

# Stochastische Analyse von Preferential Attachment-Modellen

## Diplomarbeit

von

Christoph te Kampe

eingereicht am  
Institut für Mathematik  
Goethe-Universität Frankfurt am Main

Betreuer: Prof. Dr. Ralph Neininger

April 2009

# Erklärung

Hiermit erkläre ich, die vorliegende Diplomarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben.

Frankfurt am Main, 2. April 2009

Christoph te Kampe

# Zusammenfassung

Die vorliegende Arbeit untersucht ausgewählte Eigenschaften von „Preferential Attachment“-Graphen. Darunter verstehen wir eine Klasse komplexer zufälliger Graphen, die mit einer vorgegebenen Konfiguration gestartet werden und anschließend mit jedem Zeitschritt um eine Ecke und  $m \in \mathbb{N}$  Kanten wachsen. Die Wachstumsregeln sind so gestaltet, dass eine neue Ecke ihre Kanten bevorzugt an Ecken sendet, die bereits mit vielen anderen Ecken verbunden sind, woraus sich die Bezeichnung „Preferential Attachment“ (PA) ableitet.

Die Arbeit stellt zunächst heuristisch die Eigenschaft der Skalenfreiheit von PA-Modellen vor und bespricht anschließend einen Beweis zu dieser These. Weiter betrachten wir den Durchmesser von PA-Graphen und untersuchen das Verhalten bei Anwachsen des Graphen. Wir erkennen, dass der Durchmesser bei wachsendem Graphen deutlich langsamer wächst, was wir als „Small-World“-Phänomen bezeichnen. Die zentralen Aussagen und Beweise orientieren sich an den Arbeiten von Remco van der Hofstad, der die bekannten PA-Modelle um einen Parameter  $\delta$  erweitert hat. Damit ist es möglich, sowohl logarithmische als auch doppelt-logarithmische Schranken für den Durchmesser zu erhalten.

# Inhaltsverzeichnis

<b>Einleitung</b>	<b>ii</b>
<b>1 Definition der Modelle</b>	<b>1</b>
1.1 Das Barabási-Albert-Modell . . . . .	1
1.2 Der zusammenhängende Graph . . . . .	3
1.3 Unabhängige Kanten . . . . .	4
<b>2 Gradfolge und Skalenfreiheit</b>	<b>5</b>
2.1 Die Gradfolge eines PA-Modells . . . . .	5
2.2 Die Konzentration der Gradfolge . . . . .	8
2.3 Erwartungswert der Gradfolge . . . . .	10
2.4 Beweis von Theorem 1 . . . . .	17
<b>3 Durchmesser von PA-Modellen</b>	<b>18</b>
3.1 Eine obere Schranke des Durchmessers . . . . .	18
3.2 Eine untere Schranke des Durchmessers . . . . .	28
3.3 Doppelt logarithmische obere Schranke des Durchmessers . . . . .	36
3.3.1 Der Durchmesser des Kerns . . . . .	36
3.3.2 Die Verbindung von Kern und Peripherie . . . . .	44
3.3.3 Beweis von Theorem 4 . . . . .	48
<b>4 Technische Beweise und Lemmata</b>	<b>50</b>
4.1 Das zweite Moment der Gradfolge . . . . .	50
4.2 Ecken mit hohem Grad . . . . .	52
4.3 Der Durchmesser eines multinomialen Graphen . . . . .	54
4.4 Frühe Ecken haben große Grade . . . . .	55
<b>Literatur</b>	<b>63</b>

# Einleitung

Gegenstand der vorliegenden Arbeit ist eine Klasse von komplexen zufälligen Graphen, die als „Preferential Attachment-Graphen“ bezeichnet werden (PA-Graphen). Die zugrunde liegenden Modelle zeichnen sich dadurch aus, dass sie mit einer vorgegebenen Konfiguration gestartet und dann mit jedem Zeitschritt um eine Ecke und  $m \in \mathbb{N}$  Kanten erweitert werden. Die Besonderheit der Modelle besteht darin, dass eine neue Ecke ihre Kanten bevorzugt an Ecken sendet, die bereits mit vielen anderen Ecken verbunden sind. Die vorgestellten Modelle liefern eine Erklärung für das Auftreten von Phänomenen, die mit der Theorie der klassischen Erdős-Renyi-Graphen bisher nicht erklärt werden konnten, so z.B. dem „Small World“-Phänomen oder der Skalenfreiheit eines Netzwerks.

Barabási und Albert haben eine einfache Version eines PA-Modells bereits in [BA99] vorgestellt. Bollobás et al. präzisierten das Modell und lieferten in [BRST01] einen Beweis für die Power Law-Eigenschaft der Gradfolge. In [BR04] zeigen Bollobás und Riordan dann, dass der Durchmesser der PA-Modelle asymptotisch  $\log t / \log \log t$  ist. Die vorliegende Diplomarbeit orientiert sich an [Hof07] und [HH08]. Darin werden die genannten Modelle um einen Parameter  $\delta$  erweitert, wodurch der Exponent in der Gradfolge so variiert werden kann, dass man sowohl logarithmische als auch doppelt logarithmische Schranken für den Durchmesser erhält.

Die hier besprochenen Modelle hat Remco van der Hofstad im Mai 2007 in [HH07] vorgestellt. Eine überarbeitete Version wurde zwischenzeitlich in „The Annals of Applied Probability“ zur Veröffentlichung eingereicht (vgl. [HH08]). Das Verhalten des Durchmessers von PA-Graphen für  $t \rightarrow \infty$  ist noch nicht abschließend beschrieben. Bei Abgabe der Diplomarbeit im März 2009 wurde mit [DHH09] eine weitere Arbeit vorgestellt, deren wesentliche Resultate zu Beginn von Kapitel 3 wiedergegeben werden.

In dieser Arbeit befassen wir uns mit drei leicht unterschiedlichen Mo-

dellen. Im ersten Kapitel werden die Modelle zunächst definiert und ihre Graphen bezüglich der wesentlichen Eigenschaften beschrieben. Wir geben jeweils eine Ausgangskonfiguration an sowie Regeln, nach denen die Modelle wachsen sollen. Alle drei Modelle folgen dem Prinzip des „preferential attachment“. Die Wachstumsregeln bilden die Grundlage für die weiteren Beweise und werden im Verlauf der Arbeit ständig benutzt werden.

Im zweiten Kapitel erläutern wir den Begriff der *Gradfolge* eines PA-Modells. Wir stellen eine Heuristik von Barabási und Albert vor, die zu der Vermutung führt, dass die Gradfolge einer „Power-Law“ der Form  $c_m k^{-\tau}$  folgt. Wir definieren dann eine Wahrscheinlichkeitsverteilung  $\{p_k\}_{k=1}^{\infty}$  mit eben dieser Eigenschaft und zeigen, dass die Gradfolge des PA-Modells stochastisch gegen diese Folge konvergiert. Als Ergebnis des Kapitels können wir die Eigenschaft der Skalenfreiheit für das Modell verifizieren.

Im dritten Kapitel untersuchen wir die Modelle auf das Phänomen der „Small Worlds“. Die unterschiedlichen Aussagen des Kapitels gehen auf die Variation des Exponenten  $\tau$  zurück. Im ersten Abschnitt zeigen wir, dass der Durchmesser des Graphen mit hoher Wahrscheinlichkeit nach oben beschränkt wird durch eine Konstante mal  $\log t$ , falls  $\tau \in (2, \infty)$ . Im zweiten Abschnitt wählen wir für  $\delta \geq 0$  bzw.  $\tau \in [3, \infty)$ , dass der Durchmesser des Graphen mit hoher Wahrscheinlichkeit nach unten beschränkt wird durch eine Konstante mal  $\frac{\log t}{\log \log t}$ . Im letzten Abschnitt setzen wir  $\delta \in (-m, 0)$  und reduzieren den Exponenten der Gradfolge damit auf  $\tau \in (2, 3)$ . In diesem Fall und mit  $m \geq 2$  können wir eine doppelt logarithmische obere Schranke des Durchmessers zeigen. D.h. der Durchmesser des Graphen wird mit hoher Wahrscheinlichkeit nach oben beschränkt durch eine Konstante mal  $\log \log t$ .

Im vierten Kapitel finden sich technische Lemmata und Beweise, die zugunsten der Übersicht nicht im Hauptteil untergebracht werden.

**Notation:** Für eine Folge von Ereignissen  $\{E_t\}_{t \geq 1}$ , schreiben wir dass  $E_t$  mit hoher Wahrscheinlichkeit eintritt, wenn  $\lim_{t \rightarrow \infty} \mathbb{P}(E_t) = 1$ . Mit Bezug auf die überwiegend englischen Quellen der Arbeit benutzen hierfür die Abkürzung whp (*with high probability*).

Des weiteren wird an mehreren Stellen auf die Gammafunktion verwiesen. Dabei sind die folgende Eigenschaften für uns von Bedeutung:

Sei  $\Gamma : \mathbb{R}_+^* \rightarrow \mathbb{R}$  mit  $t \mapsto \int_0^\infty x^{t-1} e^{-x} dx$ . Es gilt

$$\Gamma(n+1) = n! \text{ für alle } n \in \mathbb{N} \quad (0.0.1)$$

$$\Gamma(t+1) = t\Gamma(t) \text{ für alle } t \in \mathbb{R} \quad (0.0.2)$$

$$\frac{\Gamma(t)}{\Gamma(t-a)} = t^a(1+o(1)) \text{ für } t \rightarrow \infty, a \in \mathbb{R} \text{ und } t > a. \quad (0.0.3)$$

Einen Beweis findet man z. B. in [AE99], S. 100ff.

# Kapitel 1

## Definition der Modelle

Die Modelle, die wir untersuchen, liefern uns eine *Folge von Graphen*  $\{G_m(t)\}$ , die für  $t \geq 1$  oder  $t \geq 2$  jeweils einen Graphen mit  $t$  Ecken und  $mt$  Kanten liefert, für ein gegebenes  $m \in \mathbb{N}$ . Wir betrachten drei Variationen von PA-Modellen, die wir im folgenden mit (a), (b) und (c) bezeichnen. Ein wesentliches Unterscheidungsmerkmal ist dabei die Frage, ob eine Ecke mit sich selbst verbunden werden kann, und ob wir infolge dessen einen zusammenhängenden Graphen erreichen oder nicht. Des weiteren unterscheiden wir die Modelle darin, ob alle Kanten einer Ecke mit der gleichen Wahrscheinlichkeit auf die vorhanden Ecken verteilt werden, oder ob sich die Wahrscheinlichkeiten nach jeder der  $m$  Kanten ändern sollen. Doch gehen wir nun in die Details:

### 1.1 Das Barabási-Albert-Modell

Das erste Modell mit dem wir uns beschäftigen, geht auf Barabási und Albert zurück und wurde erstmals in [BA99] beschrieben. In unserem Fall wurde das Modell von Remco van der Hofstad<sup>1</sup> um einen Parameter  $\delta$  erweitert, den wir in späteren Abschnitten noch näher beschreiben wollen. Wir wählen zunächst  $m = 1$  und konstruieren das Modell für allgemeines  $m$  anschließend. Zum Zeitpunkt  $t = 1$  bestehe  $G_1(1)$  aus einer einzelnen Ecke, die über eine Kante mit sich selbst verbunden ist. Immer wenn eine Kante eine Ecke mit sich selbst verbindet, wollen wir von einer *Schlaufe* sprechen.

Die Ecken des Graphen bezeichnen wir mit  $1, 2, \dots$ , so dass die Ecken von  $G_m(t)$  durch die Menge  $\{1, 2, \dots, t\}$  bezeichnet sind. Weiter bezeichnen wir

---

<sup>1</sup>Remco van der Hofstad, Eindhoven University of Technology, Department of Mathematics and Computer Science



mit  $d_i(t)$  den Grad einer Ecke  $i$ , also die Anzahl von Kanten, die zu  $i$  inzident sind. Eine Schlaufe erhöhe dabei den Grad einer Ecke um 2. Der Parameter des Modells sei  $\delta \geq -1$ . Um von  $G_1(t)$  zu  $G_1(t+1)$  zu gelangen, lassen wir das Modell nun wie folgt wachsen. Für  $t > i$  beschreibe  $\{t \rightarrow i\}$  das Ereignis, dass sich die Ecke  $t$  mit der Ecke  $i$  verbinde. Wir fügen nun eine Ecke  $t+1$  hinzu, von der eine neue Kante ausgeht. Das Ende dieser Kante soll sich mit Wahrscheinlichkeit

$$\mathbb{P}(t+1 \rightarrow i | G_{1,\delta}(t)) = \frac{1 + \delta}{t(2 + \delta) + (1 + \delta)}, \quad i = t+1, \quad (1.1.1)$$

mit  $t+1$  selbst verbinden (Schlaufe) und mit Wahrscheinlichkeit

$$\mathbb{P}(t+1 \rightarrow i | G_{1,\delta}(t)) = \frac{d_i(t) + \delta}{t(2 + \delta) + (1 + \delta)}, \quad i = 1, \dots, t, \quad (1.1.2)$$

mit einer der bestehenden Ecken des Graphen. Wegen  $\sum_{i=1}^t (d_i(t) + \delta) = t(2 + \delta)$  summieren sich diese Wahrscheinlichkeiten zu 1.

Das Modell  $\{G_m(t)\}_{t=1}^{\infty}$  für  $m > 1$  leiten wir wie folgt aus dem Modell für  $m = 1$  ab: Wir starten mit  $G_1(mt)$  mit  $\delta' = \frac{\delta}{m}$  und bezeichnen die Ecken in  $G_1(mt)$  mit  $1', \dots, (mt)'$ . Dann fassen wir die Ecken  $1', \dots, m'$  in  $G_1(mt)$  zu Ecke 1 in  $G_m(t)$  zusammen und fortlaufend die Ecken  $(m+1)', \dots, (2m)'$  in  $G_1(mt)$  als Ecke 2 in  $G_m(t)$ . Allgemeiner gesagt: Die Ecken  $((j-1)m+1)', \dots, (jm)'$  in  $G_1(mt)$  bilden gerade die Ecke  $j$  in  $G_m(t)$ . Bestehen nun für  $i, j \in \mathbb{N}$  Kanten zwischen einer der Ecken  $((j-1)m+1)', \dots, (jm)'$  und einer der Ecken  $((i-1)m+1)', \dots, (im)'$  setzen wir diese Kanten auch zwischen  $i$  und  $j$ , was zu Schleifen führt wenn  $i = j$ .

Alternativ kann man das Modell für  $m \geq 1$  formulieren, wenn man auf die einzelnen Kanten abstellt. Wir bezeichnen mit  $d_i(e, t)$  den Grad der  $i$ -ten Ecke, nachdem die  $e$ -te Kante von  $t+1$  verbunden wurde. Wir starten mit einer Ecke 1, die  $m$  Verbindungen zu sich selbst hat. Im Schritt von  $t$  nach  $t+1$  fügt man eine neue Ecke mit  $m$  Kanten hinzu, wobei die Wahrscheinlichkeit, dass die  $e$ -te Kante der Ecke  $t+1$  mit einer vorhandenen Ecke  $i$  verbunden wird, durch

$$\mathbb{P}(t+1 \rightarrow i | G_{e-1,m}(t)) = \frac{d_i(e-1, t) + \delta}{t(m + \delta) + (e + \delta)}, \quad i = 1, \dots, t, \quad (1.1.3)$$

gegeben ist und die Wahrscheinlichkeit, dass die  $e$ -te Kante mit  $t+1$  selbst verbunden wird, durch

$$\mathbb{P}(t+1 \rightarrow i | G_{e-1,m}(t)) = \frac{e + \delta}{t(m + \delta) + (e + \delta)}, \quad i = t+1. \quad (1.1.4)$$

Im Vergleich zu den nachfolgenden Modellen halten wir zwei Eigenschaften fest:

- Dieses Modell ordnet dem Ereignis, dass sich eine Ecke mit sich selbst verbindet, eine Wahrscheinlichkeit größer Null zu, womit der entstehende Graph nicht notwendigerweise zusammenhängend sein muss.
- Die Wahrscheinlichkeiten für die Bindung an eine der vorhandene Ecken ändern sich mit jeder neuen Kante.

Wählen wir  $\delta = 0$  erhalten wir das ursprüngliche Barabási-Albert-Modell aus [BRST01]. Im weiteren Verlauf der Arbeit werden wir von diesem Modell als *Modell (a)* sprechen.

## 1.2 Der zusammenhängende Graph

Das zweite Modell unterscheidet sich von Modell (a) dadurch, dass wir keine Schleifen erlauben wollen. Wir beginnen wiederum mit der Formulierung für  $m = 1$ . Für  $t = 1$  wäre eine Schleife unumgänglich, weshalb wir das Modell gleich für  $t = 2$  starten.  $G_1(2)$  besteht aus den Ecken 1 und 2, die über zwei Kanten miteinander verbunden sind. Für  $t > 2$  definieren wir, bedingt auf  $G_1(t)$ , die Wachstumsregel nach  $G_1(t + 1)$  wie folgt. Für  $\delta \geq -1$  sei

$$\mathbb{P}(t + 1 \rightarrow i | G_1(t)) = \frac{d_i(t) + \delta}{t(2 + \delta)}, \quad i = 1, \dots, t. \quad (1.2.1)$$

Auch hier nutzen wir das Modell für  $m = 1$  um das Modell für  $m > 1$  zu formulieren, und zwar exakt so, wie wir es für Modell (a) gemacht haben. Eine alternative Formulierung analog (1.1.3) lautet hier

$$\mathbb{P}(t + 1 \rightarrow i | G_{e-1,m}(t)) = \frac{d_i(e - 1, t) + \delta}{t(m + \delta)}, \quad i = 1, \dots, t. \quad (1.2.2)$$

Die wesentlichen Eigenschaften sind hier:

- Das Modell liefert für alle  $m \geq 1$  einen zusammenhängenden Graphen, denn wenn der Graph für  $m = 1$  zusammenhängend ist, dann auch der für  $m > 1$ .
- Die Wahrscheinlichkeiten für die Bindung an eine der vorhandene Ecken ändern sich mit jeder neuen Kante.

Im weiteren Verlauf der Arbeit sprechen wir von diesem Modell als *Modell* (b).

**Bemerkung:** In den Modellen (a) und (b) für  $m > 1$  haben wir  $\delta' = \frac{\delta}{m}$  gerade so gewählt, dass in dem resultierenden Graphen  $G_m(t)$ , den wir aus den gruppierten Ecken von  $G_1(mt)$  konstruiert haben, die Endpunkte der Kanten einer neuen Ecke gerade abhängig vom Grad der vorhandenen Ecken plus der Konstante  $\delta$  gewählt wurden.

### 1.3 Unabhängige Kanten

Das dritte Modell beschreiben wir direkt für allgemeines  $m \geq 1$ . Bedingt auf  $G_m(t)$  wählen wir die Endpunkte der  $m$  Kanten von  $t + 1$  unabhängig davon, wieviele der  $m$  Kanten dieser Ecke wir schon verteilt haben. Wir beginnen mit  $G_m(2)$  und den Ecken 1 und 2, die durch  $2m$  Kanten miteinander verbunden sind. Da die Endpunkte der Kanten unabhängig ausgewählt werden, können wir die Definition von  $\{G_m(t)\}_{t \geq 2}$  für  $m \geq 1$  in einem Schritt angeben. Für  $1 \leq e \leq m$  ist

$$\mathbb{P}(e\text{-te Kante von } t + 1 \text{ ist verbunden mit } i | G_m(t)) = \frac{d_i(t) + \delta}{t(2m + \delta)}, \quad (1.3.1)$$

$i = 1, \dots, t$ . Die wesentlichen Eigenschaften dieses Modells, das wir als *Modell* (c) bezeichnen wollen, sind:

- Das Modell liefert für alle  $m \geq 1$  einen zusammenhängenden Graphen.
- Für alle Kanten einer neuen Ecke sind die Wahrscheinlichkeiten, sich an eine der bereits vorhandenen Ecke zu binden, gleich.

**Bemerkung:** Für  $m = 1$  sind die Modelle (b) und (c) identisch.

# Kapitel 2

## Gradfolge und Skalenfreiheit

### 2.1 Die Gradfolge eines PA-Modells

Mit  $\{p_k(t)\}_{k=1}^{\infty}$  bezeichnen wir im folgenden die *Gradfolge* eines PA-Modells. Dabei beschreibt die Zufallsvariable  $p_k(t)$  für jedes  $k \in \mathbb{N}$  die relative Anzahl von Ecken, die zum Zeitpunkt  $t$  gerade Grad  $k$  haben. Genauer definieren wir

$$p_k(t) = \frac{1}{t} \sum_{i=1}^t \mathbb{1}_{\{d_i(t)=k\}}. \quad (2.1.1)$$

Barabási und Albert haben 1999 durch Simulationen am Computer bereits darauf hingewiesen, dass die Verteilung der Grade einer „Power Law“ folgt (vgl. [BA99]). Abweichend zu Abschnitt 1.1 wurde dort definiert, dass sich eine neue Ecke  $i + 1$  mit der Wahrscheinlichkeit  $P(t + 1 \rightarrow i | G_m(t)) = \frac{d_i(t)}{\sum_j d_j(t)}$  mit einer bestehenden Ecke  $i$  verbindet, was in unserem Modell der Wachstumsregel (1.1.2) für  $\delta = 0$  entspricht. In [BA99] betrachten die Autoren den Grad  $d_i$  einer Ecke  $i$  als stetige Funktion der Zeit, und erhalten so

$$\frac{\partial d_i}{\partial t} = m \cdot \frac{d_i}{\sum_j d_j} = \frac{d_i}{2t}, \quad (2.1.2)$$

da  $\sum_j d_j(t) = 2mt$ . Die Lösung dieser Differentialgleichung ist

$$d_i(t) = m \left( \frac{t}{t_i} \right)^{\frac{1}{2}}, \quad (2.1.3)$$

wobei  $t_i$  der Zeitpunkt sein soll, zu dem die Ecke  $i$  hinzugefügt wurde. Umformen liefert

$$\mathbb{P}(d_i(t) > k) = P(t_i < t \frac{m^2}{k^2}) = \frac{m^2}{k^2}, \quad (2.1.4)$$

wobei wir benutzt haben, dass die Ecken auf  $[0, t]$  gleichmäßig hinzugefügt werden und somit für  $t_0 \leq t$  gilt,  $\mathbb{P}(t_i < t_0) = \frac{t_0}{t}$ . Differenzieren nach  $k$  liefert dann

$$\mathbb{P}(d_i(t) = k) = \frac{2m^2}{k^3}. \quad (2.1.5)$$

Wir definieren nun für  $m \geq 1$  und  $\delta > -m$  eine Wahrscheinlichkeitsverteilung  $\{p_k\}_{k=1}^\infty$  durch

$$p_k := \frac{(2 + \frac{\delta}{m})\Gamma(k + \delta)\Gamma(m + 2 + \delta + \frac{\delta}{m})}{\Gamma(m + \delta)\Gamma(k + 3 + \delta + \frac{\delta}{m})}. \quad (2.1.6)$$

Zunächst ist zu zeigen, dass es sich bei  $\{p_k\}_{k=1}^\infty$  tatsächlich um eine Wahrscheinlichkeitsverteilung auf  $\mathbb{N}$  handelt:

**Beweis:** Aus der Definition von  $p_k$  folgt direkt, dass  $p_k \geq 0$  für alle  $k \in \mathbb{N}$ . Weiter ist zu zeigen, dass  $\sum_{k=1}^\infty p_k = 1$ . Mehrfaches Anwenden von (0.0.2) liefert

$$\frac{\Gamma(k + a)}{\Gamma(k + b)} = \frac{1}{b - a - 1} \left( \frac{\Gamma(k + a)}{\Gamma(k - 1 + b)} - \frac{\Gamma(k + 1 + a)}{\Gamma(k + b)} \right). \quad (2.1.7)$$

Nun setzen wir  $a = \delta$ ,  $b = 3 + \delta + \frac{\delta}{m}$  und erhalten für (2.1.6)

$$\begin{aligned} p_k &= \frac{\Gamma(m + 2 + \delta + \frac{\delta}{m})}{\Gamma(m + \delta)} \cdot \frac{\Gamma(k + \delta)}{\Gamma(k + 3 + \delta + \frac{\delta}{m})} \cdot (2 + \frac{\delta}{m}) \\ &= \frac{\Gamma(m + 2 + \delta + \frac{\delta}{m})}{\Gamma(m + \delta)} \left( \frac{\Gamma(k + \delta)}{\Gamma(k + 2 + \delta + \frac{\delta}{m})} - \frac{\Gamma(k + 1 + \delta)}{\Gamma(k + 3 + \delta + \frac{\delta}{m})} \right). \end{aligned} \quad (2.1.8)$$

