this behaviour is also absent in the affine models of Example 2.7

2.2 Example applications of Theorems 2.2 and 2.4

In this section we give four further examples to which Theorems 2.2 and 2.4 apply, corresponding to the best understood models of ultrasmall networks in the mathematical literature.

2.2.1 Affine preferential attachment

The first example is of preferential attachment type and is discussed using Theorem 2.3. The remaining three examples are of fitness type and are discussed using Theorem 2.5. This PA-type models are studied in the work of Hooghiemstra, van der Hofstad and co-authors. We base our discussion on the paper [DHH10], where three qualitatively similar models are considered, see also [Hof13] for a survey. We focus on the first model studied in [DHH10], which is most convenient to define, the two variants can be treated with the same method.

Example 2.7. The model depends on two parameters, an integer $\mathfrak{m} \geq 1$ and a real $\delta > -\mathfrak{m}$. Roughly speaking, in every step a new vertex is added to the network and connected to $\mathfrak{m}$ existing vertices with a probability proportional to their degree plus δ. Note that in the case $\mathfrak{m} = 1$ the network has the metric structure of a tree, making this a degenerate case of less interest. The case studied by Bollobás and Riordan [BR04] corresponds to $\delta = 0$ and $\mathfrak{m} \geq 2$ and leads to a network with $\tau = 3$ and typical distance $\frac{\log N}{\log\log N}$, so that it lies outside the class of ultrasmall networks.

We first generate a dynamic network model $(\mathscr{G}_N)$ for the case $\mathfrak{m} = 1$. By $Z[n, N]$, $n \leq N$, we denote the degree of vertex n in $\mathscr{G}_N$. We use the convention that self-loops add two towards the degree of the vertex to which they are attached.

- $\mathscr{G}_1$ consists of a single vertex, labelled 1, with one self loop.

- In each further step, given $\mathscr{G}_N$, we insert one new vertex, labelled $N + 1$, and one new edge into the network such that the new edge connects the new vertex to vertex $m \in [N]$ with probability

$$\mathbb{P}\{m \leftrightarrow N + 1 \,|\, \mathscr{G}_N\} = \frac{Z[m, N] + \delta}{N(2 + \delta) + 1 + \delta},$$

  or to itself with probability

$$\frac{1 + \delta}{N(2 + \delta) + 1 + \delta}.$$

To generalise the model to arbitrary values of $\mathfrak{m}$, we take the graph $\mathscr{G}'_{\mathfrak{m}N}$ constructed using parameters $\mathfrak{m}' = 1$ and $\delta' = \frac{\delta}{\mathfrak{m}}$, and merge vertices $\mathfrak{m}(k-1)+1, \ldots, \mathfrak{m}k$ in the graph $\mathscr{G}'_{\mathfrak{m}N}$ into a single vertex denoted k, keeping all edges. The degree sequence for the case $\mathfrak{m} = 1$ is studied in [Bol+01], the general case is covered by [Hof13, Theorem 8.2]. Using martingale methods one obtains convergence of the empirical degree distribution to the distribution

$$\mu_k = \begin{cases} 0, & \text{if } 0 \leq k < \mathfrak{m}, \\ \left(2 + \frac{\delta}{\mathfrak{m}}\right) \frac{\Gamma(k+\delta)}{\Gamma(\mathfrak{m}+\delta)} \frac{\Gamma(\mathfrak{m}+2+\delta(1+\frac{1}{\mathfrak{m}}))}{\Gamma(k+3+\delta(1+\frac{1}{\mathfrak{m}}))}, & \text{if } k \geq \mathfrak{m}. \end{cases}$$

Applying Stirling's' approximation we deduce that $(\mu_k)_{k \geq 0}$ is a power law with exponent $\tau = 3 + \frac{\delta}{\mathfrak{m}}$, hence we expect the model to be in the ultrasmall range if and only if $-\mathfrak{m} < \delta < 0$.

Proposition 2.8. *For independent, uniformly chosen vertices U and V in the largest connected component of the preferential attachment model with parameters $\mathfrak{m} \geq 2$, $-\mathfrak{m} < \delta < 0$, we have*

$$d_N(U, V) \sim 4 \frac{\log \log N}{-\log\left(1 + \frac{\delta}{\mathfrak{m}}\right)} \qquad \text{in probability.}$$

Remark 2.9. Note that since $\tau = 3 + \frac{\delta}{\mathfrak{m}}$, the constant is the same as in Theorem I.

The upper bound follows from the arguments in [DHH10], see the remark following Theorem 1.6 there. It is not explicitly derived in this paper, because the authors focus on the graph diameter. We give an outline of the derivation of the bound for the typical distance below, which is very similar to the arguments we use in the proof of Theorem 2.6. The paper [DHH10] leaves the problem of finding a matching lower bound open. We resolve this problem by verifying Assumption 2.2 for $\gamma = \left(2 + \frac{\delta}{\mathfrak{m}}\right)^{-1}$ and applying Theorem 2.3.

Proof of Proposition 2.8. For the lower bound, we look at $\mathfrak{m} = 1$ first. In this case, we have, for $1 \leq m < n \leq N$,

$$\mathbb{P}\{m \leftrightarrow n\} = \frac{\mathbb{E}Z[m, n-1] + \delta}{n(2 + \delta) - 1}. \tag{2.2}$$

It is easy to see that

$$\mathbb{E}\big[Z[m, n] + \delta \mid Z[m, n-1]\big] = \big(Z[m, n-1] + \delta\big) \frac{n(2 + \delta)}{n(2 + \delta) - 1},$$

and hence

$$\mathbb{E}\big[Z[m, n] + \delta\big] = (1 + \delta) \frac{\Gamma(n+1)\Gamma\big(m - \frac{1}{2+\delta}\big)}{\Gamma\big(n + \frac{1+\delta}{2+\delta}\big)\Gamma(m)}.$$

In particular there exist constants $0 < c < C$ such that

$$c\left(\frac{n}{m}\right)^{\frac{1}{2+\delta}} \leq \mathbb{E}Z[m, n] \leq C\left(\frac{n}{m}\right)^{\frac{1}{2+\delta}} \qquad \text{for all } 1 \leq m < n.$$

Combining this with (2.2) yields, for $\gamma = \frac{1}{2+\delta}$ and a suitable $\kappa_1 > 0$, that

$$\mathbb{P}\{m \leftrightarrow n\} \leq \frac{C\left(\frac{n}{m}\right)^{\gamma} + \delta}{n(2 + \delta) - 1} \leq \kappa_1 n^{\gamma - 1} m^{-\gamma} \quad \text{for all } 1 \leq m < n. \tag{2.3}$$

To verify Assumption 2.2, following [DHH10, Lemma 2.1], we observe that, if $\mathfrak{m} \geq 1$, for distinct vertices $v_0, \ldots, v_l$, all events of the form $\{v_{j-1} \leftrightarrow v_j \leftrightarrow v_{j+1}\}$ with $j \in \{1, \ldots, l-1\}$ and $v_j < v_{j-1}, v_{j+1}$, and all events $\{v_{j-1} \leftrightarrow v_j\}$ which are not part of these, are non-positively correlated, in the sense that the probability of all of them occurring is smaller than the product of the probabilities. In particular, if $\mathfrak{m} = 1$ at most one outgoing connection can be made per vertex and therefore the events are even mutually exclusive. Recalling also (2.3) it remains to show that for $m < v, w$,

$$\mathbb{P}\{v \leftrightarrow m \leftrightarrow w\} \leq \kappa_2 v^{\gamma - 1} w^{\gamma - 1} m^{-2\gamma}, \tag{2.4}$$

for some finite constant $\kappa_2 > 0$. To this end we let $\{(Z_n^{(k,m)})_{n \geq m} : k, m \in \mathbb{N}\}$ denote the collection of right-continuous Markov jump processes starting at $Z_{m-}^{(k,m)} = k$, jumping instantly at time m and subsequently at integer time-steps following the rule

$$\mathbb{P}\{Z_n^{(k,m)} = Z_{n-}^{(k,m)} + 1 \mid Z_{n-}^{(k,m)}\} = \frac{Z_{n-}^{(k,m)} + \delta}{n(2+\delta) - \delta} = 1 - \mathbb{P}\{Z_n^{(k,m)} = Z_{n-}^{(k,m)} \mid Z_{n-}^{(k,m)}\}.$$

Note that $(Z[m,n])_{n \geq m} = (Z_n^{(1,m)})_{n \geq m}$ in law and that, for $m < n$, the event $\{m \leftrightarrow n\}$ corresponds to $\{\Delta Z_n^{(1,m)} = 1\}$, where we write $\Delta Z_n^{(k,m)} := Z_n^{(k,m)} - Z_{n-}^{(k,m)}$. Note also that $Z_n^{(k_0,m)}$ is stochastically dominated by $Z_n^{(k,m)}$ for $k \geq k_0$. Hence, for $m < n_1 < n_2$,

$$\begin{aligned}
\mathbb{E}\big[Z_{n_2}^{(2,m)} \mid \Delta Z_{n_1}^{(2,m)} = 1\big] &= \sum_{j=2}^{n_2-m+2} \sum_{k=2}^{n_1-m+1} j\, \mathbb{P}\{Z_{n_2}^{(2,m)} = j \mid Z_{n_1-}^{(2,m)} = k, \Delta Z_{n_1}^{(2,m)} = 1\} \\
&\qquad \times \mathbb{P}\{Z_{n_1-}^{(2,m)} = k \mid \Delta Z_{n_1}^{(2,m)} = 1\} \\
&\leq \sum_{j=2}^{n_2-m+2} \sum_{k=2}^{n_1-m+1} \frac{j\, \mathbb{P}\{Z_{n_2}^{(k+1,n_1)} = j\}\,(k+\delta)\,\mathbb{P}\{Z_{n_1-}^{(2,m)} = k\}}{(n_1(2+\delta) + 1 + \delta)\,\mathbb{P}\{\Delta Z_{n_1}^{(2,m)} = 1\}} \\
&= \sum_{k=2}^{n_1-m+1} \frac{(k+\delta)\,\mathbb{P}\{Z_{n_1-}^{(2,m)} = k\}\,\mathbb{E}Z_{n_2}^{(k+1,n_1)}}{(n_1(2+\delta) + 1 + \delta)\,\mathbb{P}\{\Delta Z_{n_1}^{(2,m)} = 1\}}.
\end{aligned}$$

As in the derivation of (2.3) the expectation in the last line can be bounded from above by $c_0(k+1)n_2^\gamma n_1^{-\gamma}$, for some $c_0 > 0$. Similarly, we obtain $\mathbb{P}\{\Delta Z_{n_1}^{(2,m)} = 1\} \geq c_1 n_1^{\gamma-1} m^{-\gamma}$ and

$$\mathbb{E}\big[(Z_{n_1-}^{(2,m)})^2\big] \leq c_2\, m^{-\frac{2}{2+\delta}} n_1^{\frac{2}{2+\delta}},$$

for further constants $c_1, c_2 > 0$. Summarising, we obtain

$$\mathbb{E}\big[Z_{n_2}^{(2,m)} \mid \Delta Z_{n_1}^{(2,m)} = 1\big] \leq c_3\, n_2^\gamma n_1^{-2\gamma} m^\gamma \sum_{k=2}^{m-n_1+1} k^2 \mathbb{P}\{Z_{n_1-}^{(2,m)} = k\} \leq c_4\, n_2^\gamma m^{-\gamma},$$

for some $c_3, c_4 > 0$, and this establishes (2.4). Finally, passing from $\mathfrak{m} = 1$ to general $\mathfrak{m}$ can be achieved by a simple union bound.

For the upper bound we work directly in the graph $\mathcal{G}_{2N}$ with $\mathfrak{m} \geq 2$ and $\delta \in (-\mathfrak{m}, 0)$. Using the terminology of [DHH10], we define the *core* of $\mathcal{G}_{2N}$ to be

$$\text{core}_N = \big\{ m \in [N] : Z[m,N] \geq (\log N)^\sigma \big\},$$

where $\sigma > -\frac{\mathfrak{m}}{\delta}$. [DHH10, Theorem 3.1] states that the diameter of the core in $\mathcal{G}_{2N}$ is bounded by $(4 + o(1)) \log \log N |\log(1 + \frac{\delta}{\mathfrak{m}})|^{-1}$, thus all we need to show is that, for fixed $\varepsilon > 0$, a uniformly chosen vertex $V \in [\lfloor(2-\varepsilon)N\rfloor]$ can be connected to the core using no more then $o(\log \log N)$ edges in $\mathcal{G}_{2N}$. This is done in two steps.

For the first step we explore the neighbourhood of V in $\mathcal{G}_M$, for $M = \lfloor(2-\varepsilon)N\rfloor$, until we

find a vertex w with degree $Z[w, N] \geq u_0$, where u_0 will be determined below. Denote by $\Gamma_k, k \geq 0$ the set of all vertices in $\mathscr{G}_M$ that can be reached from V using exactly k different edges from $\mathscr{G}_M$. If we fix $u \in \mathbb{N}$ and set $T_u^{(V)} = \min\{k : \Gamma_k \cap \{n : Z[n, N] \geq u\} \neq \emptyset\}$, then we wish to verify that we can find a large constant $C_{u,\varepsilon} > 0$, such that

$$\mathbb{P}\{T_u^{(V)} > C_{u,\varepsilon}\} < \varepsilon, \tag{2.5}$$

if N is sufficiently large. This can be done straightforwardly along the lines of the proof of [DHH10, Theorem 3.6]. It is shown there that the neighbourhood exploration around V using only *outgoing* edges does w.h.p. not encounter more than one circle in the first few steps and is therefore similar to an $\mathfrak{m}$-ary tree. The sum of all degrees in the tree is of the same order as its size. It follows, using this fact and that the sum of the degrees of the vertices in $\{v : Z[v, N] \geq u\}$ is large, that there is w.h.p. a two-step connection between the tree and the set $\{v : Z[v, N] \geq u\}$, provided the tree is grown up to a sufficiently large height which depends only on u, but not on N.

The second step is now to show that any vertex w satisfying $Z[w, N] \geq u_0$, for sufficiently large u_0, can be joined to the core by using $O(\log\log\log N)$ edges. To this end we apply [DHH10, Lemma A.1], which provides lower bounds for the sum of all degrees of vertices exceeding a certain minimum degree, like in the proof of [DHH10, Proposition 3.3], where a path connecting high degree vertices is constructed by using low degree vertices from the set $\{\lceil(2-\varepsilon)N\rceil, \ldots, 2N\}$. We obtain that for any vertex $j \in \mathscr{M} = \{\lceil(2-\varepsilon)N\rceil, \ldots, 2N\}$ and any vertex $a \in [N]$ with $Z[a, N] \geq u_a$,

$$\mathbb{P}\{j \leftrightarrow a, j \leftrightarrow b \text{ for some } b \text{ with } Z[b, N] \geq u_b | \mathscr{G}_M\} \geq \frac{c u_a u_b^{-(1+\frac{\delta}{\mathfrak{m}})}}{N},$$

for a positive constant c, with probability exceeding $1 - o(N^{-1})$ and as long as both u_a and u_b do not exceed a small power of N. Hence, with the same probability,

$$\mathbb{P}\{\nexists j \in \mathscr{M} : j \leftrightarrow a, j \leftrightarrow b \text{ for some } b \text{ with } Z[b, N] \geq u_b | \mathscr{G}_M\}$$
$$\leq \left(1 - \frac{c u_a u_b^{-(1+\frac{\delta}{\mathfrak{m}})}}{N}\right)^{\#\mathscr{M}} \leq \exp\left(2c\varepsilon u_a u_b^{-(1+\frac{\delta}{\mathfrak{m}})}\right). \tag{2.6}$$

Starting from the initial vertex w with $Z[w, N] \geq u_0$ and defining for $k \geq 1$,

$$u_{k+1} = \left(\frac{\varepsilon c u_k}{(\log(k+1) - \frac{1}{2}\log\varepsilon)}\right)^{\frac{1}{1+\frac{\delta}{\mathfrak{m}}}}, \tag{2.7}$$

it is straightforward to check that, for $u_a = u_k$ and $u_b = u_{k+1}$, the right hand side of (2.6) equals $\varepsilon(k+1)^{-2}$. Summing over these error bounds, we therefore obtain that (2.7) defines an increasing sequence $(u_k)_{k=0}^K$ of lower bounds on degrees at time N, for which we have

assured that with probability at least $1 - \frac{\pi^2}{6}\varepsilon$ there is a path of length $2K$ which alternates between high degree vertices and vertices from $\mathcal{M}$ and connects w to a vertex of degree u_K. The recursive definition (2.7) implies that

$$\log u_K \geq \frac{1}{\left(1 + \frac{\delta}{\mathfrak{m}}\right)^K} (\log u_0 - C_\varepsilon),$$

for some large $C_\varepsilon > 0$, thus, if $u_0 = \exp(2C_\varepsilon)$, we can connect w to a vertex belonging to core_N by choosing $K \geq D_{\sigma,\varepsilon} \log\log\log N$, where $D_{\sigma,\varepsilon} > 0$ depends only on σ and ε. Fixing $u = \exp(2C_\varepsilon)$ in (2.5) and starting the above construction in $u_0 = u$, we obtain that for a uniformly chosen vertex $V \in \mathcal{G}_{2N}$,

$$\mathbb{P}\{d_{2N}(V, \mathrm{core}_N) > 2D_{\sigma,\varepsilon} \log\log\log N + C_{e^{2C_\varepsilon},\varepsilon}\} \leq \left(2 + \frac{\pi^2}{6}\right)\varepsilon,$$

if N is sufficiently large, showing that the diameter of core_N is the dominating contribution to typical distances in $\mathcal{G}_{2N}$. $\qquad\square$

2.2.2 Inhomogeneous random graphs of rank one

The next two examples belong to the wide class of inhomogeneous random graphs of rank one, whose essential feature is the independence between different edges. The following model of a *random graph with given expected degrees* is studied in the work of Chung and Lu, see [CL02] or [CL06] for a survey.

Example 2.10. In its general form the model depends on a triangular scheme $w_1^{(N)}, \ldots, w_N^{(N)}$ of positive weights, where the weight $w_i^{(N)}$ plays the role of the expected degree of vertex i in $\mathcal{G}_N$. The model is defined by the following two requirements:

- for every pair (i, j) with $1 \leq i \neq j \leq N$ the events $\{i \leftrightarrow j\}$ are independent,

- for every pair (i, j) with $1 \leq i \neq j \leq N$ we have

$$\mathbb{P}\{i \leftrightarrow j\} = \frac{w_i^{(N)} w_j^{(N)}}{\ell_N} \wedge 1, \qquad \text{where } \ell_N := \sum_{i=1}^{N} w_i^{(N)}.$$

Proposition 2.11. *For independent, uniformly chosen vertices U and V in the largest connected component of the expected degree random graph with weights satisfying*

$$c\left(\frac{N}{i}\right)^\gamma \leq w_i^{(N)} \leq C\left(\frac{N}{i}\right)^\gamma \quad \text{for all } 1 \leq i \leq N,$$

for some $\gamma > \frac{1}{2}$ and constants $0 < c \leq C$, we have

$$d_N(U, V) \sim 2\frac{\log\log N}{\log\left(\frac{\gamma}{1-\gamma}\right)} \qquad \text{in probability.}$$

30

Proof. The upper bound is Proposition 1.15. For the lower bound we have to check Assumption 2.4. Note that, using the upper bound on the weights,

$$\mathbb{P}\{i \leftrightarrow j\} \leq \frac{w_i^{(N)} w_j^{(N)}}{\ell_N} \leq C^2 \frac{N^{2\gamma}}{\ell_N} (ij)^{-\gamma}.$$

From the lower bound on the weights we get that $\ell_N \geq cN$, for some $c > 0$, and hence $\mathbb{P}\{i \leftrightarrow j\} \leq \kappa N^{2\gamma-1} i^{-\gamma} j^{-\gamma}$ for a suitable κ. Using the independence assumption we see that Assumption 2.4 holds, and the lower bound follows from Theorem 2.5. $\qquad\square$

The second example of this section is the *conditionally Poissonian random graph* as studied in the work of Norros and Reittu, see [NR06].

Example 2.12. The model is based on drawing an independent, identically distributed sequence $W_1, W_2, \dots$ of positive capacities. Conditional on this sequence, the dynamical network model is constructed as follows:

- $\mathcal{G}_1$ consists of a single vertex, labelled 1, and no edges.

- Given $\mathcal{G}_N$, to obtain $\mathcal{G}_{N+1}$ we insert one new vertex, labelled $N+1$, and independently for any $m \in [N]$ we introduce a random number of edges between $N+1$ and m according to a Poisson distribution with parameter

$$\frac{W_i W_{N+1}}{L_{N+1}} \qquad \text{for } L_n := \sum_{k=1}^n W_k.$$

Then we remove each edge in $\mathcal{G}_N$ independently with probability $1 - \frac{L_N}{L_{N+1}}$.

Recall that having possibly several edges between two vertices has no relevance for the typical distances in the giant component. In order to be in the ultrasmall regime we require the law of the capacities to be power laws with exponent $2 < \tau < 3$.

Proposition 2.13. *Assume that the capacities in the conditionally Poissonian random graph satisfy*

$$\mathbb{P}\{W_1 > x\} = x^{1-\tau}(c + o(1)) \quad \text{*for all sufficiently large x,*}$$

where $2 < \tau < 3$ and $c > 0$ is constant. For independent, uniformly chosen vertices U and V in the largest connected component we have

$$d_N(U, V) \sim 2 \frac{\log\log N}{-\log(\tau - 2)} \qquad \text{*in probability.*}$$

Remark 2.14. The upper bound is proved in [NR06, Theorem 4.2], where it is also shown that a giant component exists. For the lower bound we verify Assumption 2.4 for $\gamma = \frac{1}{\tau-1}$ and apply Theorem 2.5.

Proof. We check that Assumption 2.4 holds conditionally with high probability, given the capacities. For fixed N we put the capacities in decreasing order

$$W_N^{(1)} > W_N^{(2)} > \cdots > W_N^{(N)}$$

and relabel the vertices so that the jth vertex has weight $W_N^{(j)}$. It follows from the definition of the model (see [NR06, Proposition 2.1] for a formal derivation) that the number of edges between vertices i and j in $\mathscr{G}_N$ is Poisson distributed with parameter $\frac{W_N^{(i)} W_N^{(j)}}{L_N}$. As the edges are conditionally independent we only have to verify that, given $\varepsilon > 0$ there exists $\kappa > 0$ such that

$$1 - \exp\left(-\frac{W_N^{(i)} W_N^{(j)}}{L_N}\right) \le \kappa N^{2\gamma-1} i^{-\gamma} j^{-\gamma} \quad \text{for all } 1 \le i < j \le N, \tag{2.8}$$

with probability $\ge 1 - 2\varepsilon$. By the law of large numbers L_N is of order N, so that it suffices to establish $W_N^{(i)} \le \kappa \left(\frac{N}{i}\right)^{\gamma}$ for all $1 \le i \le N$. To this end, we denote by $S_N^{(i)}$ the number of potential values exceeding $\kappa \left(\frac{N}{i}\right)^{\gamma}$. The random variable $S_N^{(i)}$ is binomially distributed with parameters N and $p := \mathbb{P}\{W_1 > \kappa \left(\frac{N}{i}\right)^{\gamma}\} \le c(\kappa) \frac{i}{N}$, where $c(\kappa) \to 0$ for $\kappa \to \infty$. By Bernstein's inequality, see e.g. [Ben62, (8)],

$$\mathbb{P}\{S_N^{(i)} > 2i\} \le \exp\left[\frac{-i^2}{2\text{Var}(S_N^{(i)}) + \frac{2i}{3}}\right] \le e^{-\frac{3}{8}i}, \quad \text{if } c(\kappa) < 1.$$

Hence we may choose M large enough so that $\sum_{i=M}^{\infty} \exp(-\frac{3}{8}i) < \varepsilon$, ensuring that with probability exceeding $1 - \varepsilon$ we have $W_N^{(2i)} \le \kappa \left(\frac{N}{i}\right)^{\gamma}$ for all $i \ge M$. It remains to give bounds on $W_N^{(1)}, \ldots, W_N^{(2M)}$. By a standard Poisson approximation result, e.g. [Res08, Proposition 3.21], we note that for any $1 \le i \le 2M$, we have that $S_N^{(i)}$ converges weakly to a Poisson distribution with parameter $\lambda := \lim_{N\to\infty} N\mathbb{P}\{W_1 > \kappa \left(\frac{N}{i}\right)^{\gamma}\} \le 2c(\kappa)M$, and hence, by choosing κ large, we can ensure that for large N, we have $\sum_{i=1}^{2M} \mathbb{P}\{S_N^{(i)} > i\} \le \varepsilon$, which completes the proof. $\qquad\square$

2.2.3 Configuration model

Another model which falls in the universality class of fitness models are random networks with fixed degree sequence, also known as the *configuration model*. This model is well studied and very detailed results on average distances in the case of power laws with exponent $\tau \in (2,3)$ are obtained, in particular by van der Hofstad et al. in [HHZ07]. The idea behind this class of models is to enforce a particular power-law exponent by fixing the degree sequence of the network in a first step.

Example 2.15. We choose a sequence $D_1, D_2, \ldots$ of independent and identically distributed random variables with values in the non-negative integers. For given N we assume that

$$L_N := \sum_{j=1}^{N} D_j$$

is even, which may be achieved by replacing D_N by $D_N - 1$ if necessary. Thus, with $D_1, \dots, D_N$ given, we construct the network $\mathcal{G}_N = \mathcal{G}_N(D)$ as follows:

- To any vertex $m \in [N]$ we attach D_m half-edges or stubs.

- The L_N stubs are given an (arbitrary) order.

- We start by pairing the first stub with a (uniformly) randomly chosen other stub, and continue pairing the lowest numbered unpaired stub with a remaining randomly chosen stub until all stubs are matched.

- Any pair of stubs is connected to form an edge.

Obviously the resulting network can have self-loops and double edges, but this has no relevance for the typical distances in the giant component. In order to be in the ultrasmall regime we require the law of the degrees to be a power law with exponent $2 < \tau < 3$.

Proposition 2.16. *Assume that there exists $c > 0$ such that*

$$\mathbb{P}\{D_1 > x\} = x^{1-\tau}\left(c + o(1)\right), \quad \text{for all sufficiently large } x.$$

For independent, uniformly chosen vertices U and V in the largest connected component we have

$$d_N(U, V) \sim 2\frac{\log\log N}{-\log(\tau - 2)} \quad \text{in probability.}$$

Remark 2.17. This and much more is proved in [HHZ07, Theorem 1.2]. For an alternative approach to the lower bound we now verify Assumption 2.4 for any $\gamma < 1/(\tau - 1)$ and paths of length up to $\ell = O(\log\log N)$, which is clearly sufficient to apply Theorem 2.5.

Proof of Proposition 2.16. We observe that, given $D_1, \dots, D_N$, for pairwise disjoint vertices $v_1, \dots, v_{\ell+1}$,

$$\mathbb{P}\left\{v_\ell \leftrightarrow v_{\ell+1} \mid v_1 \leftrightarrow v_2 \leftrightarrow \cdots \leftrightarrow v_{\ell-1} \leftrightarrow v_\ell\right\} \leq \frac{D_{v_\ell} D_{v_{\ell+1}}}{L_N - 2\sum_{k=1}^{\ell} D_{v_k}},$$

where the denominator is a rough lower bound on the number of stubs unaffected by the conditioning event. In particular, $\mathbb{P}\{i \leftrightarrow j\} \leq \frac{D_i D_j}{L_N - 2D_i}$. Using the law of large numbers one can easily see that there is a $c > 0$ such that

$$L_N - 2\sum_{k=1}^{\ell} D_{v_k} \geq cN \quad \text{with high probability,}$$

for any choice of $v_1, \dots, v_\ell$, if $\ell = O(\log\log N)$. Therefore, to verify Assumption 2.4 we only need to find appropriate bounds on the degrees of given vertices, which can be achieved (using the same relabelling) by a similar argument as in Example 2.12. $\qquad\square$

Chapter 3

Main ideas of proofs

This chapter explains the main concepts and ideas behind the proofs. We strive to limit technicalities while also providing some background for the proofs of the following chapters. Central for all distance estimations is the growth of the *neighbourhood shells* $\Gamma^k(v)$ around a vertex v,

$$\Gamma^k(v) = \{i \in \mathcal{G}_N : d_N(v, i) = k\} \subset \mathcal{G}_N, \quad \text{for } k \geq 0,$$

which determine the growth of the *k-neighbourhoods* $\Gamma^{\leq k}(v) = \bigcup_{i \leq k} \Gamma^i(v)$. Rough a priori estimates of the expected growth of the neighbourhoods are the subject of Section 3.1.

For upper bounds on the distances it is necessary to determine the exact growth of the neighbourhoods. We follow [DM13] and use an *exploration process*, which successively collects information about the connected component of a fixed vertex v. It starts in v and uncovers new vertices by following edges incident to v in a well-defined order. It then follows edges incident to the newly discovered vertices and so on. The crucial step is then to couple this exploration process to the exploration process in a suitably defined tree, which turns out to be the INT of Definition 1.11. The precise setup is rather intricate and takes up most of Chapter 4. An outline of the main ideas is given in Section 3.2.

The coupling of exploration processes fails as soon as the neighbourhoods become too large – a true network contains cycles after all. To ensure that already explored large subsets of the preferential attachment network are connected or that very dense subsets have small diameters we use some tools which are adaptations of techniques from classical random graph theory and briefly discussed in Section 3.3.

3.1 Truncated first and second moments

The fact that power law distributions with $\tau \leq 3$ have infinite variance is reflected in the random networks by the difficulties that we encounter when we try to establish bounds for the growth of the k-neighbourhoods. A natural approach to avoid these difficulties is to

truncate the range of the exploration.

3.1.1 A truncated first order method – the lower bound of Theorem I

The proof of the lower bounds in Theorems 2.3 and 2.5 is based on a *constrained* or *truncated first moment method*, which we now briefly explain. We start with an explanation of the (unconstrained) first moment bound and its shortcomings. Let v, w be distinct vertices of $\mathcal{G}_N$. Then, for $\delta \in \mathbb{N}$,

$$\mathbb{P}\{d_N(v,w) \le 2\delta\} = \mathbb{P}\Big(\bigcup_{k=1}^{2\delta} \bigcup_{(v_1,\dots,v_{k-1})} \{v \leftrightarrow v_1 \leftrightarrow \cdots \leftrightarrow v_{k-1} \leftrightarrow w\} \Big)$$

$$\le \sum_{k=1}^{2\delta} \sum_{(v_1,\dots,v_{k-1})} \prod_{j=1}^{k} p(v_{j-1},v_j),$$

where $(v_0,\dots,v_k)$ is any collection of pairwise distinct vertices in $\mathcal{G}_N$ with $v_0 = v$ and $v_k = w$ and, for $m,n \in \mathbb{N}$,

$$p(m,n) := \begin{cases} \kappa (m \wedge n)^{-\gamma} (m \vee n)^{\gamma-1} & \text{if Assumption 2.2 holds;} \\ \kappa\, m^{-\gamma}\, n^{-\gamma}\, N^{2\gamma-1} & \text{if Assumption 2.4 holds.} \end{cases}$$

Note that one can assign to each path $(v_0,\dots,v_k)$ the weight

$$p(v_0,\dots,v_k) := \prod_{j=1}^{k} p(v_{j-1},v_j), \tag{3.1}$$

and the upper bound is just the sum over the weights of all paths from v to w of length no more than 2δ. The shortcoming of this bound is that the paths that contribute most to the total weight are those that connect v, resp. w, quickly to vertices with extremely small indices. Since these are typically not present in the network, such paths have to be removed in order to get a reasonable estimate.

Remark 3.1. If the variance of the degrees is bounded, then the unconstrained first moment method already yields lower bounds on the typical distance of order $\log N$. Also for $\gamma = \frac{1}{2}$, no truncation is needed to obtain the lower bound stated in Theorem II.

To obtain a more accurate estimate we define a decreasing *cutoff* sequence $\ell = (\ell_k)_{k=0}^{\delta}$ of positive integers and consider a tuple of vertices $(v_0,\dots,v_n)$ as *admissible* if $v_k \wedge v_{n-k} \ge \ell_k$ for all $k \in \{0,\dots,\delta \wedge n\}$. We denote by $A_k^{(v)}$ the event that there exists a path $v = v_0 \leftrightarrow \cdots \leftrightarrow v_k$ in the network such that $v_0 \ge \ell_0,\dots,v_{k-1} \ge \ell_{k-1}$, $v_k < \ell_k$, i.e. a path that traverses the threshold sequence ℓ after exactly k steps. For fixed vertices $v,w \ge \ell_0$, the truncated first moment estimate is

$$\mathbb{P}\{d_N(v,w) \le 2\delta\} \le \sum_{k=1}^{\delta} \mathbb{P}(A_k^{(v)}) + \sum_{k=1}^{\delta} \mathbb{P}(A_k^{(w)}) + \sum_{n=1}^{2\delta} \sum_{\substack{(v_0,\dots,v_n) \\ \text{admissible}}} \mathbb{P}\{v_0 \leftrightarrow \cdots \leftrightarrow v_n\}, \tag{3.2}$$

where the admissible paths in the last sum start with $v_0 = v$ and end with $v_n = w$. By assumption,

$$\mathbb{P}\{v_0 \leftrightarrow \cdots \leftrightarrow v_n\} \leq p(v_0, \ldots, v_n)$$

so that for $v \geq \ell_0$ and $k = 1, \ldots, \delta$,

$$\mathbb{P}(A_k^{(v)}) \leq \sum_{v_1 = \ell_1}^{N} \cdots \sum_{v_{k-1} = \ell_{k-1}}^{N} \sum_{v_k = 1}^{\ell_k - 1} p(v, v_1, \ldots, v_k). \tag{3.3}$$

Given $\varepsilon > 0$ we choose $\ell_0 = \lceil \varepsilon N \rceil$ and $(\ell_j)_{j=0,\ldots,k}$ decreasing fast enough so that the first two terms on the right hand side of (3.2) together are no larger than 2ε. For $k \in \{1, \ldots, \delta\}$, set

$$\mu_k^{(v)}(u) := \mathbb{1}_{\{v \geq \ell_0\}} \sum_{v_1 = \ell_1}^{N} \cdots \sum_{v_{k-1} = \ell_{k-1}}^{N} p(v, v_1, \ldots, v_{k-1}, u),$$

and set $\mu_0^{(v)}(u) = \mathbb{1}_{\{v = u\}}$. To rephrase the truncated moment estimate in terms of μ, note that p is symmetric, so that, for all $n \leq 2\delta$ and $n^* := \lfloor n/2 \rfloor$,

$$\sum_{\substack{(v_0, \ldots, v_n) \\ \text{admissible}}} \mathbb{P}\{v_0 \leftrightarrow \cdots \leftrightarrow v_n\} \leq \sum_{v_1 = \ell_1}^{N} \cdots \sum_{v_{n^*} = \ell_{n^*}}^{N} \cdots \sum_{v_{n-1} = \ell_1}^{N} p(v, \ldots, v_{n^*}) p(v_{n^*}, \ldots, w)$$

$$= \sum_{v_{n^*} = \ell_{n^*}}^{N} \mu_{n^*}^{(v)}(v_{n^*}) \mu_{n-n^*}^{(w)}(v_{n^*}). \tag{3.4}$$

Using the recursive representation

$$\mu_{k+1}^{(v)}(n) = \sum_{m = \ell_k}^{N} \mu_k^{(v)}(m) \, p(m, n),$$

we establish upper bounds for $\mu_k^{(v)}(u)$, and use these to show that the rightmost term in (3.2) remains small if δ is chosen sufficiently small. Using the input from Assumption 2.2 and Assumption 2.4, respectively, this leads to the lower bounds for the typical distance in both Theorems 2.3 and 2.5. Detailed proofs are given in Chapter 5.

3.1.2 A truncated second order method – the upper bound of Theorem II

For the proof of Theorem II we can use the unconstrained first moment method for the lower bound. In the upper bound, however, we need a *truncated second moment method*. It is based on moment estimates for the degree evolutions $\{\mathscr{Z}[m, \cdot]\}_{m \in [N]}$ formally derived in Section 6.1. The convergence $\gamma = \lim_{k \to \infty} \Delta f(k) = \frac{1}{2}$ implies, see Section 4.1, that

$$\mathbb{E}^k f(\mathscr{Z}[m, n]) \approx f(k) \sqrt{\frac{n}{m}}, \quad \text{for large } n \in [N].$$

If f is affine, then $\approx$ can actually be replaced by $\sim$ in the above statement, which is shown Lemma 4.7. To study the deviation from the affine case systematically, we introduce

$$\psi^k(m, n) = \mathbb{E}f(\mathscr{Z}^{(k)}[m, n])\frac{\xi(m)}{\xi(n)},$$

for $m < n \in [N]$, with

$$\xi(m) = \xi(m, N) = \prod_{i=m}^{N-1}\left(1 + \frac{1}{2i}\right) = \frac{1}{f(0)}\mathbb{E}\bar{f}(\bar{\mathscr{Z}}[m, n]),$$

where $\bar{f}$ denotes the affine function $\bar{f}(k) = f(0) + \frac{1}{2}k, k \geq 0$, and $(\bar{\mathscr{Z}}[m, n])_{n \geq m}$ the corresponding degree evolution process of $m \in \mathbb{N}$.

We call $\xi(m)$ the *score* of vertex $m \in [N]$ and note that for $A \subset [N]$, the total score $\xi(A)$ is a rough estimate for the number of edges incident to A. Consequently, for fixed $k \geq 1$, the size $\#\Gamma^k(v)$ of $\Gamma^k(v)$, should be approximately equal to $\xi(\Gamma^{\leq k-1}(v))$. Given $\Gamma^{\leq k-1}(v)$, $\Gamma^k(v)$ is of course random and to show that $\#\Gamma^k(v)$ is concentrated around $\xi(\Gamma^{\leq k-1}(v))$ we need to derive bounds for $\xi^2(\Gamma^{\leq k-1}(v))$. That this is troublesome, can be seen most easily in the idealised branching random walk, which is, as already pointed out, an idealisation of the exploration around a typical vertex in the network. For simplicity, assume $v = \frac{N}{2}$, which corresponds under the preimage of π_N to an IBRW rooted in position $-\log 2$. Set $k = 1$ and recall the functions M, M^τ of Remark 1.13, then the quantity corresponding to $\xi^2(\Gamma^{\leq k-1}(v) = \xi^2\big(\Gamma^{\leq 0}\big(\frac{N}{2}\big)\big)$ in the IBRW is

$$\Sigma = \int_{-\infty}^{-\log 2} e^{-\frac{1}{2}s}\,\mathrm{d}M(-s) + \int_{-\log 2}^{0} e^{-\frac{1}{2}s}\,\mathrm{d}M^\ell(-s), \tag{3.5}$$

where $e^{-\frac{1}{2}s} = \sigma(s)$ is the score (now as a function of particles in the IBRW) of a particle of type ℓ in position $s \leq 0$. In general, M and M^τ depend on the shape of $\xi(m, n)$ and $\psi(m, n)$, where we interpret ξ as the offspring density on the exponential scale, both in the idealised setting and in the network. Similarly, ψ will have a counterpart in the idealised set up describing the deviation of the density around ξ on the polynomial scale. For the affine attachment rule $f(k) = f(0) + \frac{1}{2}k$ we obtain $\mathrm{d}M(s) = f(0)e^{-\frac{1}{2}s}\,\mathrm{d}s$, for $s \geq 0$, so ψ is constant, and we can thus directly conclude that $\Sigma = \infty$.

