

Anzahl inklusionsmaximaler Cliques in geometrischen Zufallsgraphen

Bachelorarbeit von

Niko Lemke

An der Fakultät für Informatik
Institut für Theoretische Informatik

Reviewers: TT.-Prof. Dr. Thomas Bläsius
PD Dr. Torsten Ueckerdt
Advisors: TT.-Prof. Dr. Thomas Bläsius

Time Period: 22.Dezember 2021 – 22.April 2022

Statement of Authorship

Ich versichere wahrheitsgemäß, die Arbeit selbstständig verfasst, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde sowie die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.

Karlsruhe, April 22, 2022

Zusammenfassung

Modelle zu finden und zu untersuchen, welche Graphen erzeugen, die möglichst gut mit echten Netzwerken übereinstimmen, ist Kern des Forschungsgebiet „Network Science“. Bestimmte Arten von Zufallsgraphen haben sich dafür als besonders geeignet gezeigt, verschiedene Typen von Netzwerke zu modellieren. In dieser Arbeit beschäftigen wir uns mit der Anzahl Cliques in verschiedenen Modellen zur Generierung von Zufallsgraphen.

Wir zeigen für das Modell der geometrischen Zufallsgraphen (RGG) mit erwartet konstantem Knotengrad eine Konzentrationsschranke für die Anzahl Cliques.

Außerdem zeigen wir, für eine einfache Form des GIRG-Modell für einen bestimmten Parameterbereich eine Konzentrationsschranke für die Anzahl inklusionsmaximaler Cliques.

In beiden Modellen betrachten wir zunächst Schranken für den Erwartungswert und die Varianz einer Abschätzung der Anzahl inklusionsmaximaler Cliques und zeigen dann mit der Tschebyscheffschen Ungleichung die Konzentrationsschranke.

Außerdem zeigen wir in beiden Modellen die Linearität des Erwartungswertes der Anzahl aller Cliques.

Die Linearität des Erwartungswertes der Anzahl aller Cliques zeigen wir auch in der Erweiterung des RGG-Modells um das Konzept der Temperatur.

Contents

1	Einleitung	1
1.1	Andere Arbeit zu diesem Thema	2
1.2	Mein Beitrag und Gliederung	2
2	Grundlagen	3
2.1	Graphentheoretische Grundbegriffe	3
2.2	Modelle geometrischer Zufallsgraphen	3
2.2.1	Geometrische Zufallsgraphen (RGG)	3
2.2.2	Geometrische Zufallsgraphen mit Temperatur	4
2.2.3	GIRG	4
2.3	Wahrscheinlichkeitstheoretische Grundlagen	5
2.4	Grenzwerte wichtiger Folgen, Reihen und Produkte	7
3	Geometrische Zufallsgraphen (RGG)	9
3.1	Erwartungswert der Anzahl Cliques	9
3.2	Konzentrationsschranke	11
3.2.1	Gittergröße	11
3.2.2	Nachbarschaft einer Zelle	12
3.2.3	Abschätzung der Anzahl Cliques	12
3.2.4	Erwartungswert einer Zelle	12
3.2.5	Erwartungswert aller Zellen	14
3.3	Varianz der Abschätzung	14
3.3.1	Kovarianz benachbarter Zellen	14
3.3.2	Kovarianz nicht benachbarter Zellen	15
3.3.3	Abschätzung der Varianz	16
3.3.4	Konzentrationsschranke	17
4	Geometrische Zufallsgraphen (RGG) mit Temperatur	19
4.1	Erwartungswert der Anzahl Cliques	19
4.1.1	Wahrscheinlichkeitsverteilung des Abstands	19
4.1.2	Kantenwahrscheinlichkeit	19
4.1.3	Erwartungswert	22
5	GIRG	23
5.0.1	Abschätzung der Anzahl Cliques	23
5.0.2	Erwartungswert	24
5.1	Konzentrationsschranke	25
5.1.1	Abschätzung der Anzahl inklusionsmaximaler Cliques	25
5.1.2	Varianz der Abschätzung	26
5.1.3	Konzentrationsschranke	28
6	Zusammenfassung und Ausblick	29

Bibliography	31
---------------------	-----------

1. Einleitung

Graphen sind das wichtigste Mittel zur Modellierung von Netzwerken. Es besteht also ein großes Interesse daran passende Modelle zu finden, die Graphen erzeugen, die echte Netzwerke gut abbilden, um dann z.B. effiziente Algorithmen für diese Netzwerke zu finden.