Für  $k < m$  ist  $p_k = 0$ , da jede Ecke mindestens  $m$  Kanten hat. Summieren der Wahrscheinlichkeiten liefert uns eine Teleskopsumme, so dass

$$\begin{aligned} \sum_{k \geq 1} p_k &= \sum_{k \geq m} p_k \\ &= \frac{\Gamma(m + 2 + \delta + \frac{\delta}{m})}{\Gamma(m + \delta)} \sum_{k \geq m} \frac{\Gamma(k + \delta)}{\Gamma(k + 2 + \delta + \frac{\delta}{m})} - \frac{\Gamma(k + 1 + \delta)}{\Gamma(k + 3 + \delta + \frac{\delta}{m})} \\ &= \frac{\Gamma(m + 2 + \delta + \frac{\delta}{m})}{\Gamma(m + \delta)} \cdot \frac{\Gamma(m + \delta)}{\Gamma(m + 2 + \delta + \frac{\delta}{m})} = 1. \end{aligned} \quad (2.1.9)$$

□

Wir dürfen also im folgenden von  $\{p_k\}_{k=1}^\infty$  als einer Wahrscheinlichkeitsverteilung auf  $\mathbb{N}$  sprechen. Nun wollen wir überprüfen, ob die  $p_k$  auch als Verteilung der Grade in einem  $PA$ -Modell geeignet sind. Dafür überprüfen wir

zunächst, ob die heuristisch begründete Power Law-Eigenschaft auf  $\{p_k\}_{k=1}^\infty$  zutrifft:

Wir beginnen mit dem vereinfachten Fall für  $\delta = 0$ . Dann können wir für alle  $k \geq m$  rechnen, dass

$$\begin{aligned} p_k &= \frac{2\Gamma(k)\Gamma(m+2)}{\Gamma(k+3)\Gamma(m)} = \frac{2(k-1)!(m+1)!}{(k+2)!(m-1)!} \\ &= \frac{2m(m+1)}{k(k+1)(k+2)} \sim \frac{2m^2}{k^3}. \end{aligned} \quad (2.1.10)$$

Das heißt, die Wahrscheinlichkeit, dass eine zufällig gewählte Ecke im  $PA$ -Graphen Grad  $k$  hat, nimmt mit zunehmendem  $k$  polynomiell ab. Unsere Heuristik scheint uns also auf den richtigen Weg zu bringen. Etwas allgemeiner können wir rechnen, wenn wir uns Eigenschaft (0.0.3) der Gamma-Funktion zunutze machen. Wir wählen  $t = k + \delta$  und  $a = -\tau = -(3 + \frac{\delta}{m})$ . Für  $k \rightarrow \infty$  erhalten wir aus (2.1.6)

$$\begin{aligned} p_k &= c_m(k + \delta)^{-\tau}(1 + o(1)) \\ &\sim c_m k^{-\tau}, \end{aligned} \quad (2.1.11)$$

wobei

$$\tau = 3 + \frac{\delta}{m} > 2, \quad \text{und} \quad c_m = \frac{(2 + \frac{\delta}{m})\Gamma(m + 2 + \delta + \frac{\delta}{m})}{\Gamma(m + \delta)}. \quad (2.1.12)$$

Damit haben wir gezeigt, dass  $\{p_k\}_{k=1}^\infty$  einem Potenzgesetz folgt. Wie der Leser sicher schon vermutet, wird sich  $\{p_k\}_{k=1}^\infty$  im weiteren Verlauf als die gesuchte Verteilung der Grade in einem  $PA$ -Modell herausstellen. Genauer gesagt, werden wir im restlichen Teil des Kapitels zeigen, dass  $p_k(t)$  für  $t \rightarrow \infty$  stochastisch gegen  $p_k$  konvergiert. Eine stärkere Aussage ist

**Theorem 1** (*Gradfolge eines  $PA$ -Modells*) Sei  $\delta > -m$  und  $m \geq 1$ . Dann gibt es eine Konstante  $C > 0$ , so dass für  $t \rightarrow \infty$  gilt:

$$\mathbb{P} \left( \max_k |p_k(t) - p_k| \geq C \sqrt{\frac{\log t}{t}} \right) = o(1). \quad (2.1.13)$$

Theorem 1 soll in zwei Schritten bewiesen werden. Zunächst zeigen wir in Abschnitt 2.2, dass die Gradfolge hinreichend dicht um ihren Erwartungswert konzentriert ist, was die Aussage von Proposition 2.2.1 ist. Dort erhalten wir als Zwischenergebnis

$$\mathbb{P} \left( \max_k |p_k(t) - \mathbb{E}[p_k(t)]| \geq C \sqrt{\frac{\log t}{t}} \right) = o(1). \quad (2.1.14)$$

Anschließend zeigen wir in Abschnitt 2.3, dass der Erwartungswert der Gradfolge für große  $t$  nur wenig von der in (2.1.6) definierten Verteilung abweicht, also dass gleichmäßig für alle  $k \in \mathbb{N}$

$$|\mathbb{E}[p_k(t)] - p_k| \leq \frac{C}{t}. \quad (2.1.15)$$

## 2.2 Die Konzentration der Gradfolge

Im folgenden bezeichnen wir die Anzahl der Ecken vom Grad  $k$  zum Zeitpunkt  $t$  mit

$$P_k(t) := \sum_{i=1}^t \mathbb{1}_{\{d_i(t)=k\}} = tp_k(t). \quad (2.2.1)$$

**Proposition 2.2.1** (*Konzentration der Gradfolge*) Sei  $\delta \geq -m$  und  $m \geq 1$ . Dann gibt es ein  $C \geq 0$ , so dass für  $t \rightarrow \infty$

$$\mathbb{P} \left( \max_k |P_k(t) - \mathbb{E}[P_k(t)]| \geq C\sqrt{t \log t} \right) = o(1). \quad (2.2.2)$$

Die Aussage von Proposition 2.2.1 ist also insbesondere, dass  $P_k(t)$  whp nur sehr wenig (genauer gesagt nicht mehr als  $C\sqrt{t \log t}$ ) von seinem Erwartungswert abweicht.

**Beweis:** Der maximale Grad einer Ecke ist  $m(t+1)$ . Wegen  $P_k(t) = 0$  für  $k > m(t+1)$  gilt

$$\begin{aligned} & \mathbb{P} \left( \max_k |P_k(t) - \mathbb{E}[P_k(t)]| \geq C\sqrt{t \log t} \right) \\ &= \mathbb{P} \left( \max_{k \leq m(t+1)} |P_k(t) - \mathbb{E}[P_k(t)]| \geq C\sqrt{t \log t} \right) \\ &\leq \sum_{k=1}^{m(t+1)} \mathbb{P} \left( |P_k(t) - \mathbb{E}[P_k(t)]| \geq C\sqrt{t \log t} \right). \end{aligned} \quad (2.2.3)$$

Es ist also zu zeigen, dass es ein  $C > 0$  gibt, so dass für alle  $k \leq t$  gilt

$$\mathbb{P} \left( |P_k(t) - \mathbb{E}[P_k(t)]| \geq C\sqrt{t \log t} \right) = o(t^{-1}). \quad (2.2.4)$$

Für  $n = 0, \dots, t$  bezeichne

$$M_n = \mathbb{E}[P_k(t) | G_m(n)] \quad (2.2.5)$$

die bedingte Erwartung der Anzahl von Ecken mit Grad  $k$  zum Zeitpunkt  $t$ , unter der Bedingung  $G_m(n)$ . Das heißt, wir betrachten einen Prozess, der uns für eine festes  $t$  die bedingte Erwartung von  $P_k(t)$  liefert, wobei sich die Bedingung  $G_m(n)$  mit  $n = 1, 2, \dots, t$  jeweils gemäß der Entwicklung von  $G_m$  ändert. Da der Graph  $G_m(n)$  vom Graphen  $G_m(n+1)$  abgeleitet werden kann, gilt mit der Turmeigenschaft für bedingte Erwartungswerte

$$\begin{aligned}\mathbb{E}[M_{n+1}|G_m(n)] &= \mathbb{E}[\mathbb{E}[P_k(t)|G_m(n+1)]|G_m(n)] \\ &= \mathbb{E}[P_k(t)|G_m(n)] \\ &= M_n, \quad \text{wobei } n \leq t-1.\end{aligned}\tag{2.2.6}$$

Damit erfüllt  $\{M_n\}_{n=0}^t$  die Anforderungen der bedingten Erwartungswerte für Martingale. Der Prozess  $\{M_n\}_{n=0}^t$  ist ein Doobsches Martingal. Zudem wird  $P_k(t)$  beschränkt durch die Gesamtzahl der Ecken zum Zeitpunkt  $t$ , also  $P_k(t) \leq t$ . Damit gilt

$$\mathbb{E}[|M_n|] = \mathbb{E}[M_n] = \mathbb{E}[P_k(t)] \leq t < \infty.\tag{2.2.7}$$

Als ersten Schritt zum Beweis von Proposition 2.2 halten wir daher fest, dass es sich bei dem Prozess  $\{M_n\}_{n=0}^t$  um ein Martingal bezüglich der von  $\{G_m(n)\}_{n=0}^t$  erzeugten Filtrierung handelt. Für  $M_0$  ergibt sich

$$M_0 = \mathbb{E}[P_k(t)|G_m(0)] = \mathbb{E}[P_k(t)],\tag{2.2.8}$$

da der „leere“ Graph  $G_m(0)$  keine Bedingung darstellt. Zudem ist  $M_t$  ebenfalls eindeutig bestimmt durch

$$M_t = \mathbb{E}[P_k(t)|G_m(t)] = P_k(t),\tag{2.2.9}$$

da  $P_k(t)$  von  $G_m(t)$  festgelegt wird. Als zweiten wesentlichen Schritt zum Beweis von Proposition 2.2.4 können wir nun berechnen

$$P_k(t) - \mathbb{E}[P_k(t)] = M_t - M_0.\tag{2.2.10}$$

Schritt drei erfordert zunächst das folgende Lemma.

**Lemma 2.2.2** (*Ungleichung von Azuma-Hoeffding*) Sei  $\{M_n\}_{n=0}^t$  ein Martingal, so dass für  $K_n \geq 0$  gilt

$$\mathbb{P}[|M_n - M_{n-1}| \leq K_n] = 1 \text{ für alle } n \geq 1\tag{2.2.11}$$

und  $\mu := M_0 = \mathbb{E}[M_n]$ . Dann gilt für alle  $a \geq 0$

$$\mathbb{P}[|M_n - \mu| \geq a] \leq 2 \exp\left(-\frac{a^2}{2 \sum_{i=1}^n K_i^2}\right).\tag{2.2.12}$$



Einen Beweis findet man in [Hof07], Theorem 2.22.

Wir behaupten

$$|M_n - M_{n-1}| \leq 2m \quad (2.2.13)$$

und betrachten dafür

$$M_n - M_{n-1} = \mathbb{E}[P_k(t)|G_m(n)] - \mathbb{E}[P_k(t)|G_m(n-1)] \quad (2.2.14)$$

$$= \mathbb{E}[\mathbb{E}[P_k(t)|G_m(n)] - P_k(t)|G_m(n-1)]. \quad (2.2.15)$$

Da  $G_m(n-1)$  bekannt ist, können sich neue Informationen in  $G_m(n)$  lediglich auf die  $m$  Kanten beziehen, die mit der Ecke  $n$  hinzugekommen sind. Mit dieser Information bilden wir dann den Erwartungswert der Gradfolge bezüglich der übrigen Kanten, welche mit den Ecken  $n+1, \dots, t$  hinzukommen. Einen Unterschied bei Kenntnis von  $G_m(n)$  im Gegensatz zu  $G_m(n-1)$  kann sich jedoch lediglich auf die  $m$  Kanten beziehen, die von der Ecke  $n$  stammen. Diese Anzahl ist höchstens  $m$ , so dass die Verteilung der Grade um höchstens  $2m$  abweichen kann bei Kenntnis von  $G_m(n)$  statt  $G_m(n-1)$ . Damit haben wir (2.2.13) gezeigt.

Die Ungleichung von Azuma-Hoeffding liefert dann mit  $K_n = 2m$ , dass für alle  $a > 0$

$$\mathbb{P}(|P_k(t) - \mathbb{E}[P_k(t)]| \geq a) \leq 2e^{-\frac{a^2}{8m^2t}} \quad (2.2.16)$$

also auch für  $a = C\sqrt{t \log t}$ . Wählen wir  $C^2 > 8m^2$ , gilt

$$\mathbb{P}\left(|P_k(t) - \mathbb{E}[P_k(t)]| \geq C\sqrt{t \log t}\right) \leq 2e^{-\log t \frac{C^2}{8m^2}} = o(t^{-1}). \quad (2.2.17)$$

□

## 2.3 Erwartungswert der Gradfolge

In diesem Abschnitt betrachten wir die erwartete Anzahl von Ecken in einem PA-Graphen, die genau Grad  $k$  haben. Als Bezeichnung wählen wir

$$\bar{P}_k(t) := \mathbb{E}[P_k(t)] \quad (2.3.1)$$

mit  $k \in \mathbb{N}$ . Dabei soll gezeigt werden, dass  $\bar{P}_k(t)$  nur gering von  $p_k t$  abweicht, wobei  $p_k$  definiert ist wie in (2.1.6). Diese Aussage macht die folgende Proposition:

**Proposition 2.3.1** (*Erwartungswert der Gradfolge*): Sei  $\delta > -m$  und  $m \geq 1$ . Dann gibt es eine Konstante  $C = C(\delta, m)$ , so dass für alle  $t \geq 1$  und alle  $k \in \mathbb{N}$  gilt,

$$|\bar{P}_k(t) - p_k t| \leq C. \quad (2.3.2)$$

Der Beweis von Proposition 2.3.1 lässt sich in zwei Teile gliedern, zum einen für  $m = 1$  und weiter für  $m > 1$ . Im folgenden wird nur der Beweis für  $m = 1$  besprochen. Den Beweis für  $m > 1$  findet man in [Hof07], Abschnitt 8.5.2.

**Beweis:** Zunächst leiten wir die Verteilung von  $p_k$  aus Abschnitt 2.1 her. Wir können schreiben

$$\mathbb{E}[P_k(t+1)|G_1(t)] = P_k(t) + \mathbb{E}[P_k(t+1) - P_k(t)|G_1(t)]. \quad (2.3.3)$$

Bedingt auf  $G_1$  gibt es vier Fälle, für die  $P_k(t+1) - P_k(t) \neq 0$ . Für diese unterschiedlichen Fälle können wir aus der Wachstumsregel des PA-Graphen Eintrittswahrscheinlichkeiten festlegen, die anschließend jeweils mit der Häufigkeit der möglichen Fälle im Graphen multipliziert werden. Als Ergebnis erhalten wir den bedingten Erwartungswert auf der rechten Seite von (2.3.3).

1. Wir betrachten eine zum Zeitpunkt  $t$  bereits vorhandene Ecke  $i$  mit Grad  $d_i(t) = k - 1$ , die nun mit der Ecke  $t + 1$  verbunden wird.  $P_k(t)$  erhöht sich dann um eins. Nach (1.1.2) ist die Wahrscheinlichkeit für dieses Ereignis  $\frac{k-1+\delta}{t(2+\delta)+(1+\delta)}$ .
2. Wir betrachten eine zum Zeitpunkt  $t$  bereits vorhandene Ecke  $i$  mit Grad  $d_i(t) = k$ , die nun mit der Ecke  $t + 1$  verbunden wird.  $P_{k+1}(t)$  erhöht sich dann um eins. Nach (1.1.2) ist die Wahrscheinlichkeit für dieses Ereignis  $\frac{k+\delta}{t(2+\delta)+(1+\delta)}$ .
3. Der Grad der zuletzt hinzugefügten Ecke ist eins, d.h.  $P_1(t)$  erhöht sich um eins. Das ist immer dann der Fall, wenn die Ecke  $t + 1$  nicht mit sich selbst verbunden wird. Die Wahrscheinlichkeit für dieses Ereignis ist  $1 - \frac{1+\delta}{t(2+\delta)+(1+\delta)}$ .
4. Der Grad der zuletzt hinzugefügten Ecke ist zwei, d.h. die Ecke  $t + 1$  wurde mit sich selbst verbunden. In diesem Fall erhöht sich  $P_2(t)$  um eins. Nach (1.1.1) ist die Wahrscheinlichkeit für dieses Ereignis  $\frac{1+\delta}{t(2+\delta)+(1+\delta)}$ .

Fassen wir nun diese Möglichkeiten zusammen, erhalten wir als Ausgangsgleichung, dass

$$\begin{aligned}\mathbb{E}[P_k(t+1) - P_k(t)|G_1(t)] &= \frac{k-1+\delta}{t(2+\delta) + (1+\delta)} P_{k-1}(t) \\ &\quad - \frac{k+\delta}{t(2+\delta) + (1+\delta)} P_k(t) \\ &\quad + \mathbb{1}_{\{k=1\}} \left(1 - \frac{1+\delta}{t(2+\delta) + (1+\delta)}\right) \\ &\quad + \mathbb{1}_{\{k=2\}} \frac{1+\delta}{t(2+\delta) + (1+\delta)}.\end{aligned}\tag{2.3.4}$$

Dabei sei  $k \geq 1$ . Für  $k = 0$  setzen wir

$$P_0(t) = 0.\tag{2.3.5}$$

Bilden wir nun den Erwartungswert für beide Seiten von (2.3.4), erhalten wir für  $k \geq 1$  einen rekursiven Zusammenhang von  $\bar{P}_k(t+1)$  und  $\bar{P}_k(t)$ , nämlich

$$\begin{aligned}\bar{P}_k(t+1) &= \bar{P}_k(t) + \frac{k-1+\delta}{t(2+\delta) + (1+\delta)} \bar{P}_{k-1}(t) \\ &\quad - \frac{k+\delta}{t(2+\delta) + (1+\delta)} \bar{P}_k(t) \\ &\quad + \mathbb{1}_{\{k=1\}} \left(1 - \frac{1+\delta}{t(2+\delta) + (1+\delta)}\right) \\ &\quad + \mathbb{1}_{\{k=2\}} \frac{1+\delta}{t(2+\delta) + (1+\delta)}.\end{aligned}\tag{2.3.6}$$

Wir erinnern uns, dass  $\bar{P}_k(t) = \mathbb{E}[tp_k(t)]$  und wählen nun  $p_k$  so, dass die Rekursion aus (2.3.6) erfüllt ist, also

$$p_k = \frac{k-1+\delta}{2+\delta} p_{k-1} - \frac{k+\delta}{2+\delta} p_k + \mathbb{1}_{\{k=1\}},\tag{2.3.7}$$

und per Definition  $p_0 = 0$ . Gleichung (2.3.7) aufgelöst nach  $p_k$  ergibt:

$$p_k = \frac{k-1+\delta}{k+2+2\delta} p_{k-1} + \frac{2+\delta}{k+2+2\delta} \mathbb{1}_{\{k=1\}}\tag{2.3.8}$$

Zunächst lösen wir (2.3.7) und zeigen, dass  $\{p_k\}_{k=1}^\infty$  mit der Verteilung aus (2.1.6) übereinstimmt. Für  $k = 1$  und mit  $p_0 = 0$  erhalten wir

$$p_1 = \frac{2+\delta}{3+2\delta}\tag{2.3.9}$$

und weiter für  $k > 1$

$$p_k = \frac{k-1+\delta}{k+2+2\delta} p_{k-1}. \quad (2.3.10)$$

Benutzt man (0.0.2) in der Form  $t = \frac{\Gamma(t+1)}{\Gamma(t)}$ , kann man  $p_k$  schreiben als

$$\begin{aligned} p_k &= \frac{(k-1+\delta)(k-2+\delta)\cdots(1+\delta)}{(k+2+2\delta)(k+1+2\delta)\cdots(4+2\delta)} p_1 \\ &\stackrel{(0.0.2)}{=} \frac{\frac{\Gamma(k+\delta)}{\Gamma(k-1+\delta)} \cdot \frac{\Gamma(k-1+\delta)}{\Gamma(k-2+\delta)} \cdots \frac{\Gamma(2+\delta)}{\Gamma(1+\delta)}}{\frac{\Gamma(k+3+2\delta)}{\Gamma(k+2+2\delta)} \cdot \frac{\Gamma(k+2+2\delta)}{\Gamma(k+1+2\delta)} \cdots \frac{\Gamma(5+2\delta)}{\Gamma(4+2\delta)}} \cdot \frac{2+\delta}{3+2\delta} \\ &= \frac{\Gamma(k+\delta)\Gamma(4+2\delta)}{\Gamma(k+3+2\delta)\Gamma(1+\delta)} \cdot \frac{2+\delta}{\frac{\Gamma(4+2\delta)}{\Gamma(3+2\delta)}} \\ &= (2+\delta) \frac{\Gamma(k+\delta)\Gamma(3+2\delta)}{\Gamma(1+\delta)\Gamma(k+3+2\delta)} \end{aligned} \quad (2.3.11)$$

Damit stimmt  $\{p_k\}_{k=1}^{\infty}$  mit der in (2.1.6) definierten Verteilung überein.