Analogously to the truncation of the first moment in Section 3.1.1, we need to find a deterministic *cutoff sequence* $r = (r_k)_{k=0}^{\delta}$ that prescribes lower bounds on the vertices in the shortest path between the initial vertices U, V. In the idealised setting this corresponds to replacing the integral $\int_{-\infty}^{-\log 2} e^{-\frac{1}{2}s}\,\mathrm{d}M(-s)$ in (3.5) by $\int_{r_k}^{-\log 2} e^{-\frac{1}{2}s}\,\mathrm{d}M(-s)$, where $r_k \in (-\infty, 0]$ depends on the depth k of the exploration. For the idealised branching random walk this is the problem of determining bounds on the leftmost particle in generation k, which is solved

heuristically in Section 4.5. Denoting the k-th generation of the IBRW X *killed at* r_k by X^k, the appropriate choice of $r_k, k \geq 0$, is obtained through the relation

$$S_{k-1} = \mathbb{E}\Big[\sum_{v \in X^k} \xi(v) \Big| X^0, \ldots, X^{k-1} \Big] \approx \int_{r_k}^0 \psi(s)\,\mathrm{d}s \sum_{v \in X^{k-1}} \xi(v),$$

by demanding that $(W_k)_{k \geq 0}$ given by

$$W_k = \frac{S_k}{\prod_{i=1}^k \int_{r_{i-1}}^0 \psi(s)\,\mathrm{d}s}, \quad k \geq 0,$$

be (almost) an L^2-martingale with respect to the filtration generated by the exploration. The task of transferring this approach to the network setting is undertaken in Chapter 6.

3.2 The local picture – branching processes approximation

For the upper bound proofs, we need to show that we are able to connect two randomly chosen vertices U, V. We fix V and U and then define an exploration algorithm to successively uncover the neighbourhoods of V and U up to the point where they meet or all vertices in the connected components of V or U are uncovered.

For small k, we expect the neighbourhood $\Gamma^{\leq k}(V)$ to look like a tree. In Chapter 4 we make this intuition precise and present the construction of [DM13] to couple the neighbourhood of a vertex to a random labelled tree, which can then be mapped to the INT. Both the neighbourhood and the tree are explored using a specific algorithm. We will call this the *exploration process*. We show that the exploration processes of the network and the labelled tree $\mathbb{T}(V)$ can be defined on the same probability space in such a way that, up to a large stopping time, the explored part of the network and the tree coincide.

This means we run three exploration processes in parallel – one exploration in $\mathcal{G}_N$, one in $\mathbb{T}(V)$ and one in $\pi(\mathfrak{T})$, where the latter is the killed IBRW projected onto $[N]$ using a suitable projection π and $\mathbb{T}$ is essentially an exploration of $\mathcal{G}_N$ without vertex depletion, i.e. we formally allow over-counting in the exploration process by extending the neighbourhood even if an already explored vertex is 'uncovered' again. We denote by $\mathscr{C}^l(V)$ the first l neighbours of V explored in $\mathcal{G}_N$ – note that this subgraph depends on the choice of the order of exploration and *does not* coincide with $\Gamma^{\leq l}(V)$ – and denote by $\mathbb{T}^l(V)$ and $\mathfrak{T}^l$ the corresponding subsets of $\mathbb{T}(V)$ and $\mathfrak{T}$. We start by coupling $V \in [N]$ to the random variable $s(-x_\emptyset)$ which is exponentially distributed with parameter -1 and then proceed inductively by coupling an arbitrary exploration step with the random mechanisms used to generate offspring in $\mathbb{T}(V)$ and in the IBRW. The deviation between the three explorations can be kept small up to about $c_N \approx \log N$ explored vertices and they coincide with high probability.

For $\gamma < \frac{1}{2}$, the IBRW has a well defined Malthusian parameter and the leftmost particle moves at most at linear speed. If it survives, the INT thus grows exponentially and so do the neighbourhoods $\Gamma^k(v), k \geq 0$, if k is such that the total size of the k-neighbourhood does not exceed c_N. If $\gamma \geq \frac{1}{2}$, then this is not the case any more – the INT, if it survives, grows *super-exponentially*. For $\gamma > \frac{1}{2}$, a neighbourhood size of c_N corresponds to an exploration depth of $o(\log\log N)$ generations. In fact, after only finitely many exploration steps, there are so many cycles in the network discovered that the analogy between $\Gamma^k(v)$ and a tree growing at a certain rate breaks down. Therefore the growth of the INT is not reflected in the constant in front of the double logarithm, which is not directly determined by the local approximation of the network.

For $\gamma = \frac{1}{2}$, the situation different. A size of c_N of the explored cluster is still not sufficient to determine the scaling of the shortest paths. However, the right scaling constant in front of $\frac{\log N}{\log\log N}$ is obtained from the IBRW as outlined in Section 3.1.2 above. The reason that the approach works, is that although many cycles are uncovered, there is no strong depletion effect until a neighbourhood size of about $\sqrt{N}$ is reached and the neighbourhoods grow at roughly the same rate as in the INT. This produces the upper bound of Theorem II. Therefore, interestingly, in the critical case there seems to be no real qualitative difference between shortest paths on the 'local' scale and the shortest paths on the 'global scale'.

3.3 The global picture – cores and couplings

For the upper bound of Theorem I, we run the local exploration around V until we find a vertex W of very high degree. At this point, we abandon the coupling and change the exploration strategy: we show that we can connect W via a sequence of high degree vertices to the *core* $\text{core}_N \subset \mathscr{G}_N$, a subset of $\mathscr{G}_N$ which contains vertices of particularly high degree. Performing a similar exploration for U then results in a path between U and V.

The overall approach for the upper bound of Theorem II is similar. Starting in U and V, we use an exploration scheme that has two stages. After the local stage, we abandon the coupling to the tree and estimate the size (or more precisely the score) of the growing neighbourhood shells, using the truncated second moment method described above in Section 3.1.2. Finally, we conclude that the explored parts of the graph must be connected, which we show to be the case if $\alpha = 0$ in Theorem II, or that the exploration processes have hit the core.

In both situations, we need to show that the core_N is of small diameter, more precisely we show that, for some parameter $\varepsilon > 0$, the core induces a subgraph of $\mathscr{G}_{\lceil(1+\varepsilon)N\rceil}$ which is of bounded diameter. We achieve this by using ε-*connectors*, which are a dynamical version of the sprinkling technique commonly used in classical random graph theory. The first step
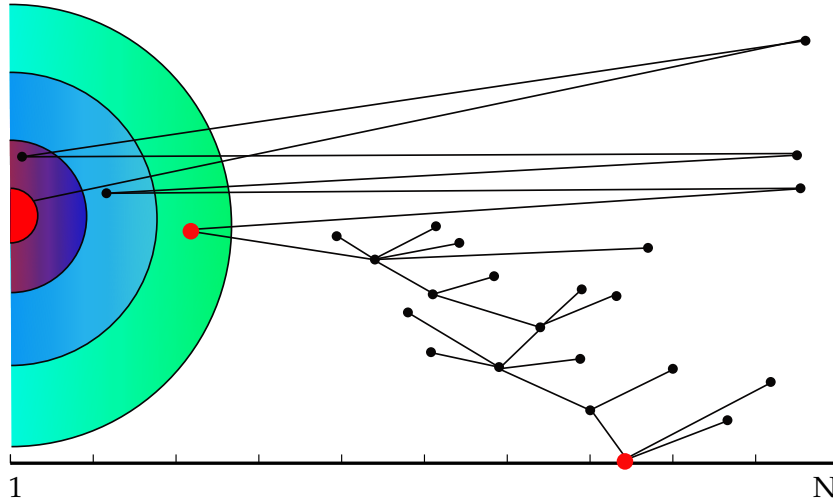
Figure 3-1: Illustration of shortest path structure in $\mathcal{G}_N$. The disks represent vertices of high degree, the lines the initial tree-like exploration. Note that *qualitatively* the picture is the same for all $\gamma \in \left[\frac{1}{2}, 1\right)$, but that the scales of the two parts of the path change dramatically at $\gamma = \frac{1}{2}$.

of the approach is a variant of the technique used in [DHH10] and works as follows: We use that our network evolves in time by considering the edges which are added to $\mathcal{G}_N$ after an additional period of length about εN. More precisely, given large $\mathcal{G}_N$ we already know which vertices have high degrees and are thus preferred by the edges emanating from the vertices $j \in \{N+1, \ldots, \lceil(1+\varepsilon)N\rceil\}$. We can use the fact that these edges are (almost) conditionally independent given $\mathcal{G}_N$. The sublogarithmic scaling renders the differences in the distances between $\mathcal{G}_{\lceil(1+\varepsilon)N\rceil}$ and $\mathcal{G}_N$ negligible.

Let $v, w \in [N]$ and $j \in \{N+1, \ldots, \lceil(1+\varepsilon)N\rceil\}$ then, conditional on $\mathcal{G}_N$,

$$\mathbb{P}\{j \leftrightarrow v, j \leftrightarrow w\} \geq \frac{f(\mathcal{Z}[v,N])f(\mathcal{Z}[w,N])}{(\lceil(1+\varepsilon)N\rceil)^2} \approx \frac{\left(\frac{N}{v}\right)^\gamma \left(\frac{N}{w}\right)^\gamma}{N^2}.$$

Since there are εN potential ε-connectors, we thus find that

$$\mathbb{P}\{d_{\lceil(1+\varepsilon)N\rceil}(v,w) \leq 2\} \approx \varepsilon \frac{\left(\frac{N}{v}\right)^\gamma \left(\frac{N}{w}\right)^\gamma}{N}, \tag{3.6}$$

which resembles the connection probabilities in an inhomogeneous random graph with weights $w_v = c\left(\frac{N}{v}\right)^\gamma$, for some $c > 0$. For these random graphs, coupling techniques to non-sparse versions of the Erdős-Rényi graph $\mathcal{G}(M, p)$, for appropriately chosen $M = M(N), p = p(N)$, are available and these couplings constitute the second step of the approach.

The couplings we utilise are based on lower bounds for the probabilities appearing in (3.6) in terms of a fixed probability p. Given a subset G of $\mathcal{G}_N$ size M, we argue that we have

sufficient independence to dominate every 2-step connection in $G \subset \mathcal{G}_{\lceil(1+\varepsilon)N\rceil}$ induced by an ε-connector by an edge in a coupled random graph $\mathcal{G}(M, p)$. For $\mathcal{G}(M, p)$, however, the exact behaviour of the diameter is well known, see Proposition B.3. Similar arguments have been used to bound distances in several non-homogeneous network models, including the proofs of Proposition 1.15 in [CL02].

Furthermore, in Chapter 5, we use a version of the relation (3.6) for $\varepsilon = 1$ to determine the connecting sequence of high degree vertices constructed in the proof of Theorem 2.6, recall also the derivation of the upper bound in Proposition 2.8. Instead of using a fixed lower bound in (3.6), we work with a bound that grows in every second step $2k$ along the path constructed outwards from W depending on k.

Chapter 4

Prerequisite results and general techniques

This chapter provides most of the tools utilised to prove the main results. The martingale techniques and the correlation bounds of Sections 4.1 and 4.2 are mostly adaptations of standard results in complex network theory. We use the score functional mentioned in Section 3.1 to describe vertex weights or expected degrees arising in the preferential attachment graph and perform some straightforward calculations for the first moment of the score. Furthermore, we establish some stochastic domination results for the degree evolutions needed for local approximation purposes and the proofs in Chapters 5 and 6.

At the heart of all our 'local' calculations in the network lies the branching process approximation detailed in Sections 4.3 and 4.4, which, up to some minimal variations, follow the exposition in [DM13]. To keep this thesis as self contained as possible, most of the elaborate construction in [DM13] is reproduced in those sections.

Section 4.5 is not strictly a prerequisite for the following chapters but provides a discussion of a more idealised version of the second moment method used to derive the upper bound of Theorem II in Chapter 6 and explains how the cutoff for the truncated second moment method is obtained.

Finally, in Section 4.6, we formalise the notion of ε-connectors, which is our adaptation of the concept of t-connectors used in [DHH10] and occupies the place in our argumentation for the preferential attachment setting that sprinkling arguments often take in classical random graph theory.

4.1 Degree evolutions and correlation bounds

We first collect some established results about the degree evolutions which will be useful later and then turn a priori estimates for the degrees into first moment bounds on the existence probability of short paths, which provides the lower bound in Theorem II.

We wish to define a number of martingales related to the degree evolutions which are instrumental in later proofs. Recall that $\{(\mathscr{Z}^{(k)}[i,n])_{n=i}^{\infty}\ i, k \in \mathbb{N}\}$ denotes the family of independent jump process starting in $k = \mathscr{Z}^{(k)}[i,i]$ at time i and satisfying

$$\mathbb{P}\{\Delta\mathscr{Z}^{(k)}[i,n] = 1 | \mathscr{Z}^{(k)}[i,n]\} = 1 - \mathbb{P}\{\Delta\mathscr{Z}^{(k)}[i,n] = 0 | \mathscr{Z}^{(k)}[i,n]\} = \frac{f(\mathscr{Z}^{(k)}[i,n])}{n}.$$

We write $\mathbb{P}^k$ and $\mathbb{E}^k$ for the distribution and expectation of the process $(\mathscr{Z}^{(k)}[i,n])_{n\geq i}$ in dependence of the initial value k. We will usually omit k in the notation if $k = 0$ or if there is no ambiguity. The same notation applies to the idealised degree evolution processes.

Remark 4.1. The distribution of the degree evolutions $(\mathscr{Z}[m,n])_{n\geq m}, m \in [N]$, in the network $\mathscr{G}_N$ is the same as that of the family $(\mathscr{Z}^{(0)}[m,n])_{n\geq m}, m \in [N]$. We can think of the jump processes above as degree evolutions started in arbitrary values. Due to the independence in the formulation of the model, we can couple $(\mathscr{G}_N)_{N\in\mathbb{N}}$ perfectly to a realisation of the ensemble $(\mathscr{Z}^{(0)}[m,n]; 1 \leq m \leq n \leq N)$, where the evolutions $(\mathscr{Z}^{(0)}[m,n])_{n\geq m}$ are mutually independent for different values of m. In this sense the random network is completely determined as a collection of independent jump processes and many of the following arguments make ample use of this fact.

Definition 4.2 (Expectation operators). Let $g : \mathbb{N}_0 \longrightarrow (0,\infty)$ then we write

$$P_{m,n}g(k) = \mathbb{E}^k g(\mathscr{Z}[m,n]), \quad k \in \mathbb{N}_0, n \geq m \in \mathbb{N},$$

and for the idealised degree evolution $(Z_t)_{t\geq 0}$, we introduce

$$P_t g(k) = \mathbb{E}^k g(Z_t), \quad k \in \mathbb{N}_0, t \in [0,\infty).$$

A fundamental monotonicity result can be derived directly before we proceed to introduce the martingales.

Lemma 4.3 (Stochastic domination II, [DM13, Lemma 2.11]). *Let $\gamma^+ = \sup_k \Delta f(k) \leq 1$ and $m \leq n \leq l$ be integers, then*

$$\mathbb{P}\{\Delta\mathscr{Z}[m,n] = 1\} \geq \mathbb{P}\{\Delta\mathscr{Z}[m,l] = 1\}.$$

Proof. It suffices to prove the statement for $l = n + 1$. Let $n \geq m$ arbitrary, then

$$\mathbb{P}\{\Delta \mathcal{Z}[m, n] = 1\} = \frac{\mathbb{E}f(\mathcal{Z}[m, n])}{n} = \frac{1}{n}\sum_{k=0}^{\infty} f(k)\mathbb{P}\{\mathcal{Z}[m, n] = k\}$$

and

$$\mathbb{P}\{\Delta \mathcal{Z}[m, n+1] = 1\} = \frac{1}{n}\sum_{k=0}^{\infty} f(k)\frac{\Delta f(k) + n}{n+1}\mathbb{P}\{\mathcal{Z}[m, n] = k\}.$$

The statement follows, since $\Delta f(k) \leq \gamma^+ \leq 1$. $\qquad\square$

Lemma 4.4 (Degree martingales I)**.** *Let $h : \mathbb{N}_0 \longrightarrow \mathbb{R}$ be a measurable function, and define a sequence $M^h = (M_1^h, M_2^h, \dots)$ of processes by*

$$M_i^h(n) = \begin{cases} h(\mathcal{Z}[i, n]) - h(\mathcal{Z}[i, i]) - \sum_{s=i}^{n-1}\frac{f(\mathcal{Z}[i,s])}{s}\Delta h(\mathcal{Z}[i, s]), & \text{if } n > i \\ 0, & \text{if } 0 \leq n \leq i. \end{cases}$$

Then M^h is a martingale with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ generated by the random variables $\{(\mathcal{Z}[i, j])_{j=0}^n, 1 \leq i \leq n\}$.

Proof. This is a special case of [BL12, Lemma 2.1]. A direct calculation yields, for $n \geq i$,

$$\begin{aligned} \mathbb{E}[M_i^h(n+1)|\mathcal{F}_n] &= \mathbb{E}[M_i^h(n) + h(\mathcal{Z}[i, n+1]) - h(\mathcal{Z}[i, n]) - \frac{f(\mathcal{Z}[i, n])}{n}\Delta h(\mathcal{Z}[i, n])|\mathcal{F}_n] \\ &= M_i^h(n) + \Delta h(\mathcal{Z}[i, n])\mathbb{P}\{\Delta\mathcal{Z}[i, n] = 1|\mathcal{Z}[i, n]\} - \frac{f(\mathcal{Z}[i, n])}{n}\Delta h(\mathcal{Z}[i, n]) \\ &= M_i^h(n), \end{aligned}$$

using the tower property of conditional expectation, the definition of M^h and the fact that $\mathcal{Z}[i, n]$ increases with probability $\frac{f(\mathcal{Z}[i,n])}{n}$. $\qquad\square$

The weight of a vertex in the network is most conveniently described by its expected indegree, recall that we call this the *score* of the vertex. Although we are only interested in $\gamma \in [0, 1)$, it is convenient to allow for larger values of γ in the definition.

Definition 4.5 (Scores)**.** Let $\gamma \in [0, 2]$, we define

$$\xi^{(\gamma)}(m, n) = \prod_{i=m}^{n-1}\left(1 + \frac{\gamma}{i}\right), \text{ for } n > m,$$

and set $\xi^{(\gamma)}(m, m) = 1$ and $\xi^{(\gamma)}(n, m) = \xi^{(\gamma)}(m, n)^{-1}$. We omit γ in the superscript, if there is no ambiguity. In a graph of fixed size N we will often write $\xi(m)$ for $\xi(m, N)$ and call $\xi(m)$ the *score* of $m \in [N]$.

Numerical bounds for ξ are of some importance in later proofs, thus we establish the following lemma.

Lemma 4.6. *Let $C = e^{\gamma\gamma_{EM}}$, then*

(i) *for all $\gamma \in [0,2]$,*

$$\xi^{(\gamma)}(m,n) \leq C\Big(\frac{n}{m}\Big)^{\gamma}, \ \ \textit{for all } m,n \in \mathbb{N};$$

(ii) *for all $\gamma \in [0,1]$,*

$$\xi^{(\gamma)}(m,n) \geq \Big(\frac{n}{m}\Big)^{\gamma}.$$

Proof. We begin with the lower bound. The statement is obvious for $\gamma \in \{0,1\}$. For $\gamma \in (0,1)$, we perform induction in $n \geq m$. Firstly, $n = m$ gives $\xi(m,m) = 1$. Secondly, applying first the induction hypotheses and then Bernoulli's inequality, we obtain

$$\xi(m,n+1)^{\frac{1}{\gamma}} = \Big(1 + \frac{\gamma}{n}\Big)^{\frac{1}{\gamma}} \xi(m,n)^{\frac{1}{\gamma}} \geq \Big(1 + \frac{\gamma}{n}\Big)^{\frac{1}{\gamma}} \Big(\frac{n}{m}\Big)$$
$$\geq \Big(1 + \frac{1}{n}\Big)\Big(\frac{n}{m}\Big) = \frac{n+1}{m},$$

which completes the induction step.

For the upper bound we recall that $\gamma_{\mathrm{EM}} = \lim_{x \to \infty} \int_1^x \frac{1}{\lfloor s \rfloor} - \frac{1}{s} \, ds$ and calculate directly, for $n \geq m \geq 1$,

$$\log\Big(\prod_{i=m}^{n-1} 1 + \frac{\gamma}{i}\Big) = \sum_{i=m}^{n-1} \log\Big(1 + \frac{\gamma}{i}\Big) \leq \gamma \sum_{i=m}^{n-1} \frac{1}{i} \leq \gamma(\log \tfrac{n}{m} + \gamma_{\mathrm{EM}}),$$

since $\log(1 + x) \leq x$ for $x \in (0,\infty)$. $\qquad \square$

We can use ξ to rescale the degree evolutions and obtain a different class of martingales.

Lemma 4.7 (Degree martingales II)**.** *Let $k \in \mathbb{N}$ and $\gamma \in (0,1)$ be fixed and $n \geq m \in \mathbb{N}$. Define*

$$X_m(n) = \frac{f(\mathcal{Z}^{(k)}[m,n])}{\xi^{(\gamma)}(m,n)} \ \ \textit{and} \ \ Y_m(n) = \frac{f(\mathcal{Z}^{(k)}[m,n])^2 + \gamma f(\mathcal{Z}^{(k)}[m,n])}{\xi^{(2\gamma)}(m,n)}, \qquad (4.1)$$

and consider the filtration generated by $(\mathcal{Z}^{(k)}[m,n])_{n \geq m}$.

(i) *If $f(l) = \gamma l + \beta$, $l \in \mathbb{N}$, then X_m and Y_m are martingales.*

(ii) *For any attachment rule f, replacing γ by γ^+ in (4.1) yields supermartingales and replacing γ by γ^- in (4.1) yields submartingales.*

(iii) *If f is concave, and γ in (4.1) is replaced by γ^-, then X_m and Y_m are submartingales. Replacing γ by $\Delta f(k)$, $k \in \mathbb{N}$, in (4.1) yields supermartingales.*

Remark 4.8.   (i) For the correlation results in the rest of the work we only need estimates for $f(\mathcal{Z}[m,n])^2$. It is not difficult to define weights approximating the higher moments of $f(\mathcal{Z}[m,n])$, the proof below is easily adjusted. In the affine case one can always explicitly calculate the weights associated with martingale sequences. See also [Hof13, Section 8], where the weights are determined for general moments in the affine PA models of Example 2.7.

(ii) The first appearance of this martingale argument to determine the weights in PA networks is in [Bol+01].

Proof of Lemma 4.7. We only give the explicit argument for the affine case (i), the assertion (ii) follows from straightforward adaptations of the calculation and so does (iii), recalling that $\Delta f(k)$ monotonically decreases towards γ^- in this case. For convenience, we suppress the dependence on the starting value k in the notation. The desired martingale property of $(X_m(n))_{n \geq m}$ for an affine attachment rule $f(k) = \gamma k + \beta$ follows directly from

$$
\begin{aligned}
\mathbb{E}[f(\mathcal{Z}[m, n+1]) | \mathcal{G}_n] &= \mathbb{E}[f(\mathcal{Z}[m, n]) + \gamma \Delta \mathcal{Z}[m, n] | \mathcal{Z}[m, n]] \\
&= \mathbb{E}[f(\mathcal{Z}[m, n]) + \gamma \mathbb{1}_{\{n+1 \to m\}} | \mathcal{Z}[m, n]] \\
&= \mathbb{E}[f(\mathcal{Z}[m, n]) | \mathcal{G}_n] + \mathbb{E}[\gamma \mathbb{1}_{\{n+1 \to m\}} | \mathcal{Z}[m, n]] \\
&= \mathbb{E}[f(\mathcal{Z}[m, n]) | \mathcal{G}_n] + \gamma \mathbb{E}\Big[\frac{f(\mathcal{Z}[m, n])}{n} \Big| \mathcal{Z}[m, n]\Big] \\
&= \Big(1 + \frac{\gamma}{n}\Big) f(\mathcal{Z}[m, n]).
\end{aligned}
\tag{4.2}
$$

The statement for $(Y_m(n))_{n \geq m}$ can be proven analogously,

$$
\begin{aligned}
\xi^{(2\gamma)}(m, n+1) &\mathbb{E}[Y_m(n+1) | \mathcal{Z}[m, n]] \\
&= \mathbb{E}[f(\mathcal{Z}[m, n+1])^2 | \mathcal{Z}[m, n+1]] + \gamma\Big(1 + \frac{\gamma}{n}\Big) f(\mathcal{Z}[m, n]),
\end{aligned}
\tag{4.3}
$$

and we determine by a calculation in the spirit of (4.2), that

$$
\mathbb{E}[f(\mathcal{Z}[m, n+1])^2 | \mathcal{Z}[m, n]] = \Big(1 + \frac{2\gamma}{n}\Big) f(\mathcal{Z}[m, n])^2 + \frac{\gamma^2}{n} f(\mathcal{Z}[m, n]).
\tag{4.4}
$$

Hence (4.3) turns into

$$
\begin{aligned}
\xi^{(2\gamma)}(m, n+1) \mathbb{E}[Y_m(n+1) | \mathcal{Z}[m, n]] &= \Big(1 + \frac{2\gamma}{n}\Big) f(\mathcal{Z}[m, n])^2 + \frac{\gamma^2}{n} f(\mathcal{Z}[m, n]) \\
&\quad + \gamma\Big(1 + \frac{\gamma}{n}\Big) f(\mathcal{Z}[m, n]) \\
&= \Big(1 + \frac{2\gamma}{n}\Big) \xi^{(2\gamma)}(m, n) Y_m(n),
\end{aligned}
$$

and dividing by $\Big(1 + \frac{2\gamma}{n}\Big) \xi^{(2\gamma)}(m, n) = \xi^{(2\gamma)}(m, n+1)$ yields the martingale property. Examining the above derivation shows that if we always bound the increments of f by

$$
\gamma^- \leq \Delta f(\mathcal{Z}[m, n]) \leq \gamma^+,
$$

then we get the same calculations with inequalities instead of equalities and obtain a super-/submartingale. □

Lemma 4.7 together with Lemma 4.6 imply the following bound on the expected degrees.

Corollary 4.9 (A priori bounds on expected degrees, [DM13, Lemma 2.7][1])**.** *For any attachment rule f and $k \in \mathbb{N}_0$, $m \le n \in \mathbb{N}$,*

$$f(k)\left(\frac{n}{m}\right)^{\gamma^-} \le \mathbb{E}^k f(\mathcal{Z}[m,n]) \le f(k) C_{4.6} \left(\frac{n}{m}\right)^{\gamma^+}.$$

The recursion (4.2) for the expected degrees from which we derive the form of the score, implies that ξ has an accessible product structure – the scores are ratios of Γ-functions – which simplifies calculations later, in particular for affine f. We collect some straightforward facts for later reference.

Observation 4.10. Fix $\gamma \in (0,1)$ and let $N \in \mathbb{N}$. Then the following relations hold:

(i) for all $l, m, n \in \mathbb{N}$,
$$\xi(l,m)\xi(m,n) = \xi(l,n);$$

(ii) for all $m, n \in \mathbb{N}$,
$$\xi^{(2\gamma)}(m,n) \le \xi^{(\gamma)}(m,n)^2 \le e^{\gamma^2 \frac{\pi^2}{6}} \xi^{(2\gamma)}(m,n);$$

(iii) for all $m, n \in \mathbb{N}$,
$$\xi(m,n) = \frac{\Gamma(n+\gamma)\Gamma(m)}{\Gamma(m+\gamma)\Gamma(n)}.$$

Lemma 4.11 (Degree Martingales III, [DM09, Lemma 2.1, Proposition 2.2])**.** *Let $u \in \mathbb{N}$, $\phi(u) = \sum_{i=0}^{u-1} \frac{1}{f(i)}$ and let*

$$M_n = \phi(\mathcal{Z}[m,n]) - \sum_{k=m}^{n-1} \frac{1}{k}, \quad n > m \in \mathbb{N}.$$

Then $(M_n)_{n=m}^{\infty}$ is a martingale w.r.t. the sequence $(\mathcal{Z}[m,n])_{n=m}^{\infty}$, which converges to a random variable M_∞ if and only if

$$\sum_{k=0}^{\infty} \frac{1}{f(k)^2} < \infty.$$

Moreover, if f is concave, then the law of M_∞ is absolutely continuous w.r.t. Lebesgue measure.

Proof. We only derive the martingale property and refer to [DM09] for the proofs of the other statements. The martingale property follows from

$$\mathbb{E}[\phi(\mathcal{Z}[m,n+1]) - \phi(\mathcal{Z}[m,n])|\mathcal{Z}[m,n]] = \frac{f(\mathcal{Z}[m,n])}{n} \frac{1}{f(\mathcal{Z}[m,n])} = \frac{1}{n}.$$

$\square$

Remark 4.12. The original statement [DM09, Lemma 2.1] is formulated for a rescaled version of the process.

We complete our collection of preliminary results with some important stochastic domination results from [DM13].

---

[1] Note that in the upper bound stated in [DM13, Lemma 2.7] the constant $C_{4.6}$ is missing.

Lemma 4.13 (Stochastic domination III, [DM13, Lemma 2.9]). *Fix $k \in \mathbb{N}$. Let $m \leq n \in \mathbb{N}$ and denote by $I = \{n_1, \ldots, n_j\}$ an ordered set of potential jump points of $(\mathcal{Z}^{(k)}[m, \cdot])$. The process $(\mathcal{Z}^{(k)}[m, n]_{n \geq m})$ conditional on the event $E_0 = \{\Delta \mathcal{Z}[m, n_i] = 0 \ \forall \ 1 \leq i \leq j\}$ is stochastically dominated by the unconditioned process under $\mathbb{P}^k$.*

Proof. For completeness we reproduce the proof given in [DM13]. Assume first that $m < n_1$. For any $k \geq 0$, we have, using the Markov property,

$$\mathbb{P}^k\{\Delta \mathcal{Z}[m, m] = 1 | E_0\} = \frac{f(k)}{m} \frac{\mathbb{P}^{k+1}\{\Delta \mathcal{Z}[m+1, n_i] = 0 \ \forall i \in \{1, \ldots, j\}\}}{\mathbb{P}^k\{\Delta \mathcal{Z}[m, n_i] = 0 \ \forall i \in \{1, \ldots, j\}\}}.$$

The last denominator can be bounded by

$$\frac{f(k)}{m} \mathbb{P}^{k+1}\{\Delta \mathcal{Z}[m+1, n_i] = 0 \ \forall i \in \{1, \ldots, j\}\} + \left(1 - \frac{f(k)}{m}\right) \mathbb{P}^k\{\Delta \mathcal{Z}[m+1, n_i] = 0 \ \forall i \in \{1, \ldots, j\}\}$$

$$\geq \mathbb{P}^{k+1}\{\Delta \mathcal{Z}[m+1, n_i] = 0 \ \forall i \in \{1, \ldots, j\}\},$$

because f is non-decreasing. Hence $\mathbb{P}^k\{\Delta \mathcal{Z}[m, m] = 1 | E_0\} \leq \frac{f(k)}{m}$, which certainly also holds if $m = n_1$ and the general result follows by induction. $\square$

Lemma 4.14 (Stochastic domination IV, [DM13, Lemma 2.10]). *Let f be concave. For integers $0 \leq k < m < n$ there is a coupling of the process $(\mathcal{Z}[m, l])_{l \geq m}$ started in $\mathcal{Z}[m, m] = k$ and conditioned on $\{\Delta \mathcal{Z}[m, n] = 1\}$ and the unconditional process $(\mathcal{Z}[m, l])_{l \geq m}$ started in $\mathcal{Z}[m, m] = k + 1$ such that for the coupled versions $(\bar{\mathcal{Z}}^{(c)}[m, l], \bar{\mathcal{Z}}^{(u)}[m, l])_{l \geq m}$ one has*

$$\Delta \bar{\mathcal{Z}}^{(c)}[l] \leq \Delta \bar{\mathcal{Z}}^{(u)}[l] + \mathbb{1}_{\{l = n\}}.$$

Thus the unconditioned process initiated in $k + 1$ dominates the process initiated at k and conditioned to have a jump at time n.

Remark 4.15. Lemma 4.14 is central to the establishment of the coupling of the exploration processes in Sections 4.3 and 4.4 and the only instance where concavity of f really is essential.

Proof of Lemma 4.14. Again, we reproduce an argument from [DM13]. We claim that

$$\mathbb{P}^k(\Delta \mathcal{Z}[m, m] = 1 | \Delta \mathcal{Z}[m, n] = 1) = \frac{f(k)}{m} \frac{\frac{1}{n} \mathbb{E}^{k+1} f(\mathcal{Z}[m+1, n])}{\frac{1}{n} \mathbb{E}^k f(\mathcal{Z}[m, n])} \leq \frac{f(k+1)}{m}, \quad (4.5)$$

which allows us to use a coupling argument. Suppose that $(\bar{\mathcal{Z}}^{(u)}[m, l])_{l \geq m}$ has the distribution of the unconditioned process started in $k + 1$ and set $N_0 = m$. Let $N_1 < N_2 < \ldots$ denote the jump times of $(\bar{\mathcal{Z}}^{(u)}[m, l])_{l \geq m}$ and let $(V_j)_{j \geq 0}$ be a sequence of uniform random variables on $[0, 1]$ which are mutually independent and independent of $(\bar{\mathcal{Z}}^{(u)}[m, l])_{l \geq m}$. Now define $(\bar{\mathcal{Z}}^{(c)}[m, l])_{l \geq m}$ as the jump process starting in k and increasing by one

- at time $N_{j+1} < n$, if

$$V_j \le \frac{f(\bar{\mathcal{Z}}^{(c)}[m, N_j])}{f(\bar{\mathcal{Z}}^{(u)}[m, N_j])} \frac{\mathbb{E}^{\bar{\mathcal{Z}}^{(c)}[m, N_j]+1} f(\bar{\mathcal{Z}}^{(c)}[N_{j+1}+1, n])}{\mathbb{E}^{\bar{\mathcal{Z}}^{(c)}[m, N_j]} f(\bar{\mathcal{Z}}^{(c)}[N_{j+1}, n])};$$ (4.6)

- at time n;

- at time $N_{j+1} > n$, if

$$V_j \le \frac{f(\bar{\mathcal{Z}}^{(c)}[m, N_{j \vee n}])}{f(\bar{\mathcal{Z}}^{(u)}[m, N_j])}.$$ (4.7)

By definition we have

$$\bar{\mathcal{Z}}^{(c)}[m, l] + 1 \le \bar{\mathcal{Z}}^{(u)}[m, l], \quad \text{for all } l = m, \dots, n-1$$

and

$$\bar{\mathcal{Z}}^{(c)}[m, l] \le \bar{\mathcal{Z}}^{(u)}[m, l], \quad \text{for all } l \ge m,$$

since the first process starts in k and can only jump at time n and at those times at which the second process jumps. Together with (4.5) it follows that the right hand sides of the inequalities (4.6) and (4.7) are bounded by 1. We can therefore check that the distribution of $(\bar{\mathcal{Z}}^{(c)}[m, l])_{l \ge m}$ is the same as that of the process $(\mathcal{Z}^{(k)}[m, l])_{l \ge m}$ starting at k and conditioned to jump at time n. Indeed, by construction and given $N_0, \dots, N_j$, for $N_j < l < n$,

$$\begin{aligned}
\mathbb{P}\{\Delta\bar{\mathcal{Z}}^{(c)}[m, l-1] = 1\} &= \mathbb{P}\{\Delta\bar{\mathcal{Z}}^{(c)}[m, N_{j+1}-1] = 1, l = N_{j+1}\} \\
&= \mathbb{P}\{\Delta\bar{\mathcal{Z}}^{(c)}[m, N_{j+1}-1] = 1 | l = N_{j+1}\}\mathbb{P}\{l = N_{j+1}\} \\
&= \frac{f(\bar{\mathcal{Z}}^{(c)}[m, N_j])}{f(\bar{\mathcal{Z}}^{(u)}[m, N_j])} \frac{\mathbb{E}^{\bar{\mathcal{Z}}^{(c)}[m, N_j]+1} f(\bar{\mathcal{Z}}^{(c)}[N_{j+1}+1, n])}{\mathbb{E}^{\bar{\mathcal{Z}}^{(c)}[m, N_j]} f(\bar{\mathcal{Z}}^{(c)}[N_{j+1}, n])} \frac{f(\bar{\mathcal{Z}}^{(u)}[m, N_j])}{l-1} \\
&= \frac{f(\bar{\mathcal{Z}}^{(c)}[m, N_j])}{l-1} \frac{\mathbb{E}^{\bar{\mathcal{Z}}^{(c)}[m, N_j]+1} f(\bar{\mathcal{Z}}^{(c)}[N_{j+1}+1, n])}{\mathbb{E}^{\bar{\mathcal{Z}}^{(c)}[m, N_j]} f(\bar{\mathcal{Z}}^{(c)}[N_{j+1}, n])},
\end{aligned}$$

and similarly for $l > n \vee N_j$

$$\mathbb{P}\{\Delta\bar{\mathcal{Z}}^{(c)}[m, l-1] = 1\} = \frac{f(\bar{\mathcal{Z}}^{(c)}[m, N_j \wedge n])}{l-1},$$

showing that the jump probabilities of both processes agree at all times $l \ge m$.

It remains to verify the claim (4.5). To this end, we fix another i.i.d. sequence $(U_j)_{j \ge 0}$ of uniform random variables on $[0, 1]$ which is independent of everything else. Now let $M_0 = m$ and $M_1 < M_2 < \dots$ denote the jump times of $(\mathcal{Z}^{(k+1)}[m, l])_{l \ge m}$ and consider the process

$(\mathcal{Y}_l)_{l \geq m}$ which is constant on $\{M_j + 1, \ldots, M_{j+1} - 1\}$ and satisfies

$$\mathcal{Y}_{M_{j+1}} = \mathcal{Y}_{M_j} + \mathbb{1}_{\left\{U_j \leq \frac{f(\mathcal{Y}_{M_j})}{f(\mathcal{Z}[m,M_j])}\right\}}. \tag{4.8}$$

By definition, $(\mathcal{Y}_l)_{l \geq m}$ has the same distribution as $(\mathcal{Z}^{(k)}[m,l])_{l \geq m}$. Since f is concave, we find that

$$\frac{f(\mathcal{Y}_{M_j})}{f(\mathcal{Z}^{(k)}[m,M_j])} \stackrel{\geq}{=} \frac{f(k) + (\mathcal{Y}_{M_j} - k)\frac{f(\mathcal{Z}^{(k)}[m,M_j]) - f(k)}{\mathcal{Z}^{(k)}[m,M_j] - k}}{f(k) + (\mathcal{Z}^{(k)}[m,M_j] - k)\frac{f(\mathcal{Z}^{(k)}[m,M_j]) - f(k)}{\mathcal{Z}^{(k)}[m,M_j] - k}}$$

and that

$$\frac{f(\mathcal{Z}^{(k)}[m,M_j]) - f(k)}{\mathcal{Z}^{(k)}[m,M_j] - k} \leq \Delta f(k),$$

which implies

$$\frac{f(\mathcal{Y}_{M_j})}{f(\mathcal{Z}[m,M_j])} \geq \frac{\mathcal{Y}_{M_j} + \frac{f(k)}{\Delta f(k)} - k}{\mathcal{Z}[m,M_j] + \frac{f(k)}{\Delta f(k)} - k}, \tag{4.9}$$

by Lemma B.4.

To $(\mathcal{Z}^{(k+1)}[m,M_j])_{j \geq 0}$ and $(\mathcal{Y}_{M_j})_{j \geq 0}$ we now couple a Pólya urn according to the following scheme: The urn process is started with an urn containing blue balls of total weight $B_0 = \beta = \frac{f(k)}{\Delta f(k)}$ and red balls of weight 1. In each step a colour is picked with probability proportional to the total weight of all balls of that colour and one ball of the picked colour is added to the urn. The total weight after j picks is $j + \beta + 1$ and we can represent the weight of the blue balls after j steps by

$$B_{j+1} = B_j + \mathbb{1}_{\left\{U_j \leq \frac{B_j}{j+\beta+1}\right\}}.$$

Now (4.8) and (4.9) imply that if blue is picked in step j, then the evolution $(\mathcal{Y}_l)_{l \geq m}$ increases at time M_j. As $(\mathcal{Z}^{(k+1)}[m,l])_{l \geq m}$ is independent of $(U_j)_{j \geq 0}$, we have, for $0 \leq i \leq l$,

$$\mathbb{E}[\mathcal{Y}_l | \mathcal{Z}^{(k+1)}[m,l] = k+1+i] - k \geq \mathbb{E}[B_i - B_0] = \frac{\beta}{1+\beta}(i+\beta+1) - \beta = i\frac{f(k)}{f(k+1)},$$

and concavity of f implies that

$$\mathbb{E}[f(\mathcal{Y}_l) | \mathcal{Z}^{(k+1)}[m,l] = k+1+i] \geq f(k)\frac{f(i+k+1)}{f(k+1)},$$

thus

$$\frac{\mathbb{E}f(\mathcal{Z}^{(k+1)}[m,l])}{\mathbb{E}f(\mathcal{Y}_l)} \leq \frac{f(k+1)}{f(k)},$$

which completes the proof of (4.5), since $\mathbb{E}f(\mathcal{Z}^{(k)}[m,l]) = \mathbb{E}f(\mathcal{Y}_l)$ and $\mathbb{E}f(\mathcal{Z}^{(k+1)}[m+1,l]) \leq \mathbb{E}f(\mathcal{Z}^{(k+1)}[m,l])$. $\qquad\square$

4.2 Untruncated first moment bounds

In this short section, we derive a statement that yields lower bounds on the typical distance in the setting of Theorem II, i.e. if $\lim_{k\to\infty}\Delta f(k) = \gamma = \frac{1}{2}$. The only assumption is a bound on the edge correlation. This is a variation of standard results, see e.g. [BR04, Lemma 4], which can also be formulated for more general subgraphs than just self-avoiding paths. The same first moment bounds can be applied for the parameter range $\gamma \in \left(0, \frac{1}{2}\right)$, which is at the bottom of the lower bound proof for the corresponding setting in [DHH10]. As we have seen in Section 3.1.1, a slight modification also leads to bounds for in range $\gamma > \frac{1}{2}$.