Das einfachste Modell eines Zufallsgraphen ist das sogenannte Erdős-Rényi-Modell [ER59]. Bei diesem Modell werden n Knoten jeweils paarweise mit einer Wahrscheinlichkeit $p \in (0, 1)$ verbunden. Die daraus resultierenden Graphen haben jedoch wenig Gemeinsamkeiten mit den meisten echten Netzwerken.

Deswegen werden meist andere Modelle, wie z.B. geometrische Zufallsgraphen (RGG) verwendet. In diesem Modell werden Knoten in der Ebene verteilt und mit Kanten verbunden wenn sie weniger als eine gewisse euklidische Distanz r voneinander entfernt sind. Die Einführung von Geometrie in das Modell hat den Effekt, dass die Kantenwahrscheinlichkeit zwischen zwei Knoten mit gemeinsamen Nachbarn höher ist als die Kantenwahrscheinlichkeit von Knoten ohne gemeinsamen Nachbarn. Diese Eigenschaft kann durch den sogenannten Clusterkoeffizient beschrieben werden. RGGs haben einen vergleichsweise hohen Clusterkoeffizient [Dal02]. Da viele echte Netzwerke einen hohen Clusterkoeffizient haben, ist dies eine wünschenswerte Eigenschaft. Dieses Modell eignet sich zum Beispiel zur Modellierung von Ad-hoc-Netzen wie in [HS95] von Huson und Sen gemacht.

Bei RGGs haben die meisten Knoten ungefähr den selben Knotengrad. Diese Homogenität ist jedoch bei vielen echten Netzwerken nicht gegeben. Viele echte Netzwerke sind sogenannte Skalenfreie Netzwerke. Das heißt die Verteilung der Knotengrade folgt ungefähr einem Power-Law. Skalenfreie Modelle mit einem hohen Clusterkoeffizient sind also besonders interessant.

Das Wahrscheinlich meist untersuchte Modell dieser Art sind hyperbolische geometrische Zufallsgraphen welche von Krioukov et al. in [KPK⁺10] eingeführt wurden. In diesem Modell werden die Knoten in einer hyperbolischen Geometrie platziert. Dieses Modell eignet sich zum Beispiel zur Modellierung des Internet-Graph [BPK10].

In dieser Arbeit untersuchen wir allerdings stattdessen eine einfache Variante sogenannter GIRGs welche von Bringmann, Keusch und Lengler in [BKL15] eingeführt wurden. Sie zeigen unter anderem, dass hyperbolische geometrische Zufallsgraphen als ein Spezialfall von GIRGs betrachtet werden können. Das GIRG-Modell hat jedoch den Vorteil, dass es mathematisch meist leichter zu untersuchen ist.

Die Anzahl inklusionsmaximaler Cliques in einem Graphen ist eine interessante Kennzahl des Graphens. Es gibt einige Arbeiten die versuchen effiziente Algorithmen für das Finden aller Inklusionsmaximaler Cliques (MCE-Problem) zu entwickeln. Für Unit Disk Graphen sei z.B. [GWG05] genannt. Die aus solchen Algorithmen gewonnenen Cliques können dann z.B. zur optimalen Ressourcen Allokation in Ad-hoc Netzwerken verwendet werden [XLN03]. Aber auch in z.B. sozialen Netzwerken, welche eher durch GIRGs modelliert werden können, haben maximale Cliques eine wichtige Bedeutung.

Es gibt aber bis jetzt kaum theoretisch fundierte Aussagen zu der Anzahl inklusionsmaximaler Cliques in diesen Modellen. Mit dieser Arbeit versuchen wir dieses Problem anzugehen.