Sei nun

$$\varepsilon_k(t) = \bar{P}_k(t) - t p_k. \quad (2.3.12)$$

Zum Beweis von Proposition 2.3.1 im Fall  $m = 1$  ist zu zeigen, dass es ein  $C = C(\delta)$  gibt, so dass

$$\max_k |\varepsilon_k(t)| \leq C. \quad (2.3.13)$$

Eine explizite Schranke  $C$  wird dabei im weiteren Verlauf des Beweises berechnet. Als Beweis führen wir eine Vollständige Induktion in  $t$ . Dazu bedarf es noch einer kurzen Vorbereitung:

Die Gleichung (2.3.7) können wir wie folgt umschreiben

$$\begin{aligned}
(1+t)p_k &= tp_k + p_k & (2.3.14) \\
&= tp_k + \underbrace{\frac{k-1+\delta}{t(2+\delta)+(1+\delta)}tp_{k-1}}_I \\
&\quad - \underbrace{\frac{k+\delta}{t(2+\delta)+(1+\delta)}tp_k + \mathbb{1}_{\{k=1\}}}_{II} \\
&\quad + \underbrace{\left(\frac{1}{2+\delta} - \frac{t}{t(2+\delta)+(1+\delta)}\right)(k-1+\delta)p_{k-1}}_{III} \\
&\quad - \underbrace{\left(\frac{1}{2+\delta} - \frac{t}{t(2+\delta)+(1+\delta)}\right)(k+\delta)p_k}_{IV}, & (2.3.15)
\end{aligned}$$

da sich jeweils  $I$  und  $III$  sowie  $II$  und  $IV$  zu Null addieren. Zur Vereinfachung definieren wir zwei Abkürzungen

$$\begin{aligned}
\kappa(t) &:= \left(\frac{1}{2+\delta} - \frac{t}{t(2+\delta)+(1+\delta)}\right)(k-1+\delta)p_{k-1} \\
&\quad - \left(\frac{1}{2+\delta} - \frac{t}{t(2+\delta)+(1+\delta)}\right)(k+\delta)p_k & (2.3.16)
\end{aligned}$$

$$\gamma(t) := -\mathbb{1}_{\{k=1\}}\frac{1+\delta}{t(2+\delta)+(1+\delta)} + \mathbb{1}_{\{k=2\}}\frac{1+\delta}{t(2+\delta)+(1+\delta)}. \quad (2.3.17)$$

Fasst man nun die Gleichungen (2.3.14) und (2.3.6) zusammen, so ergibt sich

$$\begin{aligned}
\varepsilon_k(t+1) &= \bar{P}_k(t+1) - (t+1)p_k \\
&= \left(1 - \frac{k+\delta}{t(2+\delta)+(1+\delta)}\right)\varepsilon_k(t) + \frac{k-1+\delta}{t(2+\delta)+(1+\delta)}\varepsilon_{k-1}(t) \\
&\quad + \kappa_k(t) + \gamma_k(t). & (2.3.18)
\end{aligned}$$

Durch Induktion in  $t$  wollen wir zeigen, dass  $\varepsilon_k(t)$  beschränkt ist im Sinne von (2.3.13). Wir starten den *Induktionsanfang* mit  $t = 1$ . Für  $t = 1$  besteht  $G_1(1)$  aus einer Ecke mit einer Schlaufe. Folglich ist  $\bar{P}_k(1)$  gerade 1 für  $k = 2$  und für alle anderen  $k$  Null, also

$$\bar{P}_k(1) = \mathbb{1}_{\{k=2\}}. \quad (2.3.19)$$

Da  $p_k \leq 1$ , können wir abschätzen

$$|\varepsilon_k(1)| = |\bar{P}_k(1) - p_k| \leq \max\{\bar{P}_k(1), p_k\} \leq 1 \text{ für alle } k \geq 1. \quad (2.3.20)$$

Damit haben wir den Induktionsanfang  $\max_k |\varepsilon_k(1)| \leq C$  gezeigt. Wir fahren fort mit dem *Induktionsschritt*  $t \rightarrow t + 1$ . Zunächst betrachten wir den Fall  $k = 1$ . Wegen  $\varepsilon_0(t) = P_0(t) - p_0 = 0$  (vgl. (2.3.5) und (2.3.7)), vereinfacht sich (2.3.18) zu

$$\varepsilon_1(t+1) = \left(1 - \frac{1+\delta}{t(2+\delta) + (1+\delta)}\right) \varepsilon_1(t) + \kappa_1(t) + \gamma_1(t). \quad (2.3.21)$$

Da  $\frac{1+\delta}{t(2+\delta) + (1+\delta)}$  eine Wahrscheinlichkeit ist, muss gelten

$$1 - \frac{1+\delta}{t(2+\delta) + (1+\delta)} \geq 0 \quad (2.3.22)$$

und folglich

$$|\varepsilon_1(t+1)| \leq \left(1 - \frac{1+\delta}{t(2+\delta) + (1+\delta)}\right) |\varepsilon_1(t)| + |\kappa_1(t)| + |\gamma_1(t)|. \quad (2.3.23)$$

Betrachtet man nun die ausführlichen Schreibweisen zu  $\kappa_k(t)$  und  $\gamma_k(t)$ , also (2.3.16) und (2.3.17), so sieht man nach kurzer Rechnung, dass es Konstanten  $C_\kappa = C_\kappa(\delta)$  und  $C_\gamma = C_\gamma(\delta)$  gibt, so dass für alle  $k \geq 1$  gilt

$$|\kappa_k(t)| \leq C_\kappa(t+1)^{-1} \text{ und } |\gamma_k(t)| \leq C_\gamma(t+1)^{-1}. \quad (2.3.24)$$

Wähle zum Beispiel  $C_\gamma = 1$  und  $C_\kappa = \sup_{k \geq 1} (k+\delta)p_k = (1+\delta)p_1 = \frac{(1+\delta)(2+\delta)}{3+2\delta}$ , was nach kurzer Rechnung folgt, wenn man  $p_k$  wie in (2.3.11) einsetzt. Wir setzen nun in Gleichung (2.3.23) die Induktionsvoraussetzung (2.3.13) und die Abschätzungen aus (2.3.24) ein und erhalten

$$\begin{aligned} |\varepsilon_1(t+1)| &\leq C \left(1 - \frac{1+\delta}{t(2+\delta) + (1+\delta)}\right) + (C_\kappa + C_\gamma)(t+1)^{-1} \\ &= C - \frac{1+\delta}{t(2+\delta) + (1+\delta)} C + (C_\kappa + C_\gamma)(t+1)^{-1}. \end{aligned} \quad (2.3.25)$$

Als nächstes benutzen wir, dass  $t(2+\delta) + (1+\delta) \leq (t+1)(2+\delta)$ , so dass

$$\begin{aligned} |\varepsilon_1(t+1)| &= C - \frac{1+\delta}{(t+1)(2+\delta)} C + (C_\kappa + C_\gamma)(t+1)^{-1} \\ &\leq C - (t+1)^{-1} \left(C \frac{1+\delta}{2+\delta} - (C_\kappa + C_\gamma)\right) \\ &\leq C \end{aligned} \quad (2.3.26)$$

falls

$$C \geq \frac{2+\delta}{1+\delta} (C_\kappa + C_\gamma). \quad (2.3.27)$$

Für  $k \geq 2$  benutzen wir wieder die Abschätzung (2.3.18). Für  $k \leq t(2+\delta)+1$  gilt

$$1 - \frac{k + \delta}{t(2 + \delta) + (1 + \delta)} \geq 0. \quad (2.3.28)$$

Daher wollen wir zunächst annehmen, dass diese Anforderung an  $k$  erfüllt sei und betrachten den Fall  $k > t(2 + \delta) + 2$  später. Mit (2.3.18) und (2.3.28) erhalten wir dann für  $k \geq 2$  und  $\delta > -1$

$$\begin{aligned} \varepsilon_k(t+1) &\leq \left(1 - \frac{k + \delta}{t(2 + \delta) + (1 + \delta)}\right) \varepsilon_k(t) + \frac{k - 1 + \delta}{t(2 + \delta) + (1 + \delta)} \varepsilon_{k-1}(t) \\ &\quad + \kappa_k(t) + \gamma_k(t). \end{aligned} \quad (2.3.29)$$

Wir benutzen nun wieder die Induktionsvoraussetzung  $\max_k |\varepsilon_k(t)| \leq C$  und die Abschätzungen für  $C_\kappa$  und  $C_\gamma$  aus (2.3.24) und erhalten

$$\begin{aligned} \varepsilon_k(t+1) &\leq C \left(1 - \frac{k + \delta}{t(2 + \delta) + (1 + \delta)}\right) + C \frac{k - 1 + \delta}{t(2 + \delta) + (1 + \delta)} + (C_\kappa + C_\gamma)(t+1)^{-1} \\ &= C \left(1 - \frac{1}{t(2 + \delta) + (1 + \delta)}\right) + (C_\kappa + C_\gamma)(t+1)^{-1}. \end{aligned} \quad (2.3.30)$$

Wegen  $t(2 + \delta) + (1 + \delta) \leq (t+1)(2 + \delta)$  ergibt sich daraus

$$\begin{aligned} |\varepsilon_k(t+1)| &\leq C \left(1 - \frac{1}{(t+1)(1 + \delta)}\right) C + (C_\kappa + C_\gamma)(t+1)^{-1} \\ &= C - (t+1)^{-1} \left(\frac{C}{2 + \delta} - (C_\kappa + C_\gamma)\right) \\ &\leq C, \text{ falls } C \geq (2 + \delta)(C_\kappa + C_\gamma) \text{ gilt.} \end{aligned} \quad (2.3.31)$$

Daher wählen wir

$$C = (2 + \delta)(C_\kappa + C_\gamma) \max \{1, (1 + \delta)^{-1}\}. \quad (2.3.32)$$

Abschließend betrachten wir den Fall  $k > t(2 + \delta) + 2$ :

Für  $\delta \geq -1$  ist  $k > t(2 + \delta) + 2 \geq t + 2$ . Der maximale Grad einer Ecke in  $G_1(t+1)$  ist  $t+2$ , nämlich genau dann, wenn alle Ecken mit der ersten Ecke verbunden sind. Für  $k > t(2 + \delta) + 2$  ist daher  $\bar{P}_k(t+1) = 0$  und somit

$$|\varepsilon_k(t+1)| = |\bar{P}(t+1) - (t+1)p_k| = (t+1)p_k \text{ für } k > t + 2. \quad (2.3.33)$$

Aus (2.1.11) und (2.1.12) erhalten wir für eine geeignetes  $C_p$  und gleichmäßig für alle  $k > t + 2$  und  $\delta \geq -1$

$$\begin{aligned} p_k &= \frac{(2 + \frac{\delta}{m})\Gamma(m + 2 + \delta + \frac{\delta}{m})}{\Gamma(m + \delta)} k^{-(3+\delta)} (1 + O(\frac{1}{k})) \\ &= C_p k^{-(3+\delta)} \\ &\leq C_p (t + 1)^{-(3+\delta)}. \end{aligned} \tag{2.3.34}$$

Weiter gilt

$$(1 + t)p_k \leq C_p (t + 1)^{-(2+\delta)} = o((t + 1)^{-1}). \tag{2.3.35}$$

Damit folgt für  $C \geq C_p$  die Behauptung auch für  $k > t(2 + \delta) + 2$ . Damit ist (2.3.13) auch für  $t \rightarrow t + 1$  gezeigt und somit Proposition 2.3.1 für  $m = 1$  und  $\delta > -1$  bewiesen.  $\square$

## 2.4 Beweis von Theorem 1

Abschließend setzen wir die Ergebnisse der vorherigen Abschnitte zum Beweis von Theorem 1 zusammen. Proposition 2.3.1 liefert uns, dass

$$\max_k |\mathbb{E}[P_k(t)] - p_k t| \leq C. \tag{2.4.1}$$

Durch Einsetzen in Proposition 2.2.1 erhalten wir für  $t \rightarrow \infty$

$$\mathbb{P} \left( \max_k |P_k(t) - p_k t| \geq C(1 + \sqrt{t \log t}) \right) = o(1), \tag{2.4.2}$$

und wegen  $p_k(t) = \frac{P_k(t)}{t}$ ,

$$\mathbb{P} \left( \max_k |p_k(t) - p_k| \geq \frac{C}{t} (1 + \sqrt{t \log t}) \right) = o(1). \tag{2.4.3}$$

Gleichung (2.4.3) impliziert die Aussage von Theorem 1.

$\square$



# Kapitel 3

## Durchmesser von PA-Modellen

Im vorherigen Kapitel haben wir uns mit der Verteilung der Grade in einem PA-Graphen beschäftigt und dabei die Power Law-Eigenschaft  $p_k = c_m k^{-\tau}$  als wichtiges Resultat erhalten. Den Exponenten der Power Law  $\tau = 3 + \frac{\delta}{m}$  können wir mit dem Parameter  $\delta$  so variieren, dass  $\tau \in (2, \infty)$  bzw.  $\tau \in (2, 3)$ . Im folgenden definieren wir den *Durchmesser* eines Graphen  $G$  als

$$\text{diam}(G) = \max_{i,j \in G} \{\text{dist}_G(i,j) \mid \text{dist}_G(i,j) \leq \infty\}, \quad (3.0.1)$$

wobei  $\text{dist}_G(i,j)$  den Abstand der Ecken  $i, j \in G$  bezeichne. In [DHH09] werden zwei Vermutungen aufgestellt, die noch nicht bewiesen sind:

- Für  $m \geq 1$  und  $\delta > 0$  konvergiert  $\frac{\text{diam}(G_m(t))}{\log t}$  in Wahrscheinlichkeit gegen eine positive Konstante.
- Für  $m \geq 2$  und  $\delta \in (-m, 0)$  konvergiert  $\frac{\text{diam}(G_m(t))}{\log \log t}$  in Wahrscheinlichkeit gegen eine positive Konstante.

Das folgenden Kapitel bespricht drei Beweise aus [HH08]. In Abschnitt 3.1 wird eine obere und in Abschnitt 3.2 eine untere Schranke für die erste Vermutung gezeigt. In Abschnitt 3.3 wird dann eine obere Schranke für die zweite Vermutung bewiesen. Für die letzten Ergebnisse vor Abgabe der Arbeit sei auf [DHH09] verwiesen, worin zusätzlich zu den Resultaten hier noch eine doppelt-logarithmische untere Schranke gezeigt wird, sowie eine schärfere Schranke zu Theorem 3.

### 3.1 Eine obere Schranke des Durchmessers

Im folgenden Abschnitt seien  $\delta \in (-m, \infty)$  bzw.  $\tau \in (2, \infty)$  und damit der allgemeinste Fall, den wir in diesem Kapitel betrachten wollen, denn für ein

$\delta \leq -m$  würden wir in den Wachstumsregeln der Modelle aus Kapitel 1 keine sinnvollen Wahrscheinlichkeiten erhalten.

**Theorem 2** *In den Modellen (b) und (c) sei  $m \geq 1$  und  $\delta > -m$ . Dann gibt es ein  $C = C(m, \delta) > 0$ , so dass*

$$\mathbb{P}(\text{diam}(G_m(t)) > C \log t) = o(1), \quad t \rightarrow \infty. \quad (3.1.1)$$

Mit anderen Worten ist der Durchmesser von  $G_m(t)$  in den Modellen (b) und (c) whp höchstens  $C \log t$ . Für Modell (a) lässt sich der Beweis nicht führen, da wir dort Schleifen zugelassen haben und der Graph daher nicht notwendiger Weise zusammenhängend sein muss.

**Beweis:** In Modell (c) und mit  $s_i > s_j$  schreiben wir  $s_i \rightarrow s_j$ , wenn die erste Kante von  $s_i$  mit der Ecke  $s_j$  verbunden ist. In Modell (b) schreiben wir  $s_i \rightarrow s_j$ , wenn in  $\{G_1(t)\}_{t \geq 1}$  die Ecke  $(s_i - 1)m + 1$  mit einer der Ecken  $(s_j - 1)m + 1, \dots, s_j m$  verbunden ist. In Modell (a) ist der Graph nicht notwendigerweise zusammenhängend, womit dieser Beweis scheitert. Für  $s_1 = s > s_2 > \dots > s_k = 1$  und  $\vec{s}_k = (s_1, s_2, \dots, s_k)$  schreiben wir

$$E_{\vec{s}_k} = \bigcap_{i=1}^{k-1} \{s_i \rightarrow s_{i+1}\}. \quad (3.1.2)$$

Für eine Realisierung von  $G_m(t)$  bezeichnet  $\text{dist}(s, 1)$  den eindeutigen Wert von  $k$ , so dass  $s = s_1 \rightarrow s_2 \rightarrow \dots \rightarrow s_{k-1} \rightarrow s_k = 1$ . Die Eindeutigkeit von  $k$  resultiert daraus, dass wir das Ereignis  $s_i \rightarrow s_j$  in beiden Modellen jeweils für die erste Kante definiert haben. Mit dieser Definition hat jede Ecke des Graphen einen eindeutigen Weg zur Ecke 1 mittels erster Kanten. Dieses Ereignis beschreibt gerade  $E_{\vec{s}_k}$ . Es gilt

$$\text{diam}(G_m(t)) < 2 \max_{s=1, \dots, t} \text{dist}(s, 1), \quad (3.1.3)$$

da der Abstand zwischen zwei beliebigen Punkten des Graphen geringer ist als die rechte Seite von (3.1.3). Nehmen wir nun an, dass es eine Konstante  $C'$  gibt, so dass für jedes  $1 \leq s \leq t$  und  $\varepsilon > 0$  gilt

$$\mathbb{P}(\text{dist}(s, 1) > C' \log s) \leq s^{-(1+\varepsilon)}. \quad (3.1.4)$$

Dann impliziert (3.1.4) bereits Theorem 2, denn mit (3.1.3) und  $\text{dist}(s, 1) \leq s$ , erhält man für  $C = 2C'$

$$\begin{aligned}
\mathbb{P}(\text{diam}(G_m(t)) > C \log t) &\stackrel{(3.1.3)}{\leq} \mathbb{P}\left(\max_{1 \leq s \leq t} \text{dist}(s, 1) > C' \log t\right) \quad (3.1.5) \\
&= \mathbb{P}\left(\max_{C' \log t \leq s \leq t} \text{dist}(s, 1) > C' \log t\right) \\
&\stackrel{(s \leq t)}{\leq} \mathbb{P}\left(\max_{C' \log t \leq s \leq t} \text{dist}(s, 1) > C' \log s\right) \\
&\leq \sum_{C' \log t \leq s \leq t} \mathbb{P}(\text{dist}(s, 1) > C' \log s) \\
&\stackrel{(3.1.4)}{\leq} \sum_{C' \log t \leq s \leq t} s^{-(1+\varepsilon)} = O((\log t)^{-\varepsilon}) = o(1)
\end{aligned}$$

für  $t \rightarrow \infty$ . Den restlichen Abschnitt wollen wir daher dem Beweis von Gleichung (3.1.4) widmen. Haben wir die gezeigt, folgt wegen (3.1.5) auch Theorem 2. Zum Beweis von (3.1.4) betrachten wir ein festes  $s$  und erhalten mit der Booleschen Ungleichung

$$\begin{aligned}
\mathbb{P}(\text{dist}(s, 1) > k) &= \mathbb{P}\left(\bigcup_{l>k} \{\text{dist}(s, 1) = l\}\right) \\
&= \mathbb{P}\left(\bigcup_{l>k} \bigcup_{\vec{s}_l} \bigcap_{i=1}^{l-1} \{s_i \longrightarrow s_{i+1}\}\right) \\
&\stackrel{B.U.}{\leq} \sum_{l>k} \sum_{\vec{s}_l} \mathbb{P}\left(\bigcap_{i=1}^{l-1} \{s_i \longrightarrow s_{i+1}\}\right), \quad (3.1.6)
\end{aligned}$$

wobei wir mit  $\vec{s}_l$  alle geordneten Vektoren der Länge  $l$  bezeichnen, für die  $s_1 = s$  und  $s_l = 1$  ist. Wir behaupten nun

$$\mathbb{P}(E_{\vec{s}_l}) = \mathbb{P}\left(\bigcap_{i=1}^{l-1} \{s_i \longrightarrow s_{i+1}\}\right) = \prod_{i=1}^{l-1} \mathbb{P}(s_i \longrightarrow s_{i+1}). \quad (3.1.7)$$

Die Unabhängigkeit beweisen wir mittels Induktion. Wegen der unterschiedlichen Wachstumsregeln müssen wir die Modelle (b) und (c) separat betrachten:

**Modell (c)**

$$\mathbb{P} \left( \bigcap_{i=1}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right) = \mathbb{E} \left[ \mathbb{P} \left( \bigcap_{i=1}^{l-1} \{s_i \longrightarrow s_{i+1}\} \mid G_m(s_1 - 1) \right) \right] \quad (3.1.8)$$

$$= \mathbb{E} \left[ I \left[ \bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right] \cdot \mathbb{P}(s_1 \longrightarrow s_2 \mid G_m(s_1 - 1)) \right], \quad (3.1.9)$$

da das Ereignis  $\bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\}$  messbar ist bezüglich  $G_m(s_1 - 1)$  (d. h.  $\bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\} \in \sigma(G_m(s_1 - 1))$ ).  $I[A]$  bezeichne dabei die Indikatorfunktion für das Ereignis  $A$ . Weiter gilt nach der Wachstumsregel (1.3.1)

$$\mathbb{P}(s_1 \longrightarrow s_2 \mid G_m(s_1 - 1)) = \frac{d_{s_2}(s_1 - 1) + \delta}{(2m + \delta)(s_1 - 1)}. \quad (3.1.10)$$

Insbesondere erhalten wir

$$\mathbb{P}(s_1 \longrightarrow s_2) = \mathbb{E} \left[ \frac{d_{s_2}(s_1 - 1) + \delta}{(2m + \delta)(s_1 - 1)} \right]. \quad (3.1.11)$$

Und somit

$$\mathbb{P} \left( \bigcap_{i=1}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right) \stackrel{(3.1.10)}{=} \mathbb{E} \left[ I \left[ \bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right] \frac{d_{s_2}(s_1 - 1) + \delta}{(2m + \delta)(s_1 - 1)} \right] \quad (3.1.12)$$

$$= \mathbb{P} \left( \bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right) \mathbb{E} \left[ \frac{d_{s_2}(s_1 - 1) + \delta}{(2m + \delta)(s_1 - 1)} \right], \quad (3.1.13)$$

da die Zufallsvariable  $d_{s_2}(s_1 - 1)$  nur vom Verlauf des Graphen ab dem Zeitpunkt  $s_2$  abhängt, also insbesondere von der Anzahl der Kanten die nach dem Zeitpunkt  $s_2$  mit der Ecke  $s_2$  verbunden wurden. Dies jedoch ist unabhängig vom Ereignis  $\bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\}$ , welches von den Kanten bis einschließlich dem Zeitpunkt  $s_2$  bestimmt wird. Zusammen mit (3.1.11) erhalten wir

$$\mathbb{P} \left( \bigcap_{i=1}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right) = \mathbb{P}(s_1 \longrightarrow s_2) \cdot \mathbb{P} \left( \bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right). \quad (3.1.14)$$

Wiederholtes Anwenden liefert dann (3.1.7).

**Modell (b)** Gleichung (3.1.8) gilt analog, da das Ereignis  $\bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\}$  auch in Modell (b) bezüglich  $G_m(s_1 - 1)$  messbar ist. Das Ereignis  $\{s_1 \longrightarrow s_2\}$  tritt genau dann ein, wenn in  $G_1(t)$  die Ecke  $m(s_1 - 1) + 1$  mit einer der Ecken  $m(s_2 - 1) + 1, \dots, ms_2$  verbunden ist. Nach der Wachstumsregel (1.2.1) beträgt die Wahrscheinlichkeit dafür

$$\begin{aligned} \mathbb{P}(s_1 \longrightarrow s_2 | G_m(s_1 - 1)) &= \sum_{k=m(s_2-1)+1}^{ms_2} \mathbb{P}(m(s_1 - 1) + 1 \longrightarrow k | G_1(m(s_1 - 1))) \\ &= \sum_{k=m(s_2-1)+1}^{ms_2} \frac{d_k(m(s_1 - 1)) + \delta}{(2 + \delta)(m(s_1 - 1))} \end{aligned} \quad (3.1.15)$$

und damit gilt insbesondere

$$\mathbb{P}(s_1 \longrightarrow s_2) = \mathbb{E} \left[ \frac{\sum_{k=m(s_2-1)+1}^{ms_2} d_k(m(s_1 - 1)) + \delta}{(2 + \delta)(m(s_1 - 1))} \right]. \quad (3.1.16)$$

Wir wissen, dass diese Wahrscheinlichkeit nur von der Verteilung der Kanten in  $G_1(t)$  zum Zeitpunkt  $ms_2$  abhängt und daher unabhängig von dem Ereignis  $\bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\}$  ist. Es folgt analog zur Schlussweise oben

$$\mathbb{P} \left( \bigcap_{i=1}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right) = \mathbb{P}(s_1 \longrightarrow s_2) \cdot \mathbb{P} \left( \bigcap_{i=2}^{l-1} \{s_i \longrightarrow s_{i+1}\} \right) \quad (3.1.17)$$

und durch wiederholtes Anwenden (3.1.7) für Modell (b).

Nachdem wir nun gezeigt haben, dass für  $i = 1, \dots, k - 1$  die Ereignisse  $\{s_i \longrightarrow s_{i+1}\}$  in  $E_{s_k}$  unabhängig sind, wollen wir nun deren Wahrscheinlichkeit abschätzen.