Lemma 4.16. *Let $\mathscr{G}_N = ([N], E)$ be a random graph and assume that there is an absolute constant $\kappa \geq 0$, such that, for any self-avoiding path $P = (v_0, \ldots, v_l)$,*

$$\mathbb{P}\{P \subset \mathscr{G}_N\} \leq \kappa^l \prod_{j=0}^{l-1} \mathbb{P}\{v_j \leftrightarrow v_{j+1}\} \tag{4.10}$$

and that there is a function $\Psi : \mathbb{N} \longrightarrow [0, \infty)$ such that

$$\mathbb{P}\{v \leftrightarrow w\} \leq \frac{\Psi(N)}{\sqrt{vw}}, \quad \text{for all } v, w \in [N], \tag{4.11}$$

and

$$\lim_{N\to\infty} \frac{\log \Psi(N)}{\log N} = 0. \tag{4.12}$$

Then, for uniformly chosen vertices $V, W \in \mathscr{G}_N$ and any $\delta \in (0, 1)$,

$$\lim_{N\to\infty} \mathbb{P}\left\{d_N(V, W) \geq \frac{(1-\delta)\log N}{\log\log N + \log \Psi(N)}\right\} = 1.$$

Remark 4.17. The proof below shows that the condition on Ψ is not the best possible, and that also δ may be chosen dependent on N. The lemma as stated is however sufficient to prove the lower bound of Theorem II in Section 6.3.

Proof of Lemma 4.16. Let

$$1 \leq l \leq L = L(N) = \left\lfloor \frac{(1-\delta)\log N}{\log\log N + \log \Psi(N)} \right\rfloor$$

and let $P = (v_0, \ldots, v_l)$ be self-avoiding. Then the assumptions imply that

$$\mathbb{P}\{P \subset \mathscr{G}_N\} \leq \kappa^l \prod_{j=0}^{l-1} \frac{\Psi(N)}{\sqrt{v_j v_{j+1}}} = \frac{\left(\kappa\Psi(N)\right)^l}{\sqrt{v_0 v_l}} \prod_{j=1}^{l-1} \frac{1}{v_j}.$$

Hence for $v, w \in [N]$ and $\mathscr{P}_l(v, w)$ denoting the set of all self-avoiding paths of length l from

v to w

$$\mathbb{P}\{d_N(v,w) \le L\} \le \sum_{l=1}^{L} \sum_{(v_0,\ldots,v_l)\in\mathscr{P}_l(v,w)} \frac{\left(\kappa\Psi(N)\right)^l}{\sqrt{vw}} \prod_{j=1}^{l-1} \frac{1}{v_j} \le \sum_{l=1}^{L} \frac{\left(\kappa\Psi(N)\right)^l}{\sqrt{vw}} \left(\sum_{j=1}^{N} \frac{1}{j}\right)^{l-1}$$

$$\le \frac{1}{\sqrt{vw}} \sum_{l=1}^{L} \kappa^l \Psi(N)^l (\log N)^{l-1} \le \frac{L\kappa^L \Psi(N)^L (\log N)^{L-1}}{\sqrt{vw}}. \tag{4.13}$$

For any $\varepsilon \in (0,1)$, the probability that one of the vertices U, V is smaller than $\frac{\varepsilon}{3}N$ is bounded by $\frac{2}{3}\varepsilon$ and thus using (4.13) on the complement of this event results in

$$\mathbb{P}\left\{d_N(V,W) < \frac{(1-\delta)\log N}{\log\log N + \log\Psi(N)}\right\} \le \sum_{v,w\ge\frac{\varepsilon}{3}N} \mathbb{P}\{d_N(v,w)\le L\}\mathbb{P}\{V=v,W=w\} + \frac{2\varepsilon}{3}$$

$$\le \frac{3L(\kappa\Psi(N))^L(\log N)^{L-1}}{\varepsilon N} + \frac{2\varepsilon}{3} \le \frac{3(1-\delta)(\kappa\Psi(N))^L(\log N)^L}{\varepsilon N(\log\log N + \log\Psi(N))} + \frac{2\varepsilon}{3}$$

$$\le \frac{3N^{(1-\delta)(\log\kappa+\log\log N+\log\Psi(N))/(\log\log N+\log\Psi(N))-1}}{\varepsilon(\log\log N + \log\Psi(N))} + \frac{2\varepsilon}{3}.$$

The numerator in the last line decays like a small power of N, if N is sufficiently large, so condition (4.12) guarantees that the sum is bounded by ε. $\qquad\square$

4.3 The local exploration processes

This section is devoted to presenting the details of the exploration process introduced in Section 3.2. Our eventual aim is to couple the network exploration process to an exploration process on the INT $\mathfrak{T}$, which will essentially look like the killed IBRW. To this end, we first specify the exact exploration scheme and then introduce an intermediate object, a discrete random tree on which an exploration processes can be defined that can be compared directly to the exploration process in the network.

4.3.1 Exploration of the network

We now specify the first stage, or *local stage*, of the exploration process used in our network (or in any finite graph, e.g. the tree described below), i.e. we specify the way we collect information about the neighbourhood of a particular vertex v.

Definition 4.18 (Exploration process)**.** To initiate the *exploration process* around $v \in \mathscr{G}_N$, we explore in the first step all immediate neighbours of v in the graph. To describe a general exploration step we classify the vertices in three categories:

- *veiled vertices*: vertices for which we have not yet found connections to the neighbourhood of v;

- *active vertices*: vertices for which we already know that they belong to the neighbourhood, but for which we have not yet explored all its immediate neighbours;

- *dead vertices*: vertices which belong to the neighbourhood and for which all immediate neighbours have been explored.

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

We now describe a tree $\mathbb{T}(w)$ which describes the neighbourhood of a vertex $w \in \mathcal{G}_N$ as long as it contains no cycles.

Definition 4.19 (Discrete neighbourhood tree)**.** $\mathbb{T}(w)$ is constructed inductively. Any vertex $v \in \mathbb{T}(w)$ is labelled by two parameters: its *location* $m(v) \in [N]$, and its *type* $t(v)$, an element of $\{\ell\} \cup \{m(v), \dots N\}$. The root $\emptyset$ of the tree has location $m(\emptyset) = w$ and type ℓ.

- A vertex v of type ℓ produces descendants

    - in the locations $1, \dots, m(v) - 1$ of type i with probability

$$\mathbb{P}\{v \text{ has a descendent in location } j \text{ of type } i\} = \mathbb{P}\{\Delta \mathcal{Z}[j, i-1] = 1\}; \qquad (4.14)$$

    - in the locations $m(v) + 1, \dots, N$, it independently produces descendants, all of type ℓ, such that the cumulative sum of these descendants is distributed according to the law of

$$\left(\mathcal{Z}[i, j]; i + 1 \le j \le N \right).$$

- If v is of type $t(v) = k \in [N]$, it produces descendants

    - to the left in the same way as a vertex of type ℓ;

    - to the right, independently, it produces descendants of type ℓ, in such a way that the cumulative sum of these descendants is distributed as

$$\left(\mathcal{Z}[i, j] - \mathbb{1}_{[k, \infty)} : i + 1 \le j \le N \right)$$

    conditioned on $\Delta \mathcal{Z}[i, k-1] = 1$.

That the tree resembles the neighbourhood of $w \in \mathcal{G}_N$ up to a certain stopping time is the content of the following proposition, which is a slight generalisation of the corresponding statement in [DM13], but has exactly the same proof. The result follows from a coupling between the explorations in the network and in the tree which is based on a careful analysis of an exploration step.

Proposition 4.20 (Coupling of explorations, [DM13, Proposition 5.1])**.** *Suppose that* $(c_N)_{N\in\mathbb{N}}$ *and* $(n_N)_{N\in\mathbb{N}}$ *are a sequences of positive integers with*

$$c_N = o(\log N \log\log N) \quad and \quad \lim_{N\to\infty} \frac{c_N^2}{n_N^{1-\gamma^+}} = 0,$$

then one can couple the pair $(V, \mathcal{G}_N)$ *consisting of the network and a uniformly chosen vertex* V *with* $\mathbb{T}(V)$ *such that with high probability either*

$$\mathscr{C}_N^k(V) = \mathbb{T}^k(V), \quad for\ all\ k \le c_N, \tag{4.15}$$

where $\mathscr{C}_N^k(V)$ *and* $\mathbb{T}^k(V)$ *denote the connected components uncovered in the first k steps of the exploration, or*

$$\mathscr{C}_N^k \cap [n_N] \ne \emptyset, \quad for\ some\ k \in [c_N], \quad and\ \#\mathscr{C}_N \ge c_N. \tag{4.16}$$

Remark 4.21. As c_N can be chosen to diverge as $N \to \infty$, we can potentially couple neighbourhoods of *unbounded* size. This is necessary for the argumentation in [DM13], but when we apply the coupling result in Chapters 5 and 6, it is often sufficient to look at arbitrarily large *bounded* neighbourhoods.

Before we proceed to present the proof of Proposition 4.20 given in [DM13], we need to state one more preliminary result, namely a rough a priori estimate on the size of the giant component which can be obtained using only the constant lower bound $f(k) \ge f(0)$, $k \ge 0$ and the exponential Chebyshev inequality. We omit the proof, since it is not very instructive and refer to [DM13, p. 29] for the details.

Lemma 4.22 (A priori estimate, [DM13, Lemma 4.2])**.** *Let* $(c_N)_{N\in\mathbb{N}}$ *and* $(n_N)_{N\in\mathbb{N}}$ *be sequences of positive integers satisfying*

$$c_N = o(\log N \log\log N) \quad and \quad \log n_N = o(\log N),$$

and denote by $\mathscr{C}_N(v) \subset \mathcal{G}_N$ *the connected component containing* $v \in [N]$*. Then*

$$\lim_{N\to\infty} \mathbb{P}\{\#\mathscr{C}_N(v) < c_N \text{ for any } v \in [n_N]\} = 0.$$

4.3.2 Coupling the explorations - proof of Proposition 4.20

In the following, we couple the exploration process of the network, started with a particle at position v and type ℓ, to the exploration process of the tree $\mathbb{T}(v)$, started in the root, up to a stopping time T. Before we introduce the coupling explicitly, we quote adverse events which stop the coupling.

- Whenever the exploration processes of the network revisits an active vertex we have found a cycle in the network. We call this event *(E1)* and stop the exploration so that,

before time *T*, the explored part of the neighbourhood of *v* is a tree with each node having a unique location.

- We stop once the explored part of the network differs from the explored part of the random labelled tree, calling this event *(E2)*, we shall see below how this can happen.

In cases (E1) and (E2) we say that the *coupling fails*.

Further reasons to stop the exploration are, for certain parameters $1 \leq n_N, c_N \leq N$,

- (*A*) the number of dead and active vertices exceeds c_N,

- (*B*) one vertex in $[n_N]$ is activated, and

- (*C*) there are no more active vertices left.

If we stop the exploration without (E1) and (E2) being the case, we say that the *coupling succeeds*. Once the exploration has stopped, the veiled parts of the random tree and the network may be generated independently of each other with the appropriate probabilities. Hence, if we succeed in coupling the explorations, we have coupled the random labelled tree and the network.

To distinguish both exploration processes, we use the term *descendant* for a child in the labelled random tree and the term *immediate neighbour* in the context of the neighbourhood exploration of the network. In the initial step, we explore all immediate neighbours of *v* and all the descendants of the root. Both explorations are identically distributed thus can be perfectly coupled. Suppose now that we have performed *k* steps and that we have not yet stopped the exploration. In particular, this means that both explored subgraphs coincide and that any unveiled element of the labelled random tree can be uniquely referred to by its location. We now explore the descendants and immediate neighbours of the *leftmost* active vertex, say *n*.

- We explore the descendants to the left (veiled and dead) and immediately check if they themselves have right descendants in the set of dead vertices. If we discover no dead descendants, the set of newly found left descendants is identically distributed to the immediate left neighbours in the network. Thus we can couple both explorations such that they agree in this case. Otherwise we stop the exploration due to (E2).

- We explore the descendants to the right. If the vertex *n* is *not of type ℓ*, then *n* has no right descendants that were marked as dead when *n* itself was discovered. Since we always explore the leftmost active vertex there are no new dead vertices to he right of *n*. Therefore, the explorations to the right in the network and the random labelled tree are identically distributed and we stop, if we find right neighbours in the set of active vertices due to (E1). If the vertex *n* is *of type ℓ*, then we have not gained any

information about its right descendants yet. If we find no right descendants in the set of dead vertices, it is identically distributed to the immediate right neighbours of n in the network. We stop if a right descendant is discovered that was marked as dead, corresponding to (E2), or if right descendants are discovered in the set of active vertices, corresponding to (E1).

Lemma 4.23 (Success of the coupling). *Suppose that* $(c_N)_{N\in\mathbb{N}}$, $(n_N)_{N\in\mathbb{N}}$ *are sequences of integers such that*

$$\lim_{N\to\infty}\frac{c_N^2}{n_N^{1-\gamma^+}}=0$$

Then the coupling of the exploration process satisfies

$$\lim_{N\to\infty}\sup_{v\in\{n_N+1,\dots,N\}}\mathbb{P}\{\text{coupling with initial vertex } v \text{ ends in (E1) or (E2)}\}=0$$

i.e. the coupling succeeds with high probability.

Proof. We analyse one exploration step in detail. Let $\mathfrak{a}$ and $\mathfrak{d}$ denote two disjoint subsets of $\{n_N+1,\dots,N\}$ with $\#(\mathfrak{a}\cup\mathfrak{d}) < c_N$ and $\#\mathfrak{a}\neq 0$. The exploration of the minimal vertex n in the set $\mathfrak{a}$ may only fail for one of the following reasons:

  (i) the vertex n has left descendants in $\mathfrak{d}$;

 (ii) the vertex n has left descendants which themselves have right descendants in $\mathfrak{d}$;

(iii) the vertex n has right descendants in $\mathfrak{a}\cup\mathfrak{d}$.

Indeed, if (i) and (ii) do not occur then the exploration to the left ends neither in state (E1) nor in (E2), and if (iii) does not happen the exploration to the right does not fail. Conditionally on the configuration $(\mathfrak{a},\mathfrak{d})$, the probability for the event (i) equals

$$\mathbb{P}\{\exists a\in\mathfrak{d} \text{ such that } \Delta\mathcal{Z}[a,n-1]=1\}\leq\sum_{\substack{a\in\mathfrak{d}\\a<n}}\mathbb{P}\{\Delta\mathcal{Z}[a,n-1]=1\}$$

whereas the probability for (ii) is by Lemma 4.14 equal to

$$\mathbb{P}\{\exists a\in\mathfrak{d}^c \text{ and } b\in\mathfrak{d} \text{ such that } \Delta\mathcal{Z}[a,n-1]=\Delta\mathcal{Z}[a,b-1]=1\}$$

$$\leq\sum_{\substack{a\in\mathfrak{d}^c\\a<n}}\sum_{\substack{b\in\mathfrak{d}\\b>a}}\mathbb{P}\{\Delta\mathcal{Z}[a,n-1]=\Delta\mathcal{Z}[a,b-1]=1\}$$

$$\leq\sum_{\substack{a\in\mathfrak{d}^c\\a<n}}\sum_{\substack{b\in\mathfrak{d}\\b>a}}\mathbb{P}\{\Delta\mathcal{Z}[a,n-1]=1\}\mathbb{P}^1\{\Delta\mathcal{Z}[a,b-1]=1\}$$

If the vertex n is of the type $\tau\neq\ell$, then the conditional probability of (iii) is

$$\mathbb{P}\{\exists a\in\mathfrak{a} \text{ such that } \Delta\mathcal{Z}[n,a-1]=1|\mathcal{Z}[n,\tau-1]=1,\Delta\mathcal{Z}[n,b-1]=0\forall b\in\mathfrak{d}\setminus\{\tau\}\}$$

$$\leq C \sum_{\substack{a \in \mathfrak{a} \cup \mathfrak{d} \\ a > n}} \mathbb{P}^1 \{ \Delta \mathscr{Z}[n, a-1] = 1 \},$$

using first [DM13, Lemma 2.12] (or Lemma 6.12) to obtain the constant $C > 0$ and then Lemma 4.14. If the vertex n is of type ℓ, the conditional probability of (iii) is

$$\mathbb{P}\{\exists a \in \mathfrak{a} \cup \mathfrak{d} \text{ such that } \Delta \mathscr{Z}[n, a-1] = 1\} \leq \sum_{\substack{a \in \mathfrak{a} \cup \mathfrak{d} \\ a > n}} \mathbb{P}\{\Delta \mathscr{Z}[n, a-1] = 1\}$$

Since, by Lemma 4.3, for any $a > n$,

$$\mathbb{P}^1 \{ \Delta \mathscr{Z}[n, a-1] = 1 \} \leq \mathbb{P}^1 \{ \Delta \mathscr{Z}[n_N + 1, n_N + 1] = 1 \},$$

we conclude that the probabilities of the events (i) and (iii) are bounded by

$$(2 + C) c_N \mathbb{P}^1 \{ \Delta \mathscr{Z}[n_N + 1, n_N + 1] = 1 \},$$

independently of the type τ. Moreover, the probability of (ii) is bounded by

$$c_N \mathbb{P}\{ \Delta \mathscr{Z}[1, n_N] = 1 \} \sum_{a=1}^{n-1} \mathbb{P}\{ \mathscr{Z}[a, n-1] = 1 \}.$$

The sum is the expected outdegree of vertex n, which is uniformly bounded by Corollary 4.9, and hence one of the events (i), (ii) or (iii) occurs in one step with probability less than a constant multiple of $c_N \mathbb{P}^1 \{ \mathscr{Z}[1, n_N] = 1 \}$. As there are at most c_N exploration steps until we end in one of the states (A),(B) or (C), the coupling fails due to *(E1)* or *(E2)* with a probability bounded from above by a constant multiple of

$$c_N^2 \mathbb{P}^1 \{ \Delta \mathscr{Z}[1, n_N] = 1 \} \leq C_{4.9} f(1) \frac{c_N^2}{n_N^{1 - \gamma^+}} \to 0,$$

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

Proof of Proposition 4.20. We apply the coupling of Lemma 4.23 with $(n_N)_{N \in \mathbb{N}}$ satisfying

$$\lim_{N \to \infty} \frac{\log n_N}{\log N} = 0 \text{ and } \lim_{N \to \infty} \frac{(\log N \log \log N)^2}{n_N^{1 - \gamma^+}} = 0.$$

Then, by the a priori estimate Lemma 4.22, we get that with high probability

$$\text{coupling ends in (B)} \Rightarrow \#\mathscr{C}_N(V) \geq c_N. \tag{4.17}$$

One also obtains $\lim_{N \to \infty} \max_{v=1,\dots,n_N} \mathbb{P}\{\#\mathbb{T}(v) < c_N\} = 0$ so that implication (4.17) is also valid for $\#\mathbb{T}(V)$. Since the coupling succeeds we have, with high probability,

$$\text{coupling ends in (A)} \Leftrightarrow \mathscr{C}_N^k(V) = \mathbb{T}^k(V) \text{ for all } k \in [c_N],$$

or

$$\text{coupling ends in (B)} \Leftrightarrow \#\mathscr{C}_N(V) \geq c_N \text{ and } k \in [c_N],$$

and the statement follows immediately. □

4.4 The idealised exploration process

We proceed in this section with the approximation of the local neighbourhood of a randomly chosen vertex $V \in \mathscr{G}_N$ by the idealized random tree $\mathfrak{T}$ featuring in Chapter 3. Vertices in the network $\mathscr{G}_N$ are mapped onto particles on the negative half line in such a way that the vertex with index $n \in \{1, ..., N\}$ is mapped into position $t_n - t_N$, where

$$t_n = \sum_{i=1}^{n-1} \frac{1}{i}, \ \text{ for all } n \in \mathbb{N}.$$

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

Any fixed interval $[a, b]$ fills up with more and more particle as $N \to \infty$. At the same time the age of the vertex corresponding to a particle closest to a fixed position in $[a, b]$ is increasing, which means that the probability of edges between two such vertices is decreasing. Below we see, that the combination of these two effects leads to convergence of the distribution of offspring locations on the half line. Thanks to the independence of edges with a common right endpoint, offspring to the left converge to a Poisson process by the law of small numbers, while offspring to the right converge to the point processes corresponding to the pure birth process $(Z_t : t > 0)$, if there is no dependence on previous generations.

The considerations of the previous section suggest that the only form of dependence of the offspring distribution of a vertex on previous generations, is via the relative position of its father. This information is encoded in the type of a particle, where type ℓ indicates that the father is to the left of the particle, and a numerical type τ indicates that the father is positioned τ units to the right. It should be noted that the relative positions of offspring particles only depend on the absolute position of the reproducing particle via the removal of particles whose position is not in the negative half line, and which therefore do not correspond to vertices in the network $\mathscr{G}_N$. This produces the random walk structure, which is crucial to the analysis of the underlying tree performed in [DM13]. Our main aim here is to present the derivation of the following result.

Proposition 4.24. *Suppose that* $(c_N)_{N \in \mathbb{N}}$ *is a sequence of integers with*

$$\lim_{N \to \infty} \frac{c_N}{\log N \log \log N} = 0$$

58

Then each pair $(V, \mathcal{G}_N)$ can be coupled with $\mathfrak{T}$ such that with high probability $\mathcal{C}_N(V)$ agrees with $\mathfrak{T}$ up to size c_N, in the sense that the discrete tree obtained from the first c_N exploration steps around V is the same as the tree obtained from c_N exploration steps around the root in $\pi_N^{-1}(\mathfrak{T})$.

We know from the previous section, that the neighbourhood of a vertex v in a large network is similar to the random tree $\mathbb{T}(v)$. To establish the relationship between $\mathbb{T}(V)$, for an initial vertex V chosen uniformly from $\{1, ..., N\}$ and the idealised neighbourhood tree $\mathfrak{T}$ we apply to each element of the INT $\mathfrak{T}$ the projection

$$\pi_N : (-\infty, 0] \longrightarrow \{1, ..., N\},$$

which maps onto the smallest $m \in \{1, ..., N\}$ with $t \leq -t_N + t_m$. We obtain a branching process with location parameters in $\{1, ..., N\}$, which we call π_N-projected INT. We wish to show, using a suitable coupling, that when the INT is started with a particle in $-X$, where X is standard exponentially distributed, then this projection is close to the random tree $\mathbb{T}(V)$. Again we apply the concept of an exploration process.

To this end we show that, for every particle $x = (s(x), \ell)$ with $s \leq 0$, the π_N−projected descendants of x have a similar distribution as the descendants of a vertex in location $n := \pi_N(s(x))$ in the labelled tree $\mathbb{T}(n)$. We establish couplings of both distributions and bound the probability of failure.

4.4.1 Coupling the evolution to the right for ℓ-type vertices

We fix $s \leq 0$ and $N \in \mathbb{N}$, and suppose that $m := \pi_N(s) \geq 2$. For an ℓ-type vertex in s the cumulative sum of π_N-projected right descendants is distributed as $(Z_{t_n - t_N - s})_{m \leq n \leq N}$. This distribution has to be compared with the distribution of $(\mathcal{Z}[m, n])_{m \leq n \leq N}$, which is the cumulative sum of right descendants of m in $\mathbb{T}(m)$.

Lemma 4.25. *Fix a level $T \in \mathbb{N}$. For any $s \leq 0$ with $\pi_N(s) = m \in \{2, 3, ..., N\}$ we can couple the processes $\left(Z_{t_n - t_N - s} : n \geq m\right)$ and $(\mathcal{Z}[m, n] : n \geq m)$ such that for the coupled processes $\left(\mathcal{Y}^{(1)}[n] : n \geq m\right)$ and $\left(\mathcal{Y}^{(2)}[n] : n \geq m\right)$ we have*

$$\mathbb{P}\left(\mathcal{Y}^{(1)}[n] \neq \mathcal{Y}^{(2)}[n] \text{ for some } n \leq \tau\right) \leq \left(f(0) + f(T)^2\right) \frac{1}{m-1},$$

where τ is the first time when one of the processes reaches or exceeds T.

Proof. We define the process $\mathcal{Y} = \left(\left(\mathcal{Y}^{(1)}[n], \mathcal{Y}^{(2)}[n]\right) : n \geq m\right)$ to be the Markov process with starting distribution $\mathcal{L}\left(\mathcal{Z}_{t_m - t_N - s}\right) \otimes \delta_0$ and transition kernels $p^{(n)}$ such that the first and second marginal are the respective transition probabilities of $(Z_{t_m - t_N - s} : n \geq m)$ and $(\mathcal{Z}[m, n] : n \geq m)$ and, for any integer $a \geq 0$, the law $p^{(n)}((a, a), \cdot)$ is the coupling of the laws of $Z_{\Delta t_n}$ and $\mathcal{Z}[n, n+1]$ under $\mathbb{P}^a$ provided in [DM13, Lemma 2.13]. Then the processes $\left(\mathcal{Y}^{(1)}[n] : n \geq m\right)$

are distributed as stated in the lemma. Moreover, letting σ denote the first time when they disagree, we get

$$\mathbb{P}\{\sigma \leq \tau\} = \sum_{n=m}^{\infty} \mathbb{P}\{\sigma = m\} + \sum_{n=m}^{\infty} \mathbb{P}\{\sigma = n+1 | \tau > n, \sigma > n\}$$

and, by [DM13, Lemma 2.13],

$$\mathbb{P}\{\sigma = n+1 | \tau > n, \sigma > n\} \leq \left(f(T)\frac{1}{n}\right)^2, \qquad \text{for } n \in \{m, m+1, ...\}.$$

Moreover, $\mathbb{P}\{\sigma = m\} = \mathbb{P}\{\mathscr{Y}^{(1)}[m] > 0\} = 1 - \exp[-(t_m - v)f(0)] \leq \frac{f(0)}{m-1}$. Consequently,

$$\mathbb{P}\{\sigma \leq \tau\} \leq \frac{f(0)}{m-1} + f(T)^2 \sum_{n=m}^{\infty} (f(0) + f(T)^2)\frac{1}{m-1}.$$

$\square$

4.4.2 Coupling the evolution to the left

Recall that a particle in position $s \leq 0$ produces a number of π_N-projected descendants which are Poisson-distributed at the locations $m \leq n := \pi_N(s)$ with parameter

$$\lambda := \int_{-tN+t_m-1}^{(-t_N+t_m)\wedge s} \exp[-(s-u)]\mathbb{E}f(Z_{s-u})\,du. \tag{4.18}$$

We adopt the convention that $t_0 = -\infty$. A vertex in location n in $\mathbb{T}(n)$ produces a Bernoulli distributed number of descendants in m with success probability $\mathbb{P}\{\Delta\mathscr{Z}[m, n-1] = 1\}$ for $m < n$ and no descendants for $m = n$. The following lemma provides a coupling of both distributions.

Lemma 4.26. *There exists a constant $C > 0$ such that the following holds: Let $m, n \in \mathbb{N}$ and $s \leq 0$ with $m \leq n := \pi_N(s)$ and define λ as in (4.18). If $m < n$, one can couple a $\text{Poiss}(\lambda)$-distributed random variable with $\Delta\mathscr{Z}[m, n-1]$, such that the coupled random variables $\mathscr{Y}^{(1)}$ and $\mathscr{Y}^{(2)}$ satisfy*

$$\mathbb{P}\{\mathscr{Y}^{(1)} \neq \mathscr{Y}^{(2)}\} \leq C\frac{1}{m^{1+\gamma^+}}\frac{1}{n^{1-\gamma^+}}.$$

If $m = n$, a $\text{Poiss}(\lambda)$-distributed random variable $\mathscr{Y}^{(1)}$ satisfies

$$\mathbb{P}\left(\mathscr{Y}^{(1)} \neq 0\right) \leq C\frac{1}{n}.$$

Proof. It suffices to prove the second statement for $m = n \geq 2$ note that $u \mapsto \exp[-u]\mathbb{E}f(Z_u)$ is decreasing so that

$$\lambda \leq \int_{s}^{-t_N+t_{n-1}} \exp[-(s-u)]\mathbb{E}f(Z_{s-u})\,du \leq f(0)\frac{1}{n-1}$$

which leads directly to the second statement of the lemma. Next, consider the case where $2 \le m < n$. Note that for $u \in (\, t_N + t_{m-1}, -t_N + t_m\,]$, one has $s - u \in (t_{n-1} - t_m, t_n - t_{m-1})$ which, using again that $u \mapsto \mathrm{e}^{-u}\mathbb{E}f(Z_u)$ is decreasing, implies that

$$\frac{1}{m-1}\mathrm{e}^{-t_n-t_{m-1}}\mathbb{E}f(Z_{t_n-t_{m-1}}) \le \lambda \le \frac{1}{m-1}\mathbb{E}f(Z_{t_n-t_m})].$$

Next, note that by definition of t_n we have $\log \frac{n}{m} \le t_n - t_m \le \log \frac{n-1}{m-1}$, so that

$$\left(1 - \frac{1}{m-1}\right)\frac{1}{n-1}\mathbb{E}\left[f\left(Z_{t_{n-1}-t_m}\right)\right] \le \lambda \le \left(1 + \frac{1}{m-1}\right)\frac{1}{n-1}\mathbb{E}\left[f\left(Z_{t_{n-1}it_m}\right)\right]. \tag{4.19}$$

On the other hand, $\Delta\mathscr{Z}[m, n-1]$ is a Bernoulli random variable with success probability

$$p := \frac{1}{n-1}\mathbb{E}\left[f\left(\mathscr{Z}[m, n-1]\right)\right].$$

By [DM13, Lemma A.1] it suffices to control λ^2 and $|\lambda - p|$. By [DM13, Proposition 2.14] and (4.19),

$$\left|\lambda - p\right| \le C\frac{1}{m-1}\frac{1}{n-1}\left(\mathbb{E}\left[f\left(Z_{t_{n-1}-t_m}\right)\right] + \mathbb{E}\left[f\left(\mathscr{Z}[m, n-1]\right)\right]\right), \tag{4.20}$$

and

$$\lambda^2 \le 4\left(\frac{1}{n-1}\right)^2 \mathbb{E}\left[f\left(Z_{t_{n-1}-t_m}\right)\right]^2. \tag{4.21}$$

Since $t_{n-1} - t_m \le \log \frac{n-2}{m-1}$, we get, using Corollary 4.9 and a similar statement for the idealised degree evolution, that

$$\mathbb{E}\left[f\left(Z_{t_{n-1}-t_m}\right)\right] + \mathbb{E}\left[f\left(\mathscr{Z}[m, n-1]\right)\right] \le C\left(\frac{n}{m}\right)^{\gamma^+}.$$

Recalling that $n > m \ge 2$, it is now straightforward to deduce the statement from equations (4.20) and (4.21). It remains to consider the case where $1 = m < n$. Here, we apply the same argument and $t_{n-1} \ge \log(n-1)$ to deduce that

$$\lambda \le \int_{-\infty}^{-t_N+t_1} \exp[-(s-u)]\mathbb{E}f(Z_{s-u})\mathrm{d}u \le C\int_{t_{n-1}}^{\infty}\exp[-(1-\gamma^+)u]\mathrm{d}u \le \frac{C}{1-\gamma^+}(n-1)^{\gamma^+-1},$$

while, by Lemma 4.9, $\mathbb{P}\{\Delta\mathscr{Z}[1, n-1] = 1\} \le f(0)(n-1)^{\gamma^+-1}$, therefore a Poiss$(\lambda)$-distributed random variable can be coupled with $\Delta\mathscr{Z}[1, n-1]$ so that they disagree with probability less than a constant multiple of n^{γ^+-1}. $\qquad\square$

Remark 4.27. Lemma 4.26 provides a coupling for the mechanisms with which both trees produce *left* descendants. Since the number of descendants in individual locations form an independent sequence of random variables, we can apply the coupling lemma sequentially for each location and obtain a coupling of the π_N-projected left descendants of a particle in position s and the left descendants of $n = \pi_N(s)$ in $\mathbb{T}(n)$. Indeed, under the assumptions of

Lemma 4.26, one finds a coupling of both processes such that

$$\mathbb{P}\{\text{left descendants disagree}\} \le C\frac{1}{n} + C\frac{1}{n^{1-\gamma^+}}\sum_{m=1}^{n-1}\frac{1}{m^{1+\gamma^+}} \le C'\frac{1}{n^{1-\gamma^+}},$$

where $C' > 0$ is a suitable constant.

4.4.3 Coupling the evolution to the right for particles of type $\tau \ne \ell$

We fix $s \le 0$ and $N \in \mathbb{N}$, and suppose that $m := \pi_N(s) \ge 2$. Also fix a type $\tau < -s$ with $l := \pi_N(s+\tau) > m$. The cumulative sum of π_N-projected right descendants of a particle in s of type τ (including its predecessor) is distributed according to $\left(Z_{-t_N+t_n-s} : m \le n \le N\right)$ conditioned on $\Delta\mathcal{Z}_\tau = 1$. The cumulative sum of right descendants in $\mathbb{T}(m)$ of a vertex in m of type l (including the predecessor) is distributed according to the law of $(\mathcal{Z}[m,n] : m \le n \le N)$ conditioned on $\Delta\mathcal{Z}[m,l-1] = 1$. Both processes are Markov processes and we provide a coupling of their transition probabilities.

Lemma 4.28. *There exists a constant $C > 0$ such that the following holds: Let $k \ge 0, m, n \ge 1$ be integers with $k+1 < m < n$, and let $\tau \in (t_n - t_m, t_{n+1} - t_m]$. Then the random variables $Z_{\Delta t_m}$ under $\mathbb{P}^k\{\cdot|\Delta Z_\tau = 1\}$ and $\mathcal{Z}[m,m+1]$ under $\mathbb{P}^k\{\cdot|\Delta\mathcal{Z}[m,n] = 1\}$ can be coupled such that the resulting random variables $\mathcal{Y}^{(1)}$ and $\mathcal{Y}^{(2)}$ satisfy*

$$\mathbb{P}\{\mathcal{Y}^{(1)} \ne \mathcal{Y}^{(2)}\} \le C\left(\frac{f(k)}{m}\right)^2.$$

Proof. We couple $\mathcal{Y}^{(1)}$ and $\mathcal{Y}^{(2)}$ by plugging a uniform random variable on $(0,1)$ in the generalised inverses of the respective distribution functions and conclude that

$$\mathbb{P}\{\mathcal{Y}^{(1)} \ne \mathcal{Y}^{(2)}\} = \left|\mathbb{P}\{\mathcal{Y}^{(1)} = k\} - \mathbb{P}\{\mathcal{Y}^{(2)} = k\}\right| + \mathbb{P}\{\mathcal{Y}^{(1)} \ge k+2\}.$$

The second error term is of the required order, since

$$\mathbb{P}\{\mathcal{Y}^{(1)} \ge k+2\} \le \mathbb{P}^{k+1}\{Z_{1/m} \ge k+3\} \le \left(\frac{f(k+2)}{m}\right)^2.$$

It remains to analyse the first error term. We have

$$\mathbb{P}\{\mathcal{Y}^{(2)} = k\} = 1 - f(k)\Delta t_m \frac{\mathbb{E}^{k+1}f(\mathcal{Z}[m+1,n])}{\mathbb{E}^k f(\mathcal{Z}[m,n])},$$

and representing $\left(Z_t^{[\tau]} : t \ge 0\right)$ by its compensator,

$$\mathbb{P}\left(\mathcal{Y}^{(1)} = k\right) = \exp\left\{-f(k)\int_0^{\Delta t_m}\frac{P_{\tau-u}f(k+1)}{P_{\tau-u}f(k)}\,\mathrm{d}u\right\}.$$

We need to compare

$$\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)} \text{ and } \frac{P_u f(k+1)}{P_u f(k)} \text{ for } u \in [t_n - t_{m+1}, t_{n+1} - t_m].$$

By [DM13, Proposition 2.14], one has, for $a \in \{k, k+1\}$ and sufficiently large m,

$$P_u f(a) \le P_{t_{n+1}-t_m} f(a) \le e^{\gamma^+}\left(\frac{1}{m} + \frac{1}{n}\right) P_{t_n - t_{m+1}} f(a)$$

$$\le e^{\gamma^+}\left(\frac{1}{m} + \frac{1}{n}\right)\left(1 + C\frac{f(a)}{m}\right) P_{m+1,n} f(a)$$

$$\le e^{\gamma^+}\left(\frac{1}{m} + \frac{1}{n}\right) + C\frac{f(a)}{m} P_{m+1,n} f(a),$$

for some constant $C > 0$. Conversely,

$$P_u f(a) \ge P_{t_n - t_{m+1}} f(a) \ge e^{-\frac{\gamma^+}{m}} P_{t_n - t_m} f(a) \ge e^{\frac{\gamma^+}{m}}\left(1 - C\frac{f(a)}{m}\right) P_{m,n} f(a)$$

We only need to consider large m and we may assume that $C\frac{f(k+1)}{m} \le \frac{1}{2}$, as otherwise we may choose C' large enough to ensure that the right hand side in the display of the lemma exceeds one. Then

$$P_u f(a) \ge \exp\left[\gamma^+ \frac{1}{m} - 2C\frac{f(a)}{m}\right] P_{m,n} f(a),$$

Since $\exp[-2y] \le 1 - y$ for $y \in \left[0, \frac{1}{2}\right]$. Consequently,

$$\exp\left[-\gamma^+\left(2\frac{1}{m} + \frac{1}{n}\right) - 3C\frac{f(k+1)}{m}\right]\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}$$

$$\le \frac{P_u f(k+1)}{P_u f(k)}$$

$$\le \exp\left[\gamma^+\left(2\frac{1}{m} + \frac{1}{n}\right) + 3C\frac{f(k+1)}{m}\right]\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}.$$

Recall that $\frac{P_{m+1,n}f(k+1)}{p_{m,n}f(k)}$ is uniformly bounded over all k so that we arrive at

$$\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)} - C'\frac{f(k)}{m} \le \frac{P_u f(k+1)}{P_u f(k)} \le \frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)} + C'\frac{f(k)}{m},$$

for an appropriate constant $C' > 0$. Therefore,

$$\mathbb{P}\{\mathscr{Y}^{(1)} = k\} - \mathbb{P}\{\mathscr{Y}^{(2)} = k\} \le 1 \wedge \exp\left[-f(k)\Delta t_m\left(\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)} - C'\frac{f(k)}{m}\right)\right]$$

$$-\left(1 - f(k)\Delta t_m\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}\right)$$

$$\le C'\left(\frac{f(k)}{m}\right)^2 + \frac{1}{2}\left(f(k)\Delta t_m\left(\frac{P_{m+1,n}f(k+1)}{P_{n,m}f(k)} - C'\frac{f(k)}{m}\right)\right)^2 \le C''\left(\frac{f(k)}{m}\right)^2,$$

for some suitable constant $C'' > 0$. Similarly, one finds that

$$\mathbb{P}\{\mathscr{Y}^{(2)} = k\} - \mathbb{P}\{\mathscr{Y}^{(1)} = k\} \le C'' \left(\frac{f(k)}{m} \right)^2,$$

and putting everything together yields the assertion. $\qquad\square$

From Lemma 4.28 we get the following analogue of Lemma 4.25.