1.1 Andere Arbeit zu diesem Thema

Ein Standardwerk zu geometrischen Zufallsgraphen ist das Buch von Mathew Penrose [Pen03]. Dort werden viele Eigenschaften von RGGs wie z.B. zur Konnektivität gezeigt.

Für das Thema dieser Arbeit besonders relevant ist die Aussage, dass die Anzahl inklusionsmaximaler Cliques in UDGs im Worst Case mit n Knoten $2^{\frac{n}{2}}$ groß sein kann. Ein Beweis dafür findet sich in [GWG05].

Die meiste andere Arbeit zu Cliques in RGGs konzentriert sich auf die sogenannte Cliquezahl (Größe der größten Clique) wie z.B. in [DMV13]. Dort finden sich auch Aussagen zu der erwarteten Anzahl inklusionsmaximaler Cliques bestimmter Größe.

Zu dem GIRG-Modell selbst gibt es nur wenige Arbeiten. Allerdings sind alle Aussagen zu hyperbolischen geometrischen Zufallsgraphen auch gültig für das GIRG-Modell, da hyperbolischen geometrischen Zufallsgraphen ja als ein Spezialfall von dem GIRG-Modell betrachtet werden können [BKL15]. Für hyperbolische geometrische Zufallsgraphen sei vor allem das Papier [BFK18] genannt, welches sich mit Cliques in hyperbolischen geometrischen Zufallsgraphen auseinandersetzt.

1.2 Mein Beitrag und Gliederung

Nach der Einführung der notwendigen Grundlagen in Kapitel 2, zeigen wir in Kapitel 3, dass für RGGs mit $r \in \Theta(\frac{1}{\sqrt{n}})$ sowohl der Erwartungswert der Anzahl aller Cliques in $\Theta(n)$ liegt als auch der Erwartungswert der Anzahl der inklusionsmaximalen Cliques in $\Theta(n)$ liegt.

Außerdem zeigen wir, dass die Anzahl aller Cliques und damit auch die Anzahl aller inklusionsmaximalen Cliques mit hoher Wahrscheinlichkeit ($\geq 1 - O(\frac{1}{n})$) in $O(n)$ liegt. Dies machen wir, indem wir für eine Abschätzung, der Anzahl Cliques, zeigen, dass sowohl der Erwartungswert als auch die Varianz in $O(n)$ liegen. Im Anschluss beweisen wir dann die Aussage mit der Tschebyscheffschen Ungleichung.

Desweiteren zeigen wir, dass die Ergebnisse für den Erwartungswert auch gelten, wenn wir das RGG-Modell um das Konzept der Temperatur erweitern und die Temperatur T im Intervall $(0, 1)$ liegt.

In letztem Kapitel zeigen wir für eine Vereinfachung des GIRG-Modell ohne Temperatur und mit konkreter Gewichtsfunktion $w_v = (\frac{n}{v})^{\frac{1}{\beta-1}}$, dass für Power-Law Exponent $\beta \geq 3$ der Erwartungswert, der Anzahl Cliques und damit auch der Anzahl Inklusionsmaximaler Cliques, in $O(n)$ liegt. Im Anschluss zeigen wir dann, dass die Anzahl aller inklusionsmaximalen Cliques mit hoher Wahrscheinlichkeit ($\geq 1 - O(\frac{1}{n})$) in $O(n)$ liegt. Ebenso wie im RGG zeigen wir dies, indem wir für eine Abschätzung der Anzahl inklusionsmaximaler Cliques, zeigen, dass sowohl der Erwartungswert als auch die Varianz in $O(n)$ liegen. Im Anschluss kann die Aussage dann mit der Tschebyscheffschen Ungleichung bewiesen werden.

2. Grundlagen

2.1 Graphentheoretische Grundbegriffe

Ein *ungerichteter Graph* G ist ein geordnetes Paar (V, E) , wobei die Menge V eine Menge von Knoten. Die Menge E ist die Kantenmenge bestehend aus zwei-elementigen Teilmengen von V .