**Lemma 3.1.1** *In Modell (c) gilt für alle  $s > t \geq 1$  und mit  $a = \frac{m}{2m+\delta} \in (0, 1)$ ,*

$$\mathbb{P}(s \longrightarrow t) \leq 2^a t^{-a} (s - 1)^{a-1}. \quad (3.1.18)$$

*Mit  $s, t$  und  $a$  wie oben gilt für Modell (b)*

$$\mathbb{P}(s \longrightarrow t) \leq 2^a \left( t - 1 + \frac{1}{m} \right)^{-a} (s - 1)^{a-1}. \quad (3.1.19)$$

**Beweis:** Wir beweisen die Aussage zunächst für Modell (c). Da nach Konstruktion von Modell (c) die ersten beiden Ecken per Definition miteinander verbunden sind, gilt für  $t = 1$  und  $s = 2$

$$\mathbb{P}(2 \longrightarrow 1) = 1 < 2^a \quad (3.1.20)$$

da  $a \in (0, 1)$ . Für  $s > 2$  gilt nach (3.1.11)

$$\mathbb{P}(s \longrightarrow t) = \mathbb{E} \left[ \frac{d_t(s-1) + \delta}{(2m + \delta)(s-1)} \right]. \quad (3.1.21)$$

Für  $s > 2$  und bedingt auf  $G_m(s-1)$  beschreibt die Zufallsvariable

$$X_{s,t} = d_t(s) - d_t(s-1) \quad (3.1.22)$$

gerade die zufällige Gradänderung der Ecke  $t$  im Zeitschritt von  $s-1$  nach  $s$ . Wir können  $X_{s,t}$  durch  $m$ -maliges unabhängiges Ausführen eines Bernoulli-Experiments mit Eintrittswahrscheinlichkeit  $p$  gewinnen, wobei  $X_{s,t}$ , gegeben  $d_t(s-1)$ , binomialverteilt ist zum Parameter  $m$  mit Erfolgswahrscheinlichkeit  $p = \frac{d_t(s-1) + \delta}{(2m + \delta)(s-1)}$ . Klar ist  $\mathbb{E}[X_{s,t} | d_t(s-1)] = mp$  und wir erhalten

$$\mathbb{E}[d_t(s) + \delta - (d_t(s-1) + \delta) | d_t(s-1)] = m \frac{d_t(s-1) + \delta}{(2m + \delta)(s-1)}. \quad (3.1.23)$$

Zudem ist  $\mathbb{E}[d_t(s-1) | d_t(s-1)] = d_t(s-1)$ . Es folgt

$$\begin{aligned} \mathbb{E}[d_t(s) + \delta | d_t(s-1)] &= d_t(s-1) + m \frac{d_t(s-1) + \delta}{(2m + \delta)(s-1)} \\ &= \mathbb{E}[d_t(s-1) + \delta | d_t(s-1)] \left( 1 + \frac{m}{(2m + \delta)(s-1)} \right), \end{aligned} \quad (3.1.24)$$

und insbesondere  $\mathbb{E}[d_t(s) + \delta] \geq \mathbb{E}[d_t(s-1) + \delta]$ . Wir zeigen nun induktiv

$$\mathbb{E}[d_t(s) + \delta] \leq 2^a (2m + \delta) \left( \frac{s-1}{t} \right)^a. \quad (3.1.25)$$

Als Induktionsanfang wählen wir  $s = t$ . In diesem Fall ist die linke Seite von (3.1.25) nicht größer als  $2m + \delta$ . Aus (3.1.23) und der Induktionsvoraussetzung

(3.1.25) erhalten wir

$$\begin{aligned}
\mathbb{E}[d_t(s+1) + \delta] &\stackrel{(3.1.23)}{=} \mathbb{E}[d_t(s) + \delta] \left(1 + \frac{m}{(2m + \delta)(s)}\right) \\
&\leq \underbrace{2^a(2m + \delta) \left(\frac{s-1}{t}\right)^a}_{I.V.} \left(1 + \frac{a}{s}\right) \\
&= 2^a(2m + \delta) \left(\frac{s}{t}\right)^a \underbrace{\left(1 - \frac{1}{s}\right)^a}_{\leq 1?} \left(1 + \frac{a}{s}\right) \quad (3.1.26)
\end{aligned}$$

Somit reicht es zu zeigen, dass für alle  $x = s^{-1} \in [0, 1]$  und  $a \in (0, 1)$  gilt

$$(1 - x)^a(1 + ax) \leq 1 \quad (3.1.27)$$

was man schnell sieht, wenn man für  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto (1 - x)^a(1 + ax)$  die erste Ableitung und  $f(1) = 1$  betrachtet. Tragen wir nun unsere Ergebnisse zusammen, erhalten wir für alle  $s > t \geq 1$

$$\begin{aligned}
\mathbb{P}(s \longrightarrow t) &\stackrel{(3.1.21)}{=} \frac{\mathbb{E}[d_t(s-1) + \delta]}{(2m + \delta)(s-1)} \stackrel{(3.1.23)}{\leq} \frac{\mathbb{E}[d_t(s) + \delta]}{(2m + \delta)(s-1)} \\
&\stackrel{(3.1.25)}{\leq} \frac{2^a(2m + \delta) \left(\frac{s-1}{t}\right)^a}{(2m + \delta)(s-1)} = 2^a t^{-a} (s-1)^{a-1}, \quad (3.1.28)
\end{aligned}$$

und damit die Behauptung von Lemma 3.1.1 für Modell (c).

Für Modell (b) beschränken wir uns zunächst auf  $m = 1$  und gehen analog zum Beweis für Modell (c) vor. Aus der Wachstumsregel (1.2.1) erhalten wir

$$\mathbb{P}(s \longrightarrow t | G_1(s-1)) = \frac{d_t(s-1) + \delta}{(2 + \delta)(s-1)} \quad (3.1.29)$$

und insbesondere

$$\mathbb{P}(s \longrightarrow t) = \mathbb{E} \left[ \frac{d_t(s-1) + \delta}{(2 + \delta)(s-1)} \right]. \quad (3.1.30)$$

Mithilfe einer Zufallsvariable  $X_{s,t}$  wie in (3.1.22) erhalten wir auch hier

$$\mathbb{E}[d_t(s) + \delta] = \mathbb{E}[d_t(s-1) + \delta] \left(1 + \frac{1}{(2 + \delta)(s-1)}\right), \quad (3.1.31)$$

also insbesondere  $\mathbb{E}[d_t(s) + \delta] \geq \mathbb{E}[d_t(s-1) + \delta]$  und weiter

$$\mathbb{E}[d_t(s) + \delta] \leq 2^a(2 + \delta) \left( \frac{s-1}{t} \right)^a \quad \text{mit } s \geq 2, \quad (3.1.32)$$

wobei  $m = 1$  und  $a = \frac{1}{2+\delta}$ . Damit gilt (3.1.19) auch für Modell (b) mit  $m = 1$ .

Für  $m > 1$  gilt  $\{s \rightarrow t\}$  genau dann, wenn in  $\{G_1(t)\}$  die Ecke  $m(s-1) + 1$  mit einer der Ecken  $m(t-1) + 1, \dots, mt$  verbunden ist. Mit  $a' = 1/(2 + \delta')$  und  $\delta' = \frac{\delta}{m}$  gilt dann

$$\begin{aligned} \mathbb{P}(s \rightarrow t) &= \sum_{j=m(t-1)+1}^{mt} \mathbb{P}(m(s-1) + 1 \rightarrow j) \\ (3.1.30) \quad &\stackrel{=}{=} \sum_{j=m(t-1)+1}^{mt} \frac{\mathbb{E}[d_j(m(s-1)) + \delta']}{(2 + \delta')(m(s-1))} \\ (3.1.31) \quad &\leq \sum_{j=m(t-1)+1}^{mt} \frac{\mathbb{E}[d_j(m(s-1) + 1) + \delta']}{(2 + \delta')(m(s-1))} \\ (3.1.32) \quad &\stackrel{\leq}{=} \sum_{j=m(t-1)+1}^{mt} \frac{2^{a'}(2 + \delta) \left( \frac{m(s-1)}{j} \right)^{a'}}{(2 + \delta')(m(s-1))} \\ &= 2^{a'} \sum_{j=m(t-1)+1}^{mt} j^{-a'} [m(s-1)]^{a'-1} \\ &\leq 2^{a'} m [m(t-1) + 1]^{-a'} [m(s-1)]^{a'-1} \\ &= 2^{a'} \left( t-1 + \frac{1}{m} \right)^{-a'} (s-1)^{a'-1} \quad (3.1.33) \end{aligned}$$

was uns mit  $a' = \frac{1}{2+\delta'} = \frac{m}{2m+\delta}$  die Behauptung (3.1.19) aus Lemma 3.1.1 liefert.  $\square$

Mit diesem Ergebnis können wir den Beweis von Theorem 2 fortsetzen. Wir beschränken uns hier darauf, den Beweis für Modell (c) ausführlich zu zeigen und skizzieren dann lediglich die wesentlichen Unterschiede beim Be-



weis für Modell (b).

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i=1}^{l-1}\{s_i \longrightarrow s_{i+1}\}\right) &\stackrel{(3.1.7)}{=} \prod_{i=1}^{l-1} \mathbb{P}(s_i \longrightarrow s_{i+1}) \\
&\stackrel{(3.1.18)}{\leq} \prod_{i=1}^{l-1} 2^a s_{i+1}^{-a} (s_i - 1)^{a-1} \\
&= 2^{a(l-1)} \prod_{i=1}^{l-1} \left(\frac{s_i - 1}{s_{i+1}}\right)^a \prod_{i=1}^{l-1} \frac{1}{s_i - 1} \\
&\leq 2^{a(l-1)} (s_1 - 1)^{a-1} \prod_{i=2}^{l-1} \frac{1}{s_i - 1} \\
&\leq 2^{a(l-1)} s^{a-1} \prod_{i=2}^{l-1} \frac{1}{s_i - 1}. \tag{3.1.34}
\end{aligned}$$

Ausgehend von Gleichung (3.1.6) erhalten wir damit für  $s_1 = s$  und  $s_l = 1$

$$\mathbb{P}(\text{dist}(s, 1) > k) \leq s^{a-1} \sum_{l>k} 2^{a(l-1)} \sum_{\vec{s}_l} \prod_{i=2}^{l-1} \frac{1}{s_i - 1}. \tag{3.1.35}$$

wobei wir mit  $\vec{s}_l$  über alle geordneten Vektoren  $\vec{s}_l = (s_1, \dots, s_l)$  summieren, d.h. über alle  $\vec{s}_l$ , für die  $s_1 = s > s_2 > \dots > s_l = 1$ . Summieren wir dagegen über alle ungeordneten Vektoren, für die lediglich gelten soll  $s_1 = s$  und  $s_l = 1$  erhalten wir  $(l-2)!$  mal so viele (gleiche) Summanden. Bezeichnen wir mit  $\vec{t}_l$  die ungeordneten Vektoren, reicht es, durch  $(l-2)!$  zu teilen, und es gilt

$$\mathbb{P}(\text{dist}(s, 1) > k) \leq s^{a-1} \sum_{l>k} \frac{2^{a(l-1)}}{(l-2)!} \sum_{\vec{t}_l} \prod_{i=2}^{l-1} \frac{1}{s_i - 1}. \tag{3.1.36}$$

Summieren wir über alle Vektoren  $\vec{t}_l$ , anstatt nur über die mit unterschiedlichen Koordinaten erhalten wir

$$\sum_{\vec{t}_l} \prod_{i=2}^{l-1} \frac{1}{s_i - 1} = \left(\sum_{u=1}^{s-1} \frac{1}{u}\right)^{l-2}. \tag{3.1.37}$$

Dies ist natürlich eine obere Schranke der Summe aus (3.1.36), so dass

$$\mathbb{P}(\text{dist}(s, 1) > k) \leq 2^{1-a} s^{a-1} \sum_{l>k} \frac{2^{a(l-1)}}{(l-2)!} \left(\sum_{u=1}^{s-1} \frac{1}{u}\right)^{l-2} \tag{3.1.38}$$

Nun benutzen wir  $\sum_{u=1}^{s-1} \frac{1}{u} \leq 1 + \log s$  und erhalten

$$\mathbb{P}(\text{dist}(s, 1) > k) \leq 2s^{a-1} \sum_{l>k} \frac{(2^a(1 + \log s))^{l-2}}{(l-2)!}.$$

Wenn wir den Zählindex der Summe anpassen, erhalten wir gerade die Verteilungsfunktion einer poissonverteilten Zufallsvariable  $X$  mit Parameter  $\lambda = 2^a(1 + \log s)$  an der Stelle  $X = k - 1$ , also

$$\begin{aligned} \mathbb{P}(\text{dist}(s, 1) > k) &\leq 2s^{a-1} \sum_{l \geq k-1} \frac{(2^a(1 + \log s))^l}{l!} \\ &= 2s^{a-1} \exp(2^a(1 + \log s)) \mathbb{P}(X \geq k-1) \\ &= cs^{2^a+a-1} \mathbb{P}(X \geq k-1), \end{aligned} \quad (3.1.39)$$

wobei  $c = 2 \exp(2^a)$ . Nach [JLR00] (Remark 2.5 und Remark 2.6) gilt für eine poissonverteilte Zufallsvariable  $Y$  mit Erwartungswert  $\lambda$

$$\mathbb{P}(Y > 5\lambda) = \mathbb{P}(Y > \mathbb{E}[Y] + 4\lambda) \leq \exp\left(-\frac{8\lambda^2}{\lambda + \frac{4}{3}\lambda}\right) = e^{-\frac{24\lambda}{7}}. \quad (3.1.40)$$

Wählen wir nun in (3.1.39)  $\lambda = 2^a(1 + \log s)$ , liefert das

$$\begin{aligned} \mathbb{P}(\text{dist}(s, 1) > 5 \cdot (2^a(1 + \log s) + 1)) &\leq cs^{1+a} \mathbb{P}(X \geq 5\lambda) \\ &\leq cs^{1+a} e^{-\frac{24\lambda}{7}} \\ &= cs^{1+a} e^{-\frac{24 \cdot 2^a(1 + \log s)}{7}} \\ &= \frac{cs^{1+a}}{e^{\frac{24 \cdot 2^a}{7}} s^{\frac{24 \cdot 2^a}{7}}} \\ &\leq \frac{cs^2}{e^{\frac{24}{7}} s^{\frac{24}{7}}} \\ &= \frac{c}{e^{\frac{24}{7}} s^{\frac{10}{7}}} \\ &\leq cs^{-1+\varepsilon} \end{aligned} \quad (3.1.41)$$

wobei wir benutzt haben, dass  $0 < a < 1$  und  $\varepsilon = \frac{3}{7}$  wählen. Damit haben wir den letzten Schritt zum Beweis von Gleichung (3.1.4) gezeigt und folglich Theorem 2 für Modell (c) bewiesen. Der Beweis für Modell (b) unterscheidet sich lediglich darin, dass wir in (3.1.34) die Abschätzung (3.1.19) aus Lemma 3.1.1 einsetzen.  $\square$

**Bemerkung:** Betrachten wir die Wachstumsregeln für die Modelle (a) und (b), also (1.1.1),(1.1.2) und (1.2.1), können wir unmittelbar schließen, dass (3.1.19) aus Lemma 3.1.1 auch für Modell (a) gilt. Darauf werden wir im folgenden Abschnitt beim Beweis von Lemma 3.2.2 noch zurückkommen.

## 3.2 Eine untere Schranke des Durchmessers

Das Ziel dieses Abschnitts ist es, eine untere Schranke für den Durchmesser eines PA-Graphen zu erhalten. Dabei untersuchen wir einen Pfad  $\Gamma$  der Länge  $l$ , der zwei Ecken  $s_0$  und  $s_l$  von  $G_m(t)$  miteinander verbindet. Für die einzelnen Kanten von  $\Gamma$  berechnen wir die Kantenwahrscheinlichkeiten und werden damit die Wahrscheinlichkeit abschätzen, dass ein solcher Pfad in  $G_m(t)$  existiert. Anschließend können wir mit der Anzahl aller möglichen  $\Gamma$  der Länge  $l$  die mittlere Anzahl solcher Teilgraphen errechnen. Wir werden feststellen, dass für  $l < L = \frac{\log(t-1)}{\log(3Cm^2 \log t)}$  die mittlere Anzahl der Teilgraphen gegen Null konvergiert für  $t \rightarrow \infty$ . Daraus schließen wir, dass der Durchmesser des PA-Graphens mindestens Größe  $L$  hat.

**Theorem 3** *In den Modellen (a) bis (c) seien  $m \geq 1$  und  $\delta \geq 0$ . Dann gilt für  $\varepsilon > 0$  und  $t \rightarrow \infty$ ,*

$$\mathbb{P} \left( \text{diam}(G_m(t)) < (1 - \varepsilon) \frac{\log t}{\log \log t} \right) = o(1). \quad (3.2.1)$$

*D. h. mit hoher Wahrscheinlichkeit ist der Durchmesser von  $G_m(t)$  mindestens  $(1 - \varepsilon) \frac{\log t}{\log \log t}$ .*

Für  $\delta = 0$  wurde das Theorem bereits 2004 von Bollobás und Riordan bewiesen (vgl. Theorem 1 in [BR04]). In [HH07] wird der Beweis auf  $\delta \geq 0$  erweitert werden, was wir im folgenden vorstellen. Für Modell (c) bezeichne

$$\{g(t, j) = s\}, \quad 1 \leq j \leq m, \quad (3.2.2)$$

das Ereignis, dass zum Zeitpunkt  $t$  die  $j$ -te Kante der Ecke  $t$  mit einer bereits vorhandenen Ecke  $s$  verbunden ist. In den Modellen (a) und (b) bedeutet dieses Ereignis, dass im Graphen  $G_1(mt)$  die Kante der Ecke  $m(t-1) + j$  mit einer der Ecken  $m(s-1) + 1, \dots, ms$  verbunden ist. Aus der Definition der PA-Modelle können wir folgern, dass das Eintreten dieses Ereignisses die Präferenz für die Ecke  $s$  erhöht und somit die Präferenz der Ecken  $u$  mit  $1 \leq u \leq t$ ,  $u \neq s$  relativ verringert. Intuitiv lässt sich dieser Effekt auch beschreiben, indem wir sagen, dass für  $s_1 \neq s_2, j_1 \neq j_2, t_1 \neq t_2$  die Ereignisse  $\{g(t_1, j_1) = s_1\}$  und  $\{g(t_2, j_2) = s_2\}$  negativ korreliert sind. Um dies zu beweisen, beschreibe für  $n_s \in \mathbb{N}, n_s \geq 1$

$$E_s = \bigcap_{i=1}^{n_s} \left\{ g(t_i^{(s)}, j_i^{(s)}) = s \right\} \quad (3.2.3)$$

das Ereignis, dass zum Zeitpunkt  $t_i$  die  $j_i$ -te Kante der Ecke  $t_i$  mit der bereits vorhandenen Ecke  $s$  verbunden ist. Wir beginnen zunächst mit dem

Beweis, dass für jedes  $k \geq 1$  und alle möglichen  $t_i^{(s)}, j_i^{(s)}$  die Ereignisse  $E_s$  für unterschiedliche  $s$  negativ korreliert sind:

**Lemma 3.2.1** *Für unterschiedliche  $s_1, s_2, \dots, s_k$ , gilt*

$$\mathbb{P} \left( \bigcap_{i=1}^k E_{s_i} \right) \leq \prod_{i=1}^k \mathbb{P} (E_{s_i}). \quad (3.2.4)$$

**Beweis:** Wir betrachten die Kanten im Graphen entsprechend der Reihenfolge ihres Hinzufügens und nummerieren sie von 1 bis  $mt$ . So können wir einem Ereignis  $\{g(t, j) = s\}$  seine Kantenummer  $m(t-1) + j$  zuordnen. Im folgenden führen wir eine Vollständige Induktion in der höchsten Kantenummer  $e$  der Ereignisse  $E_s$  durch. Die Induktionsvoraussetzung ist dabei, dass (3.2.4) für alle  $k$  und alle möglichen  $t_i^{(s)}, j_i^{(s)}$  gilt und  $\max_{i,s}(t_i^{(s)} - 1) + j_i^{(s)} \leq e$ .

**Induktionsstart:** Wir starten die Induktion mit  $e = 1$ . Trivialer Weise ist die Induktionsvoraussetzung hier erfüllt, da  $\bigcap_{i=1}^k E_{s_i}$  entweder leer ist oder gerade aus einem Ereignis besteht.

**Induktionsschritt:** Um die Induktion fortzusetzen, nehmen wir an, dass die Induktionsvoraussetzung für  $e - 1$  erfüllt sei, also dass (3.2.4) für alle  $k$  und alle möglichen  $t_i^{(s)}, j_i^{(s)}$  gelte, wobei  $\max_{i,s}(t_i^{(s)} - 1) + j_i^{(s)} \leq e - 1$ . Ziel ist es nun, die Aussage so zu erweitern, dass (3.2.4) gilt für alle  $k$ , alle möglichen  $t_i^{(s)}, j_i^{(s)}$  und  $\max_{i,s}(t_i^{(s)} - 1) + j_i^{(s)} \leq e$ .

Der Fall  $\max_{i,s}(t_i^{(s)} - 1) + j_i^{(s)} < e$  folgt direkt aus der Induktionsannahme. Daher reicht es,  $\max_{i,s}(t_i^{(s)} - 1) + j_i^{(s)} = e$  zu betrachten. Zunächst stellen wir fest, dass die Wahl von  $t$  und  $j$  eindeutig ist, falls gelten soll  $m(t-1) + j = e$ . In diesem Fall gibt es wiederum zwei Möglichkeiten: Entweder gibt es genau eine Wahl für  $s$  und  $t_i^{(s)}, j_i^{(s)}$ , so dass  $t_i^{(s)} = t$  und  $j_i^{(s)} = j$ , oder es gibt mindestens zwei solcher Möglichkeiten. Im zweiten Fall erhalten wir direkt  $\bigcap_{i=1}^k E_{s_i} = \emptyset$ , da die  $e$ -te Kante nur mit einer eindeutigen Ecke verbunden sein kann. Hier ist folglich nichts zu zeigen. Wir untersuchen daher den Fall einer eindeutigen Ecke  $s$  und  $t_i^{(s)} = t$  sowie  $j_i^{(s)} = j$ . Bezeichne

$$E'_s = \bigcap_{i=1: (t_i^{(s)}, j_i^{(s)}) \neq (t, j)}^{n_s} \left\{ g(t_i^{(s)}, j_i^{(s)}) = s \right\} \quad (3.2.5)$$

die Beschränkung des Ereignisses  $E_s$  auf die übrigen Kanten. Dann können

wir schreiben

$$\bigcap_{i=1}^k E_{s_i} = \{g(t, j) = s\} \cap E'_s \cap \bigcap_{i=1, s_i \neq s}^k E_{s_i}. \quad (3.2.6)$$

Per Konstruktion ist die Kantennummer der Ereignisse in  $E'_s \cap \bigcap_{i=1, s_i \neq s}^k E_{s_i}$  höchstens  $e - 1$ . Also ist

$$\mathbb{P} \left( \bigcap_{i=1}^k E_{s_i} \right) \leq \mathbb{E} \left[ I[E'_s \cap \bigcap_{i=1, s_i \neq s}^k E_{s_i}] \mathbb{P}_{e-1}(g(t, j) = s) \right], \quad (3.2.7)$$

wobei  $\mathbb{P}_{e-1}$  die bedingte Wahrscheinlichkeit für den Fall bezeichne, dass die Kanten bis einschließlich  $(e - 1)$  gegeben sind. Wir betrachten zunächst Modell (c), für das wir aus der Wachstumsregel erhalten, dass

$$\mathbb{P}_{e-1}(g(t, j) = s) = \frac{d_s(t - 1) + \delta}{(2m + \delta)(t - 1)}, \quad (3.2.8)$$

wobei

$$d_s(t - 1) = m + \sum_{(t', j'): t' \leq t-1} I[g(t', j') = s]. \quad (3.2.9)$$

Als Zwischenschritt notieren wir

$$\begin{aligned} \mathbb{P} \left( \bigcap_{i=1}^k E_{s_i} \right) &\leq \mathbb{E} \left[ \frac{m + \delta}{(2m + \delta)(t - 1)} I[E'_s \cap \bigcap_{i=1, s_i \neq s}^k E_{s_i}] \right] \\ &+ \mathbb{E} \left[ \frac{\sum_{(t', j'): t' \leq t-1} I[g(t', j') = s]}{(2m + \delta)(t - 1)} I[E'_s \cap \bigcap_{i=1, s_i \neq s}^k E_{s_i}] \right]. \end{aligned} \quad (3.2.10)$$

Jeder der Terme in (3.2.9) hat eine Kantennummer strikt kleiner  $e$  und tritt mit einer nichtnegativen multiplikativen Konstante auf. Damit können wir die Induktionsvoraussetzung auf jeden einzelnen dieser Terme anwenden und erhalten mit  $m + \delta \geq 0$

$$\begin{aligned} \mathbb{P} \left( \bigcap_{i=1}^k E_{s_i} \right) &\leq \frac{m + \delta}{(2m + \delta)(t - 1)} \mathbb{P}(E'_s) \prod_{i=1, s_i \neq s}^k \mathbb{P}(E_{s_i}) \\ &+ \frac{\sum_{(t', j'): t' \leq t-1} \mathbb{P}(\{g(t', j') = s\} \cap E'_s)}{(2m + \delta)(t - 1)} \prod_{i=1, s_i \neq s}^k \mathbb{P}(E_{s_i}). \end{aligned} \quad (3.2.11)$$

Setzen wir die Terme nun wieder zusammen, erhalten wir

$$\mathbb{P} \left( \bigcap_{i=1}^k E_{s_i} \right) \leq \mathbb{E} \left[ I[E'_s] \frac{d_s(t-1) + \delta}{(2m + \delta)(t-1)} \right] \prod_{i=1: s_i \neq s}^k \mathbb{P}(E_{s_i}), \quad (3.2.12)$$

was uns schließlich die Behauptung liefert, denn

$$\mathbb{E} \left[ I[E'_s] \frac{d_s(t-1) + \delta}{(2m + \delta)(t-1)} \right] = \mathbb{P}(E_s). \quad (3.2.13)$$

Der Beweis des Lemmas für die Modelle (a) und (b) soll hier nicht weiter ausgeführt werden. Er unterscheidet sich lediglich darin, dass wir in (3.2.8) die alternativen Formulierungen (1.1.3) und (1.2.2) benutzen und dann analog fortfahren.  $\square$

Mit dem folgenden Lemma können wir die Wahrscheinlichkeit der Ereignisse  $E_s$  abschätzen. Wir beschränken uns dabei auf  $n_s \leq 2$ :

**Lemma 3.2.2** *Für die Modelle (a) bis (c) gibt es  $M_1, M_2 \in \mathbb{R}_+$ , so dass*

(i) *für jedes  $1 \leq j \leq m$  und  $t > s$ ,*

$$\mathbb{P}(g(t, j) = s) \leq \frac{M_1}{t^{1-a} s^a}, \quad (3.2.14)$$

(ii) *für  $t_2 > t_1 > s$ , und alle  $1 \leq j_1, j_2 \leq m$  gilt*

$$\mathbb{P}(g(t_1, j_1) = s, g(t_2, j_2) = s) \leq \frac{M_2}{(t_1 t_2)^{1-a} s^{2a}}, \quad (3.2.15)$$

wobei wie bisher  $a = \frac{m}{2m+\delta}$ .