Lemma 4.29. *Fix a level $T \in \mathbb{N}$. For any $s \le 0$ and $\tau \le -s$ with $\pi_N(s) = m \in \{2, 3, ..., N\}$ and $m < 1 := \pi_N(s + \tau)$ we can couple to process $\left(Z_{t_n - t_N - s} : n \ge m \right)$ conditioned on $\Delta Z_\tau = 1$ and $(\mathscr{Z}[m, n])_{n \ge m}$ conditioned on $\mathscr{Z}[m, l - 1] = 1$ such that the coupled processes*

$$\left(\mathscr{Y}^{(1)}[n] : n \ge m \right) \quad and \quad \left(\mathscr{Y}^{(2)}[n] : n \ge m \right)$$

satisfy

$$\mathbb{P}\left(\mathscr{Y}^{(1)}[n] \ne \mathscr{Y}^{(2)}[n] \text{ for some } n \le \sigma \right) \le C \left(f(T)^2 + 1 \right) \frac{1}{m}$$

where σ is the first time when one of the processes reaches or exceeds level T and C is a positive constant.

Proof. We define the process $\mathscr{Y} = \left(\left(\mathscr{Y}^{(1)}[n], \mathscr{Y}^{(2)}[n] \right) : n \ge m \right)$ to be the Markov process with starting distribution $\mathscr{L}\left(Z_{t_m - t_N - s} | \Delta Z_\tau = 1 \right) \otimes \delta_0$ and transition kernels $p^{(n)}$ such that the first and second marginal are the conditioned transition probabilities of $\left(Z_{t_n - t_N - s} : n \ge m \right)$ and $(\mathscr{Z}[m, n] : n \ge m)$ as stated in the lemma. In the case where $n < l - 1$, we demand that, for any integer $a \ge 0$, the law $p^{(n)}((a, a,), \cdot)$ is the coupling of the laws $Z_{\Delta t_n}$ under

$$\mathbb{P}^a\{\cdot | \Delta Z_{\tau - (t_n - t_N - s)} = 1\}$$

and $\mathscr{Z}[n, n + 1]$ under

$$\mathbb{P}^a\{\cdot | \Delta \mathscr{Z}[n, l - 1] = 1\}$$

provided in Lemma 4.28. Conversely, we apply the unconditioned coupling of Lemma 4.25 for $n \ge l$. Letting ϱ denote the first time when both evolutions disagree, we get

$$\mathbb{P}\{\varrho \le \sigma\} = \sum_{n=m}^{\infty} \mathbb{P}\{\sigma \ge n, \varrho = n\} \le \mathbb{P}\{\varrho = m\} + \sum_{n=m}^{\infty} \mathbb{P}\{\varrho = n + 1 | \sigma > n, \varrho > n\}$$

and, by Lemma 4.25 and Lemma 4.28,

$$\mathbb{P}\{\varrho = n + 1 | \sigma > n, \varrho > n\} \le C_{4.28} \left(\frac{f(T)}{n} \right)^2 \text{ for } n \in \{m, m + 1, ...\} \setminus (l - 1).$$

Moreover, $\mathbb{P}\{\varrho = m\} \le \mathbb{P}^1\{Z_{t_m - t_N - s} > 0\} = 1 - \exp[-(t_m - s) f(1)]$ and

$\mathbb{P}\{\varrho = l | \varrho \geq l, \sigma \geq l\} \leq \mathbb{P}^T\{Z_{\Delta t_{l-1}} > T\} \leq f(T)\frac{1}{m}$. Consequently,

$$\mathbb{P}\{\varrho \leq \sigma\} \leq \frac{f(1)}{m-1} + \frac{f(T)}{m} + C_{4.28} f(T)^2 \sum_{n=m}^{\infty} \frac{1}{n^2} \leq C\left(f(T)^2 + 1\right)\frac{1}{m},$$

for suitable $C > 0$. $\qquad\qquad\square$

4.4.4 Proof of Proposition 4.24

We couple the labelled tree $\mathbb{T}(V)$ and the π_N-projected INT, starting with a coupling of the position of the initial vertex V and $\pi_N(-X)$, which fails with probability going to zero, by [DM13, Lemma A.2].

Again we apply the concept of an exploration process. As before we categorise vertices as *veiled*, if they have not yet been discovered, *active*, if they have been discovered, but if their descendants have not yet been explored, and *dead*, if they have been discovered and all their descendants have been explored. In one exploration step the leftmost active vertex is picked and its descendants are explored in increasing order with respect to the location parameter. We stop immediately once one of the events (A), (B) or (C) happens. Note that in that case the exploration of the last vertex might not be completed. Moreover, when coupling two explorations, we also stop in the adverse event (E) that the explored graphs disagree. In event (B), the parameters $(n_N)_{N \in \mathbb{N}}$ are chosen such that

$$\lim_{N\to\infty} \frac{\left(\log N \log\log N\right)^\zeta}{n_N} = 0 \text{ and } \lim_{N\to\infty} \frac{\log n_N}{\log N} = 0,$$

for $\zeta = \left(1 - \gamma^+\right)^{-1} \vee 3$. Noting that we never need to explore more than c_N vertices, we see from Lemma 4.25, Remark 4.27 and Lemma 4.29 that the probability of a failure of this coupling is bounded by a constant multiple of

$$c_N\left(1 + f(c_N)^2\right)\frac{1}{n_N} + c_N \frac{1}{n_N^{1-\gamma^+}} \leq \frac{c_N^3}{n_N} + \frac{c_N}{n_N^{1-\gamma^+}} \to 0.$$

Consequently, the coupling succeeds with high probability. As in the proof of Lemma 4.22 it is easy to see that, with high probability, event (B) implies that

$$\#\mathbb{T}(V) \geq c_N \text{ and } \#\mathfrak{T} \geq c_N.$$

Hence we have

$$\#\mathbb{T}(V) \wedge c_N = \#\mathfrak{T} \wedge c_N \text{ with high probability}$$

and the statement follows by combining this with Proposition 4.20.

4.5   Branching process heuristics for $\gamma = \frac{1}{2}$

For the upper bound in Theorem II, we need to show that a fixed vertex n in $\mathscr{G}_N$ is well connected. In a first step, we use the neighbourhood coupling from the previous sections up to a large finite "decoupling time" to determine whether the initial vertex lies in the giant component. After this, we associate a sequence of scores $(S_k)_{k \geq 0}$ to the neighbourhood shells $\Gamma^k = \{m : d_N(n, m) = k\}$ which roughly measures the total degree of vertices contained in Γ^k. We aim to show, that if the scores of the neighbourhoods of two vertices grow above a certain threshold, then there exists with high probability a connection between them. The typical growth rate of the scores will determine the typical distances. Possible circles in the network complicate the calculation, therefore we study the score growth in the IBRW first and use the results to adjust our parameters for the network setting. Central to argument is the *cutoff* $(r_k)_{k \geq 0}$, which we will introduce first for the idealised setting and then use in Chapter 6 in the network.

More specifically, let s_0 be a negative parameter of large absolute value, representing a position far left on the negative half line in the BRW setting (or equivalently a large score which we need to obtain in the initial local exploration) and let $\alpha \geq 0$ be the parameter that measures the deviation of the expected degrees from the linear setting. We define the cutoff sequence r as the solution to the difference equation

$$
\begin{aligned}
r_0 &= s_0 \\
-\frac{1}{2}\Delta r_k &= (\alpha + 1)\log(-r_{k+1})\log\tfrac{k+2}{k+3} + 2\log\tfrac{\log(k+2)}{\log(k+3)} - d + \left(\alpha + \frac{1}{2}\right)\log\tfrac{r_k}{r_{k+1}}, \ \text{ for } k \geq 0,
\end{aligned}
\tag{4.22}
$$

where $d = \log(2^\alpha(\alpha + 1))$. This solution exists for all $k \geq 1$ and is unique if $|s_0|$ is chosen sufficiently large, as is shown in Appendix A. The remainder of this section is devoted to a heuristic derivation of this equation.

To this end, we study the idealised score process on the INT derived from the IBRW introduced in Section 1.3.3 and several modifications of it. Let $\tilde{X}$ denote the IBRW on $\mathbb{R} \times \mathscr{S}$ initiated by a root particle located in $s_0 \in (-\infty, 0)$ with type τ_0 and killed at the origin. Above we have introduced a truncation sequence or *cutoff* $r = (r_k)_{k \geq 0}$ of deterministic points on the negative half line, which we use to kill the BRW $\tilde{X}$ also on the left. The obtained BRW is then denoted by X. For any particle $v \in \tilde{X}$, denote its position in $(-\infty, 0]$ by $s(v)$, its type by $\tau(v)$ and its graph distance from the root in the genealogical tree by $d(v)$. Let furthermore $\tilde{X}^k = \{v \in \tilde{X} : d(v) = k\}$ and $\tilde{X}^{\leq k} = \{v \in \tilde{X} : d(v) \leq k\}$.

Recall, for a particle $v = (s(v), \tau(v)) \in \mathbb{R} \times \mathscr{S}$, the function $M_v : \mathbb{R} \times \mathscr{S} \longrightarrow [0, \infty)$

$$M_v(t, \tau') = \begin{cases} \int_t^{s(v)} \mathrm{e}^{x - s(v)} \mathbb{E}[f(Z_{s(v)-x})] \, \mathrm{d}x, & \text{if } t \in (-\infty, s(v)], \tau' = s(v) - t, \\ M^\ell(t - s(v)), & \text{if } t \in (s(v), \infty), \tau(v) = \tau' = \ell, \\ M^{\tau(v)}(t - s(v)), & \text{if } t \in (s(v), \infty), \tau(v) \in [0, \infty], \tau' = \ell, \\ 0, & \text{otherwise,} \end{cases} \tag{4.23}$$

where $M^\ell(t) = \mathbb{E}[Z_t]$ and $M^\tau(t) = \mathbb{E}[Z_t | \Delta Z_\tau = 1] - \mathbb{1}_{[\tau, \infty)}(t)$, for $\tau \in [0, \infty]$, and $M^\infty(t) = \lim_{\tau \to \infty} M^\tau(t)$. M_v describes the intensity at which v reproduces into state and type space to its left and right.

Let $\mu_v(t) = \int_{\mathscr{S}} M_v(t, \tau) \, \mathrm{d}\tau$. The density $\frac{\mathrm{d}\mu_v(t)}{\mathrm{d}t}$ exists a.e. and is continuous on $\mathbb{R} \setminus \{s(v)\}$. For an attachment rule of the form $f(k) = \frac{1}{2}k + o(k)$ it turns out that it is most convenient to write

$$\mathrm{d}\mu_v(t) = \mathrm{e}^{-\frac{s(v)-t}{2}} \psi_v(t) \, \mathrm{d}t, \tag{4.24}$$

where $\psi_v(t)$ may depend on the position and type of v, is strictly positive (in fact bounded below by $f(0)$), continuous a.e., grows at most subexponentially in t and decays monotonically as a function of $\tau(v) \in [0, \infty]$ (reflecting Lemma 4.14).

Remark 4.30. For $f(k) = \frac{1}{2}k + \beta$, we get

$$\mathrm{d}\mu_v(t) = \mathbb{1}_{\{t \le s(v)\}} \beta \mathrm{e}^{\frac{t-s(v)}{2}} \, \mathrm{d}t + \mathbb{1}_{\{t \ge s(v), \tau(v) = \ell\}} \beta \mathrm{e}^{\frac{t}{2}} \, \mathrm{d}t + \mathbb{1}_{\{t \ge s(v), \tau(v) = \tau \in [0, \infty]\}} \left(\beta + \frac{1}{2} \right) \mathrm{e}^{\frac{t}{2}} \, \mathrm{d}t,$$

hence $\psi_v(t) = \beta + \frac{1}{2} \mathbb{1}_{[s(v), \infty) \times [0, \infty]}(t, \tau(v))$.

For the remainder of the section we discuss the dependence on the shape of ψ_v of the speed of the leftmost particle of the IBRW $\tilde{X}$. Let $\mathfrak{T}$ denote the corresponding INT. Note that the results in [DM13, Section 3] imply $\lim_{s_0 \to -\infty} \mathbb{P}\{\#\mathfrak{T} = \infty | s(\emptyset) = s_0\} = 1$, i.e. the survival probability of the killed random walk can be made arbitrarily large by moving the initial particle towards $-\infty$.

For the purpose of determining the growth of $\tilde{X}$, we introduce a score functional ξ on $\tilde{X}$ which associates weights to subsets of the IBRW path and resembles the functional ξ on the network. For any $v \in \tilde{X}$ and $\Gamma \subset \tilde{X}$ we set

$$\xi(v) = \mathrm{e}^{-\frac{s(v)}{2}}, \text{ and } \xi(\Gamma) = \sum_{v \in \Gamma} \xi(v). \tag{4.25}$$

Since ξ does not depend on the type of particles, we can calculate the expected score of the first generation of particles in the IBRW $\tilde{X}$, using only the densities above. We obtain

$$\mathbb{E} \sum_{v : d(v) = 1} \xi(v) \ge \int_{-\infty}^0 \mathrm{e}^{-\frac{s}{2}} \, \mathrm{d}\mu_\emptyset(s) = \mathrm{e}^{-\frac{s_0}{2}} \int_{-\infty}^0 \psi_\emptyset(s) \, \mathrm{d}s = \infty, \tag{4.26}$$

since ψ is uniformly bounded away from 0. To circumvent this degeneracy, we use the decreasing sequence $r = (r_k)_{k=1}^{\infty}$ of cutoff points on the negative half line and consider in the branching mechanism only those particles of the k-th generation of $\tilde{X}$ which have positions in $[r_k, 0]$ and kill all other particles alongside with their offspring. We denote the resulting *truncated* IBRW by $X = X_{(s_0, \tau_0)}^{(r)}$.

Let Y^k denote the empirical measure of all particle positions and types in the k-th generation X^k of X, let Y_v denote the empirical measure of the offspring $X(v)$ of a particle v and set

$$S_k = \xi(X^k) = \int_{\mathbb{R} \times \mathscr{S}} e^{-\frac{s}{2}} \, dY^k(s, \tau). \tag{4.27}$$

Let $\mathscr{E}_k$ denote the σ-field generated by the first k generations $X^{\leq k}$ of X. Then, formally,

$$\begin{aligned}
\mathbb{E}\left[S_{k+1} | \mathscr{E}_k\right] &= \sum_{v \in X^k} \mathbb{E}\xi(X(w)) = \sum_{v \in X^k} \mathbb{E} \int_{\mathbb{R} \times \mathscr{S}} e^{-\frac{s}{2}} \, dY_v(s, \tau) \\
&= \sum_{v \in X^k} \int_{r_{k+1}}^{0} e^{-\frac{s}{2}} \, d\mu_v(s) = \sum_{v \in X^k} e^{-\frac{s(v)}{2}} \int_{r_{k+1}}^{0} \psi_v(s) \, ds.
\end{aligned} \tag{4.28}$$

The integral on the right hand side of (4.28) still depends on the type of the offspring. The bounds we need are relatively coarse, so we can simplify at this stage and work with upper and lower bounds which we obtain by a modification of the branching mechanism and a stochastic domination argument.

The calculations of Chapter 6 suggest, that $\psi_v(s) \approx |s(v) - s|^{\alpha}$ is symmetric around $s(v)$ and growths polynomially. Thus we can assume that the factor $\int_{r_{k+1}}^{0} \psi_v(s) \, ds$ in (4.28) is minimal for a particle v of type ℓ in position $\frac{1}{2} r_{k+1}$ and maximal for a particle v of type 0 and position r_{k+1}. Thus, to obtain bounds for (4.28), we would like to create auxiliary branching processes $\underline{X}, \overline{X}$ in which (4.24) holds for a particle v of generation k with

$$\underline{\psi}_{k+1}(s) = \mathbb{1}_{[r_{k+1}, 0]}(s) \psi_{(\frac{1}{2} r_{k+1}, \ell)}(s)$$

or

$$\overline{\psi}_{k+1}(s) = \mathbb{1}_{[r_{k+1}, 0]}(s) \psi_{(r_{k+1}, 0)}(s),$$

respectively. Such processes can be obtained by altering the jump processes in the definition of the IBRW. We extend the notation to the associated σ-algebras, functionals and parameters in the obvious way. For the obtained modifications $\underline{S}$ and $\overline{S}$, the last integral in (4.28) does not depend on v any more. thus, equation (4.28) implies that

$$\overline{W}_k = \frac{\overline{S}_k}{\prod_{i=0}^{k} \|\overline{\psi}_k\|_1} \text{ and } \underline{W}_k = \frac{\underline{S}_k}{\prod_{i=0}^{k} \|\underline{\psi}_k\|_1}$$

are martingales w.r.t. their corresponding filtrations $\overline{\mathscr{E}}, \underline{\mathscr{E}}$, where $\|\cdot\|_1$ denotes the usual L^1-

norm.

Remark 4.31.   (i) For the linear case we have $\|\underline{\psi}_k\|_1 = -\beta r_k$ and $\|\overline{\psi}_k\|_1 = -(\beta + \gamma) r_k$.

  (ii) In the nonlinear regime where $\alpha > 0$, with

$$\psi_v(t) = |s(v) - t|^\alpha, \tag{4.29}$$

we obtain

$$\|\underline{\psi}_k\|_1 = \beta \frac{(-r_k)^{\alpha+1}}{2^\alpha(\alpha+1)} \ \text{ and } \|\underline{\psi}_k\|_2 = \beta \frac{(-r_k)^{\alpha+\frac{1}{2}}}{2^\alpha \sqrt{\alpha+1}}.$$

The martingales are non-negative, thus converge a.s., we need to make sure that they also converge in L^1. In fact we aim for L^2 convergence. We focus on $\underline{W}$ only, the other calculation being similar. We obtain

$$
\begin{aligned}
\mathbb{E}\Big[\Big(\sum_{v \in X^k} e^{-\frac{1}{2}s(v)}\Big)^2 \Big| \mathscr{E}_{k-1}\Big] &= \mathbb{E}\Big[\sum_{a,b \in X^{k-1}} \sum_{v \in X(a)} \sum_{u \in X(b)} e^{-\frac{1}{2}(s(v)+s(u))} \Big| \mathscr{E}_{k-1}\Big] \\
&= \sum_{a \in X^{k-1}} \mathbb{E}\Big[\sum_{v,w \in X(a)} e^{-\frac{1}{2}(s(v)+s(w))} \Big| \mathscr{E}_{k-1}\Big] \\
&\quad + \sum_{\substack{a,b \in X^{k-1} \\ a \neq b}} \mathbb{E}\Big[\sum_{v \in X(a)} e^{-\frac{1}{2}s(v)} \Big| \mathscr{E}_{k-1}\Big] \mathbb{E}\Big[\sum_{w \in X(b)} e^{-\frac{1}{2}s(w)} \Big| \mathscr{E}_{k-1}\Big] \\
&= \sum_{a \in X^{k-1}} \mathbb{E}\Big[\sum_{v,w \in X(a)} e^{-\frac{1}{2}(s(v)+s(w))} \Big| \mathscr{E}_{k-1}\Big] \\
&\quad + \sum_{\substack{a,b \in X^{k-1} \\ a \neq b}} e^{-\frac{s(a)+s(b)}{2}} \Big(\int_{-\infty}^0 \underline{\psi}_k(s)\,\mathrm{d}s\Big)^2,
\end{aligned}
\tag{4.30}
$$

and proceed to bound the terms in the first sum in the last line. We indicate the dependence on k of the marginal density of the offspring measure via ψ_k by writing μ_v^k. We also denote by $\mathrm{d}\mu_v^k(s,t)$ the joint density for reproduction in positions $s,t \in \mathbb{R}$ and get

$$
\begin{aligned}
\mathbb{E}\Big[\sum_{v,w \in X(a)} e^{-\frac{1}{2}(s(v)+s(w))} \Big| \mathscr{E}_{k-1}\Big] &= \mathbb{E}\Big[\sum_{v \in X(a)} e^{-s(v)} \Big| \mathscr{E}_{k-1}\Big] + 2\mathbb{E}\Big[\sum_{\substack{v,w \in X(a) \\ s(v) < s(w)}} e^{-\frac{1}{2}(v+w)} \Big| \mathscr{E}_{k-1}\Big] \\
&= \int_{-\infty}^0 e^{-s}\,\mathrm{d}\mu_a^k(s) + 2\int_{-\infty}^0 \int_s^0 e^{-\frac{1}{2}(s+t)}\,\mathrm{d}\mu_a^k(s,t).
\end{aligned}
$$

It is easy to see that by [DM13, Lemma 2.5] and concavity of f, $\mathrm{d}\mu_a(s,t) \leq \frac{f(1)}{f(0)}\mathrm{d}\mu_a(s)\,\mathrm{d}\mu_a(t)$. Hence

$$\mathbb{E}\Big[\sum_{v,w \in \Gamma(a)} e^{-\frac{1}{2}(s(v)+s(w))} \Big| \mathscr{E}_{k-1}\Big] \leq e^{-\frac{s(a)}{2}} \int_{-\infty}^0 e^{-\frac{s}{2}} \underline{\psi}_k(s)\,\mathrm{d}s + \frac{f(1)}{f(0)} e^{-s(a)} \|\underline{\psi}_k\|_1^2,$$

and Cauchy-Schwarz yields

$$\mathbb{E}\Big[\sum_{v,w \in \Gamma(a)} e^{-\frac{1}{2}(s(v)+s(w))} \Big| \mathscr{E}_{k-1}\Big] \leq e^{-\frac{s(a)}{2}} e^{-\frac{r_k}{2}} \|\underline{\psi}_k\|_2 + \frac{f(1)}{f(0)} e^{-s(a)} \|\underline{\psi}_k\|_1^2. \tag{4.31}$$

Thus

$$\mathbb{E}\Big[\Big(\sum_{v\in X^k}\mathrm{e}^{-\frac{1}{2}s(v)}\Big)^2\Big|\mathscr{E}_{k-1}\Big]\leq \underline{S}_{k-1}^2\|\underline{\psi}_k\|_1^2+\Big(\frac{f(1)}{f(0)}-1\Big)\|\underline{\psi}_k\|_1^2\sum_{a\in\Gamma_{k-1}}\xi(a)^2+\xi(r_k)\|\underline{\psi}_k\|_2\underline{S}_{k-1},\quad (4.32)$$

and

$$\mathbb{E}[\underline{W}_k^2|\mathscr{E}_{k-1}]\leq \underline{W}_{k-1}^2+\frac{\big(\frac{f(1)}{f(0)}-1\big)\|\underline{\psi}_k\|_1^2\sum_{a\in X^{k-1}}\xi(a)^2+\xi(r_k)\|\underline{\psi}_k\|_2\underline{S}_{k-1}}{\prod_{i=0}^k\|\underline{\psi}_i\|_1^2}.\quad (4.33)$$

As above we get

$$\mathbb{E}\sum_{a\in X_{k-1}}\mathrm{e}^{-s(a)}\leq \xi(r_{k-1})\|\underline{\psi}_{k-1}\|_2\mathbb{E}\underline{S}_{k-2}$$

and therefore taking expectations in (4.33) implies that $\underline{W}_k$ is bounded in L^2 if

$$\sum_{k=1}^\infty\frac{\xi(r_k)\|\underline{\psi}_k\|_2}{\prod_{i=0}^k\|\underline{\psi}_i\|_1}<\infty,\quad (4.34)$$

which gives a criterion how to chose the cutoff r and incorporate it into ψ_k. For the linear case, $\|\underline{\psi}_i\|_1=-\beta r_i$ and $\|\underline{\psi}_i\|_2=-\beta\sqrt{r_i}$. Thus, if s_0 is sufficiently small and r is the solution to the difference equation

$$
\begin{aligned}
r_0 &= s_0,\\
r_1 &= 2(-\log(-s_0)+\log\beta-\log 3-2\log\log 3),\\
-\frac{1}{2}\Delta r_i &= \log(-r_i)+\log\beta-\log\big(\tfrac{i+3}{i+2}\big)-2\log\big(\tfrac{\log(i+3)}{\log(i+2)}\big),\text{ for } i\geq 1,
\end{aligned}
$$

then

$$\frac{\xi(r_k)\|\underline{\psi}_k\|_2}{\prod_{i=0}^k\|\underline{\psi}_i\|_1}=\frac{1}{(k+2)(\log(k+2))^2}.$$

It is easy to see, that r_k decays faster than $-(2-\varepsilon)k\log k$ for any choice of $\varepsilon\in(0,2)$. An upper bound for S_k is obtained by replacing β by $\beta+\gamma$.

For the general case, we now start by fixing the condition

$$\frac{\xi(r_k)\|\underline{\psi}_k\|_2}{\prod_{i=0}^k\|\underline{\psi}_i\|_1}=\frac{1}{(k+2)(\log(k+2))^2},\quad\text{for all } k\geq 0\quad (4.35)$$

which implies the convergence (4.34). It is easy to see, that if r satisfies (4.22), then it also satisfies (4.35).

4.6 ε-connectors and the core

In this section we introduce a useful concept to establish connectivity results on a *global* scale which can be seen as a natural extension of the use of *sprinkling* arguments in classical

Random Graph Theory, see for example the proof of the emergence of the giant component in the Erdős-Rényi model [Bol01, Theorem 6.8] and also [DM13, Proposition 4.1]. It is an analogue to the use of *t-connectors* in [DHH10], which correspond to 1-connectors in our notation.

Definition 4.32 (ε-connectors)**.** Let $N \in \mathbb{N}$ and $\varepsilon > 0$. Let also $a \neq b \in [N]$ be vertices in $\mathscr{G}_N$. Then $n \in \{N+1, \dots, \lfloor (1+\varepsilon)N \rfloor\}$ is called an ε-*connector* for a and b if both $n \to a$ and $n \to b$. We write

$$\{a \overset{\varepsilon}{\leftrightarrow} b\} = \{\exists\, n \in \{N+1, \dots, (1+\varepsilon)N\} : n \text{ is an } \varepsilon\text{-connector for } a \text{ and } b\},$$

and extend the notation also to sets, i.e. if $A, B \subset \mathbb{N}$ are disjoint, then

$$\{A \overset{\varepsilon}{\leftrightarrow} B\} = \{\exists\, (a, b) \in A \times B : a \overset{\varepsilon}{\leftrightarrow} b\}.$$

Some straightforward implications that are of use later are summarised below.

Observation 4.33. Let $N \in \mathbb{N}$ and $\varepsilon > 0$, then

(i) for all $A, B \subset [N]$, $A \overset{\varepsilon}{\leftrightarrow} B$ implies $d_{\lceil (1+\varepsilon)N \rceil}(A, B) \leq 2$;

(ii) given $\mathscr{G}_N$ and $a, b \in [N]$, for all potential ε-connectors $n \in \{N+1, \dots, \lfloor (1+\varepsilon)N \rfloor\}$,

$$\mathbb{P}\{n \text{ is an } \varepsilon\text{-connector for } a \text{ and } b | \mathscr{G}_N\} \geq \frac{f(\mathscr{Z}[a, N]) f(\mathscr{Z}[b, N])}{((1+\varepsilon)N)^2},$$

which follows from the model definition and that the evolutions $\mathscr{Z}[a, \cdot], \mathscr{Z}[b, \cdot]$ are non-decreasing. Note that the bound holds for each n independently of the connection status of the other ε-connectors.

We would like to consider those subgraphs of $\mathscr{G}_{\lceil (1+\varepsilon)N \rceil}$ which are induced by connectors in couplings with the classical random graph $\mathscr{G}(n, p)$. In particular, we are interested in bounds for the diameter of certain subsets of the oldest part of the graph $\{1, \dots, M\}$, where $M = M(N) = o(N)$. To show, that the edge density is high enough to perform the coupling, we need some notation.

Definition 4.34. Fix $\varepsilon > 0$ and set

$$\xi^* = \xi^*(\gamma, \varepsilon) = \inf_k \xi^{(\gamma)}(k, \lceil (1+\varepsilon)k \rceil) \geq (1+\varepsilon)^\gamma$$

and

$$\sigma = \sigma(\gamma, \varepsilon) = \frac{\xi^* + 3}{4} - 1 > 0.$$

(I) Let f be such that $\frac{1}{2} < \gamma^- < 1$ and $M = \lfloor \sqrt{N}(\log N)^{\frac{1}{2\gamma^-}} \rfloor$. We call $v \in [M]$ a *core* vertex of $\mathscr{G}_{\lceil (1+\varepsilon)N \rceil}$ if

$$\mathscr{Z}[v, N] \geq \sqrt{\frac{N^{\gamma^-}}{\log N}} \quad \text{and} \quad \mathscr{Z}[v, \lceil (1+\varepsilon)N \rceil] - \mathscr{Z}[v, N] \geq \sigma \sqrt{\frac{N^{\gamma^-}}{\log N}}.$$

(II) Let f be as in Theorem II, with $\alpha > 0$ and $M = \lfloor (\log N)^R \rfloor$ for some $R > 0$. We call $v \in [M]$ a *core* vertex of $\mathcal{G}_{\lceil (1+\varepsilon)N \rceil}$ if

$$f(\mathcal{Z}[v, N]) \geq \frac{1}{4}\mathbb{E}\mathcal{Z}[M, N] \quad \text{and} \quad f(\mathcal{Z}[v, \lceil (1+\varepsilon)N \rceil]) - f(\mathcal{Z}[v, N] \geq \sigma f(\mathcal{Z}[M, N]).$$

Let moreover C_N denote the subset of $[M]$ containing all core vertices, and let

$$E_N = \{(v, w) \in C_N \times C_N : v \overset{\varepsilon}{\leftrightarrow} w\}$$

then we call the random subgraph $\text{core}_N = (C_N, E_N)$ of $\mathcal{G}_{\lceil (1+\varepsilon)N \rceil}$ the *core*.

Depending on the form of f, we choose M appropriately to use a coupling between the core and an Erdős-Rényi Graph $\mathcal{G}(n, p)$ for suitable n, p. Prerequisite for this coupling is the following lemma.

Lemma 4.35 (Size of core_N)**.** *Fix $\varepsilon > 0$ and f, M as in Definition 4.34.*

 (i) There exists $q(f, \varepsilon) > 0$ such that

$$\mathbb{P}\{v \in C_N\} \geq q(f, \varepsilon) \quad \text{for all } v \in [M] \text{ independently.}$$

 (ii) With high probability $\#C_N \geq M \frac{q(f,\varepsilon)}{2}$.

Proof. We prove statement (i) only for situation (II) of Definition 4.34, the argument for the other setting is given in the proof of Lemma 5.7. Independence follows from the independence of degree evolutions. Note that the condition in (II) is a condition on the growth of the degree evolution $(\mathcal{Z}[v, \cdot])$, by monotonicity it is sufficient to verify the conditions for $v = M$. Since $f(k) \geq \frac{k}{2}$, for all k, it is sufficient to demonstrate that there exists a small positive probability q, such that

$$f(\mathcal{Z}[M, N]) \geq \frac{1}{2}\mathbb{E}f(\mathcal{Z}[M, N]) \tag{4.36}$$

and

$$f(\mathcal{Z}[M, \lceil (1+\varepsilon)N \rceil]) \geq (1+\sigma)f(\mathcal{Z}[M, N]), \tag{4.37}$$

for all sufficiently large N, which is slightly stronger than the condition given in (II). To show (4.36), note that the Payley-Zygmund inequality implies that

$$\mathbb{P}\left\{f(\mathcal{Z}[M, N]) \geq \frac{1}{2}\mathbb{E}f(\mathcal{Z}[M, N])\right\} \geq \frac{1}{4}\frac{(\mathbb{E}f(\mathcal{Z}[M, N]))^2}{\mathbb{E}f(\mathcal{Z}[M, N])^2} = p(M, N).$$

The moment bounds derived in Proposition 6.1 now entail that $p(M, N) > p_1 > 0$ for some small p_1 and all sufficiently large N. Furthermore, since f is concave, $(\mathcal{Z}^{(k)}[n, s])_{s \geq n}$ dominates $(\bar{\mathcal{Z}}^{(k)}[n, s])_{s \geq n}$, the latter denoting the degree evolution associated with the linearised

attachment rule satisfying $\bar{f}(l) = f(k) + \frac{1}{2}(l-k) \leq f(l)$, $l \geq k$. Thus for some $p_2 > 0$,

$$\mathbb{P}\{f(\mathcal{Z}[M, \lceil(1+\varepsilon)N\rceil]) \geq (1+\sigma)f(\mathcal{Z}[M,N])|\mathcal{Z}[M,N] = k\}$$
$$= \mathbb{P}^k\{f(\mathcal{Z}[N, \lceil(1+\varepsilon)N\rceil]) \geq (1+\sigma)f(k)\}$$
$$\geq \mathbb{P}^k\{\bar{f}(\bar{\mathcal{Z}}[N, \lceil(1+\varepsilon)N\rceil]) \geq (1+\sigma)f(k)\}$$
$$= \mathbb{P}^k\left\{\bar{f}(\bar{\mathcal{Z}}[N, \lceil(1+\varepsilon)N\rceil]) \geq \frac{(1+\sigma)}{\xi(N, \lceil(1+\varepsilon)N\rceil)}f(k)\xi(N, \lceil(1+\varepsilon)N\rceil)\right\}$$
$$\geq \left(1 - \frac{1+\sigma}{\xi^*}\right)^2 \frac{(\mathbb{E}\bar{f}(\bar{\mathcal{Z}}[N, \lceil(1+\varepsilon)N\rceil]))^2}{\bar{f}(\mathbb{E}\bar{\mathcal{Z}}[N, \lceil(1+\varepsilon)N\rceil])^2} > p_2,$$

where we have used the Payley-Zygmund inequality and Lemma 4.7 for the last two steps. Thus (4.37) holds with probability at least p_2, independently of the value of $\mathcal{Z}[M,N]$. Since $(\mathcal{Z}[M,s])_{s \geq M}$ is a Markov process, we know that the increment

$$f(\mathcal{Z}[M, \lceil(1+\varepsilon)N\rceil]) - f(\mathcal{Z}[M,N])$$

can only depend on the past via $\mathcal{Z}[M,N]$, it follows that (4.36) and (4.37) are jointly satisfied with probability at least $q = p_1 p_2$.

For statement (ii) of the lemma, note that by (i) all $v \in [M]$ have chance of at least $q = q(f, \varepsilon)$ to be in the core and the degree evolutions are independent. Thus $\#C_N$ dominates (via an appropriate coupling) the sum $S_M = \sum_{i=1}^{M} X_i$, where $X_1, X_2, \dots$ are i.i.d. Bernoulli with parameter q. A concentration inequality (e.g. Lemma B.2) now yields

$$\mathbb{P}\left\{\#C_N < M\frac{q}{2}\right\} \leq \mathbb{P}\left\{S_M < M\frac{q}{2}\right\} \leq \exp\left(-\frac{q}{2}M\right),$$

which converges to 0 as $N \to \infty$. $\qquad\square$

We are now ready to prove that the core is, with high probability, of bounded diameter.

Proposition 4.36. *For any $\varepsilon > 0$, $core_N$ has bounded diameter with high probability, if either of the following hold*

(i) *f is affine with $1 > \gamma > \frac{1}{2}$ and $M = \lfloor\sqrt{N}(\log N)^{\frac{1}{2\gamma}}\rfloor$;*

(ii) *f is as in Theorem II with $\gamma = \frac{1}{2}, \alpha > 0$ and $M = \lfloor(\log N)^R\rfloor$ for some $R > 0$.*

Remark 4.37. The restriction to affine f in (i) is for convenience only. Since we assume f to be concave, we know that $\mathcal{G}_N$ dominates $\mathcal{G}'_N$, the graph process associated to $f'(k) = f(0) + \gamma^- k$. Since the definition of the core does only depend on $\gamma = \gamma^-$, we have that, for the coupled versions, $C'_N \subset C_N$ and $E'_N \subset E_N$, thus the diameter of $core_N$ is bounded by the diameter of $core'_N$.

Proof of Proposition 4.36. For $v, w \in C_N$ we now provide an estimate for the connection probability $\mathbb{P}\{(v,w) \in E_N\}$ which is *independent* of the occupation status of the other edges

in core$_N$. Denote

$$\psi(m,n) = \frac{\mathbb{E}f(\mathscr{Z}[m,n])}{\xi(m,n)},$$

and note that in setting (i), $\psi(m,n) = f(0)$ and in setting (ii), $\psi(m,n) \to \infty$ if $\frac{n}{m} \to \infty$, which is shown in Proposition 6.2. Fix $v \in C_N$. Since $f(\mathscr{Z}[v, \lceil(1+\varepsilon)N\rceil]) \geq (1+\sigma)f(\mathscr{Z}[v,N])$, we know that by concavity of f,

$$\mathscr{Z}[v, \lceil(1+\varepsilon)N\rceil] - \mathscr{Z}[v,N] \geq \sigma\mathscr{Z}[v,N] \geq \frac{1}{4}\mathbb{E}\mathscr{Z}[v,N].$$

Let $Z(v) = \{N < j \leq \lceil(1+\varepsilon)N\rceil : \Delta\mathscr{Z}[v,j] = 1\}$. Since $v \leq M$, the number $\#Z(v)$ of potential ε-connectors to v is thus bounded below by

$$\frac{1}{4}\mathbb{E}\mathscr{Z}[M,N] \geq \eta_\varepsilon\left(\frac{N}{M}\right)^\gamma,$$

for some $\eta_\varepsilon > 0$. Given $\mathscr{G}_N$ and $v \in C_N$, the probability that there is no ε-connector between v and w can therefore be bounded

$$\mathbb{P}\{(v,w) \notin E_N\} = \prod_{j \in Z(v)}\left(1 - \frac{f(\mathscr{Z}[w,j])}{j}\right) \leq \left(1 - \frac{f(\mathscr{Z}[w,N])}{\lceil(1+\varepsilon)N\rceil}\right)^{\eta_\varepsilon\left(\frac{N}{M}\right)^\gamma}.$$

Since $w \in C_N$, we can find $c_\varepsilon > 0$, such that

$$f(\mathscr{Z}[w,N]) \geq c_\varepsilon\psi(M,N)\left(\frac{N}{M}\right)^\gamma,$$

and thus

$$\mathbb{P}\{(v,w) \notin E_N\} \leq \left(1 - \frac{c_\varepsilon\psi(M,N)(\frac{N}{M})^\gamma}{\lceil(1+\varepsilon)N\rceil}\right)^{\eta_\varepsilon\left(\frac{N}{M}\right)^\gamma} \leq \exp\left(-\frac{c_\varepsilon\psi(M,N)\left(\frac{N}{M}\right)^\gamma}{\lceil(1+\varepsilon)N\rceil}\eta_\varepsilon\left(\frac{N}{M}\right)^\gamma\right).$$

For both possible values of M, the argument of the exponential is close to 0, for such arguments we have that $e^{-x} \leq 1 - \frac{x}{2}$. For sufficiently large N, we thus obtain

$$\mathbb{P}\{(v,w) \notin E_N\} \leq 1 - \frac{c_\varepsilon\eta_\varepsilon\psi(M,N)\left(\frac{N}{M}\right)^{2\gamma}}{2\lceil(1+\varepsilon)N\rceil}.$$

We have therefore shown, that for some small $\delta_\varepsilon > 0$, and N sufficiently large,

$$\mathbb{P}\{(v,w) \in E_N\} \geq \delta_\varepsilon\psi(M,N)N^{2\gamma-1}M^{-2\gamma}. \tag{4.38}$$

Note that, given $\mathscr{G}_N$ and C_N, this holds independently for all pairs $v, w \in C_N$. We can thus dominate core$_N$ by an Erdős-Rényi graph $\mathscr{G}(n,p)$ of (random) size $n = \#C_N$ and connection probability $p = \delta_\varepsilon\psi(M,N)N^{2\gamma-1}M^{-2\gamma}$. Theorem B.3, the diameter result for $\mathscr{G}(n,p)$, now implies the boundedness of the diameter. $\qquad\square$

Remark 4.38.   (i) The ε-connectors are easier to handle than the corresponding connec-

tions in the affine PA model, due to the conditional independence. Nevertheless, the general idea, that one has to look backwards after a large time period to gain (conditional) independence of edges is the same.