Die *Nachbarschaft* eines Knotens v bezeichnet die Menge aller Knoten, die durch eine Kante mit v verbunden sind. In dieser Arbeit werden wir jedoch meistens die Nachbarschaft von v .

Eine *Clique* ist eine Teilmenge der Knoten, für die gilt, dass jedes Knotenpaar $\{u, v\}$ der Teilmenge mit $u \neq v$ mit einer Kante verbunden ist.

Eine Clique ist *inklusionsmaximal* wenn es keinen Knoten $v \in V$ gibt, der nicht Teil der Clique ist, aber zu allen Knoten der Clique mit einer Kante verbunden ist.

2.2 Modelle geometrischer Zufallsgraphen

2.2.1 Geometrische Zufallsgraphen (RGG)

Es werden n Knoten unabhängig voneinander gleichverteilt auf dem zweidimensionalen Einheitstorus \mathbb{T}^2 platziert. Alle Knoten deren euklidische Distanz geringer ist als ein bestimmter Abstand r werden dann mit einer Kante verbunden. Der Torus \mathbb{T}^2 kann quasi als Einheitsquadrat verstanden werden, bei dem sowohl die obere und untere Seite miteinander als auch die rechte und linke Seite miteinander verbunden sind. Der Grund für die Verwendung des Torus ist, dass man den Fall, dass Knoten am Rand liegen nicht extra betrachten muss, da die Oberfläche eine kontinuierliche zusammenhängende Fläche ohne Rand ist. Formal gilt für die Knoten, dass jeweils die x - und y -Koordinate einer Gleichverteilung auf dem Intervall $[0, 1]$ folgen, wobei die x - und y -Koordinate für jeden Knoten unabhängig voneinander und auch unabhängig von den x - und y -Koordinaten aller anderen Knoten sind. Alle Knoten deren euklidische Distanz geringer ist als ein bestimmter Abstand r werden dann mit einer Kante verbunden.

Die aus diesem Modell erzeugten Graphen sind sogenannte Unit-Disk Graphen. Unit-Disk Graphen sind dadurch definiert, dass zwei Knoten genau dann mit einer Kante verbunden sind, wenn sich die Kreise mit den Knoten als Mittelpunkt und Radius $\frac{r}{2}$ schneiden.

In dieser Arbeit betrachten wir nur den Fall, dass r in $\Theta(\frac{1}{\sqrt{n}})$ liegt.

Da die Fläche des \mathbb{T}^2 genau 1 groß ist, entspricht die Kantenwahrscheinlichkeit p_{uv} zwischen beliebigen Knoten u und v , genau dem Flächeninhalt des Kreises mit Radius r um v . Da der Flächeninhalt πr^2 beträgt, und r in $\Theta(\frac{1}{\sqrt{n}})$ liegt, liegt p_{uv} in $\Theta(\frac{1}{n})$. Dies hat den Effekt, dass der erwartete Knotengrad eines jeden Knoten konstant ist, da der erwartete Knotengrad np_{uv} beträgt.

2.2.2 Geometrische Zufallsgraphen mit Temperatur

Das einfache Modell der RGGs kann um das Konzept der Temperatur erweitert werden. Die Knotenpaare deren euklidischer Distanz geringer ist als r werden weiterhin sicher mit einer Kante verbunden. Die Idee ist nun, dass auch Knoten mit einem größeren Abstand miteinander eine gewisse Wahrscheinlichkeit haben mit einer Kante verbunden zu sein und dass diese Wahrscheinlichkeit mit steigendem Abstand abnimmt. Wie stark diese Wahrscheinlichkeit abnimmt wird durch den Parameter T parametrisiert. In dieser Arbeit betrachten wir die folgende Funktion mit einem Temperaturparameter T für die Kantenwahrscheinlichkeit:

$$\Pr[uv \in E \mid \text{dist}(u, v) = x] = \begin{cases} \left(\frac{r}{x}\right)^{\frac{2}{T}}, & x > r \\ 1, & x \leq r \end{cases}$$

Die 2 im Exponent von $(\frac{r}{x})^{\frac{2}{T}}$ ist ein Normalisierungsfaktor für die Dimensionalität. Wie in der folgenden Skizze zu erkennen nimmt die Kantenwahrscheinlichkeit mit sinkendem T ab und im Grenzfall $T = 0$ entspricht dieses Modell dem RGG ohne Temperatur.