**Beweis:** Für Modell (c) können wir aus Lemma 3.1.1 direkt schließen, dass

$$\begin{aligned} \mathbb{P}(g(t, j) = s) &= \mathbb{P}(g(t, s) = 1) = \mathbb{P}(t \longrightarrow s) \\ &= 2^a s^{-a} (t-1)^{a-1} \leq M_1 \frac{t^{a-1}}{s^a}, \end{aligned} \quad (3.2.16)$$

sofern gegeben ist, dass  $M_1 \geq 2^a \sup_{t \geq 2} (\frac{t}{t-1})^{1-a} = 2$ . In den Modellen (a) und (b) soll gelten, dass

$$\begin{aligned}
\mathbb{P}(g(t, j) = s) &= \sum_{i=1}^m \mathbb{P}((m(t-1) + j \longrightarrow m(s-1) + i)) \\
&\stackrel{(3.1.19)}{\leq} \sum_{i=1}^m 2^a (m(s-1) + i)^{-a} (m(t-1) + j - 1)^{a-1} \\
&\leq m (2^a (m(s-1) + 1)^{-a} (m(t-1) + j - 1)^{a-1}) \\
&= 2^a \frac{(t-1 + \frac{j-1}{m})^{a-1}}{(s-1 + \frac{1}{m})^a} \leq M_1 \frac{t^{a-1}}{s^a}, \tag{3.2.17}
\end{aligned}$$

was wir erreichen, wenn wir  $M_1 \geq 2m^a$  wählen.

Wir zeigen nun (3.2.15) und beginnen mit Modell (c). Der Beweis für die anderen Modelle verläuft dann ähnlich. Zunächst leisten wir noch etwas Vorarbeit indem wir (3.1.23) mehrfach auf  $\mathbb{E}[(d_s(t_2 - 1) + \delta)]$  anwenden. Wir erhalten

$$\begin{aligned}
\mathbb{E}[(d_s(t_2 - 1) + \delta)] &= \mathbb{E}[(d_s(t_2 - 2) + \delta)] \left(1 + \frac{a}{t_2 - 2}\right) \\
&= \mathbb{E}[(d_s(t_1) + \delta)] \left(1 + \frac{a}{t_1}\right) \cdots \left(1 + \frac{a}{t_2 - 2}\right) \\
&= \mathbb{E}[(d_s(t_1) + \delta)] \frac{\Gamma(t_2 - 1 + a)\Gamma(t_1)}{\Gamma(t_2 - 1)\Gamma(t_1 + a)}. \tag{3.2.18}
\end{aligned}$$

Im nächsten Schritt nutzen wir  $\frac{\Gamma(t_1)}{\Gamma(t_1 + a)} \sim \frac{1}{t_1^a}$  und  $\frac{\Gamma(t_2 - 1 + a)}{(t_2 - 1)\Gamma(t_2 - 1)} \sim \frac{1}{t_2^{1-a}}$  (vgl. (0.0.3)) und können mit einem geeigneten  $M_3 > 0$  die linke Seite von (3.2.15) abschätzen durch

$$\begin{aligned}
&\mathbb{P}(g(t_1, j_1) = s, g(t_2, j_2) = s) \\
&= \mathbb{E}[\mathbb{P}(g(t_1, j_1) = s, g(t_2, j_2) = s | G_m(t_2 - 1))] \\
&= \mathbb{E} \left[ I[g(t_1, j_1) = s] \left( \frac{d_s(t_2 - 1) + \delta}{(t_2 - 1)(2m + \delta)} \right) \right] \\
&\stackrel{(3.2.18)}{=} \frac{1}{(t_2 - 1)(2m + \delta)} \frac{\Gamma(t_2 - 1 + a)\Gamma(t_1)}{\Gamma(t_2 - 1)\Gamma(t_1 + a)} \mathbb{E} [I[g(t_1, j_1) = s](d_s(t_1) + \delta)] \\
&\leq \frac{M_3}{t_1^a t_2^{1-a}} \mathbb{E} [I[g(t_1, j_1) = s](d_s(t_1) + \delta)]. \tag{3.2.19}
\end{aligned}$$

Bleibt noch  $\mathbb{E} [I[g(t_1, j_1) = s](d_s(t_1) + \delta)]$  zu berechnen:

$$\begin{aligned}
& \mathbb{E} [I[g(t_1, j_1) = s](d_s(t_1) + \delta) \mid G_m(t_1 - 1)] \\
&= \mathbb{E} [I[g(t_1, j_1) = s]d_s(t_1) - d_s(t_1 - 1) \mid G_m(t_1 - 1)] \\
&\quad + \mathbb{E} [I[g(t_1, j_1) = s](d_s(t_1 - 1) + \delta) \mid G_m(t_1 - 1)] \\
&= \sum_{j=1}^m \mathbb{E} [I[g(t_1, j_1) = s]I[g(t_1, j) = s] \mid G_m(t_1 - 1)] \\
&\quad + (d_s(t_1 - 1) + \delta)\mathbb{E} [I[g(t_1, j_1) = s] \mid G_m(t_1 - 1)] \\
&= \frac{d_s(t_1 - 1) + \delta}{(t_1 - 1)(2m + \delta)} + (m - 1) \left( \frac{d_s(t_1 - 1) + \delta}{(t_1 - 1)(2m + \delta)} \right)^2 + \frac{(d_s(t_1 - 1) + \delta)^2}{(t_1 - 1)(2m + \delta)} \\
&\leq \frac{d_s(t_1 - 1) + \delta}{(t_1 - 1)(2m + \delta)} + m \frac{(d_s(t_1 - 1) + \delta)^2}{(t_1 - 1)(2m + \delta)}, \tag{3.2.20}
\end{aligned}$$

womit insbesondere gilt

$$\mathbb{E} [I[g(t_1, j_1) = s](d_s(t_1) + \delta)] \leq \frac{\mathbb{E}[d_s(t_1 - 1) + \delta]}{(t_1 - 1)(2m + \delta)} + m \frac{\mathbb{E}[(d_s(t_1 - 1) + \delta)^2]}{(t_1 - 1)(2m + \delta)}. \tag{3.2.21}$$

Als nächstes benutzen wir eine Abschätzung des zweiten Moments der Gradfolge. Zugunsten der Übersicht haben wir das zugehörige Lemma 4.1.1 in Kapitel 4 bewiesen. Danach gibt es eine Konstante  $M_4 > 0$ , so dass

$$\mathbb{E} [(d_s(t) + \delta)^2] \leq M_4 \left( \frac{t}{s} \right)^{2a}. \tag{3.2.22}$$

Zusammen mit der Abschätzung von  $\mathbb{E}[d_s(t_1 - 1) + \delta]$  aus (3.1.25) erhalten wir damit aus (3.2.21), dass

$$\mathbb{E} [I[g(t_1, j_1) = s](d_s(t_1) + \delta)] \leq \frac{2^a(2m + \delta) \left( \frac{t_1 - 2}{s} \right)^a + mM_4 \left( \frac{t_1 - 1}{s} \right)^{2a}}{(t_1 - 1)(2m + \delta)}. \tag{3.2.23}$$

Sofern nötig passen wir  $M_4$  entsprechend an und erhalten schließlich aus (3.2.19)

$$\begin{aligned}
\mathbb{P}(g(t_1, j_1) = s, g(t_2, j_2) = s) &\leq \frac{M_3}{t_1^a t_2^{1-a}} \left[ \frac{mM_4}{(t_1 - 1)(2m + \delta)} \left( \frac{(t_1 - 1)}{s} \right)^{2a} \right] \\
&\leq \frac{M_2}{(t_1 t_2)^{1-a} s^{2a}} \tag{3.2.24}
\end{aligned}$$



für ein  $M_2 > 0$  und damit die Behauptung.  $\square$

Im folgenden nennen wir einen Pfad  $\Gamma = (s_0, s_1, \dots, s_l)$  *überschneidungsfrei*, wenn  $s_i \neq s_j$  für alle  $1 \leq i < j \leq l$ . Sei nun  $\Gamma$  ein überschneidungsfreier Pfad in  $G_m(t)$  bestehend aus den  $l + 1$  ungeordneten Ecken  $s_0, s_1, \dots, s_l$ . Da  $\Gamma$  überschneidungsfrei ist, können wir ihn als einen Kombination von  $k$  Ereignissen der Form (3.2.14) oder (3.2.15) beschreiben. Mit den beiden Lemmata 3.2.1 und 3.2.2 lässt sich dann die Wahrscheinlichkeit für die Existenz von  $\Gamma$  in  $G_m(t)$  abschätzen:

**Korollar 3.2.1** *Sei  $\Gamma = (s_0, s_1, \dots, s_l)$  ein überschneidungsfreier Pfad der Länge  $l$ . Es gilt*

$$\mathbb{P}(\Gamma \in G_m(t)) \leq (m^2 C)^l \prod_{i=0}^{l-1} \frac{1}{(s_i \wedge s_{i+1})^a (s_i \vee s_{i+1})^{1-a}}. \quad (3.2.25)$$

**Beweis:** Da  $\Gamma$  überschneidungsfrei ist, können wir schreiben

$$\{\Gamma \in G_m\} = \cap_{i=1}^k E_{s_i} \quad (3.2.26)$$

wobei wir die Ereignisse  $E_{s_i}$  definieren wie in (3.2.3) mit  $n_s \leq 1, 2$ . Es gilt also entweder

$$E_s = \{g(t, j) = s\} \quad (3.2.27)$$

für  $t > s$  und ein  $1 \leq j \leq m$  oder

$$E_s = \{g(t_1, j_1) = g(t_2, j_2) = s\} \quad (3.2.28)$$

für  $t_1, t_2 > s$  und  $1 \leq j_1, j_2 \leq m$ . Im ersten Fall können wir aus (3.2.14) schließen, dass

$$\mathbb{P}(E_s) = \mathbb{P}(g(t, j) = s) \leq \frac{M_1}{t^{1-a} s^a}, \quad (3.2.29)$$

Aus (3.2.15) folgt für den zweiten Fall

$$\mathbb{P}(E_s) = \mathbb{P}(g(t_1, j_1) = s, g(t_2, j_2) = s) \leq \frac{M_2}{t_1^{1-a} s^a t_2^{1-a} s^a}. \quad (3.2.30)$$

In beiden Fällen sind  $M_i$ ,  $i = 1, 2$  Konstante. Lemma 3.2.1 liefert dann (3.2.25) wobei sich der Faktor  $m^{2l}$  aus der Wahl von  $j \in 1, 2, \dots, m$  ergibt.  $\square$

Im letzten Schritt bauen wir die Aussage des Korollars zu Theorem 3 aus. Für natürliche Zahlen  $j > i$  gilt

$$\left(\frac{j}{i}\right)^{a-\frac{1}{2}} \leq 1 \quad (3.2.31)$$

genau dann, wenn  $a \leq \frac{1}{2}$ . Aus der Definition  $a = \frac{m}{2m+\delta}$  folgt, dass dies gerade dann zutrifft, wenn  $\delta \geq 0$  ist. Multiplizieren wir (3.2.31) mit  $(ij)^{-\frac{1}{2}}$  erhalten wir  $j^{a-1}i^{-a} \leq (ij)^{-\frac{1}{2}}$  mit  $\delta \geq 0$ . Für (3.2.25) folgt dann

$$\mathbb{P}(\Gamma \in G_m(t)) \leq (m^2 C)^l \prod_{i=0}^{l-1} \frac{1}{\sqrt{s_i s_{i+1}}}. \quad (3.2.32)$$

Als letzten Schritt wollen wir nun zeigen, dass wir aus (3.2.32) mit  $\delta \geq 0$  folgern können, dass

$$L = \frac{\log(t-1)}{\log(3Cm^2 \log t)} \quad (3.2.33)$$

eine untere Schranke für den Durchmesser von  $G_m(t)$  ist. Wir betrachten einen Pfad der Länge  $l$ , bestehend aus den Ecken  $s_0, s_1, \dots, s_l$ . Dann liefert uns (3.2.32)

$$\mathbb{P}(\Gamma \in G_m(t)) \leq (Cm^2)^l \prod_{i=0}^{l-1} \frac{1}{\sqrt{s_i s_{i+1}}} = \frac{(Cm^2)^l}{\sqrt{s_0 s_l}} \prod_{i=1}^{l-1} \frac{1}{s_i}. \quad (3.2.34)$$

Damit wird die mittlere Anzahl von Pfaden zwischen  $s_0 = t$  und  $s_l = t-1$  beschränkt durch

$$\begin{aligned} \frac{(Cm^2)^l}{\sqrt{t(t-1)}} \sum_{1 \leq s_1, \dots, s_{l-1} \leq t-2} \prod_{i=1}^{l-1} \frac{1}{s_i} &= \frac{(Cm^2)^l}{\sqrt{t(t-1)}} \left( \sum_{i=1}^{t-2} \frac{1}{s_i} \right)^{l-1} \\ &\leq \frac{(Cm^2)^l}{t-1} (\log t)^{l-1} \end{aligned} \quad (3.2.35)$$

wobei wir mit  $1 \leq s_1, \dots, s_{l-1} \leq t-2$  über alle möglichen  $\Gamma$  der Form  $(t, s_1, \dots, s_{l-1}, t-1)$  summiert haben. Für  $(2Cm^2 \log t)^l \leq t-1$  oder äquivalent, für

$$l \leq \frac{\log(t-1)}{\log((2Cm^2) \log t)} \quad (3.2.36)$$

ist

$$\frac{(Cm^2)^l}{t-1} (\log t)^{l-1} \leq \left(\frac{1}{2}\right)^l (\log t)^{-1} \longrightarrow 0. \quad (3.2.37)$$

Für  $l < L$  geht die mittlere Anzahl von Pfaden die  $t$  und  $t-1$  miteinander verbinden also gegen Null. Mit der First-Moments-Method können wir dann schließen, dass auch die Wahrscheinlichkeit der Existenz von  $\Gamma$  gegen Null geht. Für  $l > L$  dagegen ist ein solcher Graph whp vorhanden.

### 3.3 Doppelt logarithmische obere Schranke des Durchmessers

In diesem Abschnitt haben wir den Parameter  $\delta$  des PA-Modells wiederum leicht eingeschränkt. Im folgenden soll stets gelten, dass  $\delta \in (-m, 0)$ . Theorem 4 liefert uns dann eine doppelt-logarithmische obere Schranke für den Durchmesser des Graphen.

**Theorem 4** *In den Modellen (a) bis (c) sei  $m \geq 2$  und  $\delta \in (-m, 0)$ . Dann ist für jedes  $\sigma > \frac{1}{3-\tau}$  und*

$$C_G = \frac{4}{|\log(\tau - 2)|} + \frac{4\sigma}{\log m} \quad (3.3.1)$$

der Durchmesser von  $G_m(t)$  whp nach oben beschränkt durch  $C_G \log \log t$  mit  $t \rightarrow \infty$ .

Der Beweis des Theorems gliedert sich in zwei Teile. Zunächst geben wir in Theorem 5 eine Schranke für den Durchmesser des Kerns eines PA-Graphen an. Dabei verstehen wir unter dem *Kern* eines PA-Graphen diejenigen Ecken, die mindestens einen Grad in Höhe einer bestimmten Potenz von  $\log t$  haben. Im zweiten Schritt leiten wir in Theorem 6 dann eine obere Schranke für den Abstand einer Ecke mit „kleinem“ Grad und dem Kern des Graphen her. Im weiteren Verlauf wird es zweckmäßiger sein, Theorem 4 für  $2t$  statt für  $t$  zu zeigen, was jedoch nichts an den Resultaten ändert.

#### 3.3.1 Der Durchmesser des Kerns

Wir erinnern uns aus Abschnitt 2, dass

$$\tau = 3 + \frac{\delta}{m}, \quad (3.3.2)$$

so dass für  $-m < \delta < 0$  folgt  $\tau \in (2, 3)$ . In diesem gesamten Abschnitt setzen wir  $m \geq 2$ .

Zunächst führen wir einige neue Begriffe ein, die wir im Verlauf des Beweises brauchen werden: Für  $\sigma > \frac{1}{3-\tau} = -\frac{m}{\delta} > 1$  und ein PA-Modell mit Graphen  $G_m(2t)$  bezeichne

$$\text{Core}_t = \{i \in \{1, 2, \dots, t\} : d_i(t) \geq (\log t)^\sigma\}, \quad (3.3.3)$$

den *Kern* des PA-Modells. Mit anderen Worten umfasst der Kern alle Ecken des Graphen, die zum Zeitpunkt  $t$  mindestens Grad  $(\log t)^\sigma$  haben. Für einen

Graphen  $G$  mit der Eckenmenge  $\{1, 2, \dots, t\}$  und einer gegebenen Kantenmenge bezeichnen wir mit  $d_G(i, j)$  die Länge des kürzesten Wegs zwischen  $i$  und  $j$ . Weiter sei für  $A \subseteq \{1, 2, \dots, t\}$

$$\text{diam}_t(A) = \max_{i, j \in A} d_{G(t)}(i, j). \quad (3.3.4)$$

Dann bezeichnet  $\text{diam}_{2t}(\text{Core}_t)$  den Durchmesser des Kerns von  $G_m(2t)$ .

Bereits in Abschnitt 3.1 haben wir mit dem Ausdruck  $\text{dist}(i, j)$  einen Abstandsbegriff zwischen zwei Ecken  $i$  und  $j$  eingeführt. Der wesentliche Unterschied ist, dass wir mit  $d_G(i, j)$  nun den kürzesten Weg zwischen zwei Ecken benutzen, während wir bei  $\text{dist}(i, j)$  strikt die erste Kante einer Ecke benutzt haben. Für  $m = 1$  gilt natürlich  $\text{dist}(i, j) = d_G(i, j)$ .

Mit dem folgenden Theorem können wir für  $\text{diam}_{2t}(\text{Core}_t)$  eine obere Schranke angeben:

**Theorem 5** (Der Durchmesser des Kerns) *Sei  $m \geq 2$ . Für jedes  $\sigma > \frac{1}{3-\tau}$ , ist whp*

$$\text{diam}_{2t}(\text{Core}_t) \leq (1 + o(1)) \frac{4 \log \log t}{|\log(\tau - 2)|}. \quad (3.3.5)$$

Den Beweis von Theorem 5 werden wir in zwei Schritte aufteilen, Proposition 3.3.1 und Proposition 3.3.2. Wir definieren den *inneren Kern*  $\text{Inner}_t$  des Graphen  $G$  durch

$$\text{Inner}_t = \left\{ i \in \{1, 2, \dots, t\} : d_i(t) \geq t^{\frac{1}{2(\tau-1)}} (\log t)^{-\frac{1}{2}} \right\}. \quad (3.3.6)$$

Mit Proposition 3.3.1 beweisen wir die Aussage, dass der Durchmesser des inneren Kern whp nicht größer ist als 10. Anschließend untersuchen wir mit Proposition 3.3.2 den *äußeren Kern*  $\text{Outer}_t$  des Graphen, gegeben durch  $\text{Outer}_t = \text{Core}_t \setminus \text{Inner}_t$ . Wir werden zeigen, dass der Abstand des äußeren Kerns  $\text{Outer}_t$  zum inneren Kern  $\text{Inner}_t$  beschränkt wird durch eine Konstante mal  $\log \log t$ . Dies liefert uns zudem das Ergebnis, dass auch der Durchmesser des äußeren Kerns beschränkt wird durch eine Konstante mal  $\log \log t$ .

**Proposition 3.3.1** (Der Durchmesser des inneren Kerns) *Sei  $m \geq 2$ . Dann gilt whp, dass*

$$\text{diam}_{2t}(\text{Inner}_t) \leq \frac{2(\tau - 1)}{3 - \tau} + 6. \quad (3.3.7)$$

**Beweis:** Zunächst führen wir den Begriff der  $t$ -Verbindung zwischen einer Ecke  $i \in \{1, 2, \dots, t\}$  und einer Menge von Ecken  $A \subseteq \{1, 2, \dots, t\}$  ein: Sei  $A$  eine Menge von Ecken und  $i$  eine einzelne Ecke. Wir bezeichnen eine Ecke  $j \in \{t+1, t+2, \dots, 2t\}$  als  $t$ -Verbindung zwischen  $i$  und  $A$ , wenn eine der ersten beiden Kanten von  $j$  inzident zu  $i$  ist und die andere der ersten beiden Kanten von  $j$  inzident zu einer Ecke in  $A$  ist. Aus der Existenz einer  $t$ -Verbindung zwischen  $i$  und  $A$  können wir schließen, dass der Abstand zwischen  $i$  und  $A$  in  $G_m(2t)$  höchstens 2 ist.

Wir betrachten zunächst Modell (c) und darin eine Menge von Ecken  $A$  und eine einzelne Ecke  $i$  die zum Zeitpunkt  $t$  Grad  $d_i(t)$  hat. Wir berechnen nun bedingt auf  $G_m(t)$  die Wahrscheinlichkeit, dass eine Ecke  $j \in \{t+1, t+2, \dots, 2t\}$  eine  $t$ -Verbindung für  $i$  und  $A$  bildet. Dafür multiplizieren wir die Wahrscheinlichkeit, dass mindestens eine Ecke aus  $A$  mit  $j$  verbunden ist, mit der Wahrscheinlichkeit, dass  $i$  mit  $j$  verbunden ist. Schreiben wir für  $\sum_{i \in A} d_i(t) = d_A(t)$ , erhalten wir so

$$\mathbb{P}(j \text{ ist eine } t\text{-Verbindung}) = \frac{(d_A(t) + \delta |A|)(d_i(t) + \delta)}{(2t(2m + \delta))^2}, \quad (3.3.8)$$

und dies unabhängig davon, ob bereits eine andere Ecke im Graphen als  $t$ -Verbindung fungiert oder nicht. Da der Grad einer Ecke immer mindestens  $m$  beträgt, können wir schreiben  $d_i(t) + \delta \geq m + \delta > 0$  für alle  $i \leq t$ . Wegen  $\delta < 0$  erhalten wir

$$d_i(t) + \delta = d_i(t) \left(1 + \frac{\delta}{d_i(t)}\right) \geq d_i(t) \left(1 + \frac{\delta}{m}\right) = d_i(t) \frac{m + \delta}{m} \quad (3.3.9)$$

und folglich

$$d_A(t) + \delta |A| = \sum_{i \in A} d_i(t) \frac{m + \delta}{m} \geq d_A(t) \frac{m + \delta}{m}. \quad (3.3.10)$$

Zur Abkürzung setzen wir  $\eta = \frac{(m+\delta)^2}{(2m(2m+\delta))^2} > 0$ . Als Ergebnis halten wir fest, dass die Wahrscheinlichkeit, dass  $j \in \{t+1, t+2, \dots, 2t\}$  eine  $t$ -Verbindung zwischen  $i$  und  $A$  bildet, mindestens

$$\frac{\eta d_A(t) d_i(t)}{t^2} \quad (3.3.11)$$

ist, unabhängig davon, ob die anderen Ecken  $t$ -Verbindungen bilden oder nicht. Wir untersuchen nun eine Teilmenge von  $\text{Inner}_t$  der Größe  $n_t = \lfloor \sqrt{t} \rfloor$  und zeigen, dass diese whp einen beschränkten Durchmesser hat. Nach (4.2.13)

enthält  $\text{Inner}_t$  whp mindestens  $\sqrt{t}$  Ecken. Wir bezeichnen die Menge der ersten  $\lfloor \sqrt{t} \rfloor$  Ecken von  $\text{Inner}_t$  mit  $I$ . Aus (3.3.11) können wir nun schließen, dass für  $i_1, i_2 \in I$  und  $j \in \{t+1, t+2, \dots, 2t\}$ , die Wahrscheinlichkeit, dass  $j$  eine  $t$ -Verbindung für  $i_1, i_2$  ist, mindestens

$$\frac{\eta t^{\frac{1}{\tau-1}}}{t^2 \log t} \geq \frac{t^{\frac{1}{\tau-1}-2}}{(\log t)^2} =: q_t \quad (3.3.12)$$

beträgt, und zwar unabhängig davon, ob bereits andere Ecken  $t$ -Verbindungen sind oder nicht. Im folgenden benutzen wir ein Coupling und vergleichen so die Menge  $I$  und alle möglichen Paare von Ecken von  $I$ , für die  $j \in \{t+1, t+2, \dots, 2t\}$  eine  $t$ -Verbindung ist, mit einem *multinomialen* Graphen  $H_{n_t}$ .  $H_{n_t}$  habe  $n_t$  Ecken. Seine Kanten bestimmen wir wie folgt: Wir nummerieren die  $e_t = \frac{n_t(n_t-1)}{2} \sim \frac{t}{2}$  Paare aus  $I$  beliebig von 1 bis  $e_t$  und betrachten sie als  $e_t$  Zellen eines Multinomialsexperiments mit  $t$  Versuchen und den Erfolgswahrscheinlichkeiten

$$p_k = q_t \text{ für } 1 \leq k \leq e_t \text{ und } p_0 = 1 - e_t q_t. \quad (3.3.13)$$

Die  $t$  Versuche stellen wir durch Zufallsvektoren

$$N_j = (N_{j,1}, N_{j,2}, \dots, N_{j,e_t}) \quad 1 \leq j \leq t, \quad (3.3.14)$$

mit Verteilung

$$\mathbb{P}(N_j = 1_i) = q_t \text{ und } \mathbb{P}(N_j = 0) = 1 - e_t q_t \quad (3.3.15)$$

dar, wobei wir mit  $1_i$  den  $i$ -ten Einheitsvektor der Länge  $e_t$  bezeichnen und mit 0 den Nullvektor. Falls nun die Zelle  $k$  des Multinomialsexperiments nicht leer ist, also falls  $\sum_{j=1}^t N_{j,k} > 0$ , dann zeichnen wir die Kante mit der Nummer  $k$  in den Graphen  $H_{n_t}$  ein. Ist die Zelle  $k$  dagegen leer, soll die Kante  $k$  in  $H_{n_t}$  nicht vorhanden sein. Die 0-Zelle brauchen wir, um diejenigen Versuche zu zählen, die wir nicht den Zellen  $1, 2, \dots, e_t$  zugerechnet haben.