(ii) The approach used here relies on the appearance of high degrees or rather scores (i.e. "edge density"), unlike the sprinkling argument of Proposition 4.1 in [DM13], which uses the "vertex density" (since there the goal is to connect clusters of vertices which may have themselves low degrees.)

The last proposition implies that it is sufficient to show that the explored parts of the graph are connected to the core to obtain distances on the desired scale. Note also, that in the setting of Theorem II, the approach does not work if $\alpha = 0$.

Chapter 5

Typical distances for $\gamma \in \left(\frac{1}{2}, 1\right)$

This chapter is based on joint work with Steffen Dereich and Peter Mörters that has been published in [DMM12].

We first apply the truncated first moment method explained in Section 3.1 to prove the general lower bound results for preferential attachment and fitness type models, Theorem 2.3 and Theorem 2.5, in Sections 5.1 and 5.2. Then we prove the upper bound for the sublinear preferential attachment model, Theorem 2.6 in Section 5.3, using both the local exploration procedure explained in Sections 4.3 and 4.4 and the bound on the size of the core derived in Section 4.6. Finally, in Section 5.4, we deduce Theorem I.

5.1 Lower bounds for preferential attachment models - proof of Theorem 2.3

For the proof, we assume the validity of Assumption 2.2 for some $\gamma \in \left(\frac{1}{2}, 1\right)$ with a fixed constant κ. We adopt the notation laid out in the discussion at the end of Section 3.1. In particular recall the definition of $\mu_k^{(v)}$ and the key estimates (3.2), (3.3) and (3.4), which combined give

$$\mathbb{P}\{d_N(v, w) \leq 2\delta\} \leq \sum_{k=1}^{\delta} \mu_k^{(v)}[\ell_k - 1] + \sum_{k=1}^{\delta} \mu_k^{(w)}[\ell_k - 1] + \sum_{n=1}^{2\delta} \sum_{u=\ell_{n^*}}^{N} \mu_{n^*}^{(v)}(u) \mu_{n-n^*}^{(w)}(u). \qquad (5.1)$$

The remaining task of the proof is to choose $\delta \in \mathbb{N}$ and $2 \leq \ell_\delta \leq \ldots \leq \ell_0 \leq N$ which allow the required estimates for the right hand side. To do so, we will make use of the recursive representation

$$\mu_{k+1}^{(v)}(n) = \sum_{m=\ell_k}^{N} \mu_k^{(v)}(m) \, p(m, n) \quad \text{for } k \in \{0, \ldots, \delta - 1\} \text{ and } n \in [N],$$

where $\mu_0^{(v)}(n) = \mathbb{1}_{\{v=n\}}$ and

$$p(m,n) = \kappa(m \wedge n)^{-\gamma}(m \vee n)^{\gamma-1}.$$

Denote by $\bar{\mu}_k^{(v)}(m) = \mathbb{1}_{\{m \geq \ell_k\}}\mu_k^{(v)}(m)$ the truncated version of $\mu_k^{(v)}$ and regard $\mu_k^{(v)}$ and $\bar{\mu}_k^{(v)}$ as row vectors. Then

$$\mu_{k+1}^{(v)} = \bar{\mu}_k^{(v)}\mathbf{P}_N, \tag{5.2}$$

where $\mathbf{P}_N = (p(m,n))_{m,n=1,\dots,N}$. Our aim is to provide an upper bound of the form

$$\mu_k^{(v)}(m) \leq \alpha_k m^{-\gamma} + \mathbb{1}_{\{m > \ell_{k-1}\}}\beta_k m^{\gamma-1}, \tag{5.3}$$

for suitably chosen parameters $\alpha_k, \beta_k \geq 0$. Key to this choice is the following lemma.

Lemma 5.1. *Suppose that $2 \leq \ell \leq N$, $\alpha, \beta \geq 0$ and $q: [N] \longrightarrow 0, \infty)$ satisfies*

$$q(m) \leq \mathbb{1}_{\{m \geq \ell\}}(\alpha m^{-\gamma} + \beta m^{\gamma-1}), \quad \text{for all } m \in [N].$$

Then there exists a constant $c = c(\gamma, \kappa) > 1$ such that

$$q\mathbf{P}_N(m) \leq c\left(\alpha \log \tfrac{N}{\ell} + \beta N^{2\gamma-1}\right)m^{-\gamma} + \mathbb{1}_{\{m > \ell\}}c\left(\alpha \ell^{1-2\gamma} + \beta \log \tfrac{N}{\ell}\right)m^{\gamma-1}$$

for all $m \in [N]$.

Proof. One has

$$q\mathbf{P}_N(m) = \mathbb{1}_{\{m > \ell\}}\sum_{k=\ell}^{m-1} q(k)\,p(k,m) + \sum_{k=m\vee\ell}^{N} q(k)\,p(k,m)$$

$$\leq \mathbb{1}_{\{m > \ell\}}\sum_{k=\ell}^{m-1} \kappa(\alpha k^{-\gamma} + \beta k^{\gamma-1})k^{-\gamma}m^{\gamma-1} + \sum_{k=m\vee\ell}^{N} \kappa(\alpha k^{-\gamma} + \beta k^{\gamma-1})k^{\gamma-1}m^{-\gamma}$$

$$\leq \kappa\left(\alpha \sum_{k=m\vee\ell}^{N} k^{-1} + \beta \sum_{k=m\vee\ell}^{N} k^{2\gamma-2}\right)m^{-\gamma} + \mathbb{1}_{\{m > \ell\}}\kappa\left(\alpha \sum_{k=\ell}^{m-1} k^{-2\gamma} + \beta \sum_{k=\ell}^{m-1} k^{-1}\right)m^{\gamma-1}$$

$$\leq \kappa\left(\alpha \log \tfrac{N}{\ell-1} + \tfrac{\beta}{2\gamma-1}N^{2\gamma-1}\right)m^{-\gamma} + \mathbb{1}_{\{m > \ell\}}\kappa\left(\tfrac{\alpha}{1-2\gamma}(\ell-1)^{1-2\gamma} + \beta \log \tfrac{m}{\ell-1}\right)m^{\gamma-1}.$$

This immediately implies the assertion since $\ell \geq 2$ by assumption. $\qquad\square$

Proof of Theorem 2.3. We apply Lemma 5.1 iteratively. Fix $\varepsilon > 0$ small and start with

$$\ell_0 = \lceil \varepsilon N \rceil, \ \alpha_1 = \kappa(\varepsilon N)^{\gamma-1} \text{ and } \beta_1 = \kappa(\varepsilon N)^{-\gamma}.$$

Fix $v \geq \ell_0$. Then, for all $m \in [N]$,

$$\mu_1^{(v)}(m) = p(v, m) \leq \kappa \ell_0^{\gamma-1} m^{-\gamma} + \mathbb{1}_{\{m > \ell_0\}} \kappa \ell_0^{-\gamma} m^{\gamma-1}$$

$$\leq \alpha_1 m^{-\gamma} + \mathbb{1}_{\{m > \ell_0\}} \beta_1 m^{\gamma-1}.$$

Now suppose that, for some $k \in \mathbb{N}$, we have chosen α_k, β_k and an integer ℓ_{k-1} such that

$$\mu_k^{(v)}(m) \leq \alpha_k m^{-\gamma} + \beta_k m^{\gamma-1} \text{ for all } m \in [N].$$

We choose ℓ_k as an integer satisfying

$$\frac{6\varepsilon}{\pi^2 k^2} \geq \frac{1}{1-\gamma} \alpha_k \ell_k^{1-\gamma}, \tag{5.4}$$

and assume $\ell_k \geq 2$. Pick α_k, β_k such that

$$\begin{aligned} \alpha_{k+1} &\geq c_{5.1} \left(\alpha_k \log \frac{N}{\ell_k} + \beta_k N^{2\gamma-1} \right), \\ \beta_{k+1} &\geq c_{5.1} \left(\alpha_k \ell_k^{1-2\gamma} + \beta_k \log \frac{N}{\ell_k} \right). \end{aligned} \tag{5.5}$$

By the induction hypothesis we can apply Lemma 5.1 with $\ell = \ell_k$ and $q(m) = \bar{\mu}_k^{(v)}(m)$. Then, using (5.2),

$$\mu_{k+1}^{(v)}(m) \leq \alpha_{k+1} m^{-\gamma} + \mathbb{1}_{\{m > \ell_k\}} \beta_{k+1} m^{\gamma-1} \quad \text{for all } m \in [N], \tag{5.6}$$

showing that the induction can be carried forward up to the point where $\ell_k < 2$, say in step K. Summing (5.6) over $m \leq \ell_{k-1}$ and using (5.4) we obtain

$$\mu_k^{(v)}[\ell_k - 1] \leq \frac{1}{1-\gamma} \alpha_k \ell_k^{1-\gamma} \leq \frac{6\varepsilon}{\pi^2 k^2}.$$

Hence the first two terms on the right hand side in (5.1) are together smaller than 2ε. It remains to choose $\delta = \delta(N)$ as large as possible while ensuring that $\delta < K$ and

$$\lim_{N \to \infty} \sum_{n=1}^{2\delta} \sum_{u=\ell_{n^*}}^{N} \mu_{n^*}^{(v)}(u) \mu_{n-n^*}^{(w)}(u) = 0.$$

To this end assume that ℓ_k is the largest integer satisfying (5.4) and the parameters α_k, β_k are defined via equalities in (5.5). To establish lower bounds for the decay of ℓ_k we investigate the growth of $\eta_k := \frac{N}{\ell_k} > 0$.

Going backwards through the definitions yields, that for an integer $k \geq 0$ with $k + 1 < K$ and if the right hand side is less or equal to $\left(\frac{N}{3}\right)^{1-\gamma}$, one has

$$\left(\eta_{k+2}^{-1} + \frac{1}{N} \right)^{\gamma-1} \leq \frac{c_{5.1}^2 (k+2)^2}{k^2} \eta_k^\gamma + 2c_{5.1}^2 \frac{(k+2)^2}{(k+1)^2} \eta_{k+1}^{1-\gamma} \log \eta_{k+1}.$$

In particular, it follows that $K > k + 2$ in that case.

It is easy to check inductively that for any solution of this system there exist constants $b, B > 0$ (not depending on N) such that,

$$\eta_k \leq b \exp\left(B\left(\tfrac{\gamma}{1-\gamma}\right)^{\frac{k}{2}}\right) \tag{5.7}$$

for $k < K$ and, moreover, the right hand side exceeds $\left(\tfrac{N}{3}\right)^{1-\gamma}$ before step K. We now use (5.6) to estimate

$$\sum_{n=1}^{2\delta} \sum_{u=\ell_k}^{N} \mu_{n^*}^{(v)}(u)\mu_{n-n^*}^{(w)}(u) \leq 2 \sum_{k=1}^{\delta} \sum_{u=\ell_k}^{N} \left(\alpha_k u^{-\gamma} + \beta_k u^{\gamma-1}\right)^2$$

$$\leq \tfrac{4}{2\gamma-1} \sum_{k=1}^{\delta} \left(\alpha_k^2 \ell_k^{1-2\gamma} + \beta_k^2 N^{2\gamma-1}\right) \leq \tfrac{4}{2\gamma-1} \delta\left(\alpha_\delta^2 \ell_\delta^{1-2\gamma} + \beta_\delta^2 N^{2\gamma-1}\right).$$

Using (5.4) and (5.7) the first term in the last bracket can be estimated by

$$\alpha_\delta^2 \ell_\delta^{1-2\gamma} \leq \left(\delta^{-2}\tfrac{6\varepsilon}{\pi^2}(1-\gamma)\right)^2 \ell_\delta^{-1} \leq \left(\tfrac{6\varepsilon}{b\pi^2}(1-\gamma)\right)^2 \tfrac{1}{N\delta^4} \exp\left(B\left(\tfrac{\gamma}{1-\gamma}\right)^{\frac{\delta}{2}}\right).$$

Using equality in (5.5) we get $\beta_\delta \leq c_{5.1}\left(\alpha_\delta \ell_\delta^{1-2\gamma} + \alpha_\delta N^{1-2\gamma}\log\tfrac{N}{\ell_\delta}\right)$. Noting that the second term of the sum on the right hand side is bounded by a multiple of the first, we find a constant $C_1 > 0$ such that $\beta_\delta^2 N^{2\gamma-1} \leq C_1 \alpha_\delta^2 \ell_\delta^{1-2\gamma}$, and thus, for a suitable constant $C_2 > 0$,

$$\sum_{n=1}^{2\delta} \sum_{u=\ell_k}^{N} \mu_{n^*}^{(v)}(u)\mu_{n-n^*}^{(w)}(u) \leq C_2 \tfrac{1}{N\delta^3} \exp\left(B\left(\tfrac{\gamma}{1-\gamma}\right)^{\frac{\delta}{2}}\right).$$

Hence, for a suitable constant $C > 0$, choosing

$$\delta \leq \frac{\log\log N}{\log\sqrt{\tfrac{\gamma}{1-\gamma}}} - C$$

we obtain that the term we consider goes to zero as $O((\log\log N)^{-3})$. Note from (5.7) that this choice also ensures that $\ell_\delta \geq 2$. We have thus shown that

$$\mathbb{P}\{d_N(v, w) \leq 2\delta\} \leq 2\varepsilon + O\left((\log\log N)^{-3}\right),$$

whenever $v, w \geq \ell_0 = \lceil \varepsilon N \rceil$, which implies the statement of Theorem 2.3. $\qquad\square$

5.2 Lower bounds for fitness models - proof of Theorem 2.5

In this section, we assume validity of Assumption 2.4 for some $\tfrac{1}{2} < \gamma < 1$ with a fixed constant $\kappa \geq 1$. Recall again the notation and framework from the introductory Section 3.1.1. We use the same approach as in the proof of Theorem 2.3 but now we have to consider the

symmetric matrix $\mathbf{P}_N := (p(m,n))_{m,n\in[N]}$ given by

$$p(m,n) := \kappa m^{-\gamma} n^{-\gamma} N^{2\gamma-1} \text{ for } m,n \in [N]. \tag{5.8}$$

We obtain the following lemma, which is the analogue of Lemma 5.1.

Lemma 5.2. *Suppose that* $2 \le \ell \le N$ *and* $q \colon [N] \longrightarrow 0, \infty)$ *satisfies*

$$q(m) \le \mathbb{1}_{\{m \ge \ell\}} m^{\gamma-1} \ell^{-\gamma}, \quad \textit{for all } m \in [N].$$

Then, for all $m \in [N]$,

$$q\mathbf{P}_N(m) \le \kappa \, m^{-\gamma} N^{\gamma-1} \left(\frac{N}{\ell}\right)^\gamma \log \frac{N}{\ell-1}.$$

Proof. By (5.8) and the assumption on q,

$$q\mathbf{P}_N(m) = \sum_{i=1}^N q(i)p(i,m) \le \kappa m^{-\gamma} \ell^{-\gamma} N^{2\gamma-1} \sum_{i=\ell}^N \tfrac{1}{i} \le \kappa m^{-\gamma} \ell^{-\gamma} N^{2\gamma-1} \log \frac{N}{\ell-1},$$

which implies the statement of the lemma. $\qquad\square$

For fixed $\varepsilon > 0$ we first construct inductively a strictly decreasing sequence of integers (ℓ_k) by letting $\ell_0 = \lceil \varepsilon N \rceil$ and defining ℓ_{k+1} as the largest integer such that

$$\frac{\kappa}{1-\gamma} \left(\frac{\ell_{k+1}}{N}\right)^{1-\gamma} \le \frac{6\varepsilon}{\pi^2(k+1)^2} \left(\log \frac{N}{\ell_k-1}\right)^{-1} \left(\frac{\ell_k}{N}\right)^\gamma. \tag{5.9}$$

We stop once we find $\ell_k \le 1$, say in step K. Recall the definition and recursive formula for $\mu_k^{(v)}$ and let $\bar{\mu}_k^{(v)}(m) := \mathbb{1}_{\{m \ge \ell_k\}} \mu_k^{(v)}(m)$. Then $\mu_{k+1}^{(v)}(m) = \bar{\mu}_k^{(v)} \mathbf{P}_N(m)$. We now show, for $k = 1, \ldots, K-1$, that

$$\mu_k^{(v)}(m) \le \kappa \, m^{-\gamma} N^{\gamma-1} \left(\frac{N}{\ell_{k-1}}\right)^\gamma \log \frac{N}{\ell_{k-1}-1} \le m^{-\gamma} \ell_k^{\gamma-1}, \text{ for all } m \in [N]. \tag{5.10}$$

Indeed, for $k = 1$ the statement follows from (5.8) and (5.9). We then continue by induction using Lemma 5.2. Considering the truncated first moment estimate (3.2) for $\delta < K$ and our choice of $(\ell_k)_{k=0}^K$, we obtain from (5.10) that

$$\mathbb{P}\big(A_k^{(v)}\big) \le \mu_k^{(v)}[\ell_k - 1] \le \frac{\kappa}{1-\gamma} \left(\frac{\ell_k}{N}\right)^{1-\gamma} \left(\frac{N}{\ell_{k-1}}\right)^\gamma \log \frac{N}{\ell_{k-1}-1}.$$

Hence (5.9) entails that $\sum_{k=1}^\delta \mathbb{P}\big(A_k^{(v)}\big) \le \varepsilon$. The last step is to choose $\delta = \delta(N)$ as large as possible while ensuring that $\delta < K$ and

$$\sum_{n=1}^{2\delta} \sum_{u=\ell_{n^*}}^N \mu_{n^*}^{(v)}(u) \mu_{n-n^*}^{(w)}(u) \to 0 \tag{5.11}$$

goes to zero as $N \to \infty$. By (5.10) this double sum can be bounded by a constant multiple of $N^{2\gamma-2} \sum_{k=1}^\delta \ell_k^{1-2\gamma}$. To verify (5.11) we have to bound the growth of the values $\eta_k := \frac{N}{\ell_k}$. The

choice made in (5.9) implies, for $k < K$ and if the right hand side is smaller than $(\frac{N}{3})^{1-\gamma}$, that

$$\left(\eta_{k+1}^{-1} + \tfrac{1}{N}\right)^{\gamma-1} < \tfrac{\pi^2 \kappa}{1-\gamma} \tfrac{(k+1)^2}{6\varepsilon} \eta_k^\gamma \log(2\eta_k), \text{ for } k \geq 0.$$

In particular, one has $k + 1 < K$ in that case.

From this it is straightforward to verify inductively the existence of constants $b, B > 0$, which only depend on ε, κ and γ, such that

$$\eta_k \leq b \exp\left(B\left(\tfrac{\gamma}{1-\gamma}\right)^k\right), \text{ for } k < K,$$

and, moreover, the right hand side exceeds $(\frac{N}{3})^{1-\gamma}$ before step K. Hence, we may choose a suitable constant $C > 0$ such that for

$$\delta \leq \frac{\log\log N}{\log(\frac{\gamma}{1-\gamma})} - C$$

we have $\ell_\delta \geq 2$. To complete the proof, we note that

$$N^{2\gamma-2} \sum_{k=1}^{\delta} \ell_k^{1-2\gamma} \leq \frac{1}{N} \sum_{k=1}^{\delta} \eta_k^{2\gamma-1} \leq \delta\, b\, N^{B\left(\frac{\gamma}{1-\gamma}\right)^{-C}-1},$$

which implies convergence in (5.11) when C is chosen large enough.

5.3 Upper bounds for preferential attachment networks - proof of Theorem 2.6

The proof will encompass three steps: a local exploration around the randomly chosen vertices, connection of successive layers of high degree vertices in the core using 1-connectors and a coupling to a classical random graph for the vertices of highest degree. The proof is a variation of the argument in [DHH10, Section 3] for the affine preferential attachment model, the sharper bound featured in Theorem 2.6 results from our ability to identify (and work with) the INT as the weak local limit of the graph.

For convenience, we work in the graph $\mathcal{G}_{2N}$, and we thus use ε-connectors with $\varepsilon = 1$. In the case of an affine attachment rule $f(k) = \gamma k + \beta$ all calculations become completely explicit. For the remainder of this section, $\varepsilon > 0$ is a fixed but arbitrary small parameter, not to be confused with parameter for the connectors which equals 1 throughout. Depending on ε and γ, we will parametrise a sequence of *layers* $\mathcal{N}^{(k)}$, $k \in \mathbb{N}_0$, of high degree vertices in the network. For this purpose, we fix a strictly increasing sequence $u = (u_k)_{k=0}^{\infty}$ of positive integers, prescribing the minimum degrees of the vertices in those layers.

Definition 5.3. Let

$$u_0 = u_0(\gamma, \varepsilon) = \min\{n \in \mathbb{N} : \log n \geq a(\gamma, \varepsilon)\}, \tag{5.12}$$

where

$$a(\gamma, \varepsilon) = \gamma\left(\frac{\gamma}{1-\gamma} - 1\right)(\log 2 + \gamma_{\mathrm{EM}}) + \sum_{j=0}^{\infty} \left(\frac{\gamma}{1-\gamma}\right)^{-j+1}(\log\log(j+3) + c(\gamma, \varepsilon)), \tag{5.13}$$

and

$$c(\gamma, \varepsilon) = \frac{\gamma}{1-\gamma}\log \frac{8 - \frac{4\log\frac{\varepsilon}{\pi^2}}{\log 2}}{\gamma\sigma q(\gamma)}, \tag{5.14}$$

where $\sigma = \sigma_{4.34}(\gamma, 1) = \frac{\xi^* + 3}{4} - 1$ and $q(\gamma) = q_{4.35}(f, 1) > 0$ only depend on γ. Now define inductively for $k \geq 0$,

$$u_{k+1} = u_k^{\frac{\gamma}{1-\gamma}}\left(\log(k+3)\right)^{-\frac{\gamma}{1-\gamma}}\mathrm{e}^{-c(\gamma, \varepsilon)}. \tag{5.15}$$

The choice of u is rather technical, but has some useful implications which we collect for later reference.

Lemma 5.4 (Properties of u)**.** *The definition of u implies that*

(i) we have

$$\frac{u_{k+1}}{u_{k+2}} \leq 2^{-\gamma}\mathrm{e}^{-\gamma_{EM}\gamma}, \quad \text{for all } k \geq 0;$$

(ii) for all $k \geq 0$,

$$\exp\left(-\frac{\gamma\sigma q(\gamma)}{4}u_k u_{k+1}^{1-\frac{1}{\gamma}}\right) \leq \frac{1}{(k+3)^2}\frac{6\varepsilon}{\pi^2}; \tag{5.16}$$

(iii) there exists a constant $C > 0$ such that for

$$L = L(N) = \min\left\{k \in \mathbb{N} : u_k \geq 2\sqrt{\frac{N^\gamma}{\log N}}\right\} \tag{5.17}$$

and every $N \in \mathbb{N}$,

$$L \leq \frac{\log\log N}{\log\frac{\gamma}{1-\gamma}} + C. \tag{5.18}$$

Proof. For ease of notation we write $\eta_k = \log u_k$. The recursion (5.15) implies that

$$\eta_{k+1} = \frac{\gamma}{1-\gamma}\eta_k - \left(\frac{\gamma}{1-\gamma}\log\log(k+3) + c(\gamma, \varepsilon)\right),$$

and thus

$$\eta_k = \left(\frac{\gamma}{1-\gamma}\right)^k\left(\eta_0 - \sum_{j=0}^{k-1}\left(\frac{\gamma}{1-\gamma}\right)^{-j}(\log\log(j+3) + c(\gamma, \varepsilon))\right). \tag{5.19}$$

Therefore

$$\Delta\eta_k \geq \left(\frac{\gamma}{1-\gamma}\right)^k\left[\left(\frac{\gamma}{1-\gamma} - 1\right)\eta_0 - \sum_{j=0}^{k-1}\left(\frac{\gamma}{1-\gamma}\right)^{-j+1}(\log\log(j+3) + c(\gamma, \varepsilon))\right]$$

and (5.13) implies that

$$-\Delta\eta_k \le -\gamma(\log 2 + \gamma_{\mathrm{EM}})\Big(\frac{\gamma}{1-\gamma}\Big)^k$$

and the latter term is maximal for $k = 0$. The assertion (i) now follows from the observation that

$$-\Delta\eta_k \le -\gamma(\log 2 + \gamma_{\mathrm{EM}})$$

if and only if

$$\frac{u_k}{u_{k+1}} \le 2^{-\gamma}\mathrm{e}^{-\gamma_{\mathrm{EM}}\gamma}.$$

Furthermore, (ii) is equivalent to demanding that

$$u_{k+1} \le u_k^{\frac{\gamma}{1-\gamma}}\Big(\frac{\gamma\sigma q(\gamma)}{8\log(k+3) - 4\log\frac{6\varepsilon}{\pi^2}}\Big)^{\frac{\gamma}{1-\gamma}},$$

which is satisfied for all k by the choice of $c(\gamma,\varepsilon)$ in (5.14), as

$$\mathrm{e}^{-c(\gamma,\varepsilon)} = \Big(\frac{\gamma\sigma q(\gamma)}{8 - 4\frac{\log\frac{\varepsilon}{\pi^2}}{\log 2}}\Big)^{\frac{\gamma}{1-\gamma}} \le \Big(\frac{\gamma\sigma q(\gamma)}{8 - 4\frac{\log\frac{\varepsilon}{\pi^2}}{\log(k+3)}}\Big)^{\frac{\gamma}{1-\gamma}}, \quad \text{for all } k \in \mathbb{N}.$$

Finally, for (iii), note that (5.19) implies that $\eta_k \ge \log 2 + g_0\big(\frac{\gamma}{1-\gamma}\big)^k$ for some $g_0 > 0$ and all $k \ge k_0$, where k_0 does only depend on γ and ε. Thus, if

$$k \ge \frac{\log\log N}{\log\frac{\gamma}{1-\gamma}} + \frac{\log\big(\frac{\gamma}{2g_0} - \frac{\log\log N}{2g_0\log N}\big)}{\log\frac{\gamma}{1-\gamma}}$$

and N sufficiently large,

$$\eta_k \ge \log 2 + \frac{\gamma}{2}\log N - \frac{1}{2}\log\log N$$

which implies the statement. $\qquad\square$

The sequence u is now utilised to split the high degree vertices into layers.

Definition 5.5. The k-*core* of $\mathscr{G}_{2N}$ is given by

$$\mathscr{K}_N^{(k)} = \{i \in [N] : \mathbb{E}\mathscr{Z}[i, N] \ge u_k\}, \quad \text{for } k \ge 0.$$

For $k = 0$ we just speak of the *extended core*. Let furthermore $\mathscr{N}^{(k)} = \mathscr{K}^{(k)} \setminus \mathscr{K}^{(k+1)}$.

The size of the layers plays a role in the calculations below, we therefore derive some rough bounds using the inverse of $\xi^{(\gamma)}(\cdot, N)$.

Lemma 5.6 (Size of layers)**.** *For all $\mathscr{N}^{(k)} \subset V(\mathscr{G}_{2N})$ such that $\mathscr{N}^{(k+1)} \ne \emptyset$, we have*

$$\#\mathscr{N}^{(k)} \ge N(u_k^{-\frac{1}{\gamma}} - \mathrm{e}^{\gamma_{EM}}u_{k+1}^{-\frac{1}{\gamma}}).$$

Proof. Let v denote the inverse of the discrete function $\xi^{(\gamma)}(\cdot, N)$, then Lemma 4.6 entails that, for all $y \in Y = \{\xi^{(\gamma)}(i, N), i \in [N]\}$,

$$N u_k^{-\frac{1}{\gamma}} \le v(y) \le e^{\gamma \mathrm{EM}} N u_{k+1}^{-\frac{1}{\gamma}}$$

and therefore

$$\#\mathcal{N}^{(k)} = \#\{v(y), y \in Y \cap [u_k, u_{k+1}]\} \ge N y^{-\frac{1}{\gamma}} - e^{\gamma \mathrm{EM}} N y^{-\frac{1}{\gamma}}.$$

$\square$

The following lemma is a straightforward consequence of the Payley-Zygmund inequality.

Lemma 5.7. *Let $f(k) = \gamma k + \beta$, where $\beta, \gamma \in (0, 1)$ and recall that $\xi^* = \inf_{k \in \mathbb{N}} \xi^{(\gamma)}(k, 2k) \ge 2^\gamma$. The following hold for all choices $n > m \in \mathbb{N}$:*

(i) we have

$$\mathbb{P}\{\mathcal{Z}[m, n] \ge \vartheta \mathbb{E}\mathcal{Z}[m, n]\} \ge (1 - \vartheta)^2 \frac{\beta^2}{\beta^2 + \gamma}, \quad \text{for all } 0 \le \vartheta < 1;$$

(ii) we have

$$\mathbb{P}\{\mathcal{Z}[m, 2n] - \mathcal{Z}[m, n] \ge \vartheta \mathcal{Z}[m, n]\} \ge \left(1 - \frac{1 + \vartheta}{xi^*}\right)^2 \quad \text{for all } 0 \le \vartheta < \xi^* - 1.$$

Proof. From Lemma 4.7 we can infer that

$$\mathbb{E}^k f(\mathcal{Z}[m, n]) = f(k)\xi^{(\gamma)}(m, n) \tag{5.20}$$

and that

$$\begin{aligned} \mathbb{E} f(\mathcal{Z}[m, n])^2 &= (f(k)^2 + \gamma)\xi^{(2\gamma)}(m, n) - \gamma f(k)\xi^{(\gamma)}(m, n) \\ &\le (f(k)^2 + \gamma)\xi^{(2\gamma)}(m, n), \end{aligned} \tag{5.21}$$

and therefore an application of the Payley-Zygmund inequality and Lemma 4.7 yields

$$\begin{aligned} \mathbb{P}\{f(\mathcal{Z}[m, n]) \ge \vartheta \mathbb{E} f(\mathcal{Z}[m, n])\} &\ge (1 - \vartheta)^2 \frac{(\mathbb{E} f(\mathcal{Z}[m, n]))^2}{\mathbb{E} f(\mathcal{Z}[m, n])^2} \\ &\ge (1 - \vartheta)^2 \frac{\beta^2 \xi^{(\gamma)}(m, n)^2}{\beta^2 + \gamma \xi^{(2\gamma)}(m, n)} \ge (1 - \vartheta)^2 \frac{\beta^2}{\beta^2 + \gamma}, \end{aligned}$$

which proves statement (i), since $f(\mathcal{Z}[m, n]) \ge \vartheta \mathbb{E} f(\mathcal{Z}[m, n])$ if $\mathcal{Z}[m, n] \ge \vartheta \mathbb{E}\mathcal{Z}[m, n]$.

For (ii), note that for $k \in \mathbb{N}_0$,

$$\mathbb{E}\left[\frac{f(\mathcal{Z}[m, 2n])}{f(\mathcal{Z}[m, n])}; \mathcal{Z}[m, n] = k\right] = \frac{\mathbb{E}^k f(\mathcal{Z}[n, 2n])}{f(k)} \mathbb{P}\{\mathcal{Z}[m, n] = k\} = \xi^{(\gamma)}(n, 2n)\mathbb{P}\{\mathcal{Z}[m, n] = k\},$$

by using the Markov property of the evolution at time n and (5.20). Thus

$$\mathbb{E}\frac{f(\mathcal{Z}[m,2n])}{f(\mathcal{Z}[m,n])} = \sum_{k=0}^{\infty} \mathbb{E}\Big[\frac{f(\mathcal{Z}[m,2n])}{f(\mathcal{Z}[m,n])}; \mathcal{Z}[m,n] = k\Big] = \xi^{(\gamma)}(n,2n).$$

By a similar argument, using the inequality (5.21), we also obtain

$$\mathbb{E}\Big(\frac{f(\mathcal{Z}[m,2n])}{f(\mathcal{Z}[m,n])}\Big)^2 \le \xi^{(2\gamma)}(n,2n).$$

We also note that for (the linear interpolation of) affine f

$$f(xk) \le x f(k), \quad \text{for all } x \ge 1,$$

and together with one more application of the Payley-Zygmund inequality and Lemma 4.7 we can conclude the proof,

$$\begin{aligned}
\mathbb{P}\{\mathcal{Z}[m,2n] - \mathcal{Z}[m,n] \ge \vartheta \mathcal{Z}[m,n]\} &= \mathbb{P}\{\mathcal{Z}[m,2n] \ge (1+\vartheta)Zz[m,n]\} \\
&= \mathbb{P}\{f(\mathcal{Z}[m,2n]) \ge f((1+\vartheta)\mathcal{Z}[m,n])\} \\
&\ge \mathbb{P}\{f(\mathcal{Z}[m,2n]) \ge (1+\vartheta)f(\mathcal{Z}[m,n])\} \\
&\ge \Big(1 - \frac{1+\vartheta}{\xi^{(\gamma)}(n,2n)}\Big)^2 \frac{\xi^{(\gamma)}(n,2n)^2}{\xi^{(2\gamma)}(n,2n)} \ge \Big(1 - \frac{1+\vartheta}{\xi^{(\gamma)}(n,2n)}\Big)^2.
\end{aligned}$$

$\square$

Definition 5.8. Recall that $\sigma = \frac{\xi^* + 3}{4} - 1$. We call a vertex $v \in \mathcal{N}^{(k)}$ *good* if both $\mathcal{Z}[v,N] \ge \frac{1}{2}u_k$ and $\mathcal{Z}[v,2N] - \mathcal{Z}[v,N] \ge \sigma u_k$

Note that the definition of the layers is such that for sufficiently large $k \ge L$, with L as in Lemma 5.4, the good vertices in $\mathcal{N}^{(k)}$ are in core$_N$.

The next step is to bound from below the probability that a vertex is good. Note also that goodness of a vertex is a condition on the degree evolution, and the corresponding events are therefore independent.

Lemma 5.9. *The probability that $v \in \mathcal{N}^{(k)}$ is good is uniformly bounded away from 0, i.e. there exists $q(\gamma) > 0$ such that*

$$\mathbb{P}\{v \text{ is good}\} \ge q(\gamma), \quad \text{for all } N \in \mathbb{N}, k \text{ with } \mathcal{N}^{(k)} \ne \emptyset \text{ and } v \in \mathcal{N}^{(k)}.$$

Proof. We calculate

$$
\begin{aligned}
\mathbb{P}\{v \text{ is good}\} &= \sum_{x \geq \frac{1}{2} u_k} \mathbb{P}\{\mathcal{Z}[v, 2N] - \mathcal{Z}[v, N] \geq \sigma u_k, \mathcal{Z}[v, N] = x\} \\
&= \sum_{x \geq \frac{1}{2} u_k} \mathbb{P}\Big\{\mathcal{Z}[v, 2N] \geq \Big(1 + \frac{\sigma u_k}{\mathcal{Z}[v, N]}\Big)\mathcal{Z}[v, N]\Big|\mathcal{Z}[v, N] = x\Big\}\mathbb{P}\{\mathcal{Z}[v, N] = x\} \\
&= \sum_{x \geq \frac{1}{2} u_k} \mathbb{P}^x\Big\{\mathcal{Z}[v, 2N] \geq \Big(1 + \frac{\sigma u_k}{x}\Big)x\Big\}\mathbb{P}\{\mathcal{Z}[v, N] = x\} \\
&\geq \sum_{x \geq \frac{1}{2} u_k} \mathbb{P}^x\Big\{f(\mathcal{Z}[v, 2N]) \geq \Big(1 + \frac{\sigma u_k}{x}\Big)f(x)\Big\}\mathbb{P}\{\mathcal{Z}[v, N] = x\} \\
&\geq \sum_{x \geq \frac{1}{2} u_k} \Big(1 - \frac{1 + \frac{\sigma u_k}{x}}{\xi^{(\gamma)}(N, 2N)}\Big)\frac{\xi^{(\gamma)}(N, 2N)^2}{\xi^{(2\gamma)}(N, 2N)}\mathbb{P}\{\mathcal{Z}[v, N] = x\}
\end{aligned}
$$

using the Markov property and the Payley-Zygmund inequality. The restriction $x \geq \frac{1}{2} u_k$ and the choice of σ imply the result. $\qquad\square$

We are now set up to begin the main part of the proof. We determine especially short paths connecting good vertices in the extended core. The following proposition states that the distance between a uniformly chosen initial vertex $V \in \mathscr{C}_{2N}$ and a good vertex in the extended core is with high probability bounded. The proof relies in part on the coupling of the exploration procedures detailed in Sections 4.3 and 4.4 and the properties of the INT mentioned in Section 1.3.3.

Proposition 5.10 (Distance of periphery and extended core)**.** *Let $V \in \mathscr{G}_{2N}$ be a uniformly chosen random vertex. Then, for every $\varepsilon > 0$, there is a $K = K(\varepsilon) \in \mathbb{N}$ such that*

$$\mathbb{P}\{\mathscr{C}_{2N}^K(V) \text{ contains a good vertex}\} \geq p(f) - \varepsilon.$$

Proof. We use the couplings of Chapter 4 to be able to work with the INT. We fix some notation first by defining the sequence of events

$$E_k = E_k^{(V)} = \{\text{the coupling fails before step } k + 1\} \quad \text{for } k \in \mathbb{N}.$$

Now observe that the sets $\mathcal{N}^{(k)}$ have disjoint pre-images under the projection π_{2N} and that the goodness of a vertex in the network is (under the coupling) equivalent to a condition on the number and positions of offspring of the corresponding particle in $\mathfrak{T}$. We may thus also speak of *good particles* in the INT, but recall that this definition depends on ε. We are only interested in the situation where the coupling succeeds. For any $k \in \mathbb{N}$, let $\mathfrak{T}^k$ denote the

first k vertices explored in $\mathfrak{T}$, then

$$
\begin{aligned}
\mathbb{P}\{\mathscr{C}_{2N}^K(V) \text{ contains a good vertex}\} & \\
&\geq \mathbb{P}\{\mathscr{C}_{2N}^K(V) \text{ contains a good vertex}|E_K^C\} - \mathbb{P}(E_K) \\
&= \mathbb{P}\{\mathfrak{T}^K \text{ contains a good particle}\} - \mathbb{P}(E_K) \\
&\geq p(f)\mathbb{P}\{\mathfrak{T}^K \text{ contains a good particle}|\#\mathfrak{T} = \infty\} - \mathbb{P}(E_K).
\end{aligned}
\tag{5.22}
$$

We first establish that the INT survives if and only if it contains infinitely many particles left of any bound $-L$ on the negative real line, i.e.

$$
\#\mathfrak{T} = \infty \quad \text{if and only if} \quad \forall L < \infty : \#(\mathfrak{T} \cap (-\infty, -L]) = \infty \quad \text{almost surely.} \tag{5.23}
$$

Only one implication is non-trivial. To see that it holds, we look at the killed BRW generating $\mathfrak{T}$ and note that, almost surely, every generation has only finitely many particles. Assume that the walk survives but has only finitely many particles left of $-L$. In this case there must be a generation $R < \infty$ after which no offspring in $(-\infty, -L]$ is produced. Survival implies that there is at least one particle x_i in every generation $R + i$ for all $i \in \mathbb{N}$. Each of those particles has a positive probability $p_i \geq p \geq \mu((-\infty, -L) > 0$ to produce offspring left of $-L$, where μ denotes the intensity measure of the Poisson offspring distribution to the left. Thus

$$
\mathbb{P}\{\text{no } x_i \text{ has offspring in } (-\infty, -L]\} \leq \lim_{I \to \infty} (1 - p)^I = 0.
$$

Thus $R = \infty$ almost surely which contradicts our assumption.