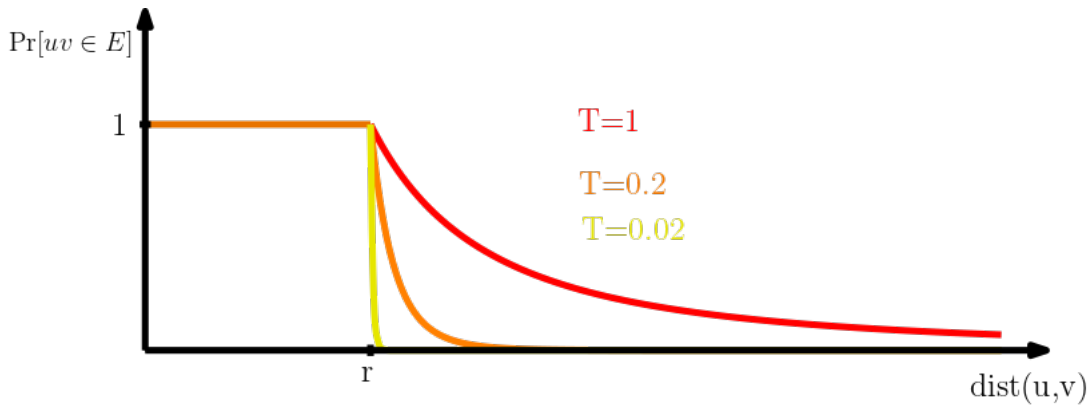


Figure 2.1: Kantenwahrscheinlichkeit in Abhängigkeit der Distanz der Knoten von verschiedenen Parametern T

2.2.3 GIRG

In dieser Arbeit werden wir nur die simpelste Form eines GIRG ohne Temperatur, mit konkreter Gewichtsfunktion und im Eindimensionalen betrachten. Im Gegensatz zu RGGs platzieren wir die Knoten nicht auf dem zweidimensionalen Torus Einheitsstorus \mathbb{T}^2 sondern auf dem eindimensionalen Einheitsstorus \mathbb{T}^1 , welcher dem Einheitskreis entspricht. Wir erhalten ihn, wenn wir beide Enden des Intervalls $[0,1]$ miteinander verbinden. Als Distanzfunktion zwischen zwei Knoten nehmen wir die Länge des kürzeren Kreisabschnitts zwischen den Knoten.

Der wirklich Entscheidende Unterschied ist zu RGGs ist jedoch, dass in diesem Modell jeder Knoten u der Knotenmenge $V = \{1, \dots, n\}$ ein Gewicht w_u zugeordnet bekommt und

der Abstand bis zu dem zwei Knoten miteinander verbunden sind von den Gewichten der beiden Knoten abhängt. Die Summe aller Gewichte bezeichnen wir mit W . Zwei Knoten werden genau dann mit einer Kante verbunden wenn die Distanz zwischen den beiden Knoten kleiner gleich $\frac{w_u w_v}{W}$ ist. Als Gewichtsfunktion wählen wir $w_u = (\frac{u}{n})^{\frac{1}{\beta-1}}$ mit $\beta > 2$. Diese Funktion erfüllt das Power-Law, was heißt, dass die Anzahl an Knoten mit einem Gewicht von mindestens w proportional zu $w^{1-\beta}$ ist.

Außerdem gilt, dass die Summe aller Gewichte W in $O(n)$ liegt. Dies lässt sich mit Lemma 2.3 zeigen. In dieser Arbeit werden wir nur den Fall Power-Law Exponent $\beta \geq 3$ betrachten.

2.3 Wahrscheinlichkeitstheoretische Grundlagen

In dieser Kapitel werden die wichtigsten wahrscheinlichkeitstheoretischen Grundlagen erläutert. Die meisten Aussagen finden sich in Standardwerken wie in [MU17].