Nach der Konstruktion von  $H_{n_t}$  und (3.3.13) ist der Abstand von zwei Ecken aus  $I$  in  $G_m(t)$  höchstens zwei mal so groß, wie der Abstand der korrespondierenden Ecken in  $H_{n_t}$ . Nach Lemma 4.3, Kapitel 4, ist der Durchmesser von  $H_{n_t}$  höchstens so groß wie der Durchmesser eines *gleichmäßigen* Erdős-Rényi-Graphen  $G(n_t, M_t)$ , mit  $n_t$  Ecken und  $M_t$  Kanten, wobei

$$M_t = \frac{1}{2} e_t (1 - (1 - q_t)^t) \quad (3.3.16)$$

gerade die mittlere Anzahl von Kanten in  $H_{n_t}$  ist. Mit [JLR00], Kapitel 1.4, kann man schließen, dass der gleichmäßige Erdős-Rényi-Graph  $G(n_t, M_t)$

asymptotisch äquivalent zu einem klassischen *binomialen* Erdős-Rényi-Graphen mit  $G(n_t, \lambda_t)$  ist, wobei wir die Kantenwahrscheinlichkeit  $\lambda_t$  definieren durch

$$\lambda_t = \frac{1}{2} \left(1 - (1 - q_t)^t\right) \sim \frac{t^{\frac{1}{\tau-1}-1}}{2(\log t)^2}. \quad (3.3.17)$$

Im nächsten Schritt wollen wir zeigen, dass  $\text{diam}(G(n_t, p_t))$  whp beschränkt ist durch  $\frac{\tau-1}{3-\tau} + 1$ . Dafür benutzen wir eine Aussage aus [Bol01] (Corollary 10.12), welche uns eine scharfe Schranke für den Durchmesser eines Erdős-Rényi-Graphen liefert: Falls für einen Erdős-Rényi-Graphen  $G(n, p)$  und  $d \in \mathbb{N}$  gilt, dass

$$p^d n^{d-1} - 2 \log n \rightarrow \infty, \quad (3.3.18)$$

und

$$p^{d-1} n^{d-2} - 2 \log n \rightarrow -\infty, \quad (3.3.19)$$

dann ist whp  $\text{diam}(G(n, p)) = d$ .

In unserem Fall ist  $n = n_t = \sqrt{t}$  und  $p = \lambda_t$ , woraus folgt, dass whp  $\frac{\tau-1}{3-\tau} < d \leq \frac{\tau-1}{3-\tau} + 1$ . Wenn der Durchmesser von  $G(n_t, \lambda_t)$  whp durch  $\frac{\tau-1}{3-\tau} + 1$  beschränkt wird, folgt wegen der  $t$ -Verbindungen, dass whp

$$\text{diam}_{2t}(I) \leq 2 \left( \frac{\tau-1}{3-\tau} + 1 \right). \quad (3.3.20)$$

Zuletzt zeigen wir, dass die Wahrscheinlichkeit, dass es keine  $t$ -Verbindung zwischen  $i \in \text{Inner}_t \setminus I$  und der Menge  $I$  gibt, lediglich klein ist. Mit hoher Wahrscheinlichkeit ist

$$d_I(t) \geq \sqrt{t} t^{\frac{1}{\tau-1}} (\log t)^{-\frac{1}{2}} \quad \text{und} \quad d_i(t) \geq t^{\frac{1}{\tau-1}} (\log t)^{-\frac{1}{2}}. \quad (3.3.21)$$

Nach (3.3.11) ist die Wahrscheinlichkeit, dass es keine  $t$ -Verbindung zwischen  $i$  und  $I$  gibt, nach oben beschränkt durch

$$\begin{aligned} \left(1 - \frac{\eta d_I(t) d_i(t)}{t^2}\right)^t &\leq \exp\left(-\eta \frac{d_I(t) d_i(t)}{t}\right) \\ &= \exp\left(-\eta \frac{t^{\frac{1}{\tau-1}-\frac{1}{2}}}{\log t}\right) = o(t^{-1}) \end{aligned} \quad (3.3.22)$$

für  $\tau < 3$ . Wir haben somit gezeigt, dass es whp eine  $t$ -Verbindung zwischen  $i$  und  $I$  gibt und daher der Abstand zwischen  $i \in \text{Inner}_t \setminus I$  und der Menge  $I$  höchstens 2 beträgt. Zusammen mit (3.3.20) erhalten wir die Behauptung von Proposition 3.3.1.  $\square$

**Proposition 3.3.2** (Der Abstand zwischen äußerem und innerem Kern) *Sei  $m \geq 2$ . Mit hoher Wahrscheinlichkeit gilt, dass*

$$\max_{i \in \text{Outer}_t} \min_{j \in \text{Inner}_t} d_{G_m(2t)}(i, j) \leq \frac{2 \log \log t}{|\log(\tau - 2)|}. \quad (3.3.23)$$

Anders formuliert sagt uns die Proposition, dass im Graphen  $G_m(2t)$  der innere Kern  $\text{Inner}_t$  von jeder Ecke des äußeren Kerns  $\text{Outer}_t$  whp mit nicht mehr als  $\frac{2 \log \log t}{|\log(\tau - 2)|}$  Kanten erreicht werden kann.

**Beweis:** Der Beweis gliedert sich in drei Schritte. Zunächst definieren wir Mengen  $\mathcal{N}^{(k)}$ ,  $k \in \mathbb{N}$ , in denen wir Ecken mit einem Grad von mindestens  $u_k$  zusammenfassen. Dabei sei  $\mathcal{N}^{(1)} = \text{Inner}_t$ . Anschließend zeigen wir, dass man bei geeigneter Wahl der  $u_k$  whp nicht mehr als zwei Schritte benötigt, um von  $\mathcal{N}^{(k-1)}$  nach  $\mathcal{N}^{(k)}$  zu gelangen. Abschließend definieren wir ein  $k^* \in \mathbb{N}$ , so dass  $\text{Outer}_t \subseteq \mathcal{N}^{(k^*)}$ , was schließlich zur Behauptung führen wird.

Wir erinnern uns, dass wir definiert hatten

$$\text{Outer}_t = \text{Core}_t \setminus \text{Inner}_t. \quad (3.3.24)$$

Weiter setzen wir

$$\mathcal{N}^{(1)} = \text{Inner}_t = \{i : d_i(t) \geq u_1\} \quad (3.3.25)$$

wobei

$$u_1 = l_t = t^{\frac{1}{2(\tau-1)}} (\log t)^{-\frac{1}{2}}. \quad (3.3.26)$$

Wir definieren nun rekursiv eine Folge  $u_k$ ,  $k \geq 2$ , so dass für jede Ecke  $i \in \{1, 2, \dots, t\}$  mit einem Grad von mindestens  $u_k$  die Wahrscheinlichkeit, dass es keine  $t$ -Verbindung zwischen  $i$  und der Menge

$$\mathcal{N}^{(k-1)} = \{j : d_j(t) \geq u_{k-1}\} \quad (3.3.27)$$

gibt, bedingt auf  $G_m(t)$  lediglich sehr klein ist. Denn nach (3.3.22) und Lemma 4.2.1 ist diese Wahrscheinlichkeit höchstens

$$\begin{aligned} \exp\left(-\eta \frac{d_{\mathcal{N}^{(k-1)}}(t) d_i(t)}{t}\right) &= \exp\left(-\eta \frac{\sum_{j: d_j(t) \geq u_{k-1}} d_j(t) u_k}{t}\right) \\ &\stackrel{(4.2.1)}{=} \exp\left(-\eta \frac{Bt(u_{k-1})^{2-\tau} u_k}{t}\right). \end{aligned} \quad (3.3.28)$$

Definieren wir nun

$$u_k = D \log t (u_{k-1})^{\tau-2} \quad (3.3.29)$$



ergibt sich daraus

$$\begin{aligned} \exp\left(-\eta \frac{Bt(u_{k-1})^{2-\tau} u_k}{t}\right) &= \exp(-\eta BD \log t) \\ &= o(t^{-1}) \end{aligned} \quad (3.3.30)$$

wenn wir  $D > (\eta B)^{-1}$  wählen. Das folgende Lemma soll uns helfen  $u_k$  näher zu bestimmen:

**Lemma 3.3.3** (Bestimmung von  $u_k$ ) *Für jedes  $k \in \mathbb{N}$  ist*

$$u_k = D^{a_k} (\log t)^{b_k} t^{c_k} \quad (3.3.31)$$

wobei

$$a_k = \frac{1 - (\tau - 2)^{k-1}}{3 - \tau}, \quad b_k = \frac{1 - (\tau - 2)^{k-1}}{3 - \tau} - \frac{(\tau - 2)^{k-1}}{2}, \quad (3.3.32)$$

$$c_k = \frac{(\tau - 2)^{k-1}}{2(\tau - 1)}. \quad (3.3.33)$$

**Beweis:** Wir bestimmen  $a_k$ ,  $b_k$  und  $c_k$  rekursiv. Für  $k = 1$  erhalten wir aus (3.3.26)

$$a_1 = 0, \quad b_1 = -\frac{1}{2}, \quad c_1 = \frac{1}{2(\tau - 1)}. \quad (3.3.34)$$

Benutzen wir (3.3.29), können wir  $a_k$ ,  $b_k$  und  $c_k$  jeweils auf ihre Vorgänger  $a_{k-1}$ ,  $b_{k-1}$  und  $c_{k-1}$  zurückführen und erhalten

$$a_k = 1 + (\tau - 2)a_{k-1}, \quad b_k = 1 + (\tau - 2)b_{k-1}, \quad c_k = (\tau - 2)c_{k-1}, \quad (3.3.35)$$

was man schnell sieht, wenn man schreibt

$$\begin{aligned} u_k &= D \log t (u_{k-1})^{\tau-2} \\ &= D \log t (D^{a_{k-1}} (\log t)^{b_{k-1}} t^{c_{k-1}})^{\tau-2}. \end{aligned} \quad (3.3.36)$$

Wiederholtes Anwenden von (3.3.35) liefert schließlich (3.3.32).  $\square$

**Lemma 3.3.4** (Konnektivität zwischen  $\mathcal{N}^{(k-1)}$  und  $\mathcal{N}^{(k)}$ ) *Seien  $m, k \geq 2$ . Die Wahrscheinlichkeit, dass es ein  $i \in \mathcal{N}^{(k)}$  gibt, das in  $G_m(2t)$  einen Abstand von mehr als zwei von  $\mathcal{N}^{(k-1)}$  hat, ist  $o(t^{-1})$ .*

**Beweis:** Nach Lemma 4.2 gilt mit einer Wahrscheinlichkeit von mehr als  $1 - o(t^{-1})$ ,

$$\sum_{i \in \mathcal{N}^{(k-1)}} d_i(t) \geq Bt[u_{k-1}]^{2-\tau}. \quad (3.3.37)$$

Nehmen wir an, (3.3.37) sei eingetreten. Mithilfe von (3.3.22) können wir, bedingt auf  $G_m(t)$ , die Wahrscheinlichkeit berechnen, dass es ein  $i \in \mathcal{N}^{(k)}$  gibt, für das es keine  $t$ -Verbindung zwischen  $i$  und  $\mathcal{N}^{(k-1)}$  gibt. Mit der Booleschen Ungleichung erhalten wir dann

$$te^{-\eta B[u_{k-1}]^{2-\tau} u_k} = te^{-\eta B D \log t} = o(t^{-1}), \quad (3.3.38)$$

wobei wir (3.3.29) benutzt haben und  $D > 2(\eta B)^{-1}$  wählen.  $\square$

Wir setzen nun den Beweis von Proposition 3.3.2 fort. Wir setzen

$$k^* = \left\lfloor \frac{\log \log t}{|\log(\tau - 2)|} \right\rfloor. \quad (3.3.39)$$

Nach Lemma 3.3.4 ist der Abstand zwischen  $\mathcal{N}^{(k^*)}$  und  $\text{Inner}_t$  höchstens  $2k^*$ . Es ist also zu zeigen, dass

$$\text{Outer}_t \subseteq \mathcal{N}^{(k^*)} \quad (3.3.40)$$

oder mit der Definition von  $\text{Outer}_t$  (vgl. (3.3.3)) und  $\mathcal{N}^{(k^*)}$

$$\{i : d_i(t) \geq (\log t)^\sigma\} \subseteq \{i : d_i(t) \geq u_{k^*}\}, \quad (3.3.41)$$

was gerade der Fall ist, wenn  $(\log t)^\sigma \geq u_{k^*}$  für alle  $\sigma > \frac{1}{3-\tau}$ . Nach Lemma 3.3.3 ist

$$u_{k^*} = D^{a_{k^*}} (\log t)^{b_{k^*}} t^{c_{k^*}}. \quad (3.3.42)$$

Für  $k^* + 1$  rechnen wir

$$\begin{aligned} a_{k^*+1} &= \frac{1 - \exp\left(\frac{\log \log t}{|\log(\tau-2)|} \log(\tau-2)\right)}{3-\tau} = \frac{1 - \frac{1}{\log t}}{3-\tau} \xrightarrow{t \rightarrow \infty} \frac{1}{3-\tau} \\ b_{k^*+1} &= \frac{1 - \frac{1}{\log t}}{3-\tau} - \frac{\exp\left(\frac{\log \log t}{|\log(\tau-2)|} \log(\tau-2)\right)}{2} = \frac{1 - \frac{1}{\log t}}{3-\tau} - \frac{1}{2 \log t} \xrightarrow{t \rightarrow \infty} \frac{1}{3-\tau} \\ c_{k^*+1} &= \frac{1}{2(\tau-1) \log t} \xrightarrow{t \rightarrow \infty} 0 \end{aligned} \quad (3.3.43)$$

Insbesondere ist  $t^{c_{k^*}} = O(1) = (\log t)^{o(1)}$ ,  $(\log t)^{b_{k^*}} = (\log t)^{\frac{1}{3-\tau}+o(1)}$  und  $D^{a_{k^*}} = (\log t)^{o(1)}$  und damit

$$u_{k^*} = (\log t)^{\frac{1}{3-\tau}+o(1)} \text{ für } t \rightarrow \infty. \quad (3.3.44)$$

Wählen wir  $t$  hinreichend groß, ist  $\sigma \geq \frac{1}{3-\tau} + o(1)$ , womit wir Proposition 3.3.2 bewiesen haben.  $\square$

**Beweis:** (*Theorem 5*) Wir führen nun unsere Resultate zusammen zum Beweis von Theorem 5. In Proposition 3.3.1 haben wir gesehen, dass der Durchmesser des inneren Kerns  $\text{Inner}_t$  whp kleiner ist als 10, sofern  $\sigma > \frac{1}{3-\tau}$ . Proposition 3.3.2 liefert uns, dass whp jede Ecke des inneren Kerns  $\text{Inner}_t$  von jeder Ecke des äußeren Kerns  $\text{Outer}_t$  mit nicht mehr als  $k^* = \frac{2 \log \log t}{|\log(\tau-2)|}$  Schritten erreicht werden kann. Insgesamt brauchen wir also whp in  $\text{Core}_t = \text{Inner}_t \cup \text{Outer}_t$  nicht mehr als  $10 + 2k^*$  Kanten um zwei beliebige Ecken des Kerns  $\text{Core}_t$  zu verbinden. Das ist die Aussage von Theorem 5.  $\square$

### 3.3.2 Die Verbindung von Kern und Peripherie

In diesem Abschnitt erweitern wir die Resultate aus dem vorherigen Abschnitt und betrachten den Abstand zwischen den Ecken, die nicht im Kern  $\text{Core}_t$  liegen und dem Kern.

**Theorem 6** (Die Verbindung von Kern und Peripherie) *Sei  $m \geq 2$ . Für jedes  $\sigma > \frac{1}{3-\tau}$  ist der maximale Abstand zwischen einer beliebigen Ecke des Graphen und  $\text{Core}_t$  in  $G_m(2t)$  whp nach oben beschränkt durch  $2\sigma \frac{\log \log t}{\log m}$ .*

Wir beweisen Theorem 6 in zwei Schritten. Den ersten bildet Proposition 3.3.5, nach der für zwei Ecken  $i, j \in \{1, 2, \dots, t\}$  der Abstand  $d_{G_m(2t)}(i, j)$  beschränkt ist durch eine Konstante mal  $\log \log t$ . Das heißt,  $\text{diam}_{2t}(G_m(t))$  ist beschränkt durch eine Konstante mal  $\log \log t$ . Der zweiten Schritt, Proposition 3.3.8, zeigt uns dann, dass der Abstand zwischen einer beliebigen Ecke aus  $\{t+1, t+2, \dots, 2t\}$  und einer beliebigen Ecke aus  $\{1, 2, \dots, t\}$  whp beschränkt ist durch eine andere Konstante mal  $\log \log t$ .

**Proposition 3.3.5** (Die Verbindung der Hälfte der Peripherie mit dem Kern) *Für jedes  $\sigma > \frac{1}{3-\tau}$  ist whp der maximale Abstand zwischen einer beliebigen Ecke aus  $\{1, 2, \dots, t\}$  und  $\text{Core}_t$  in  $G_m(2t)$  nach oben beschränkt durch  $\sigma \frac{\log \log t}{\log m}$ .*

**Beweis:** Wir beginnen mit einer Ecke  $i \in \{1, 2, \dots, t\}$  und wollen zeigen, dass die Wahrscheinlichkeit, dass der Abstand zwischen  $i$  und  $\text{Core}_t$  mehr

als  $C \log \log t$  ist, mit  $C = \frac{\sigma}{\log m}$ , lediglich klein ist, genauer gesagt sogar nur  $o(t^{-1})$ . Das beweist unsere Behauptung.

Wir erkunden die Nachbarschaft von  $i$  wie folgt: Von  $i$  ausgehend setzen wir  $m \geq 2$  Kanten. Von jeder der so erreichten  $m$  benachbarten Ecken setzen wir wieder jeweils  $m$  Kanten, jedoch nur zu Ecken, die wir bisher nicht erkundet haben. So fahren wir fort. Den so gestarteten Prozess nennen wir  $k$ -*Erkundungsbaum*, wenn die erkundeten Ecken einen Abstand von bis zu  $k$  von der Ausgangsecke  $i$  haben.

Sofern wir niemals zwei Kanten mit der gleichen Ecke verbinden, können wir in  $k$  Schritten  $m^k$  Ecken erreichen. Ziel ist es nun, den Baum soweit wachsen zu lassen, bis er mindestens eine Ecke aus  $\text{Core}_t$  enthält. Das Ereignis, dass eine Kante zu einer Ecke führt, die bereits Teil des Erkundungsbaums ist, nennen wir *Kollision*. Mit steigenden  $k$  steigt auch die Wahrscheinlichkeit einer Kollision. Wie wir jedoch im folgenden zeigen werden, ist die Wahrscheinlichkeit gering, dass in dem  $k$ -Erkundungsbaum vor dem Erreichen des Kerns eine Ecke gibt, die mehr als zwei Kollisionen hat.

**Lemma 3.3.6** (Eine Schranke für die Wahrscheinlichkeit mehrerer Kollisionen) *Sei  $m \geq 2$ ,  $\delta \in (-m, 0)$ ,  $C = \frac{\sigma}{\log m}$ ,  $l \geq 1$ ,  $b \in (0, 1]$ . Wir wählen  $k \leq C \log \log t$ . Für jede Ecke  $i \in \{1, 2, \dots, t\}$  ist die Wahrscheinlichkeit, dass der in  $i$  gestartete  $k$ -Erkundungsbaum mindestens  $l$  Kollisionen hat, bevor er  $\text{Core}_t \cup \{j : j \leq t^b\}$  erreicht, nach oben beschränkt durch*

$$\left( (\log t)^d t^{-b} \right)^l = (\log t)^{dl} t^{-bl}, \quad (3.3.45)$$

für  $\delta > 0$ .

**Beweis:** Wir wählen  $i \in \{\lceil t^b \rceil + 1, \lceil t^b \rceil + 2, \dots, t\}$  und betrachten seinen  $k$ -Erkundungsbaum  $\mathcal{T}_i^{(k)}$ . Da wir alle Kanten des Baums erst nach dem Zeitpunkt  $t^b$  hinzugefügt haben, ist der Nenner in den Wachstumsregeln der Modelle (vgl. Abschnitte 1.1 bis 1.3) mindestens  $t^b$ . Außerdem hat jede Ecke, bevor der  $k$ -Erkundungsbaum  $\text{Core}_t$  erreicht, einen Grad kleiner als  $(\log t)^\sigma$ , denn bei einem höheren Grad wäre  $\mathcal{T}_i^{(k)} \cap \text{Core}_t \neq \emptyset$  und damit der Kern bereits erreicht. Daher ist die Wahrscheinlichkeit für das im Lemma beschriebene Ereignis für  $l = 1$  höchstens

$$\sum_{v \in \mathcal{T}_i^{(k)}} \frac{d_v(t) + \delta}{t^b} \leq \sum_{v \in \mathcal{T}_i^{(k)}} \frac{(\log t)^\sigma}{t^b} \leq \frac{m^k (\log t)^\sigma}{t^b}, \quad (3.3.46)$$

wobei wir benutzt haben, dass  $\delta < 0$  und  $|\mathcal{T}_i^{(k)}| \leq m^k$ . Für allgemeines  $l$  lautet die obere Schranke dann

$$\left( \frac{m^k (\log t)^\sigma}{t^b} \right)^l. \quad (3.3.47)$$

Wenn  $k = C \log \log t$  mit  $C = \frac{\sigma}{\log m}$ , erhalten wir schließlich, dass  $m^{lk} = (\log t)^{l\sigma}$ , was man leicht sieht, wenn man die Definition der allgemeinen Potenz benutzt (siehe z.B. [AE02], S. 296). Somit gilt die Behauptung in Lemma 3.3.6 für  $d = 2\sigma$ .  $\square$

Im nächsten Schritt zeigen wir, dass whp  $\{j : j \leq t^b\}$  eine Teilmenge des Kerns ist:

**Lemma 3.3.7** (Frühe Ecken haben große Grade) *Sei  $m \geq 1$ . Dann gibt es ein  $b > 0$ , so dass whp*

$$\min_{j \leq t^b} d_j(t) \geq (\log t)^\sigma, \quad (3.3.48)$$

für ein  $\sigma > \frac{1}{3-\tau}$ . Also gilt insbesondere whp  $\{j : j \leq t^b\} \subseteq \text{Core}_t$ .