We can use (5.23) now to show that for all $\delta > 0$, there is $K(\delta)$ such that

$$
\mathbb{P}\{\mathfrak{T}^{K(\delta)} \text{ contains a good particle} |\#\mathfrak{T} = \infty\} \geq 1 - \delta. \tag{5.24}
$$

Let $y = y(\varepsilon, \gamma) < 0$ denote the least upper bound of $\pi_{2N}^{-1}(\mathcal{N}^0)$. For $m \in \mathbb{N}$, set

$$
p_{m,k} = \mathbb{P}\{\mathfrak{T}^k \text{ contains a good particle} |\#\mathfrak{T}^k \cap (-\infty, y) \geq m, \#\mathfrak{T} = \infty\}.
$$

By (5.23) we can fix for every m some $k(m) \in \mathbb{N}$ such that

$$
\mathbb{P}\{\#\mathfrak{T}^{k(m)} \cap (-\infty, y] \geq m|\#\mathfrak{T} = \infty\} \geq 1 - \frac{\delta}{2}.
$$

Every particle left of y has a small positive probability, uniformly bounded from below by some $q > 0$, to be good, thus

$$
M = M(\delta) = \min\left\{m \in \mathbb{N} : p_{m,k(m)} \geq 1 - \frac{\delta}{2}\right\}
$$

is finite. Combining these estimates yields

$$\mathbb{P}\{\mathfrak{T}^{k(M)} \text{ contains a good particle}|\#\mathfrak{T} = \infty\}$$
$$\geq p_{M,k(M)}\mathbb{P}\{\#\mathfrak{T}^{k(M)} \cap (-\infty, y) \geq M|\#\mathfrak{T} = \infty\}$$
$$\geq (1 - \frac{\delta}{2})^2 \geq 1 - \delta,$$

and setting $K(\delta) = k(M)$ shows (5.24). Setting $\delta = \frac{\varepsilon}{2p(f)}$ in (5.24), we obtain from (5.22) that

$$\mathbb{P}\{\mathscr{C}_{2N}^K(V) \text{ contains a good vertex}\}$$
$$\geq p(f)(1 - \delta) - \mathbb{P}\{E_K^C\} \geq p(f) - \frac{\varepsilon}{2} - \frac{\varepsilon}{2},$$

if N is sufficiently large, since K does not depend on N and we can use the coupling result Proposition 4.24. □

Remark 5.11. In this proof, we could also have used the version of the coupling lemma in [DM13]. In fact, the much weaker demand that $\mathscr{C}_{2N}(V)$ has $\mathfrak{T}$ as its local weak limit is sufficient.

At the time when the successful exploration stops, the score of the explored part of the graph is at most finite (i.e. it does can be bounded independently of N). This is useful for the remainder of the proof since it implies that conditioning on the information obtained from the degree evolutions during the initial part of the exploration does not have a large effect on the connection probabilities in the extended core of the network.

Lemma 5.12. *Let $K \in \mathbb{N}$. Conditional on successful completion of the initial exploration in step $K \in \mathbb{N}$, for every $\varepsilon > 0$, there exists $S(\varepsilon)$ such that*

$$\mathbb{P}\{\xi(\mathscr{C}^K(V)) \leq S(\varepsilon)\} \geq 1 - \frac{\varepsilon}{2}.$$

Proof. The coupling succeeds, so we can argue using the INT. Denote by $\mathfrak{T}^K$ the explored part of the INT corresponding to $\mathscr{C}_{2N}^K(V)$. We can first find a large constant $M(K)$ such that

$$\mathbb{P}\{\#\mathfrak{T}^K \leq M(K)\} \geq 1 - \frac{\varepsilon}{2}$$

and then fix a position $y = y(M(K))$ on the negative real line such that

$$\mathbb{P}\{\mathfrak{T}^K \text{ contains a particle left of } y|\#\mathfrak{T}^K \leq M(K)\} \leq \frac{\varepsilon}{2}.$$

Conditional on success of the coupling, we therefore get by applying Markov's inequality

$$\mathbb{P}\{\xi(\mathscr{C}^K(V)) \geq S\}$$

$$\leq \mathbb{P}\{\xi(\mathscr{C}^K(V)) \geq S, \#\mathfrak{T}^K \leq M(K), \mathfrak{T}^K \text{ contains a particle left of } y\} + \frac{\varepsilon}{2}$$

$$\leq \mathbb{P}\{\xi(\mathscr{C}^K(V)) \geq S \,|\, \#\mathfrak{T}^K \leq M(K), \mathfrak{T}^K \text{ contains a particle left of } y\} + \frac{\varepsilon}{2}$$

$$\leq \frac{\mathbb{E}\{\xi(\mathscr{C}^K(V)) \,|\, \#\mathfrak{T}^K \leq M(K), \mathfrak{T}^K \text{ contains a particle left of } y\}}{S} + \frac{\varepsilon}{2}$$

$$\leq \frac{M(K)\xi(\pi_{2N}(y(M(K))))}{S} + \frac{\varepsilon}{2}.$$

The assertion is obtained by choosing S large enough, since $\xi(\pi_{2N}(y(M(K))))$ can be uniformly bounded in N. $\qquad\square$

We shall now see that along the good vertices, we can find with high probability a short path into the core.

Proposition 5.13 (Distance of the extended core to the core)**.** *Let $\varepsilon > 0$ and assume that there is a good vertex $W \in \mathcal{K}_N^{(0)}$ in which the local exploration around V ends after at most $K(\varepsilon)$ steps. There is a constant $C = C(\varepsilon, \gamma) < \infty$ such that*

$$\mathbb{P}\Big\{d_{2N}(W, core_N) > 2\frac{\log\log N}{\log\frac{\gamma}{1-\gamma}} + C\Big\} \leq \varepsilon,$$

if N is sufficiently large.

Proof. We aim to construct a path from W to the core using 1-connectors to successively connect a good vertex in layer $\mathcal{N}^{(k)}$ to a good vertex in layer $\mathcal{N}^{(k+1)}$. This means every connection step corresponds to 2 edges in the constructed path. We start at step $k = 0$ with the good vertex W, we can assume that $W \in \mathcal{N}^{(0)}$ otherwise we reduce the total number of steps which only makes the path shorter. Assume the procedure has succeeded up to step k, i.e. we have constructed a path of length $2k$ to a good vertex $v \in \mathcal{N}^{(k)}$. Let $\mathscr{F}_k$ denote the σ-algebra generated by all degree evolutions used in the construction so far and let $\mathscr{C}_0(V)$ denote the part of the neighbourhood around V explored at the time at which W was discovered. Let furthermore $w \in \mathcal{N}^{(k+1)}$ be any vertex and $j \in \{N+1, \dots, 2N\}$ with $\Delta\mathscr{Z}[v, j] = 1$ be the potential next vertices on the constructed path. Denote by $\Gamma^>(w)$ the immediate right neighbours of w and note that conditional on $\mathscr{F}_k$ the probability of the event

$$E_0 = E_0^{(w)} = \{\Gamma^>(w) \cap \mathscr{C}_0(V) = \emptyset\}$$

can be made larger than $1 - \frac{q(\gamma)}{2}$ uniformly in w, since $\xi(\mathscr{C}_0(V)) \leq S(\varepsilon)$,

$$\mathbb{P}\{E_0\} \geq 1 - \sum_{i \in \mathscr{C}_0(V)} \frac{\mathbb{E}f(\mathscr{Z}[w, i])}{i} \geq 1 - \frac{\xi(\mathscr{C}_0(V))}{(2N)^{1-\gamma}},$$

using that $1-\gamma < \gamma$ and recalling that in the affine case $\xi(w,i) = \mathbb{E}\mathcal{Z}[w,i]$. We are interested in the probability that j is connected to w and that w is good.

$$\mathbb{P}\{\Delta\mathcal{Z}[w,j] = 1, w \text{ is good}|\mathscr{F}_k\}$$
$$\geq (\mathbb{P}\{w \text{ is good}\} - \mathbb{P}\{E_0^C\})\mathbb{P}\{\Delta\mathcal{Z}[w,j] = 1|w \text{ is good}\}$$
$$\geq \frac{q(\gamma)}{2}\mathbb{P}\{\Delta\mathcal{Z}[w,j] = 1|w \text{ is good}\},$$

if N is sufficiently large, since the indegree evolution $\mathcal{Z}[w,\cdot]$ is independent of $\mathscr{F}_k$ and by Lemma 5.9. Furthermore,

$$\mathbb{P}\{\Delta\mathcal{Z}[w,j] = 1|w \text{ is good}\} \geq \mathbb{P}\{\Delta\mathcal{Z}[w,j] = 1\} \geq \frac{f(u_{k+1})}{2N}$$

and this bound holds independently of the status of all other potential 1-connectors between v and $\mathcal{N}^{(k+1)}$. Using this independence we can get a bound on the probability that we fail to connect in the next step by estimating the probability that there is no 1-connector to a good vertex in $\mathcal{N}^{(k+1)}$ among the right neighbours $\Gamma^{>}(v)$ of v.

$$\mathbb{P}\{\nexists \text{ good } w \in \mathcal{N}^{(k+1)}: w \overset{1}{\leftrightarrow} v\} \leq \left(1 - q(\gamma)\frac{f(u_{k+1})}{2N}\right)^{\sigma u_k \# \mathcal{N}^{(k+1)}}$$
$$\leq \exp\left(-q(\gamma)\frac{f(u_{k+1})}{2}\sigma u_k(u_{k+1}^{-\frac{1}{\gamma}} - e^{\gamma \text{EM}}u_{k+2}^{-\frac{1}{\gamma}})\right),$$

using first the goodness of v and then Lemma 5.6. Applying now first $f(k) \geq \gamma k$ and then Lemma 5.4 (i) and (ii), we can bound the last expression further

$$\exp\left(-q(\gamma)\frac{f(u_{k+1})}{2}\sigma u_k(u_{k+1}^{-\frac{1}{\gamma}} - e^{\gamma \text{EM}}u_{k+2}^{-\frac{1}{\gamma}})\right) \leq \exp\left(-\frac{q(\gamma)\gamma\sigma}{2}u_k u_{k+1}^{1-\frac{1}{\gamma}}\left(1 - e^{\gamma \text{EM}}\left(\frac{u_k+1}{u_{k+2}}\right)^{\frac{1}{\gamma}}\right)\right)$$
$$\leq \exp\left(-\frac{\gamma\sigma q(\gamma)}{4}u_k u_{k+1}^{1-\frac{1}{\gamma}}\right) \leq \frac{6\varepsilon}{\pi^2}\frac{1}{(k+3)^2}.$$

The cumulative probability of failure can be bounded by summing up the right hand side for every stage k,

$$\frac{6\varepsilon}{\pi^2}\sum_{k=0}^{\infty}\frac{1}{(k+3)^2} < \frac{6\varepsilon}{\pi^2}\sum_{k=0}^{\infty}\frac{1}{k^2} = \varepsilon$$

and the length of the path featured in the proposition follows from Lemma 5.4 (iii). □

Proof of Theorem 2.6. We fix $\varepsilon > 0$. Let $U, V \in \mathscr{C}_{2N}$ be uniformly chosen vertices, then Proposition 5.10 implies that there is a large number $K = K(\varepsilon)$ such that

$$\mathbb{P}\{\mathscr{C}_{2N}^K(V) \text{ and } \mathscr{C}_{2N}^K(U) \text{ contain good } W^{(v)}, W^{(w)} \in \mathscr{K}_N^{(0)}\} \geq 1 - \varepsilon, \quad \text{if } N \text{ is sufficiently large.}$$

By Proposition 5.13, both $W^{(v)}$ and $W^{(w)}$ can be connected to core_N in less than $2\frac{\log\log N}{\log\frac{\gamma}{1-\gamma}} + C(\varepsilon,\gamma)$ steps with probability exceeding $1 - 2\varepsilon$, if N is sufficiently large. Finally, the diameter $\text{diam}_{2N}(\text{core}_N)$ is bounded by a constant $D = D(\gamma)$ with probability at least $1 - \varepsilon$ by Proposi-

tion 4.36. Combining all these bounds, we arrive at

$$\mathbb{P}\left\{d_{2N}(U,V) > 4\frac{\log\log N}{\log\frac{\gamma}{1-\gamma}} + 2C + 2K + D\right\} \le 4\varepsilon, \quad \text{if } N \text{ is sufficiently large,}$$

and thus

$$d_{2N}(U,V) \le 4\frac{\log\log N}{\log\frac{\gamma}{1-\gamma}} + O(1) \quad \text{with high probability.}$$

$\square$

5.4 Proof of Theorem I

Proof. We start with the lower bound. First note that, for $v < w \in [N]$,

$$\mathbb{P}\{v \leftrightarrow w\} = \frac{\mathbb{E}f(\mathcal{Z}[v,w-1])}{w-1}. \tag{5.25}$$

An estimate for the expectation can be obtained by fixing first $\varepsilon > 0$ and then using the concavity of f to fix k such that, for all $n \ge k$, we have $f(n) \le f(k) + (\gamma+\varepsilon)(n-k)$. Note that $\mathbb{E}f(\mathcal{Z}[v,w-1]) \le \mathbb{E}^k f(Z[v,w-1])$. Lemmas 4.7(iv) and 4.6 now imply that

$$\mathbb{E}f(\mathcal{Z}[v,w-1]) \le f(k)\xi^{(\gamma+\varepsilon)}(v,w) \le f(k)C_{4.6}\left(\frac{w}{v}\right)^{\gamma+\varepsilon} = C_1\left(\frac{w}{v}\right)^{\gamma+\varepsilon}, \tag{5.26}$$

For the constant $C_1 = f(k)C_{4.6}$. We now use (5.26) to verify Assumption 2.2 for $\gamma + \varepsilon$. For $v < w \in [N]$, all events $\{v \leftrightarrow w\}$ with different values of v are independent. Hence $\mathbb{P}\{v_0 \leftrightarrow \cdots \leftrightarrow v_n\}$ can be decomposed into factors of the form $\mathbb{P}\{v_{j-1} \leftrightarrow v_j \leftrightarrow v_{j+1}\}$ with $v_j < v_{j-1}, v_{j+1}$ and factors of the form $\mathbb{P}\{v_{j-1} \leftrightarrow v_j\}$ for the remaining edges. It remains to estimate factors of the latter form. We may assume $v < u < w$ and get

$$\mathbb{P}\{u \leftrightarrow v \leftrightarrow w\} = \frac{\mathbb{E}[f(\mathcal{Z}[v,u-1])f(\mathcal{Z}[v,w-1])]}{(u-1)(w-1)}.$$

Arguing as in the derivation of (5.26) we get, for a suitable constant $C_2 > 0$,

$$\mathbb{E}[f(\mathcal{Z}[v,w-1]) | \mathcal{Z}[v,u-1] = k] \le C_2 f(k) w^{\gamma+\varepsilon} u^{-\gamma-\varepsilon}.$$

Hence

$$\mathbb{E}[f(\mathcal{Z}[v,u-1])f(\mathcal{Z}[v,w-1])] \le C_2 \mathbb{E}[f(Z[v,u-1])^2] w^{\gamma+\varepsilon} u^{-\gamma-\varepsilon},$$

and, using a similar argument as above, we obtain $C_3 > 0$ such that

$$\mathbb{E}[f(\mathcal{Z}[v,u-1])^2] \le C_3 u^{2\gamma+\varepsilon} v^{-2\gamma-\varepsilon}.$$

Summarising, we obtain a constant $C_4 > 0$ such that

$$\mathbb{P}\{u \leftrightarrow v \leftrightarrow w\} \leq C_4 u^{\gamma-1+\varepsilon} v^{2\gamma-\varepsilon} w^{\gamma-1+\varepsilon},$$

as required for Assumption 2.2. Since ε was chosen arbitrarily, Theorem 2.3 implies that

$$d_N(V, W) \geq (4 + o(1)) \frac{\log \log N}{\log \frac{\gamma}{1-\gamma}}, \quad \text{with high probability as } N \to \infty.$$

If f is affine, we can verify Assumption 2.2 directly for γ, and then Theorem 2.3 yields a deviation of constant order, as stated in Theorem I.

The upper bound in the affine case is Theorem 2.6. For general concave f, we note that

$$f(k) \geq f(0) + \gamma k = \bar{f}(k)$$

for all k and thus the stochastic domination of Observation 1.7 yields that $\mathcal{G}_N$ and $\bar{\mathcal{G}}_N$ can be coupled such that the distances in $\mathcal{G}_N$ are bounded by the distances in $\bar{\mathcal{G}}_N$. Applying Theorem 2.6 to $\bar{\mathcal{G}}_N$ then yields the bound. □

Chapter 6

Typical distances for $\gamma = \frac{1}{2}$

This chapter is based on recent joint work with Steffen Dereich and Peter Mörters and the results have not yet been published elsewhere.

In this chapter we combine the techniques provided in Chapter 4 with some more specific calculations to prove Theorem II. First, in Propositions 6.1 and 6.2, we derive bounds for the expected degrees of the vertices in the network given the particular form of attachment rule featured in Theorem II. Then we adapt the second moment discussion of Section 4.5 from the BRW setting to preferential attachment graphs in Section 6.2. Finally, we prove Theorem II in Section 6.3 by combining all these results with the exploration and sprinkling techniques of Chapter 4.

The proof works along the following lines: We work outwards from the randomly chosen vertices U, V using a *depth first* exploration to determine whether they lie in the giant component and to accumulate a high initial score. We then carry on with the explorations but move one step away from the tree picture and allow some circles in the explored part of the graph, using a *breadth first* exploration. We show that the score of the exploration grows at a rate which implies Theorem II until it reaches a certain threshold at which we stop exploring. Finally, we use ε-connectors to show that the distance between the explored parts of the graph is negligible, in particular we use Proposition 4.36(ii), i.e. that the set of old vertices in $\mathcal{G}_N$ has a small diameter, if the parameter α in Theorem II is strictly positive. If $\alpha = 0$, we argue slightly differently with the aim to show that the (explored) components to which the initial vertices belong are directly connected via ε-connectors.

The following section provides the bounds for the edge density around a given vertex, which are needed to show that the score grows at the rate postulated in Theorem II.

6.1 Moment bounds for $f(k) = \frac{1}{2}k + \frac{\alpha k}{2\log k} + o\big(\frac{k}{(\log k)^{1+\eta}}\big)$

Below we derive lower and upper bounds for $\mathbb{E}^k f(\mathcal{Z}[m,n])$. For notational convenience, denote $f(k) = \frac{1}{2}k + g(k)$, where g satisfies

$$g(k) = \frac{\alpha}{2}\frac{k}{\log k} + o\Big(\frac{k}{(\log k)^{1+\eta}}\Big), \quad \text{for some } \eta > 0, \tag{6.1}$$

where $\alpha \geq 0$ is the parameter of interest. In the linear setting, the score of a vertex is proportional to its expected degree, recall that, by Lemma 4.7, for $n > m \in [N]$, and fixed k,

$$\xi(n,m) = \frac{\mathbb{E}^k f(\mathcal{Z}[m,n])}{f(k)} = \prod_{i=m}^{n-1}\Big(1 + \frac{1}{2i}\Big) \approx \sqrt{\frac{n}{m}}.$$

Recall that, for the deviation from the linear case, we use the notation

$$\psi^k(m,n) = \frac{\mathbb{E}^k f(\mathcal{Z}[m,n])}{\xi(m,n)}.$$

Determining the order of ψ^k will be the first step towards the proof of Theorem II.

Proposition 6.1 (First and second moment upper bound)**.** *Let f be an attachment rule with $f(k) = \frac{1}{2}k + g(k)$, where $g : \mathbb{N}_0 \longrightarrow (0,\infty)$ is concave and satisfies (6.1). Then, for any $k \in \mathbb{N}$, there exists constant $C(k), C'(k) < \infty$ depending only on f and k, such that uniformly for all pairs $m, n \in \mathbb{N}$ with $n \geq m$,*

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \leq C(k)\sqrt{\frac{n}{m}}\big(\log \tfrac{n}{m} \vee 1\big)^\alpha$$

and

$$\mathbb{E}^k f(\mathcal{Z}[m,n])^2 \leq C'(k)\frac{n}{m}\big(\log \tfrac{n}{m} \vee 1\big)^{2\alpha}.$$

Proposition 6.2 (First moment lower bound)**.** *Let f, g be as in Proposition 6.1. Then, for any $k \in \mathbb{N}$ there exists a constant $c > 0$, only dependent on f, such that uniformly for all pairs $m, n \in \mathbb{N}$ with $n > m$,*

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \geq c\sqrt{\frac{n}{m}}\big(\log \tfrac{n}{m} \vee 1\big)^\alpha.$$

To prove Propositions 6.1 and 6.2 we need two auxiliary statements, which are straightforward consequences of our restrictions on f.

Lemma 6.3. *Let f be a concave attachment rule with $\gamma = \frac{1}{2}$. There is an integer $K = K(f) > 0$, such that for all start times $m \in \mathbb{N}$ and initial degrees $k > K \in \mathbb{N}$,*

$$\mathrm{e} \vee \sqrt{\frac{n}{m}} \leq \mathbb{E}^k f(\mathcal{Z}[m,n]) \leq f(k)\mathrm{e}^{\Delta f(0)\gamma_{EM}}\Big(\frac{n}{m}\Big)^{\Delta f(0)},$$

for all $n \geq m \in \mathbb{N}$, where γ_{EM} denotes the Euler-Mascheroni constant.

Proof. This is a corollary of Lemma 4.9. By concavity $\gamma^+ = \Delta f(0)$ and thus the upper bound follows from the upper bound there. For the lower bound, note that concavity implies that $\gamma^- = \gamma = \frac{1}{2}$ and thus the lower bound in Lemma 4.9 yields

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \geq f(k)\sqrt{\frac{n}{m}},$$

i.e. $K(f) = \min\{k \in \mathbb{N} : f(k) > e\}$ has the required property. $\qquad\square$

We next obtain bounds for the function $\phi(x) = \sum_{i=0}^{x-1} \frac{1}{f(i)}$ featuring in Lemma 4.11 and its inverse.

Lemma 6.4 (Bounds for ϕ). *Let f satisfy condition* (6.1) *with $\alpha, \eta > 0$. Then,*

(i) *the linear interpolation $\phi^{-1} : \left[\frac{1}{f(0)}, \infty\right) \longrightarrow [0, \infty)$ of the inverse of ϕ exists and is strictly monotone, in particular, for $x \geq \frac{1}{f(0)}$ and $k \in \mathbb{N}$,*

$$\phi^{-1}(x) \geq k, \quad \text{if } x \geq \phi(k);$$

(ii) *there are constants $c, C \in (0, \infty)$, only depending on f, such that, for all $k \in \mathbb{N}$,*

$$\frac{1}{f(0)} \vee \big(2\log_+ k - 2\alpha \log\log_+ k - c\big) \leq \phi(k) \leq 2\log_+(k-1) - 2\alpha \log\log_+(k-1) + C,$$

where $\log_+(k) = \log k \vee 1$ and $\log\log_+(k) = \log\log k \vee 1$.

Proof. For (i) note, that the attachment rule f is positive and strictly increasing, which implies that $\Delta\phi = \frac{1}{f} > 0$ is strictly decreasing. Thus ϕ is concave and strictly increasing, hence its inverse is well defined, convex, strictly increasing and

$$\phi^{-1}(x) = k, \quad \text{if } x = \sum_{i=0}^{k-1} \frac{1}{f(i)},$$

and the claimed monotonicity is inherited by the linear interpolation.

To show (ii), we use our assumption on the asymptotics of g, to establish the existence of $k_0 = k_0(f) > e^2$, such that, for all $k > k_0$,

$$\frac{k}{2} + \frac{\alpha k}{2\log k} - \frac{k}{(\log k)^{(1+\eta)}} \leq f(k) \leq \frac{k}{2} + \frac{\alpha k}{2\log k} + \frac{k}{(\log k)^{(1+\eta)}}.$$

From these inequalities we can deduce that there is a constant $C' > 0$ and $k_1 > k_0$ such that for all $k \geq k_1$

$$\frac{2}{k} - \frac{2\alpha}{k\log k} - \frac{C'}{k(\log k)^{1+\eta}} \leq \frac{1}{f(k)} \leq \frac{2}{k} - \frac{2\alpha}{k\log k} + \frac{C'}{k(\log k)^{1+\eta}}. \tag{6.2}$$

We set $\Sigma(\eta) = \sum_{k=2}^{\infty} \frac{1}{k(\log k)^{1+\eta}}$, then start with the lower bound in (ii). Summation over the first inequality in (6.2) yields, for $k > k_1$,

$$\phi(k) = \sum_{i=0}^{k-1} \frac{1}{f(i)} \geq -C'\Sigma(\eta) + \sum_{i=0}^{k_1} \frac{1}{f(i)} + 2\sum_{i=k_1}^{k-1} \frac{1}{i} - 2\alpha \sum_{i=k_1}^{k-1} \frac{1}{i\log i},$$

in which we can bound the latter two sums by the integrals $\int_{k_0}^{k-1} s^{-1}\,\mathrm{d}s$ and $\int_{\mathrm{e}}^{k} (s\log s)^{-1}\,\mathrm{d}s$ and we arrive at

$$
\begin{aligned}
\phi(k) \geq &-C'\Sigma(\eta) + \sum_{i=0}^{k_1} \frac{1}{f(i)} + 2\int_{k_1}^{k-1} \frac{1}{s}\,\mathrm{d}s - 2\alpha \int_{\mathrm{e}}^{k} \frac{1}{s\log s}\,\mathrm{d}s \\
= &-C'\Sigma(\eta) + \sum_{i=0}^{k_0} \frac{1}{f(i)} - 2\log\left(\tfrac{k}{k-1}\right) + 2\log\left(\tfrac{k}{k_1}\right) - 2\alpha \log\log k \\
\geq &-C'' + 2\log k - 2\alpha \log\log k,
\end{aligned}
$$

for all $k > k_1$ and a large constant $C'' > 0$. The desired bound now holds for all k, since ϕ is bounded from below by $\frac{1}{f(0)}$ and increasing.

For the upper bound of (ii) we proceed in a similar fashion and sum over the second inequality in (6.2), to obtain for all $k > k_1$ and some constant $D > 0$,

$$\phi(k) \leq D + 2\sum_{i=k_1}^{k-1} \frac{1}{i} - 2\alpha \sum_{i=k_1}^{k-1} \frac{1}{i\log i}.$$

This fact, together with an approximation by integrals for the remaining sums similar to the above yields the existence of a constant $D' > 0$ such that

$$\phi(k) \leq D' + 2\int_{k_1}^{k-1} \frac{1}{s}\,\mathrm{d}s - 2\alpha \int_{k_1}^{k-1} \frac{1}{s\log s}\,\mathrm{d}s$$

for all $k > k_1$. Thus we can find another constant $C > 0$, depending only on D' and k_1, such that

$$\phi(k) \leq C + 2(\log(k-1) \vee 1) - 2\alpha(\log\log(k-1) \vee 1)$$

for all $k \in \mathbb{N}$. $\qquad\square$

We now combine these results to derive the desired bounds for $\mathbb{E}^k f(\mathcal{Z}[m,n])$.

Proof of Proposition 6.1. We start with the first moment. Since f is non-decreasing, stochastic domination implies that it is sufficient to derive the bound for sufficiently large $k > k_1$, as $\mathbb{E}^k f(\mathcal{Z}[m,n]) \leq \mathbb{E}^{k_1} f(\mathcal{Z}[m,n])$ for all $k \leq k_1$ and $m \leq n \in \mathbb{N}$. Hence we are allowed to choose $k_1 = k_1(f)$ large enough to make use of the asymptotic properties of f. We will use this to implement another simplification, namely that we are allowed to make explicit assumptions on the shape of g. To this end, first define k_1 to be the smallest integer larger than e^2 such

that both

$$g(k) < \frac{\alpha k}{2 \log k} + \frac{k}{2(\log k)^{1+\eta}}, \quad \text{for all } k \geq k_1, \tag{6.3}$$

and

$$\mathbb{E}^{k_1} f(\mathcal{Z}[m,n]) \geq \sqrt{\frac{n}{m}} \vee e \tag{6.4}$$

hold. The existence of such k_1 is entailed by the assumption on g and Lemma 6.3. Let now $\bar{f}$ be given by

$$\bar{f}(k) = \begin{cases} f(k), & \text{if } k \leq k_1, \\ \frac{1}{2}k + \frac{\alpha k}{2 \log k} + \frac{k}{2(\log k)^{1+\eta}}, & \text{if } k > k_1; \end{cases}$$

then (6.3) implies that $\bar{f} \geq f$ on $\mathbb{N}_0$. Denoting by $\bar{\mathcal{Z}}[m,\cdot]$ the degree evolutions corresponding to the jump rate $\bar{f}$, another stochastic domination argument yields

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \leq \mathbb{E}^k \bar{f}(\bar{\mathcal{Z}}[m,n]) \quad \text{for all } k \in \mathbb{N}, \, m \leq n \in \mathbb{N}.$$

For the attachment rule $\bar{f}$ we are now able to perform explicit calculations and obtain the desired bound, which will also hold for f. Thus, we can assume without loss of generality for the remainder of the argument that $f = \bar{f}$.

We observe that for all $k > k_1$,

$$\Delta f(k) \leq \frac{1}{2} + \frac{\alpha}{2 \log k} + \frac{1}{2(\log k)^{1+\eta}},$$

which implies the existence of $k_0 > k_1$ such that for all $k \geq k_0$

$$\Delta f(k) \leq \frac{1}{2} + \frac{\alpha}{2 \log f(k)} + \frac{1}{(\log f(k))^{1+\eta}}. \tag{6.5}$$

We are now ready to formulate the recursion argument which will produce the explicit upper bound. Let $k \geq k_0$ be fixed and let $m \in \mathbb{N}$. We can decompose

$$f(\mathcal{Z}[m,n]) = f(k) + \sum_{s=m}^{n-1} \frac{f(\mathcal{Z}[m,s])}{s} \Delta f(\mathcal{Z}[m,s]) + M_m^f(n),$$

where M_m^f is the martingale of Lemma 4.4. Taking expectations we obtain the relation

$$\mathbb{E}^k f(\mathcal{Z}[m,n+1]) = \mathbb{E}^k f(\mathcal{Z}[m,n]) + \mathbb{E}^k \frac{f(\mathcal{Z}[m,n]) \Delta f(\mathcal{Z}[m,n])}{n}. \tag{6.6}$$

Since $k \geq k_0 > k_1$ we can apply (6.5) to obtain

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \Delta f(\mathcal{Z}[m,n]) \leq \frac{1}{2} \mathbb{E}^k f(\mathcal{Z}[m,n]) + \frac{\alpha}{2} \mathbb{E}^k \frac{f(\mathcal{Z}[m,n])}{\log f(\mathcal{Z}[m,n])} + \mathbb{E}^k \frac{f(\mathcal{Z}[m,n])}{(\log \mathcal{Z}[m,n])^{1+\eta}}.$$

The functions $x \mapsto \frac{x}{(\log x)^{1+s}}, s \geq 0$, are concave on (e^2, ∞), thus we can apply Jensen's inequal-

ity to the second and third term in the sum and obtain

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \Delta f(\mathcal{Z}[m,n]) \leq \frac{1}{2} \mathbb{E}^k f(\mathcal{Z}[m,n]) + \frac{\alpha}{2} \frac{\mathbb{E}^k f(\mathcal{Z}[m,n])}{\log \mathbb{E}^k f(\mathcal{Z}[m,n])} + \frac{\mathbb{E}^k f(\mathcal{Z}[m,n])}{(\log \mathbb{E}^k f(\mathcal{Z}[m,n]))^{1+\eta}}.$$

Applying this bound to the right hand side of (6.6) yields, after division by $\mathbb{E}^k f(\mathcal{Z}[m,n])$,

$$\frac{\mathbb{E}^k f(\mathcal{Z}[m,n+1])}{\mathbb{E}^k f(\mathcal{Z}[m,n])} \leq 1 + \frac{1}{2n} + \frac{\alpha}{2n \log \mathbb{E}^k f(\mathcal{Z}[m,n])} + \frac{1}{n(\log \mathbb{E}^k f(\mathcal{Z}[m,n]))^{1+\eta}}. \tag{6.7}$$

We can apply (6.4) to bound the denominator in last two terms from below by $n\left(1 \vee \log \frac{n}{m}\right)$ and $n\left(1 \vee \log \frac{n}{m}\right)^{1+\eta}$, respectively, to obtain

$$\frac{\mathbb{E}^k f(\mathcal{Z}[m,n+1])}{\mathbb{E}^k f(\mathcal{Z}[m,n])} \leq 1 + \frac{1}{2n} + \frac{\alpha}{n(1 \vee \log \frac{n}{m})} + \frac{2^{1+\eta}}{n(1 \vee \log \frac{n}{m})^{1+\eta}}. \tag{6.8}$$

Iterating both sides of (6.8) in n then yields

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \leq f(k) \prod_{i=m}^{n-1} \left(1 + \frac{1}{2i} + \frac{\alpha}{i(1 \vee \log \frac{i}{m})} + \frac{2^{1+\eta}}{i(1 \vee \log \frac{i}{m})^{1+\eta}}\right)$$

and using the inequality $1 + x \leq e^x$ we get

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \leq f(k) \exp\left(\sum_{i=m}^{n-1} \frac{1}{2i} + \sum_{i=m}^{n-1} \frac{\alpha}{i(1 \vee \log \frac{i}{m})} + \sum_{i=m}^{n-1} \frac{2^{1+\eta}}{i(1 \vee \log \frac{i}{m})^{1+\eta}}\right),$$

which implies

$$\mathbb{E}^k f(\mathcal{Z}[m,n]) \leq f(k) \exp\left[\frac{1}{2} \sum_{i=m}^{n-1} \frac{1}{i} + \alpha\left(\sum_{i=m}^{\lceil em \rceil - 1} \frac{1}{i} + \sum_{i=\lceil em \rceil}^{n-1} \frac{1}{i \log \frac{i}{m}}\right)\right]$$
$$\times \exp\left(\sum_{i=m}^{\lceil em \rceil - 1} \frac{2^{1+\eta}}{i} + \sum_{i=\lceil em \rceil}^{n-1} \frac{2^{1+\eta}}{i(\log \frac{i}{m})^{1+\eta}}\right). \tag{6.9}$$

The exponential of $\frac{1}{2} \sum_{i=m}^{n-1} \frac{1}{i}$ sum is always less than $C'\sqrt{\frac{n}{m}}$, for some constant C'. For the second expression in the exponential we observe that $\sum_{i=m}^{\lceil em \rceil - 1} \frac{1}{ik} \leq \frac{11}{6k}$ and for some absolute constant C'',

$$\sum_{i=\lceil em \rceil}^{n-1} \frac{1}{i \log \frac{i}{m}} \leq \int_{em}^{n} \frac{\frac{1}{m}}{\frac{s}{m} \log \frac{s}{m}} \, ds + C'' = \int_{e}^{\frac{n}{m}} \frac{1}{x \log x} \, dx + C'' = \log\log \frac{n}{m} + C''.$$

Finally, the second factor on the right hand side of (6.9) is bounded by a constant $D > 0$, due to the convergence of the series $\sum_{i=2}^{\infty} \frac{1}{i(\log i)^{1+\eta}}$. Applying all these estimates to (6.9) we arrive at

$$\mathbb{E}^k \mathcal{Z}[m,n] \leq f(k) e^{\frac{11\alpha}{6k} + C''} C' \sqrt{\frac{n}{m}} D\left(1 \vee \log \frac{n}{m}\right)^{\alpha},$$

proving the desired bound for $C(k, f) = f(k)e^{\frac{11\alpha}{6k} + C''}C'D$.

It remains to deduce the bound for the second moment. We apply the same argument as before, but apply Lemma 4.4 with the function f^2, i.e. we obtain

$$\mathbb{E}^k f(\mathscr{Z}[m, n])^2 = f(k)^2 + \sum_{s=m}^{n-1} \mathbb{E}^k \frac{f(\mathscr{Z}[m, s])\Delta f(\mathscr{Z}[m, s])^2}{s}.$$

Since f is non decreasing, we find that $\Delta f(k)^2 \le f(k+1)2\Delta f(k)$ and thus

$$\mathbb{E}^k f(\mathscr{Z}[m, n])^2 \le f(k)^2 + \sum_{s=m}^{n-1} \mathbb{E}^k \frac{2f(\mathscr{Z}[m, s] + 1)^2 \Delta f(\mathscr{Z}[m, s])}{s} = E(m, n).$$

The function $E(m, n)$ can be bounded in the same fashion as the first moment, we obtain

$$\frac{E(m, n+1)}{E(m, n)} \le 1 + \frac{1}{n} + \frac{2\alpha}{n(1 \vee \log \frac{n}{m})} + \frac{2^{2+\eta}}{n(1 \vee \log \frac{n}{m})^{1+\eta}},$$

which implies, as above, that

$$E(m, n) \le C'(k) \frac{n}{m} \left(\log \frac{n}{m}\right)^{2\alpha},$$

and therefore the second moment bound holds.

$\square$

Finally we can derive the lower bound.

Proof of Proposition 6.2.   First we consider the case $\alpha = 0$, in which the assertion follows directly from Lemma 6.3. Now assume $\alpha > 0$. We focus on the lower bound for $k = 0$ and begin with the observation that the concavity condition implies that

$$\mathbb{E}f(\mathscr{Z}[m, n]) \ge f(0) + \frac{1}{2}\mathbb{E}\mathscr{Z}[m, n]. \tag{6.10}$$

We will use the bounds on ϕ derived in Lemma 6.4, to derive bounds for $\mathbb{E}\mathscr{Z}[m, n]$, based on the fact (see Lemma 4.11) that

$$\phi(\mathscr{Z}[m, n]) = \sum_{i=m}^{n-1} \frac{1}{i} + M_n,$$

where $(M_n)_{n \ge m}$ is a martingale. Clearly,

$$\mathbb{E}\phi(\mathscr{Z}[m, n]) = \sum_{i=m}^{n-1} \frac{1}{i}$$

and, using concavity of ϕ, Jensen's inequality implies that

$$\phi(\mathbb{E}\mathcal{Z}[m,n]) \geq \sum_{i=m}^{n-1} \frac{1}{i},$$

which yields, since ϕ is strictly increasing,

$$\mathbb{E}\mathcal{Z}[m,n] \geq \phi^{-1}\Big(\sum_{i=m}^{n-1} \frac{1}{i}\Big), \tag{6.11}$$

in which, as before, we use the notation ϕ^{-1} to represent the strictly increasing linear interpolation of the inverse of ϕ. We wish to use Lemma 6.4(i) to establish a lower bound for the right hand side, i.e. it is sufficient to show

$$\sum_{i=m}^{n-1} \frac{1}{i} \geq \phi(K(m,n)),$$

for a suitably chosen $K(m,n)$. By Lemma 6.4(ii), we can determine a constant $C_{6.4}$ such that

$$\phi(k) \leq C_{6.4} + 2\log(k-1) - 2\alpha\log\log(k-1),$$

for all $k \in \mathbb{N}$. Let

$$K(m,n) = \Big\lceil \sqrt{\frac{n}{m}} \big(\log \tfrac{n}{m}\big)^{\alpha} \Big\rceil,$$

then

$$
\begin{aligned}
\phi(K(m,n)) &\leq C_{6.4} + \log \tfrac{n}{m} + 2\alpha\log\log \tfrac{n}{m} - 2\alpha\log\big(\tfrac{1}{2}\log \tfrac{n}{m} + \alpha\log\log \tfrac{n}{m}\big) \\
&\leq C_{6.4} + \log \tfrac{n}{m} + 2\alpha\log\log \tfrac{n}{m} - 2\alpha\log\big(\tfrac{1}{2}\log \tfrac{n}{m}\big) \\
&\leq 2\alpha\log 2 + C_{6.4} + \log \tfrac{n}{m} \\
&\leq C' + \sum_{i=m}^{n-1} \tfrac{1}{i},
\end{aligned}
\tag{6.12}
$$

where C' is a constant depending only on f. Thus, (6.12) implies that $\sum_{i=m}^{n-1} \frac{1}{i} \geq \phi(K(m,n))$ and therefore

$$\phi^{-1}\Big(\sum_{i=m}^{n-1} \tfrac{1}{i}\Big) \geq \Big\lceil \sqrt{\frac{n}{m}}\big(\log \tfrac{n}{m}\big)^{\alpha} \Big\rceil \geq \sqrt{\frac{n}{m}}\big(\log \tfrac{n}{m}\big)^{\alpha}.$$

Combining this with (6.10) and (6.11) yields

$$\mathbb{E}f(\mathcal{Z}[m,n]) \geq f(0) + \tfrac{1}{2}\sqrt{\frac{n}{m}}\big(\log \tfrac{n}{m}\big)^{\alpha}, \tag{6.13}$$

for all n, m with $n > me$. Hence the desired bound holds with $c = \frac{f(0)}{e^{\alpha}}$ for all $m > n$. The bound for $k > 0$ follows by stochastic domination. $\qquad\square$

The bounds obtained will now be used to estimate the growth of the score of the exploration.