Erwartungswert

Der Erwartungswert einer Zufallsvariable gibt an welchen Wert eine Zufallsvariable im Mittel annimmt. Sei I eine Indexmenge. Sei X eine diskreten Zufallsvariable, die die Werte x_i mit $i \in I$, annehmen kann. Dann ist der Erwartungswert von X definiert als

$$\mathbb{E}[X] = \sum_{i \in I} x_i \Pr[X = x_i].$$

Linearität des Erwartungswertes

Seien X_1, \dots, X_n beliebige Zufallsvariablen. Seien a_1, \dots, a_n beliebige Konstanten. Sei $X = \sum_{i=1}^n a_i X_i$ Dann gilt

$$\mathbb{E}[X] = \sum_{i=1}^n a_i \mathbb{E}[X_i].$$

Multiplikationsformel des Erwartungswertes

Seien X_1, \dots, X_n beliebige unabhängige Zufallsvariablen. Sei $X = \prod_{i=1}^n X_i$ Dann gilt

$$\mathbb{E}[X] = \prod_{i=1}^n \mathbb{E}[X_i].$$

Varianz

Die Varianz ist ein Maß für die Streuung einer Zufallsvariable. Sie ist die mittlere quadratische Abweichung einer Zufallsvariable von ihrem Mittelwert. Die Varianz einer beliebigen Zufallsvariable ist definiert als

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Kovarianz

Die Kovarianz ist ein Maß für den monotonen Zusammenhang zweier Zufallsvariablen. Die Kovarianz zweier Zufallsvariablen X und Y mit einer gemeinsame Wahrscheinlichkeitsverteilung lautet

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y].$$

Mithilfe der Kovarianz lässt sich ein andere Ausdruck für die Varianz finden.

Lemma 2.1. Seien X_1, \dots, X_n beliebige Zufallsvariablen. Sei $C = \sum_{k=1}^n X_k$ ihre Summe.

Dann gilt $\mathbb{V}[C] = \sum_{k=1}^n \sum_{i=1}^n \text{Cov}(X_k, X_i)$.

Proof. Aus der Definition der Varianz und Kovarianz und einfachen Termumformungen folgt die Gleichung.

$$\begin{aligned}
 \mathbb{V}[C] &= \mathbb{E}[C^2] - \mathbb{E}[C]^2 \\
 &= \mathbb{E}\left[\left(\sum_{k=1}^n X_k\right)^2\right] - \mathbb{E}\left[\sum_{k=1}^n X_k\right]^2 \\
 &= \mathbb{E}\left[\sum_{k=1}^n \sum_{i=1}^n X_k X_i\right] - \left(\sum_{k=1}^n \mathbb{E}[X_k]\right)^2 \\
 &= \sum_{k=1}^n \sum_{i=1}^n \mathbb{E}[X_k X_i] - \sum_{k=1}^n \sum_{i=1}^n \mathbb{E}[X_k] \mathbb{E}[X_i] \\
 &= \sum_{k=1}^n \sum_{i=1}^n \mathbb{E}[X_k X_i] - \mathbb{E}[X_k] \mathbb{E}[X_i] \\
 &= \sum_{k=1}^n \sum_{i=1}^n \text{Cov}(X_k, X_i)
 \end{aligned}$$

□

Tschebyscheffsche Ungleichung

Die Tschebyscheffsche Ungleichung gibt mit Hilfe der Varianz eine Schranke für die Wahrscheinlichkeit, dass eine Zufallsvariable mehr als ein vorgegebenen Schwellenwert von ihrem Mittelwert abweicht. Für eine beliebige Zufallsvariable X und einen beliebigen Schwellenwert α gilt

$$\Pr[|X - \mathbb{E}[X]| < \alpha] \geq 1 - \frac{\mathbb{V}[X]}{\alpha^2}.$$

Cauchy-Schwarz Ungleichung für Erwartungswerte

Für zwei beliebige Zufallsvariablen X und Y mit einer gemeinsamen Wahrscheinlichkeitsverteilung gilt

$$\mathbb{E}[XY] \leq \sqrt{\mathbb{E}[X^2] \mathbb{E}[Y^2]}.$$

Gesetz der totalen Wahrscheinlichkeit

Sei X eine diskrete Zufallsvariable mit Wahrscheinlichkeitsfunktion f_X und Y eine stetige Zufallsvariable mit Wahrscheinlichkeitsdichte f_Y . Dann gilt

$$f_X(x) = \int f_Y(y) f_X(x | Y = y) dy.$$

Die hier gezeigte Version ist nur eine bestimmte Variante dieses Gesetzes.