Wir verschieben den Beweis auf Abschnitt 4.4 und fahren gleich fort mit dem Beweis von Proposition 3.3.5:

Kombinieren wir nun die Lemmata 3.3.6 und 3.3.7. Wir berechnen die Wahrscheinlichkeit, dass es ein  $i \in \{1, 2, \dots, t\}$  gibt, für das der Erkundungsbaum  $\mathcal{T}_i^{(k)}$  mindestens  $l$  Kollisionen hat, bevor er auf den Kern trifft. Mit der Booleschen Ungleichung gilt

$$\sum_{i=1}^t (\log t)^{dl} t^{-bl} = (\log t)^{2\sigma l} t^{-bl+1} = o(1), \quad (3.3.49)$$

falls  $l > \frac{1}{b}$ . Wenn der  $k$ -Erkundungsbaum auf den Kern trifft, sind wir nach Theorem 5 fertig. Trifft der  $k$ -Erkundungsbaum von einer Ecke  $i$  dagegen nicht auf den Kern, hat aber auch weniger als  $l$  Kollisionen, dann hat der Baum mindestens  $m^{k-l}$  Ecken. Denn, wenn es höchstens  $l$  Kollisionen gibt, erhalten wir die minimale Größe des Baums indem wir die  $l$  Ecken identifizieren und ihren Ursprung. Die Größe des „belaubten“ Baums ist dann mindestens  $m^{k-l}$ .

Wenn  $k = C \log \log t$  mit  $C = \frac{\sigma}{\log m}$  ist die Anzahl der Ecken des Baums mindestens  $(\log t)^{\sigma+o(1)}$ . Das gesamte Gewicht des Kerns ist nach 4.2.1 mindestens

$$\sum_{i \in \text{Core}_t} (d_i(t) + \delta) \geq Bt(\log t)^{-(\tau-2)\sigma}. \quad (3.3.50)$$

Die Wahrscheinlichkeit, dass es keine  $t$ -Verbindung zwischen dem  $k$ -Erkundungsbaum und dem Kern gibt, können wir mit (3.3.22) nach oben abschätzen durch

$$\exp\left(\frac{\eta B t (\log t)^{-(\tau-2)\sigma} (\log t)^{\sigma+o(1)}}{t}\right) = o(t^{-1}), \quad (3.3.51)$$

falls wir  $B$  hinreichend groß wählen, da  $\sigma > \frac{1}{3-\tau}$ . Damit ist der Beweis komplett.  $\square$

**Proposition 3.3.8** *Sei  $m \geq 2$ . Für  $\sigma > \frac{1}{3-\tau}$  ist der maximale Abstand zwischen jeder Ecke und  $\text{Core}_t \cup \{1, 2, \dots, t\}$  in  $G_m(t)$  whp nach oben beschränkt durch  $\frac{\sigma \log \log t}{\log m}$ .*

**Beweis:** Sei  $k = \left\lfloor \frac{\sigma \log \log t}{\log m} \right\rfloor - 1$ . Wir lassen den  $k$ -Erkundungsbaum erneut wachsen und zwar nun von den Ecken  $i \in \{t+1, t+2, \dots, 2t\}$ .

Nach Lemma 3.3.6 für  $b = 1$  ist die Wahrscheinlichkeit, dass es eine Ecke gibt, deren  $k$ -Erkundungsbaum mehr als  $l = 2$  Kollisionen enthält, bevor er auf eine Ecke aus der Menge  $\text{Core}_t \cup \{1, 2, \dots, t\}$  trifft, nach oben beschränkt durch  $t^{-2}(\log t)^{d_1}$  für  $d_1$  hinreichend groß. Sobald der  $k$ -Erkundungsbaum eine Ecke aus  $\text{Core}_t \cup \{1, 2, \dots, t\}$  enthält, benutzen wir Proposition 3.3.5 und Theorem 5 und müssen nichts weiter zeigen. Falls dies nicht der Fall ist und es mehr als zwei Kollisionen im Baum gibt, dann gibt es mindestens  $m_k = (m-1)m^{k-1}$  Ecken in  $\{t+1, t+2, \dots, 2t\}$  mit einem Abstand von genau  $k$  zur Ursprungsecke des Baums. Wir bezeichnen diese Ecken mit  $i_1, \dots, i_{m_k}$  und den  $k$ -Erkundungsbaum mit  $\mathcal{T}_i^{(k)}$ . Wir können schreiben

$$\begin{aligned} & \mathbb{P}\left(\nexists j \in \{1, 2, \dots, m_k\} \text{ so dass } i_j \rightarrow \{1, 2, \dots, t\} \mid \mathcal{T}_i^{(k)}\right) \\ &= \prod_{j=1}^{m_k} \mathbb{P}\left(i_j \nrightarrow \{1, 2, \dots, t\} \mid i_s \nrightarrow \{1, 2, \dots, t\} \forall s < j, \mathcal{T}_i^{(k)}\right). \end{aligned} \quad (3.3.52)$$

Wir wissen, das unabhängig davon, wie die Ecken in  $G_m(2t)$  verteilt sind, für jedes  $s \in \{t+1, t+2, \dots, 2t\}$  gilt, dass

$$\frac{\sum_{i=1}^t (d_i(s) + \delta)}{(2m + \delta)s} \geq \frac{\sum_{i=1}^t (d_i(t) + \delta)}{(2m + \delta)s} \geq \frac{1}{2}. \quad (3.3.53)$$

Das heißt, für jede Ecke  $i_j$  im Rand von  $\mathcal{T}_i^{(k)}$ , ist die Wahrscheinlichkeit, dass sie direkt mit  $\{1, 2, \dots, t\}$  verbunden ist, mindestens  $\frac{1}{2}$ . Als Ergebnis

erhalten wir, dass gleichmäßig für alle  $t, i$  und  $j$  gilt

$$\mathbb{P}\left(i_j \leftrightarrow \{1, 2, \dots, t\} \mid i_s \leftrightarrow \{1, 2, \dots, t\} \forall s < j, \mathcal{T}_i^{(k)}\right) \leq \frac{(2m + \delta)t}{(2m + \delta)2t} = \frac{1}{2}. \quad (3.3.54)$$

Damit erhalten wir, dass

$$\mathbb{P}\left(\nexists j = 1, \dots, m_k, \text{ so dass } i_j \leftrightarrow \{1, 2, \dots, t\} \mid \mathcal{T}_i^{(k)}\right) \leq 2^{-m_k}. \quad (3.3.55)$$

Da  $m_k = \frac{m-1}{m}(\log t)^\sigma$ , mit  $\sigma = \frac{1}{3-\tau} > 1$ , ist  $2^{-m_k} = o(t^{-1})$ . Damit liegt whp jede Ecke  $i \in \{t+1, t+2, \dots, 2t\}$  in einem Abstand von  $k+1$  von  $\{1, 2, \dots, t\}$ .  $\square$

Schließlich führen wir unsere bisherigen Ergebnisse zusammen und komplettieren den Beweis zu Theorem 6:

**Beweis:** Proposition 3.3.8 liefert uns, dass whp jede Ecke in  $G_m(2t)$  einen maximalen Abstand von  $k+1 = \left\lfloor \frac{\sigma \log \log t}{\log m} \right\rfloor$  von  $\text{Core}_t \cup \{1, 2, \dots, t\}$  hat. Weiter ist die Aussage von Proposition 3.3.5, dass whp jede Ecke der Menge  $\{1, 2, \dots, t\}$  höchstens einen Abstand von  $k+1$  vom Kern  $\text{Core}_t$  hat. Insgesamt liegt also whp jede Ecke in  $G_m(2t)$  in einem Abstand von  $2(k+1)$  vom Kern.  $\square$

### 3.3.3 Beweis von Theorem 4

Nun müssen wir noch die Aussagen der Abschnitte 3.3.1 und 3.3.2 zusammenfügen und können den Beweis zu Theorem 4 abschließen. Wir behalten im Gedächtnis, dass alle Ereignisse des Abschnitts whp eintreten, also ihre Wahrscheinlichkeit für  $t \rightarrow \infty$  gegen 1 geht.

**Beweis:** Nach Theorem 6 beträgt in  $G_m(2t)$  der

- Abstand einer beliebigen Ecke zum Kern höchstens  $\frac{2\sigma \log \log t}{\log m}$ .

Nach Theorem 5 ist der

- Durchmesser des Kerns nach oben beschränkt durch  $\frac{4 \log \log t}{|\log(\tau-2)|}$ .

Wir können beispielsweise zwei Ecken  $i, j$  aus der Peripherie verbinden, indem wir zunächst von  $i$  in maximal  $\frac{2\sigma \log \log t}{\log m}$  Schritten zum Kern springen, uns anschließend im Kern mit höchstens  $\frac{4 \log \log t}{|\log(\tau-2)|}$  Schritten zu einer geeigneten

Ecke bewegen, von der aus wir schließlich wieder in  $\frac{2\sigma \log \log t}{\log m}$  Schritten die Ecke  $j$  erreichen. In der Summe benötigen wir nicht mehr als

$$C_G = \frac{4}{|\log(\tau - 2)|} + \frac{4\sigma}{\log m} \quad (3.3.56)$$

Schritte. Liegen  $i$  oder  $j$  oder beide bereits im Kern, verkürzt sich unser Weg entsprechend. Damit ist Theorem 4 bewiesen.  $\square$



# Kapitel 4

## Technische Beweise und Lemmata

In diesem Kapitel sind die Beweise zu einigen technischen Lemmata zusammengefasst, die der Übersicht halber nicht in den jeweiligen Abschnitten des Hauptteils untergebracht sind.

### 4.1 Das zweite Moment der Gradfolge

**Lemma 4.1.1** Für  $M_4 > 0$  gilt

$$\mathbb{E} [(d_s(t) + \delta)^2] \leq M_4 \left(\frac{t}{s}\right)^{2a}. \quad (4.1.1)$$

**Beweis:** Wir beginnen mit Modell (c). Wir können  $\mathbb{E} [(d_s(t_1 + 1) + \delta)^2]$  rekursiv wie folgt berechnen:

$$\begin{aligned} & \mathbb{E} [(d_s(t_1 + 1) + \delta)^2 \mid G_m(t_1)] \\ &= \mathbb{E} [(d_s(t_1 + 1) - d_s(t_1) + d_s(t_1) + \delta)^2 \mid G_m(t_1)] \\ &= \mathbb{E} [(d_s(t_1 + 1) - d_s(t_1))^2 \mid G_m(t_1)] \\ & \quad + 2(d_s(t_1) + \delta)\mathbb{E} [(d_s(t_1 + 1) - d_s(t_1)) \mid G_m(t_1)] + (d_s(t_1) + \delta)^2 \\ &= \mathbb{E} \left[ \left( \sum_{j=1}^m I[g(t_1, j)] \right)^2 \mid G_m(t_1) \right] + \frac{2m(d_s(t_1) + \delta)^2}{(2m + \delta)t_1} + (d_s(t_1) + \delta)^2 \\ &= m \left( \frac{d_s(t_1) + \delta}{(2m + \delta)t_1} \right) - m \left( \frac{d_s(t_1) + \delta}{(2m + \delta)t_1} \right)^2 \\ & \quad + m^2 \left( \frac{d_s(t_1) + \delta^2}{(2m + \delta)t_1} \right)^2 + \frac{2m(d_s(t_1) + \delta)^2}{(2m + \delta)t_1} + (d_s(t_1) + \delta)^2. \end{aligned} \quad (4.1.2)$$

Daraus erhalten wir nun

$$\begin{aligned}\mathbb{E} [(d_s(t_1 + 1) + \delta)^2] &= \left(1 + \frac{m^2 - m}{(2m + \delta)^2 t_1^2} + \frac{2m}{(2m + \delta)t_1}\right) \mathbb{E} [(d_s(t_1) + \delta)] \\ &\quad + \frac{m}{(2m + \delta)t_1} \mathbb{E} [(d_s(t_1) + \delta)^2].\end{aligned}\quad (4.1.3)$$

Zur Abkürzung definieren wir

$$q_s(t_1) = \mathbb{E} [(d_s(t_1) + \delta)^2], \quad a = \frac{m}{2m + \delta}, \quad (4.1.4)$$

$$e_s(t_1) = \mathbb{E} [(d_s(t_1) + \delta)], \quad b = \frac{1}{2m + \delta}. \quad (4.1.5)$$

Benutzen wir nun (3.1.23) und (4.1.3) ergibt sich daraus

$$e_s(t_1 + 1) = e_s(t_1) \left(1 + \frac{a}{t_1}\right) \quad (4.1.6)$$

$$q_s(t_1 + 1) = \left(1 + \frac{a(a - b)}{t_1^2} + \frac{2a}{t_1}\right) q_s(t_1) + \frac{a}{t_1} e_s(t_1). \quad (4.1.7)$$

Im folgenden schreiben wir  $q_s(t+1) = (1+a(t))q_s(t) + r_s(t)$ , wobei wir  $a(k) = \frac{a(a-b)+2ak}{k^2}$  und  $r_s(t) = \frac{a}{t}e_s(t)$  setzen. Anschließend wenden wir sukzessive (4.1.7) an und erhalten

$$\begin{aligned}q_s(t) &= (1 + a(t-1))q_s(t-1) + r_s(t-1) \\ &= q_s(s) \prod_{k=s}^{t-1} (1 + a(k)) + \sum_{k=s}^{t-1} r_s(k) \prod_{l=k+1}^{t-1} (1 + a(l))\end{aligned}\quad (4.1.8)$$

Da  $a(k) = \frac{a(a-b)+2ak}{k^2} \leq \frac{2a}{k}$  und  $a \geq b > 0$  können wir schreiben

$$\begin{aligned}\prod_{k=s}^{t-1} (1 + a(k)) &\leq \exp \left\{ \sum_{k=s}^{t-1} a(k) \right\} \leq \exp \left\{ \sum_{k=s}^{t-1} \frac{2a}{k} \right\} \\ &\leq \exp \left\{ \int_{s-1}^{t-1} \frac{2a}{x} dx \right\} = \left( \frac{t-1}{s-1} \right)^{2a}.\end{aligned}\quad (4.1.9)$$

Damit haben wir den ersten Summanden aus (4.1.8) abgeschätzt. Mit (3.1.25) ist

$$\begin{aligned}e_s(t) &= \mathbb{E}[d_s(t) + \delta] \leq 2^a (2m + \delta) \left( \frac{t-1}{s} \right)^a \\ &\leq M \left( \frac{t}{s} \right)^a\end{aligned}\quad (4.1.10)$$

Wenden wir nun (4.1.10) auf  $\prod_{l=k+1}^{t-1} (1 + a(l))$  an und setzen  $r_s(t) = \frac{a}{t} e_s(t)$ , erhalten wir für den zweiten Summanden

$$\begin{aligned}
\sum_{k=s}^{t-1} r_s(k) \prod_{l=k+1}^{t-1} (1 + a(l)) &\stackrel{(4.1.9)}{\leq} \sum_{k=s}^{t-1} r_s(k) \left( \frac{t-1}{k} \right)^{2a} \leq (t-1)^{2a} \sum_{k=s}^{t-1} \frac{r_s(k)}{k^{2a}} \\
&\leq a(t-1)^{2a} \sum_{k=s}^{t-1} \frac{e_s(k)}{k^{1+2a}} \leq aM \frac{t^{2a}}{s^a} \sum_{k=s}^{t-1} \frac{1}{k^{1+a}} \\
&\stackrel{(4.1.10)}{\leq} aM \frac{t^{2a}}{s^a} \sum_{k=s}^{\infty} \frac{1}{k^{1+a}} \leq M_5 \left( \frac{t}{s} \right)^{2a}. \quad (4.1.11)
\end{aligned}$$

Setzen wir nun in (4.1.8) die Ergebnisse aus (4.1.10) und (4.1.11) ein, erhalten wir

$$q_s(t) \leq q_s(s) \left( \frac{t-1}{s-1} \right)^{2a} + M_5 \left( \frac{t}{s} \right)^{2a} \leq M_4 \left( \frac{t}{s} \right)^{2a}. \quad (4.1.12)$$

□

## 4.2 Ecken mit hohem Grad

**Lemma 4.2.1** *Sei  $l = l(t) \rightarrow \infty$  für  $t \rightarrow \infty$  and  $l(t) \leq u_1$ . Dann gibt es eine Konstante  $B > 0$ , so dass mit einer Wahrscheinlichkeit größer als  $1 - o(t^{-1})$  gilt*

$$\sum_{i: d_i(t) \geq l} d_i(t) \geq B t l^{2-\tau}. \quad (4.2.1)$$

**Beweis:** Sei  $P_{\geq l}(t) = \#\{i \leq t : d_i(t) \geq l\}$  die Anzahl von Ecken zum Zeitpunkt  $t$ , mit einem Grad von mindestens  $l$ . Dann können wir schreiben

$$\sum_{i: d_i(t) \geq l} d_i(t) \geq l P_{\geq l}(t). \quad (4.2.2)$$

Nun wollen wir  $P_{\geq l}(t)$  untersuchen. In [DEHH99] finden wir für Modell (c) verschiedene Aussagen zur Asymptotik von  $P_{\geq l}(t)$ , die wir hier aufgreifen wollen. Zudem geben wir Anmerkungen, wie sich die gleichen Aussagen für die Modelle (a) und (b) zeigen lassen. Im ersten Schritt wird gezeigt, dass es ein  $B_1$  gibt, so dass gleichmäßig für alle  $l$  gilt

$$\mathbb{P} \left( |P_{\geq l}(t) - \mathbb{E}[P_{\geq l}(t)]| \geq B_1 \sqrt{t \log t} \right) = o(t^{-1}). \quad (4.2.3)$$

Diese Aussage haben wir bereits mit Proposition 2.2.1 gezeigt. Hier erhalten wir eine Abschätzung für die Konzentration der Anzahl von Ecken mit einem

Grad von mindestens  $l$ . Der Beweis folgt einer Arbeit von Bollobás et al. (vgl. [BRST01], Beweis von Theorem 1). Die Aussage gilt für alle Modelle (a) bis (c). Im zweiten Schritt bezeichne

$$P_l(t) = \# \{i \leq t : d_i(t) = l\} \quad (4.2.4)$$

die Anzahl der Ecken mit Grad  $l$ . Weiter sei  $p_k$  gegeben durch

$$p_k = \frac{\frac{\delta}{m} \Gamma(k + \delta) \Gamma(m + \delta + \theta)}{\Gamma(m + \delta) \Gamma(k + 1 + \delta + \theta)}, \quad k \geq m, \quad (4.2.5)$$

mit  $\theta = 2 + \frac{\delta}{m}$ . In Kapitel 2 hatten wir bereits  $p_k \sim k^{-\tau}$  gezeigt für  $\tau = \frac{\delta}{m}$ , was wir auch hier wieder benutzen wollen. Weiter wissen wir aus Proposition 2.3 für Modell (a), dass es eine Konstante  $B_2$  gibt, so dass

$$\sup_{l \geq 1} |\mathbb{E}[P_l(t)] - tp_l| \leq B_2. \quad (4.2.6)$$

Für Modell (c) wird dies in [DEHH99] gezeigt. Für Modell (b) kann man den Beweis aus [Hof07], Kapitel 8 entsprechend anpassen. Allgemeinere Resultate dieser Art, mit einer schärferen Grenze für (4.2.6) findet man auch in [HW06].

Aus (4.2.3) haben wir gesehen, dass sich  $P_{\geq l}(t)$  whp bis auf eine Differenz von  $B_1 \sqrt{t \log t}$  an  $\mathbb{E}[P_{\geq l}(t)]$  nähert. Weiter folgt aus (4.2.6), dass whp  $\mathbb{E}[P_l(t)] \geq tp_l - B_2$ . Mit einer Wahrscheinlichkeit größer  $1 - o(t^{-1})$  gilt somit

$$\begin{aligned} P_{\geq l}(t) &\stackrel{(4.2.3)}{\geq} \mathbb{E}[P_{\geq l}(t)] - B_1 \sqrt{t \log t} \\ &\geq \mathbb{E}[P_{\geq l_t}(t)] - \mathbb{E}[P_{\geq 2l_t}(t)] - B_1 \sqrt{t \log t} \\ &\stackrel{(4.2.6)}{\geq} \sum_{l=l_t}^{2l_t-1} [tp_l - B_2] - B_1 \sqrt{t \log t} \\ &\geq B_3 t l_t^{1-\tau} - B_2 l_t - B_1 \sqrt{t \log t}, \end{aligned} \quad (4.2.7)$$

wobei wir im letzten Schritt benutzt haben, dass  $p_{l_t} \sim l_t^{-\tau}$ . Nun wollen wir ein  $l_t$  finden, so dass  $B_3 t l_t^{1-\tau}$  deutlich größer ist als  $B_2 l_t$  und  $B_1 \sqrt{t \log t}$ . Es soll also zum einen gelten

$$t l_t^{1-\tau} \gg l_t, \quad (4.2.8)$$

was wir umformen können zu  $l_t \ll t^{\frac{1}{\tau}}$ , und weiter

$$t l_t^{1-\tau} \gg \sqrt{t \log t}, \quad (4.2.9)$$

was wir umformen können zu  $l_t \ll t^{\frac{1}{2(\tau-1)}} (\log t)^{-\frac{1}{2(\tau-1)}}$ . Für  $\tau > 2$  ist  $\frac{1}{\tau} \geq \frac{1}{2(\tau-1)}$  und  $(\log t)^{-\frac{1}{2(\tau-1)}} \geq (\log t)^{-\frac{1}{2}}$ , weshalb

$$l_t \leq t^{\frac{1}{2(\tau-1)}} (\log t)^{-\frac{1}{2}} \equiv u_1 \quad (4.2.10)$$

den Bedingungen genügt. Bei dieser Wahl von  $l_t$  sind  $B_2 l_t$  und in  $B_1 \sqrt{t \log t}$  in (4.2.7) klein im Vergleich zu  $B_3 t l_t^{1-\tau}$  und für ein geeignetes  $B$  können wir schreiben, dass mit einer Wahrscheinlichkeit größer  $1 - o(t^{-1})$  gilt

$$P_{\geq l_t}(t) \geq B t l_t^{1-\tau} \quad (4.2.11)$$

und mit (4.2.2) gilt mit einer Wahrscheinlichkeit größer  $1 - o(t^{-1})$

$$\sum_{i: d_i(t) \geq l} d_i(t) \geq B t l_t^{2-\tau}. \quad (4.2.12)$$

Für  $l_t$  wie in (4.2.10) und mit der Definition von  $\text{Inner}_t$  wie in (3.3.6) aus Abschnitt 3.3.1 gilt dann whp

$$\#\text{Inner}_t = P_{\leq t}(t) \geq B t l_t^{1-\tau} \gg \sqrt{t}, \quad (4.2.13)$$

was wir im Beweis zu Proposition 3.3.1 benutzen werden.  $\square$

## 4.3 Der Durchmesser eines multinomialen Graphen

**Lemma 4.3.1** *Sei  $H_{n_t}$  ein multinomialer Graph wie im Beweis zu Proposition 3.3.1. Dann ist der Durchmesser von  $H_{n_t}$  whp nach oben beschränkt durch den Durchmesser eines gleichmäßigen Erdős-Rényi-Graphen  $G_n(n_t, M_t)$ , wobei*

$$M_t = \frac{1}{2} e_t (1 - (1 - q_t)^t). \quad (4.3.1)$$

**Beweis:** Es ist zu zeigen, dass die zufällige Anzahl von Kanten  $M_{n_t}$  die deterministische Anzahl  $M_t$  whp dominiert. Dies kann man aus der Chebyshevschen Ungleichung wie folgt schließen:

Aus der Definition des multinomialen Graphen, und mit  $e_t = \frac{n_t(n_t-1)}{2}$  ist

$$M_{n_t} = \sum_{i=1}^{e_t} I \left[ \left\{ \sum_{j=1}^t N_{j,i} > 0 \right\} \right]. \quad (4.3.2)$$

Daraus folgt nach umfangreicher Rechnung und mithilfe einer Taylorentwicklung

$$\mathbb{E}[M_{n_t}] = e_t(1 - (1 - q_t)^t) = 2M_t, \quad (4.3.3)$$

und

$$\text{Var}(M_{n_t}) \sim tq_t e_t \sim \mathbb{E}[M_{n_t}] = 2M_t, \quad (4.3.4)$$

so dass die Varianz von gleicher Ordnung ist wie das erste Moment. Wenden wir nun die Chebyshevsche Ungleichung an, erhalten wir

$$\begin{aligned} \mathbb{P}(M_{n_t} < M_t) &\leq \mathbb{P}(|M_{n_t} - \mathbb{E}[M_{n_t}]| > M_t) \\ &\leq \frac{\text{Var}(M_{n_t})}{M_t^2} \sim \frac{4 \text{Var}(M_{n_t})}{\mathbb{E}[M_{n_t}]^2} \rightarrow 0. \end{aligned} \quad (4.3.5)$$

□

## 4.4 Frühe Ecken haben große Grade

Im folgenden beweisen wir Lemma 3.3.7:

**Beweis:** Wir untersuchen das Problem zunächst für Modell (a). Die Erweiterung auf Modell (b) ist einfach und erfolgt gleich im Anschluss. Der Beweis für Modell (c) ist etwas komplizierter.