6.2 Truncated second moments

In this section we study a particular phase of the network exploration scheme in $\mathcal{G}_N$. We assume that we have already determined (up to a small error) in the initial Phase (I) of local exploration around a finite number $V_1, V_2, \ldots, V_j$ of initial vertices, whether these initial vertices are contained in $\mathcal{C}_N$. We therefore start with an initial subset $\Gamma_0 \subset \mathcal{G}_N$, which is a forest with total score $\xi(\Gamma_0) \geq s_0$, and corresponds to the explored part of the largest connected component after local exploration around the typical vertices $V_1, V_2, \ldots, V_j$. By $\mathcal{E}_0$ we denote the The σ-field generated by the exploration during Phase (I). Note that here the parameter s_0 controls the probability that the initial vertices are not in $\mathcal{C}_N$ and is independent of N.

Phase (II) now consists of a truncated breadth first exploration starting from Γ_0, which we will not couple directly to the exploration of a tree as we did in Phase (I). Our aim is, to show that in Phase (II) the accumulated score of the discovered vertices grows in a similar way as the score in the truncated IBRWs in Section 4.5, provided the exploration is not stopped according to certain criteria we will define below. We start by setting up the necessary notation.

Recall the truncation sequence r we obtained in Section 4.5 via the equation

$$
\begin{aligned}
r_0 &= s_0 \\
-\frac{1}{2}\Delta r_k &= (1+\alpha)\log(-r_k) + \frac{1}{2}\log\frac{r_{k+1}}{r_k} - 2\log\frac{k+4}{k+3} + \log\frac{c}{\sqrt{1+\alpha}}, \quad \text{for } k \geq 0,
\end{aligned}
\tag{6.14}
$$

This equation has a unique solution r for all sufficiently large initial values, as is shown in Appendix A. For the discrete model, we are interested in the π_N-projection of r, which we will denote (slightly abusing notation) also by r.

Definition 6.5. Let
$$
r = r(s_0, N) = (r_k)_{k \in \mathbb{N}_0} \subset [N]
$$
denote the π_N-projection of the solution to (6.14) with initial condition $-s_0$. We are going to use r to control the first moment of the total score of the neighbourhood shells around a given set. Define inductively the (truncated) *k-th neighbourhood shells* of $\Gamma_0 \subset [N]$ for $k \geq 1$ by
$$
\Gamma^k = \{i \in [N] : d_N(\Gamma^{k-1}, i) = 1\} \setminus \{1, \ldots, r_k - 1\},
$$
the (truncated) *k-neighbourhood* by
$$
H^k = \bigcup_{i=0}^{k} \Gamma^k,
$$

and the *undiscovered vertices before exploration stage* $k \in \mathbb{N}$ by

$$U^k = \{r_k, \dots, N\} \setminus H^{k-1}.$$

For $m \in [N]$ and $k \geq 1$, we set furthermore

$$\Gamma[m] = \{n \in [N] : d_N(m, n) = 1\} \text{ and } \Gamma_k[m] = \Gamma[m] \cap U^k$$

and note that $\Gamma_k[m]$ is the disjoint union of

$$\Gamma_k^<[m] = \Gamma_k[m] \cap \{1, \dots, m-1\} \text{ and } \Gamma_k^>[m] = \Gamma_k[m] \cap \{m+1, \dots, N\}.$$

The filtration generated by the exploration procedure is denoted by

$$\mathcal{E}_0 = \sigma(\Gamma_0) \text{ and } \mathcal{E}_{k+1} = \sigma(\mathcal{E}_k \cup \{\sigma(\mathbb{1}_{m \mapsto n}), m \in \Gamma_k, n \in U_k\}), \text{ for } k \in \mathbb{N}.$$

Remark 6.6. r_k is roughly equal to $Ne^{-(2\alpha+2)k \log k}$, if $N \gg k \gg 0$, as can be seen from the calculations in Appendix A. However, due to the projection, we have $r_k = 1$ for all sufficiently large k.

In Phase (II), we therefore uncover in every step a whole neighbourhood shell, which is in contrast to Phase (I) where each step corresponded to the exploration of a single vertex. We also expect to encounter some circles in the network during Phase (II). We count special types of circles, in which during one exploration step multiple edges to the same vertex are discovered. Showing that the number of these circles is small, enables us to deduce that the score of the network exploration behaves almost like the score in a tree.

Definition 6.7. Let $H \subset H^{k-1}$ and define, for $i \geq 2$,

$$Y_k^i(H) = \#\{m \in U^k : \exists \text{ distinct } m_1, \dots, m_i \in H \text{ with } \Delta \mathcal{Z}[m, m_s] = 1 \, \forall 1 \leq s \leq i\},$$

the number of *i-stars induced by H* at stage k.

Remark 6.8. The choice of ℓ_0 is motivated by our estimates below, which show that i-stars almost never occur in the network for $i \geq 4$.

Additionally, we introduce the notion of thin sets, which will be useful in describing the extent of the explored part if the graph in terms of the score.

Definition 6.9. Set
$$\ell_0 = \ell_0(f) = \min \left\{ \ell \in \mathbb{N} : \frac{f(4)}{\ell^{\Delta f(4)}} < 1 \right\}.$$

For $N \in \mathbb{N}$, we call disjoint pairs of vertex sets $I_0, I_1 \subset \{\ell_0, \dots, N\}$ *thin* if

$$\xi(I_0)\xi(I_1) \leq \frac{1}{2} \frac{N}{\psi^4(\ell_0, N)}.$$

Lemma 6.10. *Let α be as in* (6.1), *then*

$$\liminf_{n\to\infty} \frac{(\log n)^\alpha}{\psi^4(\ell_0, n)} = \chi > 0.$$

Proof. This is a direct consequence of Proposition 6.1 □

The expected number of i-stars induced by a subset H of vertices can be bounded in terms of the score $\xi(H)$. In turn, given that the exploration has not yet uncovered a 4-star, we can bound the probability, that the explored set H during stage k of the exploration (i.e. $H \subset H^k$) paired with the set $\{r_{k+1}\}$ ceases to be thin. Thus we define the stopping criteria for the exploration in terms of thin sets and i-stars.

Definition 6.11. We introduce two N-dependent stopping times

$$S = \min\{k \in \mathbb{N} : \text{a 4-star is discovered during exploration step } k\},$$

and

$$T = \min\left\{ k \in \mathbb{N} : \xi(H^k) > s_0 \frac{\sqrt{N}}{\psi^4(\ell_0, N)} \right\} \wedge k^*,$$

where

$$k^* = \min\left\{ k : \xi(r_{k+1}) > \frac{1}{2s_0} \frac{\sqrt{N}}{\psi^4(\ell_0, N)} \right\}.$$

Recalling that $r_k \approx N e^{-(2\alpha+2)k\log k}$, we can conclude that $k^* \leq \left(\frac{1}{2\alpha+2} + \varepsilon\right)\frac{\log N}{\log\log N}$ if N is sufficiently large. Also observe, that the stopping time T basically marks the stage at which $(\{r_{T+1}\}, H^T)$ ceases to be thin. We can establish upper bounds on the jump probabilities of the degree evolutions of the unexplored vertices at a given stage $k < S \wedge T$ of the exploration.

Lemma 6.12 (Modification of [DM13, Lemma 2.12.]). *Fix $N \in \mathbb{N}$ and let $(\{m\}, I_0)$ be thin. Let furthermore $I_1 \subset \{m, \ldots, N\}$ be disjoint of I_0 with $\#I_1 \leq 4$. Set*

$$E_i = \{\Delta\mathcal{Z}[m, j] = \delta_{1i} \text{ for all } j \in I_i\}, \quad \text{for } i \in \{0, 1\}. \tag{6.15}$$

Then there is a constant $C > 0$, only dependent on f such that

$$\mathbb{P}\{\Delta\mathcal{Z}[m, n] = 1 | E_1, E_0\} \leq C\mathbb{P}\{\Delta\mathcal{Z}[m, n] = 1 | E_1\}, \quad \text{for all } n \in \{m, \ldots, N-1\}.$$

Proof. We have

$$\mathbb{P}\{\Delta\mathcal{Z}[m, n] = 1 | E_0, E_1\} \leq \frac{\mathbb{P}\{\Delta\mathcal{Z}[m, n] = 1 | E_1\}}{\mathbb{P}\{E_0 | E_1\}},$$

so it is sufficient to bound the denominator $\mathbb{P}\{E_0 | E_1\}$ uniformly from below. Since $\#I_1 \leq 4$, we get by Lemma 4.14,

$$\mathbb{P}\{\Delta\mathcal{Z}[m, n] = 0 \,\forall n \in I_0 | \Delta\mathcal{Z}[m, n] = 1 \,\forall n \in I_1\} \geq \mathbb{P}^4\{\Delta\mathcal{Z}[m, n] = 0 \,\forall n \in I_0\}. \tag{6.16}$$

Denoting $i = \min I_0$, we obtain

$$\mathbb{P}^4\{\Delta\mathcal{Z}[m,n] = 0 \; \forall n \in I_0\} = \mathbb{P}^4\{\Delta\mathcal{Z}[m,n] = 0 \; \forall n \in I_0 \setminus \{i\} | \Delta\mathcal{Z}[m,i] = 0\}$$
$$\times \mathbb{P}^4\{\Delta\mathcal{Z}[m,i] = 0\}$$
$$\geq \mathbb{P}^4\{\Delta\mathcal{Z}[m,n] = 0 \; \forall n \in I_0 \setminus \{i\}\}\mathbb{P}^4\{\Delta\mathcal{Z}[m,i] = 0\},$$

using Lemma 4.13. Iteration yields

$$\mathbb{P}^4\{\Delta\mathcal{Z}[m,n] = 0 \; \forall n \in I_0\} \geq \prod_{n \in I_0} \mathbb{P}^4\{\Delta\mathcal{Z}[m,n] = 0\}$$
$$= \prod_{n \in I_0} \left(1 - \frac{\mathbb{E}^4 f(\mathcal{Z}[m,n])}{n}\right), \tag{6.17}$$

and inserting (6.17) into (6.16) yields

$$\mathbb{P}\{\Delta\mathcal{Z}[m,n] = 0 \; \forall n \in I_0 | \Delta\mathcal{Z}[m,n] = 1 \; \forall n \in I_1\} \geq \prod_{n \in I_0} \left(1 - \frac{\mathbb{E}^4 f(\mathcal{Z}[m,n])}{n}\right). \tag{6.18}$$

By Lemma 4.9 and the definition of ℓ_0,

$$\frac{\mathbb{E}^4 f(\mathcal{Z}[m,n])}{n} \leq f(4) n^{\Delta f(4)-1} m^{-\Delta f(4)} \leq \frac{f(4)}{\ell_0^{\Delta f(4)}} < 1,$$

hence there is a constant $c = c(f) > 1$ such that for all $\ell_0 \leq m \leq n$,

$$-\log\left(1 - \frac{\mathbb{E}^4 f(\mathcal{Z}[m,n])}{n}\right) \leq \frac{\mathbb{E}^4 f(\mathcal{Z}[m,n])}{n} + c\left(\frac{\mathbb{E}^4 f(\mathcal{Z}[m,n])}{n}\right)^2 < 2c\frac{\mathbb{E}^4 f(\mathcal{Z}[m,n])}{n}.$$

Thus, taking the logarithm in (6.18) and using Observation 4.11, we can bound, using the monotonicity properties of ψ and Lemma 4.6,

$$-\log\mathbb{P}\{E_0 | E_1\} \leq 2c \sum_{n \in I_0} \frac{\mathbb{E}^4 f(\mathcal{Z}[m,n])}{n} = 2c \sum_{n \in I_0} \frac{\psi^4(m,n)\xi(m)}{n\xi(n)}$$
$$\leq 2c\frac{\xi(m)}{N} \sum_{n \in I_0} \xi(n)\psi^4(\ell_0,n) \leq 2c\frac{\xi(m)\psi^4(m,N)\xi(I_0)}{N},$$

and the last expression is uniformly bounded by c since (m, I_0) is thin. $\qquad\square$

The next step is to derive bounds on the expected number of i-stars, given that S and T have not occurred yet.

Lemma 6.13. *Let $y_k^i(H) = \mathbb{E}[Y_k^i(H)|\mathscr{E}_k]$, for $i \geq 2$ and $H \subset H^{k-1}$, then there is a constant $C(i) > 0$ such that*

$$y_k^i(H) \leq C(i)\psi^0(r_k,N)^i N^{-\frac{i}{2}}\xi(H)^i r_k^{1-\frac{i}{2}},$$

conditional on $T \wedge S > k - 1$.

Proof. Let $k-1 < T \wedge S$, $m \in U^k$ and let I_0, I_1 denote, as in Lemma 6.12, sets of jump or non-jump times of $\mathcal{Z}[m, \cdot]$ representing already explored vertices. Let also, similarly to Lemma 6.12, $E_0, E_1 \in \mathcal{E}_{k-1}$ denote the corresponding events. Denote by

$$J = \{n_1, \ldots, n_i\} \subset \{m, \ldots, N\} \setminus (I_0 \cup I_1)$$

the potential jump times of $\mathcal{Z}[m, \cdot]$, then we can use the independence of the degree evolutions and Lemma 6.12 to obtain

$$\mathbb{P}\{\Delta\mathcal{Z}[m, n_l] = 1 \ \forall l \in J | \mathcal{E}_{k-1}\} = \mathbb{P}\{\Delta\mathcal{Z}[m, n_l] = 1 \ \forall l \in J | E_0, E_1, T \wedge S \geq k\}$$
$$\leq C_{6.12}\mathbb{P}^4\{\Delta\mathcal{Z}[m, n_l] = 1 \ \forall l \in J\}.$$

Proposition 6.1 now yields in conjunction with Lemma 4.6

$$\mathbb{P}^4\{\Delta\mathcal{Z}[m, n_l] = 1 \ \forall l \in J\} \leq C_{6.1}(4) \prod_{l=1}^{i} \frac{\psi^{l-1}(m, n_l - 1)\xi(m, n_l - 1)}{n_l - 1}$$
$$\leq C_{6.1}(4)C^i \log\left(\tfrac{N}{r_k}\right)^{\alpha i}(Nm)^{-\frac{i}{2}} \prod_{l=1}^{i} \xi(n_l),$$

for some constant $C > 0$ and we conclude

$$\mathbb{P}\{\Delta\mathcal{Z}[m, n_l] = 1 \ \forall l \in J | \mathcal{E}_{k-1}\} \leq C_{6.12}C_{6.1}(4)C^i \log\left(\tfrac{N}{r_k}\right)^{\alpha i}(Nm)^{-\frac{i}{2}} \prod_{l=1}^{i} \xi(n_l). \tag{6.19}$$

Summing over all combinations $n_1 < \cdots < n_l \in H \subset H^{k-1}$ in (6.19) yields

$$\sum_{n_1 < \cdots < n_l \in H} \mathbb{P}\{\Delta\mathcal{Z}[m, n_l] = 1 \ \forall l \in J | \mathcal{E}_{k-1}\}$$
$$\leq C_{6.12}C_{6.1}(4)C^i \log\left(\tfrac{N}{r_k}\right)^{\alpha i}(Nm)^{-\frac{i}{2}} \sum_{n_1 < \cdots < n_l \in H} \prod_{l=1}^{i} \xi(n_l)$$
$$\leq C_{6.12}C_{6.1}(4)C^i \left(\log\left(\tfrac{N}{r_k}\right)\right)^{\alpha i}(Nm)^{-\frac{i}{2}}\xi(H)^i,$$

and summation over $m = r_k, \ldots, N$ yields the existence of another constant $C'(i) > 0$ with

$$y_k^i(H) \leq C'(i)\left(\log\left(\tfrac{N}{r_k}\right)\right)^{\alpha i} N^{-\frac{i}{2}}\xi(H)^i r_k^{1-\frac{i}{2}}.$$

$\square$

Remark 6.14. The bound implies that

$$y_k^2(H) = O\left(\frac{\psi^1(r_k, N)^2 \xi(H)^2}{N}\right)$$

and that, for $i > 2$,

$$y_k^i(H) = O\Big(\frac{r_k^{1-\frac{i}{2}}\psi^{i-1}(r_k,N)^i \xi(H)^i}{N^{\frac{i}{2}}}\Big),$$

suggesting that for $i = 2$ and $\xi(H)$ around $\sqrt{N}$ the number of circles discovered does not vanish, which illustrates why we do not expect to be able to uphold the tree coupling of the local exploration.

We can now use Lemma 6.13 to show that stars are rare as long as the score stays low.

Lemma 6.15. *With high probability,*

$$T \wedge S = T.$$

Proof. Assume T has not occurred by time $k-1$, then, by Lemma 6.13

$$\mathbb{P}\{Y_k^{(i)}(H^{k-1}) \geq 1|\mathscr{E}_{k-1}\} \leq \mathbb{E}[Y_k^{(i)}(H^{k-1})|\mathscr{E}_{k-1}]$$
$$\leq C_{6.13}(i) r_k^{1-\frac{i}{2}}\psi^0(r_k,N)^i \xi(H^{k-1})^i N^{-\frac{i}{2}},$$

and $\xi(H^{k-1}) \leq s_0 \frac{sqrtN}{\psi^4(\ell_0,N)}$ implies that

$$\mathbb{P}\{Y_k^{(i)}(H^{k-1}) \geq 1|\mathscr{E}_{k-1}\} \leq C_{6.13}(i) s_0^i r_k^{1-\frac{i}{2}} \frac{\psi^0(r_k,N)^i}{\psi^4(\ell_0,N)^i}.$$

Choosing $i = 4$ yields

$$\mathbb{P}\{Y_k^{(4)}(H^{k-1}) \geq 1|\mathscr{E}_{k-1}\} \leq C r_k^{-1} \frac{\psi^0(r_k,N)^4}{\psi^4(\ell_0,N)^4} \leq C' r_k^{-1}$$

for suitably chosen constants $C, C' > 0$. Summation over k and recalling that r_k decays and $k \leq T \leq k^* = O\big(\frac{\log N}{\log\log N}\big)$ gives

$$\mathbb{P}\{S \leq T\} \leq \mathbb{E}[\sum_{k=1}^{k^*}\mathbb{P}\{S = k|\mathscr{E}_{k-1}\}] \leq C'\sum_{k=1}^{k^*} r_k^{-1} \leq C''\frac{\log N}{r_{k^*}\log\log N}.$$

Observe that $\xi(r_k*) \leq \frac{1}{2s_0\frac{\sqrt{N}}{\psi^4(\ell_0,N)}}$ implies $r_{k^*} \geq c\log N$ for some $c > 0$. Thus $\mathbb{P}\{S \leq T\}$ vanishes as $N \to \infty$.

$\square$

We can now start to estimate first and second moment of $S_k = \xi(\Gamma^k)$, $k = 1,\dots,T-1$ assuming that S has not yet occurred. Before we state our results, we need one more auxiliary correlation bound for the conditioned degree evolutions. Analogously to Lemma 6.12, we wish to use also a lower bound for the jump probabilities of the conditioned process.

Lemma 6.16 (Lower bound for conditioned jump probabilities)**.** *Fix $N \in \mathbb{N}$, let $(\{m\}, I_0)$ be*

thin and let $I_1 \subset \{m, \ldots, N\}$ be disjoint of I_0 with $\#I_1 \leq 3$. Then, for E_0, E_1 as in (6.15),

$$\mathbb{P}\{\Delta\mathcal{Z}[m,n] = 1 | E_0, E_1\} \geq \frac{1}{2}\mathbb{P}\{\Delta\mathcal{Z}[m,n] = 1 | E_1\} \quad \text{for all } n \in \{m, \ldots, N-1\} \setminus I_0.$$

Proof. Let $n \in \{m, \ldots, N-1\} \setminus I_0$. We have

$$
\begin{aligned}
\mathbb{P}\{\Delta\mathcal{Z}[m,n] = 1, \Delta\mathcal{Z}[m,k] = 0 \ \forall k \in I_0 | E_1\} &= \mathbb{P}\{\Delta\mathcal{Z}[m,n] = 1 | E_1\} \\
&\quad - \mathbb{P}\{\Delta\mathcal{Z}[m,n] = 1, \ \exists k \in I_0 : \Delta\mathcal{Z}[m,k] = 1 | E_1\} \\
&\geq \mathbb{P}\{\Delta\mathcal{Z}[m,n] = 1 | E_1\} - \sum_{k \in I_0} \mathbb{P}\{\Delta\mathcal{Z}[m,n] = \Delta\mathcal{Z}[m,k] = 1 | E_1\}.
\end{aligned}
\tag{6.20}
$$

The last sum can be rewritten

$$
\begin{aligned}
\sum_{k \in I_0} \mathbb{P}\{\Delta\mathcal{Z}[m,n] =& \Delta\mathcal{Z}[m,k] = 1 | E_1\} \\
&= \mathbb{P}\{\Delta\mathcal{Z}[m, n = 1 | E_1]\} \sum_{k \in I_0} \mathbb{P}\{\Delta\mathcal{Z}[m,k] = 1 | E_1, \Delta\mathcal{Z}[m,n] = 1\}.
\end{aligned}
\tag{6.21}
$$

Applying Lemmas 4.14 and 4.6 yields, for all $k \in I_0$,

$$
\begin{aligned}
\mathbb{P}\{\Delta\mathcal{Z}[m,k] = 1 | E_1, \Delta\mathcal{Z}[m,n] = 1\} &\leq \mathbb{P}^4\{\Delta\mathcal{Z}[m,k] = 1\} \\
&= \frac{\mathbb{E}^4 f(\mathcal{Z}[m,k])}{k} = \frac{\psi^4(m,k)\xi(m,k)}{k} \\
&\leq \frac{\psi^4(\ell_0, N)\xi(m,k)}{k} = \frac{\psi^4(\ell_0, N)\xi(m)\xi(k)}{N}.
\end{aligned}
\tag{6.22}
$$

Inserting (6.22) into (6.21) in combination with (6.20) yields

$$\mathbb{P}\{\Delta\mathcal{Z}[m,n] = 1, \Delta\mathcal{Z}[m,k] = 0 \ \forall k \in I_0 | E_1\} \geq \mathbb{P}\{\Delta\mathcal{Z}[m, n = 1 | E_1]\}\left(1 - \frac{\xi(m)\xi(I_0)}{N}\right),$$

and using that $(\{m\}, I_0)$ are thin yields the statement. □

We are now ready to state our bounds for the expected growth of the score.

Proposition 6.17. *Let f satisfy the assumptions of Theorem II. There are constants $c, C > 0$, depending only on f, and $N_0 \in \mathbb{N}$, such that for $S_k = \xi(\Gamma^k)$*

$$\mathbb{E}[S_{k+1} | \mathcal{E}_k] \geq c\left(\log\frac{N}{r_{k+1}}\right)^{1+\alpha} S_k, \tag{6.23}$$

and

$$\mathbb{E}[S_{k+1} | \mathcal{E}_k] \leq C\left(\log\frac{N}{r_{k+1}}\right)^{1+\alpha} S_k, \tag{6.24}$$

almost surely for all $N \geq N_0$ and $k \leq T \wedge S - 1$.

Proof. Observe that

$$\mathbb{E}[S_{k+1}|\mathscr{E}_k] \geq \mathbb{E}\Big[\sum_{u \in \Gamma^k} \big(\xi(\Gamma^<_{k+1}[u]) + \xi(\Gamma^>_{k+1}[u])\big)|\mathscr{E}_k\Big]$$
$$- \mathbb{E}\Big[\sum_{u \in \Gamma^k} \sum_{\substack{t \in \Gamma^k: \\ t < u}} \big(\xi(\Gamma^<_{k+1}[u] \cap \Gamma[t]) + \xi(\Gamma^>_{k+1}[u] \cap \Gamma[t])\big)|\mathscr{E}_k\Big], \quad (6.25)$$

where the second term accounts for double counting due to circles in the network. If the graph was a tree, the second term could be omitted (which yields an upper bound, see below). A lower bound for the second term in (6.25) is obtained by estimating

$$\mathbb{E}\Big[\sum_{u \in \Gamma^k} \sum_{\substack{t \in \Gamma^k: \\ t < u}} \big(\xi(\Gamma^<_{k+1}[u] \cap \Gamma[t]) + \xi(\Gamma^>_{k+1}[u] \cap \Gamma[t])\big)|\mathscr{E}_k\Big] \quad (6.26)$$

$$\leq \mathbb{E}\Big[\sum_{\substack{u,t \in \Gamma^k: \\ t < u}} \xi(\Gamma^>_{k+1}[u] \cap \Gamma^>_{k+1}[t])|\mathscr{E}_k\Big] + \mathbb{E}\Big[\sum_{\substack{u,t \in \Gamma^k: \\ t < u}} \xi(\Gamma^<_{k+1}[u] \cap \Gamma^>_{k+1}[t])|\mathscr{E}_k\Big] \quad (6.27)$$

$$+ \mathbb{E}\Big[\sum_{\substack{u,t \in \Gamma^k: \\ t < u}} \xi(\Gamma^<_{k+1}[u] \cap \Gamma^<_{k+1}[t])|\mathscr{E}_k\Big]. \quad (6.28)$$

The two first terms (6.27) are easier to bound, we have

$$\mathbb{E}[\xi(\Gamma^>_{k+1}[t] \cap \Gamma^<_{k+1}[u])|\mathscr{E}_k] \leq \sum_{i=t+1}^{u-1} \xi(i)\mathbb{P}\{\Delta\mathscr{Z}[t,i] = \Delta\mathscr{Z}[i,u] = 1|\mathscr{E}_k\}$$
$$= \sum_{i=t+1}^{u-1} \xi(i)\mathbb{P}\{\Delta\mathscr{Z}[t,i] = 1|\mathscr{E}_k\}\mathbb{P}\{\Delta\mathscr{Z}[i,u] = 1|\mathscr{E}_k\}, \quad (6.29)$$

using the independence of indegree evolutions, and similarly

$$\mathbb{E}[\xi(\Gamma^>_{k+1}[u] \cap \Gamma^>_{k+1}[t])|\mathscr{E}_k] \leq \sum_{i=u}^{N} \xi(i)\mathbb{P}\{\Delta\mathscr{Z}[t,i] = \Delta\mathscr{Z}[u,i] = 1|\mathscr{E}_k\}$$
$$= \sum_{i=u}^{N} \xi(i)\mathbb{P}\{\Delta\mathscr{Z}[t,i] = 1|\mathscr{E}_k\}\mathbb{P}\{\Delta\mathscr{Z}[u,i] = 1|\mathscr{E}_k\}. \quad (6.30)$$

Only the last term (6.28) contains dependencies, since there are jumps of the same indegree evolution involved, we obtain

$$\mathbb{E}[\xi(\Gamma^<_{k+1}[u] \cap \Gamma^<_{k+1}[t])|\mathscr{E}_k] \leq \sum_{i=r_{k+1}}^{t} \xi(i)\mathbb{P}\{\Delta\mathscr{Z}[i,t] = \Delta\mathscr{Z}[i,u] = 1|\mathscr{E}_k\} \quad (6.31)$$

In the expression (6.31) we need to bound, for active $v \in \Gamma^k$ and unexplored $x_1, x_2 \in U^k$, with

$v < x_1 < x_2$,

$$\mathbb{P}\{\Delta\mathcal{Z}[v,x_1] = \Delta\mathcal{Z}[v,x_2] = 1|\mathscr{E}_k\} = \mathbb{P}\{\Delta\mathcal{Z}[v,x_2] = 1|\Delta\mathcal{Z}[v,x_1] = 1,\mathscr{E}_k\}$$
$$\times \mathbb{P}\{\Delta\mathcal{Z}[v,x_1] = 1|\mathscr{E}_k\}$$
$$\leq C_{6.12}^2 \mathbb{P}^4\{\Delta\mathcal{Z}[v,x_2] = 1\}\mathbb{P}^3\{\Delta\mathcal{Z}[v,x_1] = 1\} \tag{6.32}$$
$$= \frac{C_{6.12}^2 \mathbb{E}^4 f(\mathcal{Z}[v,x_2])\mathbb{E}^3 f(\mathcal{Z}[v,x_1])}{x_2 x_1},$$

where we have used first Lemma 6.12, noting that before time S, no explored vertex has more than 3 right neighbours and that the score of H^k is low, and then Lemma 4.14. Applying (6.32) to (6.31) and using that, for all $i \in \Gamma^k, j \in [N]$ and $I \in \{3,4\}$,

$$\psi^I(i,j) \leq \psi^I(r_k, N) =: \overline{\psi}_k^I(N)$$

now yields a bound for the term (6.26),

$$\mathbb{E}\Big[\sum_{u \in \Gamma^k} \sum_{\substack{t \in \Gamma^k: \\ t < u}} \big(\xi(\Gamma_{k+1}^<[u] \cap \Gamma[t]) + \xi(\Gamma_{k+1}^>[u] \cap \Gamma[t])\big) \Big| \mathscr{E}_k \Big]$$
$$\leq \sum_{u \in \Gamma^k} \sum_{\substack{t \in \Gamma^k: \\ t < u}} C_{6.12}^2 \overline{\psi}_k^3(N)\Big(\overline{\psi}_k^4(N)\sum_{i=r_{k+1}}^{t} \frac{\xi(i,N)\xi(i,u)\xi(i,t)}{ut}$$
$$+ \overline{\psi}_k^3(N)\sum_{i=t+1}^{u-1} \frac{\xi(i,N)\xi(t,i)\xi(i,u)}{iu} + \overline{\psi}_k^3(N)\sum_{i=u}^{N} \frac{\xi(i,N)\xi(t,i)\xi(u,i)}{i^2}\Big).$$

We recall that $\xi(m,n) \leq C_{4.6}\sqrt{\frac{n}{m}}$ and hence there is $B > 0$, such that

$$\mathbb{E}\Big[\sum_{u \in \Gamma^k} \sum_{\substack{t \in \Gamma^k: \\ t < u}} \big(\xi(\Gamma_{k+1}^<[u] \cap \Gamma[t]) + \xi(\Gamma_{k+1}^>[u] \cap \Gamma[t])\big) \Big| \mathscr{E}_k \Big]$$
$$\leq \sum_{u \in \Gamma^k} \sum_{\substack{t \in \Gamma^k: \\ t < u}} B\overline{\psi}_k^3(N)\Big(\overline{\psi}_k^4(N)\sum_{i=r_{k+1}}^{t} \frac{\xi(i,N)}{i\sqrt{ut}}$$
$$+ \overline{\psi}_k^3(N)\sum_{i=t+1}^{u-1} \frac{\xi(i,N)}{i\sqrt{ut}} + \overline{\psi}_k^3(N)\sum_{i=u}^{N} \frac{\xi(i,N)}{i\sqrt{ut}}\Big)$$
$$\leq B^2 \overline{\psi}_k^3(N)\overline{\psi}_k^4(N)\sum_{u \in \Gamma^k}\sum_{t \in \Gamma^k}\sqrt{\frac{N}{ut}}\sum_{i=r_{k+1}}^{N} i^{-\frac{3}{2}}$$
$$= \frac{B^2 \overline{\psi}_k^3(N)}{N}\overline{\psi}_k^4(N)\Big(\sum_{u \in \Gamma^k}\sqrt{\frac{N}{u}}\Big)\Big(\sum_{t \in \Gamma^k}\sqrt{\frac{N}{t}}\Big)\Big(\sum_{i=r_{k+1}}^{N}\frac{\sqrt{N}}{i^{\frac{3}{2}}}\Big),$$

on which we can use the bounds in the definition of T for S_k and $\xi(r_{k+1})$, to obtain, for some

appropriately chosen constants $C', C > 0$,

$$\mathbb{E}\Big[\sum_{u \in \Gamma^k} \sum_{\substack{t \in \Gamma^k: \\ t < u}} \xi(\Gamma_{k+1}[u] \cap \Gamma[t]) \, \Big| \, \mathscr{E}_k \Big] \leq \frac{C' \overline{\psi}_k^3(N) \psi_k^4(N)}{N} \overline{\xi}(\Gamma^k) \xi(r_{k+1}) S_k \leq C S_k. \tag{6.33}$$

It remains to bound the first term on the right hand side of (6.25). We estimate now, using Lemma 6.16 and the definition of ξ while setting $\psi = \psi^0$, for $u \in \Gamma^k$,

$$\mathbb{E}[\xi(\Gamma_{k+1}[u])|\mathscr{E}_k] = \sum_{j=r_{k+1}}^{r_k - 1} \xi(j) \mathbb{P}\{\Delta \mathscr{Z}[j, u] = 1\} + \sum_{\substack{j \in U^k \\ r_k \leq j < u}} \xi(j) \mathbb{P}\{\Delta \mathscr{Z}[j, u] = 1 | \mathscr{E}_k\}$$

$$+ \sum_{\substack{j \in U^k \\ u < j \leq N}} \xi(j) \mathbb{P}\{\Delta \mathscr{Z}[u, j] = 1 | \mathscr{E}_k\}$$

$$\geq c_{6.16}\Big(\sum_{\substack{j \in U^k \\ r_{k+1} \leq j < u}} \xi(j) \frac{\xi(j) \psi(j, u)}{\xi(u) u} + \sum_{\substack{j \in U^k \\ u < j \leq N}} \xi(j) \frac{\xi(u) \psi(u, j)}{\xi(j) j} \Big)$$

$$\geq c' \xi(u)\Big(\sum_{\substack{j \in U^k \\ r_{k+1} \leq j < u}} \xi^2(j) \frac{\psi(j, u)}{N} + \sum_{\substack{j \in U^k \\ u < j \leq N}} \frac{\psi(u, j)}{j} \Big)$$

for a small constant $c' > 0$. Set now

$$\underline{\psi}(j) = \underline{\psi}_{k,u}(j) = \begin{cases} \frac{1}{j}\big(1 \vee \log \frac{u \vee j}{u \wedge j}\big)^\alpha & \text{for } j \in \{r_{k+1}, \ldots, N\} \\ 0, & \text{otherwise,} \end{cases}$$

and let $\underline{\psi}(A) = \sum_{a \in A} \underline{\psi}_{k,u}(a)$ for $A \subset [N]$. The form of $\underline{\psi}$ is due to Proposition 6.2. Thus we obtain, for some constant $d = d(f) > 0$, which does not depend on u, k or N,

$$\mathbb{E}[\xi(\Gamma_{k+1}[u])|\mathscr{E}_k] \geq d\xi(u)\underline{\psi}(U^k). \tag{6.34}$$

Since $\underline{\psi}$ is non-negative, we can bound the latter term from below by using

$$\underline{\psi}(U^k) \geq \underline{\psi}([N]) - \underline{\psi}(H^k).$$

To this end calculate

$$\underline{\psi}([N]) = \sum_{j=r_{k+1}}^{\lfloor \frac{u}{e} \rfloor} \frac{1}{j}\big(\log \frac{u}{j}\big)^\alpha + \sum_{\lceil \frac{u}{e} \rceil}^{\lfloor eu \rfloor} \frac{1}{j} + \sum_{j=\lceil eu \rceil}^{N} \frac{1}{j}\big(\log \frac{j}{u}\big)^\alpha$$

$$\geq 2 + c'' + \frac{1}{\alpha + 1}\Big(\log \frac{u}{r_{k+1}}\Big)^{\alpha+1} + \big(\log \frac{N}{u}\big)^{\alpha+1}\Big)$$

$$\geq 2 + c'' + \frac{1}{(\alpha + 1)2^\alpha}\big(\log \frac{N}{r_{k+1}}\big)^{\alpha+1},$$

where c'' is an approximation error and the last inequality follows from convexity of $x \mapsto$

$x^{\alpha+1}$. Thus (6.34) implies

$$\mathbb{E}[\xi(\Gamma_{k+1}[u])|\mathscr{E}_k] \geq d\xi(u)\log\tfrac{N}{r_{k+1}})^{\alpha+1}\Big(\frac{1}{(\alpha+1)2^{\alpha}} - \frac{\psi(H^k)}{\big(\log\tfrac{N}{r_{k+1}}\big)^{\alpha+1}}\Big). \tag{6.35}$$

Note furthermore, that before T has occurred, for all k,

$$\sum_{j\in H^k} \frac{1}{j}\big(1\vee\log\tfrac{u\vee j}{u\wedge j}\big)^{\alpha} \leq \big(\log\tfrac{N}{r_k}\big)^{\alpha}\sum_{j\in H^k}\frac{1}{j} \leq \big(\log\tfrac{N}{r_k}\big)^{\alpha+1},$$

and hence (6.35) yields

$$\mathbb{E}[\xi(\Gamma_{k+1}[u])|\mathscr{E}_k] \geq e\xi(u)\big(\log\tfrac{N}{r_{k+1}}\big)^{\alpha+1}, \tag{6.36}$$

for a suitably chosen constant $e > 0$. Applying (6.36) and (6.33) in (6.25), we conclude

$$\mathbb{E}[S_{k+1}|\mathscr{E}_k] \geq \big(e\log\big(\tfrac{N}{r_{k+1}}\big)^{\alpha+1} - C_{(6.33)}\big)S_k, \tag{6.37}$$

and the proof of the lower bound is finished by adjusting the constants (recall that r_0 can be made small by adjusting s_0).

For the upper bound a similar, slightly more straightforward calculation shows that also for $u\in\Gamma^k$

$$\mathbb{E}[\xi(\Gamma_{k+1}[u])\mathscr{E}_k] \leq D\psi^J(r_{k+1},N)\xi(u)\log\tfrac{N}{r_{k+1}}$$

under the conditions of the proposition, for some constant D. Thus applying Proposition 6.1 yields

$$\mathbb{E}[S_{k+1}|\mathscr{E}_k] \leq C\big(\log\tfrac{N}{r_{k+1}}\big)^{1+\alpha}S_k,$$

for some $C = C(f)$ which finishes the proof. $\qquad\square$

To prove concentration of the score around its mean, we also need to bound the variance of the score growth.

Proposition 6.18. *Let f satisfy the assumptions of Theorem II and r be defined via 6.14. Then there are constants $C, C' > 0$ such that*

$$\mathbb{E}[S_{k+1}^2|\mathscr{E}_k] \leq \mathbb{E}[S_{k+1}|\mathscr{E}_k]^2 + \big(\log\tfrac{N}{r_{k+1}}\big)^{2\alpha}S_k^2 + C'\sum_{u\in\Gamma^k}\xi(u)^2\big(\log\tfrac{N}{u}\big)^{2\alpha+2},$$

almost surely for all $k < T\wedge S$.