Satz des totalen Erwartungswertes

Sei X eine Zufallsvariable und B eine beliebige Bedingung und $\neg B$ das entsprechende Komplementärereignis. Dann gilt

$$\mathbb{E}[X] = \mathbb{E}[X | B] \Pr[B] + \mathbb{E}[X | \neg B] \Pr[\neg B]$$

2.4 Grenzwerte wichtiger Folgen, Reihen und Produkte

Lemma 2.2. Jede Folge der Art $(1 + O(\frac{1}{n}))^{\Theta(n)}$ bzw $(1 - O(\frac{1}{n}))^{\Theta(n)}$ liegt in $\Theta(1)$.

Proof. Nach Definition der O-Notation existieren für ausreichend große n jeweils Konstanten α und β das die Folgenden Ungleichungen gelten.

$$\begin{aligned} 1 &< (1 + O(\frac{1}{n}))^{\Theta(n)} < (1 + \frac{\alpha}{n})^{\beta n} \\ 1 &> (1 - O(\frac{1}{n}))^{\Theta(n)} > (1 - \frac{\alpha}{n})^{\beta n} \end{aligned}$$

Die Folge $(1 + \frac{\alpha}{n})^{\beta n}$ ist monoton steigend und konvergiert gegen $e^{\alpha\beta}$. Die Folge $(1 - \frac{\alpha}{n})^{\beta n}$ ist monoton steigend und konvergiert gegen $e^{-\alpha\beta}$. Somit liegen beide Folgen in $\Theta(1)$. \square

Lemma 2.3. [BFG⁺] Die Reihe $\sum_{i=1}^n i^{-\frac{1}{\beta-1}}$ liegt in $\Theta(n^{\frac{\beta-2}{\beta-1}})$

Proof. Da $i^{-\frac{1}{\beta-1}}$ monoton sinkend ist gilt

$$\begin{aligned} \sum_{i=1}^n i^{-\frac{1}{\beta-1}} &\leq 1 + \int_{i=1}^n i^{-\frac{1}{\beta-1}} di \\ &= 1 + \frac{\beta-1}{\beta-2} (n^{\frac{\beta-2}{\beta-1}} - 1) = O(n^{\frac{\beta-2}{\beta-1}}) \end{aligned}$$

und

$$\begin{aligned} \sum_{i=1}^n i^{-\frac{1}{\beta-1}} &\geq n^{-\frac{1}{\beta-1}} + \int_{i=1}^n i^{-\frac{1}{\beta-1}} di \\ &\geq \frac{\beta-1}{\beta-2} n^{\frac{\beta-2}{\beta-1}} + n^{-\frac{1}{\beta-1}} = O(n^{\frac{\beta-2}{\beta-1}}) \end{aligned}$$

\square

Lemma 2.4. Das Produkt $\prod_{u=1}^{v-1} 1 + \Theta\left(\frac{n^{\frac{3-\beta}{\beta-1}}}{(uv)^{\frac{1}{\beta-1}}}\right)$ mit $\beta \geq 3$ kann für alle $v \leq n$ durch einen Term beschränkt werden der in $O(1)$ liegt.