**Modell (a):** Wir beginnen mit der Aussage für  $m = 1$  und einem festen  $\delta > -1$ . Wir zeigen durch Induktion in  $j$ , dass für  $m = 1$  und für alle  $t > i$

$$\mathbb{P}(d_i(t) = j) \leq C_j \frac{\Gamma(t)\Gamma(i + \frac{1+\delta}{2+\delta})}{\Gamma(t + \frac{1+\delta}{2+\delta})\Gamma(i)} \quad (4.4.1)$$

wobei wir  $C_j$  im Verlauf des Beweises näher bestimmen werden. Wir starten mit  $j = 1$  und berechnen die Wahrscheinlichkeit, dass sich bis zu einem beliebigen Zeitpunkt  $t > i$  keine weitere Ecke mit  $i$  verbunden hat, also  $\mathbb{P}(d_i(t) = 1)$ . Die Wachstumsregel (1.1.1) aus Kapitel 1 liefern uns für dieses Ereignis

$$\begin{aligned} \mathbb{P}(d_i(t) = 1) &= \prod_{s=i+1}^t \left( 1 - \frac{1 + \delta}{(2 + \delta)(s - 1) + (1 + \delta)} \right) \\ &= \prod_{s=i+1}^t \left( \frac{s - 1}{s - 1 + \frac{1+\delta}{2+\delta}} \right) = \frac{\Gamma(t)\Gamma(i + \frac{1+\delta}{2+\delta})}{\Gamma(t + \frac{1+\delta}{2+\delta})\Gamma(i)}, \end{aligned} \quad (4.4.2)$$

womit wir die Induktion gestartet haben, wenn wir  $C_1 = 1$  setzen.

Um die Induktion zu entwickeln, sei  $s < t$  der letzte Zeitpunkt, zu dem sich eine Ecke mit  $i$  verbunden hat. In dem Term den wir im nächsten Schritt aufsummieren wollen bezeichnet

- (I) die Wahrscheinlichkeit, dass eine Ecke  $i$  zum Zeitpunkt  $s - 1$  gerade Grad  $j - 1$  hat,
- (II) die Wahrscheinlichkeit, dass sich die neue Ecke  $s$  mit  $i$  verbindet,
- (III) die Wahrscheinlichkeit, dass sich nach dem Zeitpunkt  $s$  keine weitere Ecke mehr mit  $i$  verbindet.

Insgesamt summieren wir von  $s = i + j - 1$  bis  $t$ , also über alle Zeitpunkte  $s$ , in denen die Ecke  $i$  den Grad  $j$  erreichen kann. Wir erhalten

$$\mathbb{P}(d_i(t) = j) = \sum_{s=i+j-1}^t \underbrace{\mathbb{P}(d_i(s-1) = j-1)}_{(I)} \cdot \underbrace{\frac{j-1+\delta}{(2+\delta)(s-1)+1+\delta}}_{(II)} \cdot \underbrace{\mathbb{P}(d_i(t) = j | d_i(s) = j)}_{(III)}. \quad (4.4.3)$$

Mit der Induktionsvoraussetzung (4.4.1) erhalten wir für (I)

$$\mathbb{P}(d_i(s-1) = j-1) \leq C_{j-1} \frac{\Gamma(s-1)\Gamma(i + \frac{1+\delta}{2+\delta})}{\Gamma(s-1 + \frac{1+\delta}{2+\delta})\Gamma(i)}. \quad (4.4.4)$$

Analog zu (4.4.2) berechnen wir für (III)

$$\begin{aligned} \mathbb{P}(d_i(t) = j | d_i(s) = j) &= \prod_{q=s+1}^t \left( 1 - \frac{j+\delta}{(2+\delta)(q-1) + (1+\delta)} \right) \\ &= \prod_{q=s+1}^t \left( \frac{q-1 - \frac{j-1}{2+\delta}}{q-1 + \frac{1+\delta}{2+\delta}} \right) = \frac{\Gamma(t - \frac{j-1}{2+\delta})\Gamma(s + \frac{1+\delta}{2+\delta})}{\Gamma(t + \frac{1+\delta}{2+\delta})\Gamma(s - \frac{j-1}{2+\delta})}. \end{aligned} \quad (4.4.5)$$

Nun setzen wir diese Ergebnisse in (4.4.3) ein und erhalten

$$\begin{aligned} \mathbb{P}(d_i(t) = j) &\leq \sum_{s=i+j-1}^t C_{j-1} \frac{\Gamma(s-1)\Gamma(i + \frac{1+\delta}{2+\delta})}{\Gamma(s-1 + \frac{1+\delta}{2+\delta})\Gamma(i)} \\ &\quad \frac{j-1+\delta}{(2+\delta)(s-1)+1+\delta} \cdot \frac{\Gamma(t - \frac{j-1}{2+\delta})\Gamma(s + \frac{1+\delta}{2+\delta})}{\Gamma(t + \frac{1+\delta}{2+\delta})\Gamma(s - \frac{j-1}{2+\delta})}. \end{aligned} \quad (4.4.6)$$

Wir nutzen als nächstes die Eigenschaft der Gamma-Funktion, dass für alle  $x \in \mathbb{R}_+^*$  gilt  $x\Gamma(x) = \Gamma(x+1)$ . Dann ist

$$\Gamma\left(s-1 + \frac{1+\delta}{2+\delta}\right) ((2+\delta)(s-1) + (1+\delta)) = (2+\delta)\Gamma\left(s + \frac{1+\delta}{2+\delta}\right) \quad (4.4.7)$$

und wir erhalten

$$\mathbb{P}(d_i(t) = j) \leq C_{j-1} \frac{j-1+\delta}{2+\delta} \frac{\Gamma(i + \frac{1+\delta}{2+\delta})}{\Gamma(i)} \frac{\Gamma(t - \frac{j-1}{2+\delta})}{\Gamma(t + \frac{1+\delta}{2+\delta})} \sum_{s=i+j-1}^t \frac{\Gamma(s-1)}{\Gamma(s - \frac{j-1}{2+\delta})}. \quad (4.4.8)$$

Falls  $l+b, l+1+a > 0$  und  $a-b+1 > 0$ , können wir rechnen

$$\begin{aligned} \sum_{s=l}^t \frac{\Gamma(s+a)}{\Gamma(s+b)} &= \frac{1}{a-b+1} \left[ \frac{\Gamma(t+1+a)}{\Gamma(t+b)} - \frac{\Gamma(t+1+a)}{\Gamma(l+b)} \right] \\ &\leq \frac{1}{a-b+1} \frac{\Gamma(t+1+a)}{\Gamma(t+b)} \end{aligned} \quad (4.4.9)$$

Wenden wir nun (4.4.9) mit  $a = -1$ ,  $b = -\frac{j-1}{2+\delta}$ ,  $l = i+j-1$ , so dass  $a-b+1 = \frac{j-1}{2+\delta} > 0$ , wenn  $j > 1$ , erhalten wir

$$\begin{aligned} \mathbb{P}(d_i(t) = j) &\leq C_{j-1} \frac{j-1+\delta}{2+\delta} \frac{\Gamma(i + \frac{1+\delta}{2+\delta})}{\Gamma(i)} \frac{\Gamma(t - \frac{j-1}{2+\delta})}{\Gamma(t + \frac{1+\delta}{2+\delta})} \frac{1}{\frac{j-1}{2+\delta}} \frac{\Gamma(t)}{\Gamma(t - \frac{j-1}{2+\delta})} \\ &= C_{j-1} \frac{j-1+\delta}{j-1} \frac{\Gamma(i + \frac{1+\delta}{2+\delta})}{\Gamma(i)} \frac{\Gamma(t)}{\Gamma(t + \frac{1+\delta}{2+\delta})}. \end{aligned} \quad (4.4.10)$$

Gleichung (4.4.10) liefert dann die Induktionsvoraussetzung, wenn wir

$$C_j = \frac{j-1+\delta}{2+\delta} C_{j-1}. \quad (4.4.11)$$

setzen. Damit ist der Beweis für Modell (a) komplett.

**Modell (b):** Für Modell (b) ist der Beweis ähnlich. Die Induktionsvoraussetzung lautet hier

$$\mathbb{P}(d_i(t) = j) \leq C_j \frac{\Gamma(t - \frac{1+\delta}{2+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i - \frac{1+\delta}{2+\delta})}, \quad (4.4.12)$$



wobei  $C_j$  wiederum im Verlauf näher bestimmt wird. In Modell (b) liefert uns die Wachstumsregel für  $t \geq i$

$$\begin{aligned}\mathbb{P}(d_i(t) = 1) &= \prod_{s=i+1}^t \left(1 - \frac{1 + \delta}{(2 + \delta)(s - 1)}\right) \\ &= \prod_{s=i+1}^t \left(\frac{s - 1 - \frac{1+\delta}{2+\delta}}{s - 1}\right) = \frac{\Gamma(t - \frac{1+\delta}{2+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i - \frac{1+\delta}{2+\delta})},\end{aligned}\quad (4.4.13)$$

womit wir die Induktion starten. Weiter ist

$$\begin{aligned}\mathbb{P}(d_i(t) = j | d_i(s) = j) &= \prod_{q=s+1}^t \left(1 - \frac{j + \delta}{(2 + \delta)(q - 1)}\right) \\ &= \frac{\Gamma(t - \frac{j+1}{2+\delta})\Gamma(s)}{\Gamma(t)\Gamma(s - \frac{j+1}{2+\delta})}.\end{aligned}\quad (4.4.14)$$

Wir wenden nun die Induktionsvoraussetzung (4.4.12) in der Form  $\mathbb{P}(d_i(s - 1) = j - 1) \leq C_{j-1} \frac{\Gamma(s-1 - \frac{1+\delta}{2+\delta})\Gamma(i)}{\Gamma(s-1)\Gamma(i - \frac{1+\delta}{2+\delta})}$  an und summieren über alle Zeitpunkte  $s$ , in denen die Ecke  $i$  den Grad  $j$  erreichen kann. Wir erhalten

$$\mathbb{P}(d_i(t) = j) \leq \sum_{s=i+j-1}^t C_{j-1} \frac{\Gamma(s-1 - \frac{1+\delta}{2+\delta})\Gamma(i)}{\Gamma(s-1)\Gamma(i - \frac{1+\delta}{2+\delta})} \frac{j-1+\delta}{(s-1)(2\delta)} \frac{\Gamma(t - \frac{j+1}{2+\delta})\Gamma(s)}{\Gamma(t)\Gamma(s - \frac{j+1}{2+\delta})}.\quad (4.4.15)$$

Dann ist

$$\begin{aligned}\mathbb{P}(d_i(t) = j) &\leq C_{j-1} \frac{j-1+\delta}{2\delta} \frac{\Gamma(i)}{\Gamma(i - \frac{1+\delta}{2+\delta})} \frac{\Gamma(t - \frac{j+\delta}{2+\delta})}{\Gamma(t)} \sum_{s=i+j-1}^t \frac{\Gamma(s-1 - \frac{1+\delta}{2+\delta})}{\Gamma(s - \frac{j+\delta}{2+\delta})} \\ &= C_{j-1} \frac{j-1+\delta}{2\delta} \frac{\Gamma(i)}{\Gamma(i - \frac{1+\delta}{2+\delta})} \frac{\Gamma(t - \frac{j+\delta}{2+\delta})}{\Gamma(t)},\end{aligned}\quad (4.4.16)$$

was uns mit  $C_1 = 1$  und  $C_j = C_{j-1} \frac{j-1+\delta}{j-1}$  die Behauptung liefert.

**Modell (c):** In den Beweisen zu den Modellen (a) und (b) haben wir das Verhältnis zwischen  $G_m(t)$  und  $G_1(t)$  betrachtet, wodurch es ausreichend war, den Fall  $m = 1$  zu untersuchen. Für  $m = 1$  konnte dort mit jedem Zeitschritt höchstens eine neue Kante hinzugefügt werden. Diese Strategie scheitert jedoch bei Modell (c). Somit müssen wir unser Argument zunächst anpassen.

Wir wollen nun ein  $e_i(t)$  definieren, so dass zum einen  $e_i(t) \leq d_i(t)$  und des weiteren  $e_i(t)$  mit jedem Zeitschritt höchstens um eins wächst. Wir definieren  $e_i(t)$  rekursiv und beginnen mit  $e_i(i) = d_i(i) = m$ . Angenommen wir hätten bereits gezeigt, dass  $d_i(t) = e_i(t) + r_i(t)$ , mit  $r_i(t) \geq 0$ , dann gehen wir weiter nach  $t + 1$  indem wir  $e_i(t)$  genau dann um eins erhöhen, wenn die erste Kante der Ecke  $t + 1$  mit der Ecke  $i$  verbunden wird, was mit einer Wahrscheinlichkeit von  $\frac{e_i(t)+\delta}{(2m+\delta)t}$  eintreten soll. Mit einer Wahrscheinlichkeit von  $\frac{r_i(t)}{(2m+\delta)t}$ , bleibt  $e_i(t + 1) = e_i(t)$  und wir erhöhen  $r_i(t)$  um eins. Für die anderen  $m - 1$  Kanten erhöhen wir  $r_i(t)$  mit einer Wahrscheinlichkeit  $\frac{d_i(t)+\delta}{(2m+\delta)t}$ .

Etwas ungenau könnte man sagen, dass wir den Grad der Ecke  $i$  in einen „Grad der ersten Kanten“ und einen „Grad der restlichen  $m - 1$  Kanten“ aufspalten, wobei wir  $e_i(t)$  auch die  $m$  Kanten zugeschlagen haben, die von der Ecke  $i$  selbst ausgehen und zudem die Wahrscheinlichkeit für eine Erhöhung von  $e_i(t)$  mit  $\frac{e_i(t)+\delta}{(2m+\delta)t}$  etwas kleiner ist.

Es gilt  $e_i(t + 1) \leq d_i(t + 1)$ , falls  $e_i(t) \leq d_i(t)$ , denn die Differenz zwischen  $e_i(t)$  und  $d_i(t)$  ist gerade  $r_i(t)$ , das für wachsende  $t$  monoton steigt. Außerdem wissen wir von  $e_i(t + 1)$ , dass es entweder gleich  $e_i(t)$  oder gleich  $e_i(t) + 1$  ist, wobei letzteres mit der Wahrscheinlichkeit

$$\mathbb{P}(e_i(t + 1) = j | e_i(t) = j - 1) = \frac{j - 1 + \delta}{(2m + \delta)t} \quad (4.4.17)$$

eintritt. Von nun an gehen wir ähnlich vor wie in den Beweisen zu den Modellen (a) und (b), jedoch betrachten wir nun die Zufallsvariable  $e_i(t)$ . Die Induktionsvoraussetzung lautet nun

$$\mathbb{P}(e_i(t) = j) \leq C_j \frac{\Gamma(t - \frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i - \frac{m+\delta}{2m+\delta})}, \quad (4.4.18)$$

wobei  $C_j$  auch hier im weiteren Verlauf bestimmt wird. Nach der Wachstumsregel für  $e_i(t)$  gilt für alle  $t \geq i$

$$\begin{aligned} \mathbb{P}(e_i(t) = m) &= \prod_{s=i+1}^t \left( 1 - \frac{m + \delta}{(2m + \delta)(s - 1)} \right) = \prod_{s=i+1}^t \left( \frac{s - 1 - \frac{m+\delta}{2m+\delta}}{s - 1} \right) \\ &= \frac{\Gamma(t - \frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i - \frac{m+\delta}{2m+\delta})}, \end{aligned} \quad (4.4.19)$$

womit die Induktion gestartet ist wenn wir  $C_m = 1$  setzen. Weiter ist

$$\mathbb{P}(e_i(t) = j | e_i(s) = j) = \prod_{q=s+1}^t \left( 1 - \frac{j + \delta}{(2m + \delta)(q - 1)} \right) = \frac{\Gamma(t - \frac{j+\delta}{2m+\delta})\Gamma(s)}{\Gamma(t)\Gamma(s - \frac{j+\delta}{2m+\delta})}. \quad (4.4.20)$$

Für  $j > m$  setzen wir die Induktion fort, indem wir die Induktionsvoraussetzung in der Form

$$\mathbb{P}(e_i(s-1) = j-1) \leq C_j \frac{\Gamma(s-1 - \frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(s-1)\Gamma(i - \frac{m+\delta}{2m+\delta})} \quad (4.4.21)$$

einsetzen. Wir erhalten

$$\begin{aligned} \mathbb{P}(e_i(t) = j) &\leq \\ &\sum_{s=i+j-1}^t C_{j-1} \left( \frac{\Gamma(s-1 - \frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(s-1)\Gamma(i - \frac{m+\delta}{2m+\delta})} \right) \left( \frac{j-1+\delta}{(2m+\delta)(s-1)} \right) \frac{\Gamma(t - \frac{j+\delta}{2m+\delta})}{\Gamma(t)\Gamma(s - \frac{j+\delta}{2m+\delta})} \\ &= C_{j-1} \frac{j-1+\delta}{2m+\delta} \frac{\Gamma(i)}{\Gamma(i - \frac{m+\delta}{2m+\delta})} \frac{\Gamma(t - \frac{j+\delta}{2m+\delta})}{\Gamma(t)} \sum_{s=i+j-1}^t \frac{\Gamma(s-1 - \frac{m+\delta}{2m+\delta})}{\Gamma(s - \frac{j+\delta}{2m+\delta})} \\ &= C_{j-1} \frac{j-1+\delta}{j-m} \frac{\Gamma(i)}{\Gamma(i - \frac{m+\delta}{2m+\delta})} \frac{\Gamma(t - \frac{j+\delta}{2m+\delta})}{\Gamma(t)}, \end{aligned} \quad (4.4.22)$$

so dass wir nun  $C_j$  bestimmen, indem wir für  $C_m = 1$  setzen und für alle  $j > m$

$$C_j = \frac{j-1+\delta}{j-m} C_{j-1}. \quad (4.4.23)$$

Wir erinnern uns, dass die Modelle (a) und (b) für allgemeines  $m \geq 1$  aus dem Modell für  $m = 1$  konstruiert werden, indem wir  $\delta' = \delta/m$  gesetzt hatten und wir jeweils Gruppen von  $m$  Ecken betrachtet haben. Die Ecke  $i$  in  $G_m(t)$  bestimmt sich dann aus den Ecken  $(i-1)m+1, \dots, im$  in  $G_1(t)$ . Insbesondere ist der Grad einer Ecke  $i$  in  $G_m(t)$  nach unten beschränkt durch den Grad der Ecke  $im$  in  $G_1(mt)$ . Folglich dürfen wir für die Modelle (a) und (b) die Resultate auf  $m \leq 1$  erweitern, indem wir  $\delta' = \frac{\delta}{m}$  wählen. Für alle drei Modelle können wir dann schreiben

$$\mathbb{P}(d_i(t) = j) \leq C_j \frac{\Gamma(t - a_1)\Gamma(i + a_2)}{\Gamma(t + a_2)\Gamma(i - a_1)}, \quad (4.4.24)$$

mit  $j \geq m \geq 1$  und  $c_j > 0$  fest. wobei für Modell (a)

$$a_1 = 0 \quad a_2 = \frac{m+\delta}{2m+\delta} \quad (4.4.25)$$

und für (b) und (c)

$$a_1 = \frac{m + \delta}{2m + \delta} \qquad a_2 = 0. \qquad (4.4.26)$$

Für alle drei Modelle gilt

$$C_j \leq 1 \leq j^{p-1} \qquad (4.4.27)$$

für ein  $p \geq 1$ . Mit (4.4.24) ist dann

$$\begin{aligned} \mathbb{P}(d_i(t) \leq j) &= \sum_{l=1}^j \mathbb{P}(d_i(t) = l) \\ &\leq \sum_{l=1}^j l^{p-1} \frac{\Gamma(t - a_1)\Gamma(i + a_2)}{\Gamma(t + a_2)\Gamma(i - a_1)} \\ &\leq j^p \frac{\Gamma(t - a_1)\Gamma(i + a_2)}{\Gamma(t + a_2)\Gamma(i - a_1)}. \end{aligned} \qquad (4.4.28)$$

Nutzen wir Eigenschaft 0.0.3 der Gammafunktion, erhalten wir für große  $t$  und  $i$

$$\mathbb{P}(d_i(t) \leq j) \leq j^p t^{-(a_1+a_2)} i^{a_1+a_2} (1 + o(1)). \qquad (4.4.29)$$

Nun führen wir unsere Resultate zum Beweis von Lemma 4.4 zusammen. Wir wählen  $0 < b < \frac{a_1+a_2}{a_1+a_2+1} = \frac{m+\delta}{3m+2\delta}$  und wenden die Boolesche Ungleichung (BU) auf die Ecken  $i \leq t^b$  an, so dass

$$\begin{aligned} \mathbb{P}(\exists i \leq t^b : d_i(t) \leq (\log t)^\sigma) &\stackrel{(BU)}{\leq} \sum_{i=1}^{t^b} \mathbb{P}(d_i(t) \leq (\log t)^\sigma) \\ &\leq (\log t)^{\sigma p} \frac{\Gamma(t - a_1)}{\Gamma(t + a_2)} \sum_{i=1}^{t^b} \frac{\Gamma(i + a_2)}{\Gamma(i - a_1)} \\ &\leq (\log t)^{\sigma p} t^{-(a_1+a_2)} \sum_{i=1}^{t^b} i^{a_1+a_2} \\ &\leq \frac{(\log t)^{\sigma p}}{t^{b(a_1+a_2+1)-(a_1+a_2)}} \xrightarrow{t \rightarrow \infty} 0. \end{aligned} \qquad (4.4.30)$$

□

# Literaturverzeichnis

- [AE99] H. Amann and J. Escher. *Analysis II*. Birkhäuser Verlag, Berlin, (1999).
- [AE02] H. Amann and J. Escher. *Analysis I*. Birkhäuser Verlag, Berlin, (2002).
- [BA99] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, (1999).
- [Bol01] B. Bollobás. *Random Graphs*, volume 73 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, (2001).
- [BR04] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, (2004).
- [BRST01] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures and Algorithms*, 18(3):279–290, (2001).
- [DEHH99] M. Deijfen, H. van den Esker, R. van der Hofstad, and G. Hooghiemstra. A preferential attachment model with random initial degrees. *Science*, 286:509–512, (1999).
- [DHH09] S. Dommers, R. van der Hofstad, and G. Hooghiemstra. Diameters in preferential attachment models. (March 20, 2009). [http://www.win.tue.nl/~rhofstad/DiamPA\\_DHH\\_fin.pdf](http://www.win.tue.nl/~rhofstad/DiamPA_DHH_fin.pdf).
- [Dur07] R. Durrett. *Random graph dynamics*. Cambridge University Press, Cambridge, (2007).
- [HH07] R. van der Hofstad and G. Hooghiemstra. Diameters in preferential attachment models. *Preprint*, (May 11, 2007). <http://www.win.tue.nl/~rhofstad/DiamPRATfin.pdf>.

- [HH08] R. van der Hofstad and G. Hooghiemstra. Diameters in preferential attachment models. Submitted to *The Annals of Applied Probability*, (April 25, 2008). [http://arxiv.org/PS\\_cache/arxiv/pdf/0705/0705.4153v1.pdf](http://arxiv.org/PS_cache/arxiv/pdf/0705/0705.4153v1.pdf).
- [Hof07] R. van der Hofstad. Random graphs and complex networks. Lecture Notes. TU Eindhoven, (2007). <http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf>.
- [HW06] O. Hagberg and C. Wiuf. Convergence properties of the degree distribution of some growing network models. *Bull. Math. Biol.*, 68:1275 – 1291, (2006).
- [JLR00] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, (2000).
- [Kle06] A. Klenke. *Wahrscheinlichkeitstheorie*, volume 1. Springer Verlag, Berlin, (2006).