Proof. Note that, for $k \geq 0$,

$$
\begin{aligned}
S_{k+1}^2 &\leq \left(\sum_{u \in \Gamma^k} \xi(\Gamma_{k+1}[u]) \right)^2 = \sum_{u,v \in \Gamma^k} \xi(\Gamma_{k+1}[u]) \xi(\Gamma_{k+1}[v]) \\
&= \sum_{u,v \in \Gamma^k} (\xi(\Gamma_{k+1}^<[u]) + \xi(\Gamma_{k+1}^>[u]))(\xi(\Gamma_{k+1}^<[v]) + \xi(\Gamma_{k+1}^>[v])) \\
&= \sum_{u,v \in \Gamma^k} \xi(\Gamma_{k+1}^<[u]) \xi(\Gamma_{k+1}^<[v]) + \sum_{u,v \in \Gamma^k} \xi(\Gamma_{k+1}^<[u]) \xi(\Gamma_{k+1}^>[v]) \\
&\quad + \sum_{u,v \in \Gamma^k} \xi(\Gamma_{k+1}^>[u]) \xi(\Gamma_{k+1}^<[v]) + \sum_{u,v \in \Gamma^k} \xi(\Gamma_{k+1}^>[u]) \xi(\Gamma_{k+1}^>[v]).
\end{aligned}
\tag{6.38}
$$

Since we are interested in the expectations conditional on $\mathscr{E}_k$, we determine which of the terms on the right hand side of (6.38) contain dependencies. Since the outdegree of a fixed vertex is independent of its indegree, we obtain that $\xi(\Gamma^<[u]_{k+1})$ and $\xi(\Gamma^>[v]_{k+1})$ are independent and so are $\xi(\Gamma^<[v]_{k+1})$ and $\xi(\Gamma^>[u]_{k+1})$. Also, $\xi(\Gamma^>[v]_{k+1})$ and $\xi(\Gamma^>[u]_{k+1})$ are independent, unless $u = v$. Finally, $\Gamma^<[v]_{k+1}$ and $\Gamma^<[u]_{k+1}$ both contain information about the indegree evolutions of vertices $m < u \wedge v$ and are thus dependent. We first treat the case of $\xi(\Gamma^>[u]_{k+1})^2$, for $u \in \Gamma^k$, and $k \leq T \wedge S$

$$
\begin{aligned}
\mathbb{E}\left[\xi(\Gamma_{k+1}^>[u])^2 \Big| \mathscr{E}_k \right] &= \mathbb{E}\left[\sum_{a,b \in \Gamma_{k+1}^>[u]} \xi(a)\xi(b) \Big| \mathscr{E}_k \right] \\
&= 2 \sum_{a=u+1}^N \sum_{b=a+1}^N \xi(a)\xi(b) \mathbb{P}\{\Delta\mathscr{Z}[u,a-1] = \Delta\mathscr{Z}[u,b-1] = 1 | \mathscr{E}_k\} \\
&\leq 2C_{6.12} \sum_{a=u+1}^N \sum_{b=a+1}^N \xi(a)\xi(b) \mathbb{P}\{\Delta\mathscr{Z}[u,a-1] = 1 | \mathscr{E}_k\} \mathbb{P}^1\{\Delta\mathscr{Z}[u,b-1] = 1 | \mathscr{E}_k\} \\
&= 2C_{6.12} \sum_{a=u+1}^N \sum_{b=a+1}^N \xi(a)\xi(b) \mathbb{P}\{\Delta\mathscr{Z}[u,a-1] = 1 | \mathscr{E}_k\} \mathbb{P}\{\Delta\mathscr{Z}[u,b-1] = 1 | \mathscr{E}_k\} \\
&\quad + 2C_{6.12} \sum_{a=u+1}^N \Big(\xi(a) \mathbb{P}\{\Delta\mathscr{Z}[u,a-1] = 1 | \mathscr{E}_k\} \\
&\quad \times \sum_{b=a+1}^N \xi(b)(\mathbb{P}^1\{\Delta\mathscr{Z}[u,b-1] = 1 | \mathscr{E}_k\} - \mathbb{P}\{\Delta\mathscr{Z}[u,b-1] = 1 | \mathscr{E}_k\}) \Big) \\
&= \mathbb{E}\left[\xi(\Gamma_{k+1}^>[u]) \Big| \mathscr{E}_k \right]^2 \\
&\quad + c \sum_{a=u+1}^N \sum_{b=a+1}^N \xi(a)\xi(b)\psi^0(u,a)\psi^0(u,b) \frac{\xi(u,a)\xi(u,b)}{ab},
\end{aligned}
\tag{6.39}
$$

for some constant $c > 0$, where we have used that $\frac{\psi^1}{\psi^0}$ is uniformly bounded by Proposition 6.1 and 6.2. Using that $\xi(u,a)\xi(a,N) = \xi(u,N) = \xi(u)$, and $\psi^0(x,y) \leq c_{6.1}\left(\log\frac{y}{x}\right)^\alpha \vee 1$ we

obtain

$$\sum_{a=u+1}^{N} \sum_{b=a+1}^{N} \xi(a)\xi(b)\psi^0(u,a)\psi^0(u,b)\frac{\xi(u,a)\xi(u,b)}{ab}$$

$$\leq c_{6.1}^2 \xi(u)^2 \Big(\sum_{a=u}^{N} \frac{1}{a}(\log \tfrac{a}{u} \wedge 1)^\alpha \sum_{b=a+1}^{N} \frac{1}{b}(\log \tfrac{b}{u} \wedge 1)^\alpha \Big)$$

$$\leq c_{6.1}^2 \xi(u)^2 \Big(\sum_{a=u}^{N} \frac{1}{a}(\log \tfrac{a}{u} \wedge 1)^\alpha \Big(c' + 1 + \int_{eu}^{N} \frac{1}{b}(\log \tfrac{b}{u})^\alpha \, db \Big) \Big)$$

$$\leq c_{6.1}^2 \xi(u)^2 \big(c' + \big(\log \tfrac{N}{u}\big)^{\alpha+1} \big)\Big(c' + 1 + \int_{eu}^{N} \frac{1}{a}(\log \tfrac{a}{u})^\alpha \, da \Big)$$

$$\leq \xi(u)^2 c'' \big(\log \tfrac{N}{u} \wedge 1\big)^{2\alpha+2},$$

where c', c'' are suitably chosen constants. Inserting this into (6.39) yields

$$\mathbb{E}\Big[\xi(\Gamma_{k+1}^{>}[u])^2 \Big| \mathscr{E}_k\Big] \leq \mathbb{E}\Big[\xi(\Gamma_{k+1}^{>}[u]) \Big| \mathscr{E}_k\Big]^2 + \xi(u)^2 C'' \big(\log \tfrac{N}{u} \wedge 1\big)^{2\alpha+2}, \tag{6.40}$$

for some $C'' > 0$. The other case with dependencies which we have to consider are the random variables $\Gamma_{k+1}^{<}[u], \Gamma_{k+1}^{<}[v]$. We note that

$$\mathbb{P}\{\Delta \mathscr{Z}[a,u] = \Delta \mathscr{Z}[b,v] = 1 | \mathscr{E}_k\} = \mathbb{P}\{\Delta \mathscr{Z}[a,u] = 1 | \mathscr{E}_k\}\mathbb{P}\{\Delta \mathscr{Z}[b,v] = 1 | \mathscr{E}_k\},$$

unless $a = b < u \wedge v$ in which case, as above,

$$\mathbb{P}\{\Delta \mathscr{Z}[a,u] = \Delta \mathscr{Z}[b,v] = 1 | \mathscr{E}_k\} \leq C_{6.12} \mathbb{P}\{\Delta \mathscr{Z}[a, u \wedge v] = 1 | \mathscr{E}_k\}\mathbb{P}^1\{\Delta \mathscr{Z}[a, u \vee v] = 1 | \mathscr{E}_k\}.$$

A consideration similar to the first case now yields

$$\mathbb{E}\Big[\xi(\Gamma_{k+1}^{<}[u])\xi(\Gamma_{k+1}^{<}[v]) \Big| \mathscr{E}_k\Big] = \mathbb{E}[\xi(\Gamma_{k+1}^{<}[u]) | \mathscr{E}_k]\mathbb{E}[\xi(\Gamma_{k+1}^{<}[u]) | \mathscr{E}_k]$$

$$+ D' \sum_{a=r_{k+1}}^{u \wedge v} \xi(a)^2 \frac{\xi(a,u)\xi(a,v)}{uv} \big(\log \tfrac{u}{a} \vee 1\big)^\alpha \big(\log \tfrac{v}{a} \vee 1\big)\alpha, \tag{6.41}$$

for some constant $D' > 0$. Applying Lemma 4.6 and summing over all $u, v \in \Gamma^k$ in (6.41) yields

$$\sum_{u,v \in \Gamma^k} \mathbb{E}\Big[\xi(\Gamma_{k+1}^{<}[u])\xi(\Gamma_{k+1}^{<}[v]) \Big| \mathscr{E}_k\Big] \leq \sum_{u,v \in \Gamma^k} \mathbb{E}[\xi(\Gamma_{k+1}^{<}[u]) | \mathscr{E}_k]\mathbb{E}[\xi(\Gamma_{k+1}^{<}[u]) | \mathscr{E}_k]$$

$$+ D \big(\log \tfrac{N}{r_{k+1}}\big)^{2\alpha} \sum_{u,v \in \Gamma^k} \xi(u)\xi(v), \tag{6.42}$$

for another appropriately chosen constant $D > 0$. Combining (6.40) and (6.42) with (6.38) now concludes the proof.

$\square$

To conclude the argument about the growth of the score, we note that if $T < k^*$, then S_T is larger than $s_0 \frac{\sqrt{N}}{\psi^4(\ell_0, N)}$. If this is not the case, then we use the previous result to show that

S_{k^*} is concentrated around its expectation.

Proposition 6.19. *With high probability, as $N \to \infty$,*

$$S_{T \wedge S} \geq s_0 \frac{\sqrt{N}}{\psi^4(\ell_0, N)}.$$

Proof. We have already established, that $S \wedge T = T$ with high probability. The case where $T < k^*$ satisfies the lower bound of the proposition by definition and we only need to focus on the case $T = k^*$. Let $0 \leq k \leq k^*$, then an application of the conditional Chebychev inequality yields

$$
\begin{aligned}
\mathbb{P}\Big\{ |S_{k+1} - \mathbb{E}[S_{k+1}|\mathscr{E}_k]| &> \frac{1}{2}\mathbb{E}[S_{k+1}|\mathscr{E}_k]|\mathscr{E}_k \Big\} \\
&\leq \frac{C_{6.18}\big(\log \frac{N}{r_{k+1}}\big)^{2\alpha} S_k^2 + C'_{6.18}\sum_{u \in \Gamma^k}\xi(u)^2\big(\log \frac{N}{u} \vee 1\big)^{2\alpha+2}}{\frac{1}{4}\mathbb{E}[S_{k+1}|\mathscr{E}_k]^2},
\end{aligned}
\tag{6.43}
$$

using the second moment estimate of Proposition 6.18. Analogously to (4.33), we obtain that the dominating term on the right hand side of (6.43) is $\big(\log \frac{N}{r_{k+1}}\big)^{2\alpha} S_k^2 \mathbb{E}[S_{k+1}|\mathscr{E}_k]^{-2}$. Taking expectations in (6.43) and recalling that r_k decays like $Ne^{-k\log k}$ yields

$$\sum_{k=0}^{\infty} \mathbb{P}\Big\{ |S_{k+1} - \mathbb{E}[S_{k+1}|\mathscr{E}_k]| > \frac{1}{2}\mathbb{E}[S_{k+1}|\mathscr{E}_k] \Big\} < \infty.$$

This implies $\frac{S_k}{S_{k-1}} \leq \frac{1}{2}(\log \frac{N}{r_k})^{\alpha+1}$ only finitely many times as $k \to \infty$. Therefore we can find a small constant δ such that

$$S_{k^*} \geq \delta c_{(6.23)}^{k^*} \prod_{i=0}^{k^*} \big(\log \tfrac{N}{r_i}\big)^{1+\alpha}$$

with high probability. The choice of r_k and $k^* \geq \eta \frac{\log N}{\log\log N}$ for some $\eta > 0$ now imply that the right hand side exceeds $\sqrt{N}$ if N is sufficiently large, which concludes the argument and also implies that $S \wedge T < k^*$ with high probability. $\qquad\square$

6.3 Proof of Theorem II

We start with the lower bound, which is a straightforward application of Lemma 4.16. Using the bound on the expected degrees provided in Proposition 6.1, we can apply Lemma 4.16 with $\Psi(N) = (\log N)^\alpha$ and have thus proven the following statement.

Proposition 6.20 (Lower bound on distances)**.** *Let f be an attachment rule of the form*

$$f(k) = \frac{1}{2}k + \frac{\alpha k}{2\log k} + h(k),$$

where $\alpha \geq 0$ and

$$\lim_{k \to \infty} \frac{h(k)(\log k)^{1+\eta}}{k} = 0,$$

for some $\eta > 0$. Then, for uniformly chosen vertices $V, W \in \mathscr{G}_N$ and any $\delta \in (0,1)$,

$$\lim_{N \to \infty} \mathbb{P}\left\{ d_N(V, W) \geq \frac{(1-\delta)\log N}{(1+\alpha)\log\log N} \right\} = 1.$$

The upper bound is more involved. The dominating part of the distance is covered by the exploration phase after the initial local exploration. Before we can apply the results on the score growth from the previous section, we need to make sure that we can start the breath first search in a favourable configuration and that high final score implies that we have w.h.p. found the core of the network. The whole *exploration scheme* therefore has three parts

- Phase (I)– local exploration around V, W until either the component is explored or we have accumulated a score of at least $s_0 = S(\varepsilon)$;

- Phase (II) – breadth first exploration to accumulate a high score S_T;

- Phase (III) – connecting components of high score in $\mathscr{G}_N$ by short paths in $\mathscr{G}_{(1+\varepsilon)N}$.

To find a path between two uniformly chosen vertices we first establish that they are connected. This amounts to showing that they are in the giant component which can be inferred from the coupling of the local exploration processes to the INT.

Proposition 6.21 (Local exploration)**.** *Fix $\varepsilon > 0$. Depending on the value of α, we fix a threshold $S(\varepsilon)$,*

(i) if $\alpha > 0$, let $S(\varepsilon) = \frac{-16(1+8\varepsilon)\log\varepsilon}{\sigma_{4.34} q_{4.35}(f,\varepsilon)\chi_{6.10}}$,

(ii) if $\alpha = 0$, let $S(\varepsilon) = -\log(\varepsilon)(1+2\varepsilon)\psi^4(\ell_0, N)$.

Let $V, W \in \mathscr{G}_N$ be independently and uniformly chosen vertices. In both cases (i) and (ii), there exist N-independent constants $K = K(\varepsilon), r_0 = r_0(\varepsilon)$ such that

$$\lim_{N \to \infty} \mathbb{P}\left\{ \min\left(\xi(\mathscr{C}^K(V)), \xi(\mathscr{C}^K(W))\right) \geq S \right\} = p(f)^2,$$

where, for $k \geq 0$, $\mathscr{C}^k(v) = \mathscr{C}^k(v, \varepsilon)$ denotes the first k vertices discovered in the local exploration around $v \in \mathscr{G}_N$ truncated at r_0.

Proof. For the untruncated exploration the statement is implied by the local weak convergence of $\mathscr{C}^k$ to the projected INT, which is a consequence of Proposition 4.20. For every fixed k, and $\delta > 0$ we can also find $r < 0$ such that $\mathbb{P}\{\mathfrak{T}^k \cap (-\infty, r]\} < \delta$. As $\#\mathscr{C}^K < \infty$ for the untruncated exploration, we can therefore use this local convergence also to infer the statement for the truncated version. □

Proof of Theorem II. Only the upper bound remains to be shown. Let $\varepsilon > 0$. We start local explorations in the uniformly chosen vertices $V, W \in \mathscr{C}_N$. By Proposition 6.21, we reach in

both explorations a minimum score of $S_0 = S\left(\frac{\varepsilon}{8}\right)$ after at most $K\left(\frac{\varepsilon}{8}\right)$ exploration steps with probability exceeding $1 - \frac{\varepsilon}{4}$. We denote the explored part at the time the score S_0 is reached by $H_0^{(1)}, H_0^{(2)}$ and start the second exploration stage, in which we only look at the scores of the two explorations. We know by Proposition 6.19 that, if N is sufficiently large, the explorations will end with probability exceeding $1 - \frac{\varepsilon}{4}$ after $T < k^*$ steps, by which point

$$\xi(H_T^{(i)}) \geq S_0 \frac{\sqrt{N}}{\psi^4(\ell_0, N)}, \quad i = 1, 2. \tag{6.44}$$

We now distinguish two cases.

Firstly, if $\alpha > 0$, then we have collectd no information about the degree evolutions of vertices in $[M]$, for $M = \lfloor (\log N)^{2\alpha} \rfloor$ during the whole exploration, due to the truncation. Therefore we can apply Proposition 4.36 to deduce that, for sufficiently large N, with probability exceeding $1 - \frac{\varepsilon}{4}$, the subgraph $\mathrm{core}_N = (C_N, V_N) \subset \mathcal{G}_{(1+\varepsilon)N}$ is of bounded diameter D. We will show now, that $\mathbb{P}\{H_T^{(i)} \overset{\varepsilon}{\hookleftarrow} \mathrm{core}_N\} \geq 1 - \frac{\varepsilon}{4}$, if N is large. Let $H \subset [N]$ and $i \in \{N, \ldots, (1+\varepsilon)N\}$, then for every $u \in H$,

$$\mathbb{P}\{\Delta \mathcal{Z}[u, i] = 0\} \leq 1 - \frac{\mathbb{E}f(\mathcal{Z}[u, N])}{(1+\varepsilon)N} \leq \mathrm{e}^{-\frac{f(0)\xi(u)}{(1+\varepsilon)N}}$$

and thus

$$\mathbb{P}\{\Delta \mathcal{Z}[u, i] = 0 \text{ for all } u \in H\} \leq \mathrm{e}^{-\frac{f(0)\xi(H)}{(1+\varepsilon)N}}, \quad i \in \{N, \ldots, (1+\varepsilon)N\}. \tag{6.45}$$

Applying Lemma 4.35, we obtain that with probability exceeding $1 - \frac{\varepsilon}{8}$, denoting $\sigma = \sigma_{4.34}$ and $q = q_{4.35}(f, \varepsilon)$,

$$\#\{i \in \{N, \ldots, (1+\varepsilon)N\} : i \leftarrow C_N\} \geq \frac{\sigma q M \mathbb{E}\mathcal{Z}[M, N]}{8} \geq \frac{\sigma q M f(0)\xi(M)}{8},$$

if N is sufficiently large. Combining this with (6.45) yields

$$1 - \mathbb{P}\{H \overset{\varepsilon}{\hookleftarrow} \mathrm{core}_N\} \leq \exp\left(-\frac{f(0)^2 \sigma q M \xi(H)\xi(M)}{8(1+\varepsilon)N}\right) + \frac{\varepsilon}{8},$$

and applying the latter inequality to $H = H_T^{(i)}, i = 1, 2$ and using (6.44) we obtain

$$
\begin{aligned}
1 - \mathbb{P}\{H_T^{(i)} \overset{\varepsilon}{\hookleftarrow} \mathrm{core}_N\} &\leq \exp\left(-\frac{f(0)^2 \sigma q M \xi(H_T^{(i)})\xi(M)}{8(1+\varepsilon)N}\right) + \frac{\varepsilon}{8} \\
&\leq \exp\left(-\frac{f(0)^2 \sigma q M S_0 \sqrt{N}\xi(M)}{8\psi^4(\ell_0, N)(1+\varepsilon)N}\right) + \frac{\varepsilon}{8} \\
&\leq \exp\left(-\frac{f(0)^2 \sigma q M S_0 N}{8\psi^4(\ell_0, N)\sqrt{M}(1+\varepsilon)N}\right) + \frac{\varepsilon}{8} \\
&\leq \exp\left(-\frac{f(0)^2 \sigma q \chi_{6.10} S_0}{16(1+\varepsilon)}\right) + \frac{\varepsilon}{8} \leq \frac{\varepsilon}{4},
\end{aligned}
$$

using the lower bound

$$S_0 \geq S\left(\tfrac{\varepsilon}{8}\right) \geq \frac{-16(1+\varepsilon)\log\left(\tfrac{\varepsilon}{8}\right)}{\sigma_{4.34}\, q_{4.35}\left(f, \tfrac{\varepsilon}{8}\right)\chi_{6.10}}$$

in the last step and this holds for both $H_T^{(1)}$ and $H_T^{(2)}$ independently. Thus, we have shown that with probability exceeding $1-\varepsilon$,

$$d_{(1+\varepsilon)N}(U, V) \leq 4 + D + k^* + 2K\left(\tfrac{\varepsilon}{8}\right),$$

if N is sufficiently large.

Secondly, if $\alpha = 0$, we cannot conclude anything about the diameter of the core, but we can argue as above to obtain that with probability exceeding $1 - \tfrac{\varepsilon}{2}$,

$$1 - \mathbb{P}\{H_T^{(1)} \overset{\varepsilon}{\leftrightarrow} H_T^{(2)} | \mathscr{G}_N\} \leq \exp\left(-\frac{\xi(H_T^{(1)})\xi(H_T^{(2)})\varepsilon}{(1+\varepsilon)^2 N}\right) \leq e^{-\frac{S_0^2 \varepsilon}{(1+\varepsilon)^2 \psi^4(\ell_0, N)}} \leq \frac{\varepsilon}{2},$$

since $S_0 \geq S\left(\tfrac{\varepsilon}{2}\right)$, for all sufficiently large N.

In both cases we have shown that w.h.p. the dominating part of the distance is bounded by k^*, thus the upper bound is shown by taking $\varepsilon \to 0$. $\qquad\square$

Appendix A

Properties of the truncation sequence

The cutoff sequence $r = (r_k)_{k=0}^{\infty}$ used in the proof of Theorem II is introduced in Section 4.5 as solution to the difference equation

$$
\begin{aligned}
r_0 &= s_0 \\
-\frac{1}{2}\Delta r_k &= (1+\alpha)\log(-r_k) + \frac{1}{2}\log\frac{r_{k+1}}{r_k} - 2\log\frac{k+4}{k+3} + \log\frac{c}{\sqrt{1+\alpha}}, \quad \text{for } k \geq 0,
\end{aligned}
\tag{A.1}
$$

where $c > 0, \alpha \geq 0$ and the initial condition $s_0 < 0$ is a negative parameter of large absolute value.

Lemma A.1. *The equation* (A.1) *has a unique, strictly decreasing solution* $(r_k)_{k=0}^{\infty}$*, if the initial value* s_0 *is chosen sufficiently small.*

Proof. Note that

$$
-2\log\frac{k+4}{k+3} + \log\frac{c}{\sqrt{1+\alpha}} \geq \log\frac{9c}{16\sqrt{1+\alpha}} = d,
$$

thus (A.1) has a unique strictly decreasing solution if

$$
\begin{aligned}
s_0 &= s_0 \\
-\frac{1}{2}\Delta s_k &= (1+\alpha)\log(-s_k) + \frac{1}{2}\log\frac{s_{k+1}}{s_k} + d, \quad k \geq 0,
\end{aligned}
\tag{A.2}
$$

has a solution in which case $s_k \geq r_k$ for all k. We will argue inductively. Setting $k = 1$ in (A.2) yields

$$
-\frac{1}{2}(s_1 - s_0) = (1+\alpha)\log(-s_0) + \frac{1}{2}\log\frac{s_1}{s_0} + d,
$$

which can be formally rewritten as

$$
\phi(-s_1) = -2s_0 + (2\alpha + 1)\log(-s_0) + 2d,
\tag{A.3}
$$

with $\phi(x) = x - \log x$. The right hand side of (A.3) is strictly greater than 1 if $-s_0$ is sufficiently large, and ϕ defines an increasing bijection of $(1, \infty)$ into itself. Thus s_1 exists and is strictly

negative. Furthermore,

$$-s_1 + s_0 = \log(-s_1) - s_0 + 2d + (2\alpha + 1)\log(-s_0) > 0,$$

if $-s_0$ is sufficiently large. Now assume, $s_0 > s_1 > \cdots > s_k$ are given, then s_{k+1} is obtained via

$$-s_{k+1} = \phi^{-1}(-2s_k + 2(\alpha + 1)\log(-s_k) + 2d)$$

by the same argument and $-s_{k+1} > -s_k$. $\qquad\square$

We now collect some results about the asymptotics of $(r_k)_{k \geq 0}$ here. First, observe that if r_k grows superlinearly but subexponentially, we expect, for $N \gg k \gg 0$ sufficiently large,

$$-\Delta r_k \approx 2(\alpha + 1)\log(-r_k). \tag{A.4}$$

For the proof of Theorem II, we need reasonably sharp asymptotic lower bounds on the decay of r, which is the content of the following lemma.

Lemma A.2. *Let $r = (r_k)_{k \geq 0}$ be the solution to* (A.1). *Then*

$$\liminf_{k \to \infty} \frac{-r_k}{k \log k} \geq 2(\alpha + 1).$$

Proof. Let $d = \log \frac{9c}{16\sqrt{1+\alpha}}$ and set

$$k_0 = \min\left\{ k : \log\log k > (k+1)\log\frac{k+1}{k} - \frac{d}{2(\alpha+1)} - \log(2\alpha + 1) \right\}.$$

Fix also

$$t_0 = -s_0 > 2(|d| + 1) \vee (2(\alpha + 1)k_0 + \log k_0 + d),$$

and let $t = (t_k)_{k=0}^{\infty}$ solve

$$\Delta t_k = 2(\alpha + 1)\log t_k + d.$$

Note that the t_k are uniquely defined, and also $-t_k \geq r_k$, for all k. We have

$$t_{k+1} = \phi(t_k)$$

with

$$\phi(x) = x + 2(\alpha + 1)\log x + d,$$

which defines a concave, non-negative and strictly increasing function on $(2|d| + 1, \infty)$. To prove the statement about r, we only need to verify

$$t_{k+1} \geq 2(\alpha + 1)(k + 1)\log(k + 1) \tag{A.5}$$

119

for all sufficiently large k. We induce induction for $k = k_0$ and obtain

$$t_{k_0} > t_1 = (-s_0) + (2\alpha + 1)\log(-s_0) + d > 2(\alpha + 1)k_0 \log k_0,$$

by choice of s_0. Assume now, that (A.5) holds for t_k, then

$$
\begin{aligned}
t_{k+1} = \phi(t_k) &\geq \phi(2(\alpha + 1)k \log k) \\
&= 2(\alpha + 1)k \log k + (2\alpha + 1)\log(2(\alpha + 1)k \log k) + d \\
&= 2(\alpha + 1)(k + 1)\log(k + 1) + 2(\alpha + 1)(\log\log k + \log(2\alpha + 1)) + d - 2(\alpha + 1)(k + 1)\log\tfrac{k+1}{k} \\
&\geq 2(\alpha + 1)(k + 1)\log(k + 1),
\end{aligned}
$$

since $k \geq k_0$, which concludes the argument.

Remark A.3. For any $\varepsilon > 0$, it can also be shown that $-r_k$ does not grow faster than $(2\alpha + 2 + \varepsilon)k \log k$.

$\square$

Using Lemma A.2 we now derive the bound for k^* featured in the proof of Theorem II.

Lemma A.4. *Define $r = (r_k)_{k=0}^{\infty}$ as the solution to (A.1). Set, for $N \in \mathbb{N}$,*

$$m_k = \pi_N(r_k) = N e^{r_k}.$$

Let $R \in \mathbb{N}_0$ be independent of N and $\psi(N) = O((\log N)^R)$. define

$$k^*(N) = \min\left\{k : \xi(m_k, N) \geq \frac{\sqrt{N}}{(\psi(N)}\right\}.$$

For any $\varepsilon \in (0, 2\alpha + 2)$, we have that $k^* \leq \frac{\log N}{(2 + 2\alpha - \varepsilon)\log\log N}$, if N is sufficiently large.

Proof. Lemma A.2 implies that if N and k are sufficiently large,

$$\xi(m_k, N) \geq k^{1 + \alpha - \frac{\varepsilon}{2}} k. \tag{A.6}$$

Fix $\varepsilon > 0$, let $C > 0$ be such that $\psi(N) \leq C(\log N)^R$, for all N, and let

$$k_0 = k_0(N) = \left\lceil \frac{\log N}{(2 + 2\alpha - \varepsilon)\log\log N} \right\rceil,$$

then

$$2\left(\alpha + 1 + \frac{\varepsilon}{2}\right)k_0 \log k_0 = \frac{2\alpha + 2 + \frac{\varepsilon}{2}}{2\alpha + 2 - \varepsilon}\frac{\log N}{\log\log N}(\log\log N(1 + o(1))),$$

thus for sufficiently large N, we obtain

$$2(\alpha + 1)k_0 \log k_0 \geq \log N - 2R\log\log N - \log C,$$

120

hence (A.6) implies that $k_0 \in \left\{ k : \xi(m_k, N) \geq \frac{\sqrt{N}}{\psi(N)} \right\}$ and thus $k^* \leq k_0$. $\qquad\square$

Remark A.5. Following Remark A.3, we also obtain that $k^*(N) \geq \frac{\log N}{\log \log N}$ for some c and N sufficiently large.

Appendix B

Further preliminaries

Proposition B.1 (A priori bounds on degrees)**.**   *(i) Let f be affine, $f(k) = \gamma k + \beta$, and set $\Xi = \sqrt{2} \frac{\gamma}{\beta} \sqrt{\sum_{k=1}^{\infty} \xi(1,k)^{-2}}$. Then, for any $\delta > 0$, w.h.p. $\mathcal{G}_n$ has no vertex of degree larger than $f(0)\xi(1,N)(\Xi + \delta)\sqrt{\log N}$.*

(ii) Let f be concave with $\gamma = \frac{1}{2}$ then for any $\delta > 0$ w.h.p. there is no vertex with degree $f(\mathcal{Z}[m,N]) \geq n^{\frac{1}{2}+\delta}$.

Proof. For (i), we would like to apply the Azuma-Hoeffding inequality to X_k. For a bound on the increment, look at

$$|X_{n+1} - X_n| = \begin{cases} \frac{\left(1+\frac{\gamma}{k}\right)f(\mathcal{Z}[1,n]) - f(\mathcal{Z}[1,n])}{\xi(1,n+1)} & \text{if } \Delta f(\mathcal{Z}[1,n]) = 0 \\ \frac{\left|\left(1+\frac{\gamma}{k}\right)f(z[1,n]) - f(\mathcal{Z}[1,n]) - \gamma\right|}{\xi(1,n+1)} & \text{otherwise} \end{cases} \leq \frac{\gamma}{\xi(1,n+1)}$$

Recall that for all n,

$$\mathbb{P}\left\{X_n \geq \beta + t\right\} \quad \leq \quad \exp -\frac{t^2}{2^{\frac{\delta}{\gamma}} \sum \xi(1,n+1)^{-2}} \tag{B.1}$$

$$\leq \quad e^{-\frac{t^2}{\Xi^2}\beta^2}. \tag{B.2}$$

And we obtain

$$\mathbb{P}\left\{f(\mathcal{Z}[1,N]) \quad \geq \quad \beta\xi(1,N)(\Xi + \delta)\sqrt{\log N}\right\} \tag{B.3}$$

$$= \quad \mathbb{P}\left\{\Xi_N \geq \beta + \beta\left((\Xi + \delta)\sqrt{\log N} - 1\right)\right\} \tag{B.4}$$

$$\leq \quad \exp\left\{-\frac{(\Xi + \delta)^2 \log N - 2\sqrt{\log N} + 1}{\Xi^2}\right\} \tag{B.5}$$

$$\leq \quad N^{-\frac{(\Xi+\delta)^2}{\Xi^2} - \frac{2}{\sqrt{\log N}}} = O\left(N^{-1}\right). \tag{B.6}$$

Since $\mathcal{Z}[1,N]$ dominates all other degree evolutions taking a union bound implies the statement.

The proof of (ii) works similarly, but we start in $k_0 = \min\left\{k : \Delta f(k) \leq \frac{1}{2} + \frac{\delta}{2}\right\}$ and then use the supermartingale $\frac{f(\mathcal{Z}[1,N])}{\xi^{\frac{1}{2}+\frac{\delta}{2}}(1,n)}$. The above derivation then applies with $t = f(K)\left(n^{\frac{\delta}{2}} - 1\right)$. $\square$

Lemma B.2 (Janson's Lemma, [Jan02]). *Let $X_1, X_2, \ldots$ be a sequence of independent Bernoulli random variables with $\mathbb{P}\{X_i = 1\} = p_i$. Denote $S_n = \sum_{i=1}^{n} X_i$ and $\mu_n = \mathbb{E}S_n$, then, for any $t \geq 0$,*

$$\mathbb{P}\{S_n \geq \mu_n + t\} \leq \exp\left(-\frac{t^2}{2\left(\mu_n + \frac{t}{3}\right)}\right),$$

and

$$\mathbb{P}\{S_n \leq \mu_n - t\} \leq \exp\left(-\frac{t^2}{2\mu_n}\right).$$

Proof. This is a consequence of Chernoff's inequality. $\square$

Proposition B.3 (Diameter of $\mathcal{G}(n, p)$, [Bol01, Corollary 10.12, first part]). *Suppose $d(n) \geq 3$ and $0 < p(n) < 1$ satisfy*

(i) $\lim_{n \to \infty} \frac{\log n}{d(n)} - 3\log\log n = \infty,$

(ii) $\lim_{n \to \infty} p(n)^{d(n)} n^{d(n)-1} - 2\log n = \infty,$

(iii) $\lim_{n \to \infty} p(n)^{d(n)-1} n^{d(n)-2} - 2\log n = -\infty;$
   then, with high probability, $\mathcal{G}(n, p)$ has diameter $d(n)$.

Proof. The general theorem for $\mathcal{G}(n, p)$ is proved in [Bol01, Theorem 10.10], however this corollary is sufficient for our purposes. The original proof was published in [Bol81]. $\square$

Lemma B.4. *Let $a, b, x, y, z > 0$ and $a \leq b$. Then*

$$\frac{x + ya}{x + za} \geq \frac{x + yb}{x + zb}$$

if and only if $z \geq y$.

Proof. Note that under the assumptions, $z \geq y$ if and only if $z(b - a) \geq y(b - a)$, which is equivalent to

$$zb + ya \geq za + yb$$

and therefore also to

$$x^2 + xzb + xya + yzab \geq x^2 + xza + xyb + yzab.$$

The last inequality can be rewritten as

$$(x + zb)(x + ya) \geq (x + za)(x + yb),$$

and we have obtained the desired statement. $\square$

Bibliography

[AB02]   Réka Albert and Albert-László Barabási. "Statistical mechanics of complex net-works". In: *Reviews of Modern Physics* 74.1 (Jan. 2002), p. 47.

[AGS08]   Krishna B. Athreya, Arka P. Ghosh, and Sunder Sethuraman. "Growth of prefer-ential attachment random graphs via continuous-time branching processes". In: *Proceedings Mathematical Sciences* 118.3 (2008), pp. 473–494.

[Ath07]   Krishna B. Athreya. "Preferential attachment in random graphs with general weight function". In: *Internet Mathematics* 4.4 (Jan. 2007), pp. 401–418.

[BA99]   Albert-László Barabási and Réka Albert. "Emergence of scaling in random net-works". In: *Science* 286.5439 (1999), pp. 509–512.

[Bar02]   Albert-László Barabási. *Linked: The new science of networks.* Perseus Publishing, Cambridge, Massachusetts, 2002.

[BB01]   Ginestra Bianconi and Albert-László Barabási. "Bose-Einstein condensation in complex networks". In: *Physical Review Letters* 86.24 (June 2001), p. 5632.

[BC78]   Edward A. Bender and E. Rodney Canfield. "The asymptotic number of labeled graphs with given degree sequences". In: *Journal of Combinatorial Theory, Se-ries A* 24.3 (1978), pp. 296–307.

[Ben62]   George Bennett. "Probability inequalities for the sum of independent random variables". In: *Journal of the American Statistical Association* 57 (1962), pp. 33–45.

[Big77]   John D. Biggins. "Martingale convergence in the branching random walk". In: *Journal of Applied Probability* (1977), pp. 25–37.

[BJR07]   Béla Bollobás, Svante Janson, and Oliver Riordan. "The phase transition in in-homogeneous random graphs". In: *Random Structures & Algorithms* 31 (2007), pp. 3–122.

[BL12]   Graham Brightwell and Malwina J. Luczak. "Vertices of high degree in the pref-erential attachment tree". In: *Electronic Journal of Probability* 17.14 (2012), 43 pages.

[Bol+01]   Béla Bollobás, Oliver Riordan, Joel Spencer, and Gabor. Tusnády. "The degree sequence of a scale-free random graph process". In: *Random Structures and Algorithms* 18.3 (2001), pp. 279–290.

[Bol01]   Béla Bollobás. *Random Graphs*. Cambridge University Press, Oct. 2001, p. 520. ISBN: 9780521797221.

[Bol81]   Béla Bollobás. "The diameter of random graphs". In: *Transactions of the American Mathematical Society* 267 (1981), pp. 41–52.

[Bor+07]   Christian Borgs, Jennifer Chayes, Constantinos Daskalakis, and Sebastien Roch. "First to market is not everything: an analysis of preferential attachment with fitness". In: *Proceedings of the thirty-ninth annual ACM-Symposium on the Theory of Computing*. San Diego, California, USA: ACM, 2007, pp. 135–144. ISBN: 978-1-59593-631-8.

[BR04]   Béla Bollobás and Oliver Riordan. "The diameter of a scale free random graph". In: *Combinatorica* 24.1 (Jan. 2004), pp. 5–34.

[CF03]   Collin Cooper and Alan Frieze. "A general model of web graphs". In: *Random Structures & Algorithms* 22.3 (2003), pp. 311–335.

[CL02]   Fan Chung and Linyuan Lu. "The average distance in a random graph with given expected degrees". In: *Proceedings of the National Academy of Sciences* 99.25 (2002), pp. 15879–15882.

[CL06]   Fan Chung and Linyuan Lu. *Complex Graphs and Networks*. Boston, MA, USA: American Mathematical Society, Aug. 2006, p. 278. ISBN: 9780821836576.

[DHH10]   Sander Dommers, Remco van der Hofstad, and Gerard Hooghiemstra. "Diameters in Preferential Attachment Models". In: *Journal of Statistical Physics* 139.1 (Apr. 2010), pp. 72–107.

[DM09]   Steffen Dereich and Peter Mörters. "Random networks with sublinear preferential attachment: degree evolutions". In: *Electronic Journal of Probability* 14.43 (2009), pp. 1222–1267.

[DM11]   Steffen Dereich and Peter Mörters. "Random networks with concave preferential attachment rule". In: *Jahresbericht der Deutschen Mathematiker Vereinigung* 113 (2011), pp. 21–40.

[DM13]   Steffen Dereich and Peter Mörters. "Random networks with sublinear preferential attachment: the giant component". In: *Annals of Probability* 41.1 (Jan. 2013), pp. 329–384.

[DMM12]   Steffen Dereich, Christian Mönch, and Peter Mörters. "Typical distances in ultrasmall random networks". In: *Advances in Applied Probability* 44.2 (June 2012), pp. 583–601.

[DO13]   Steffen Dereich and Marcel Ortgiese. "Robust analysis of preferential attachment models with fitness". In: *arXiv preprint 1302.3385* (2013).

[EM13]   M. Eckhoff and P. Mörters. "Vulnerability of robust preferential attachment networks". In: *arXiv preprint 1306.3454* (2013).

[HH08]   R. van der Hofstad and G. Hooghiemstra. "Universality for distances in power-law random graphs". In: *Journal of Mathematical Physics* 49.12 (2008), p. 125209.

[HHZ07]   R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. "Distances in random graphs with finite mean and infinite variance degrees". In: *Electronic Journal of Probability* 12 (2007), pp. 703–766.

[Hof13]   R. van der Hofstad. "Random Graphs and Complex Networks, book in preparation, TU Eindhoven". 2013. URL: http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf.

[Jan02]   S. Janson. "On concentration of probability". In: *Contemporary combinatorics* 10.3 (2002), pp. 1–9.

[JM12]   E. Jacob and P. Mörters. "Spatial preferential attachment networks: Power laws and clustering coefficients". In: *arXiv preprint 1210.3830* (2012).

[Jor13]   J. Jordan. "Geometric preferential attachment in non-uniform metric spaces". In: *Electronic Journal of Probability* 18.8 (2013), pp. 1–15.

[Mil67]   S. Milgram. "The small world problem". In: *Psychology Today* May (1967), pp. 60–67.

[NR06]   I. Norros and H. Reittu. "On a conditionally Poissonian graph process". In: *Advances in Applied Probability* 38 (2006), pp. 59–75.

[NR08]   I. Norros and H. Reittu. "Network models with a 'soft hierarchy': a random graph construction with loglog scalability". In: *Network, IEEE* 22.2 (2008), pp. 40–46.

[OS05]   R. Oliveira and J. Spencer. "Connectivity Transitions in Networks with Super-Linear Preferential Attachment". In: *Internet Mathematics* 2.2 (Jan. 2005), pp. 121–163.

[Pit94]   B. Pittel. "Note on the heights of random recursive trees and random m-ary search trees". In: *Random Structures and Algorithms* 5.2 (1994), pp. 337–347.

[Res08]   S. I. Resnick. *Extreme values, regular variation, and point processes.* Springer, New York, 2008.

[She89]   L. A. Shepp. "Connectedness of certain random graphs". In: *Israel Journal of Mathematics* 67.1 (1989), pp. 23–33.

[Szy87]   Jerzy Szymański. "On a Nonuniform Random Recursive Tree". In: *Annals of Discrete Mathematics (33) Proceedings of the International Conference on Finite Geometries and Combinatorial Structure.* Ed. by A. Barlotti, M. Biliotti, A. Cossu, G. Korchmaros, and G. Tallini. Vol. 144. North-Holland Mathematics Studies. North-Holland, 1987, pp. 297–306.

[Wat03]   D. J. Watts. *Six degree. The science of a connected age.* W. W. Norton & Co. Inc., New York, 2003.

[Yul25]   G. U. Yule. "A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, FRS". In: *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character* 213 (1925), pp. 21–87.