Proof. Durch anwenden des Logarithmus folgt

$$\log\left(\prod_{u=1}^{v-1} 1 + \Theta\left(\frac{n^{\frac{3-\beta}{\beta-1}}}{(uv)^{\frac{1}{\beta-1}}}\right)\right) = \sum_{u=1}^{v-1} \log\left(1 + \Theta\left(\frac{n^{\frac{3-\beta}{\beta-1}}}{(uv)^{\frac{1}{\beta-1}}}\right)\right).$$

Mit der bekannten Ungleichung $\log(1+x) \leq x$ für $x > 0$ kann die Summe nach oben abgeschätzt werden durch

$$\sum_{u=1}^{v-1} \Theta\left(\frac{n^{\frac{3-\beta}{\beta-1}}}{(uv)^{\frac{1}{\beta-1}}}\right) = \Theta\left(n^{\frac{3-\beta}{\beta-1}}\right) v^{-\frac{1}{\beta-1}} \sum_{u=1}^{v-1} u^{-\frac{1}{\beta-1}}$$

Nach Lemma 2.3 liegt die Summe $\sum_{u=1}^{v-1} u^{-\frac{1}{\beta-1}}$ in $\Theta(n^{\frac{\beta-2}{\beta-1}})$ und der Term vereinfacht sich zu

$$\Theta\left(n^{\frac{3-\beta}{\beta-1}} v^{-\frac{1}{\beta-1}}\right) \Theta(v^{\frac{\beta-2}{\beta-1}}) = \Theta\left(\left(\frac{n}{v}\right)^{\frac{3-\beta}{\beta-1}}\right).$$

Da $\beta \geq 3$ gilt ist der Exponent kleiner als Null. Da v kleiner als n ist ist die Basis größer als eins und damit der gesamte Term in $O(1)$. Wir haben somit auch gezeigt, dass das Produkt in $O(1)$ liegt, da $e^{O(1)}$ auch in $O(1)$ liegt.

□

3. Geometrische Zufallsgraphen (RGG)

3.1 Erwartungswert der Anzahl Cliques

In diesem Abschnitt wird der Erwartungswert der Anzahl inklusionsmaximaler Cliques des Graphens $G = (V, E)$ betrachtet. Es wird zunächst gezeigt, dass der Erwartungswert der Anzahl aller Cliques in $O(n)$ liegt. Im Anschluss wird gezeigt, dass der Erwartungswert der Anzahl aller Inklusionsmaximaler Cliques in $\Omega(n)$ liegt. Daraus folgt, dass sowohl der Erwartungswert aller Cliques als auch der Erwartungswert aller inklusionsmaximaler Cliques in $\Theta(n)$ liegt.

Lemma 3.1. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph. Sei p_{uv} die Kantenwahrscheinlichkeit. Sei F eine beliebige nicht leere Teilmenge von V , der Größe k . Dann ist die Wahrscheinlichkeit, dass F eine Clique ist, kleiner gleich $(p_{uv})^{k-1}$.*

Proof. Sei $x \in F$ ein beliebiger aber fester Knoten aus F . Die Wahrscheinlichkeit, dass F eine Clique ist, ist kleiner als die Wahrscheinlichkeit, dass alle Knoten aus $F - \{x\}$ mit x verbunden sind. Dies liegt daran, dass wenn F eine Clique ist, alle Knoten aus $F - \{x\}$ mit x verbunden sein müssen. Die Wahrscheinlichkeit, dass ein Knoten mit x verbunden ist, beträgt p_{uv} . Die Wahrscheinlichkeiten, dass die Knoten aus $F - \{x\}$ mit x verbunden sind, sind unabhängig voneinander. Da $F - \{x\}$ insgesamt $k - 1$ Knoten enthält, ist die Wahrscheinlichkeit, dass alle Knoten aus $F - \{x\}$ mit x verbunden sind, $(p_{uv})^{k-1}$. □

Theorem 3.2. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit n Knoten und einem Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei X die Anzahl Cliques des Graphens G . Dann gilt dass der Erwartungswert von X in $O(n)$ liegt.*

Proof. Aufgrund der Linearität des Erwartungswertes kann der Erwartungswert der Anzahl der Cliques als Summe der Wahrscheinlichkeiten für jede Teilmenge, dass diese Teilmenge eine inklusionsmaximale Clique ist, formuliert werden. Da es $\binom{n}{k}$ k -elementige Teilmengen F von V gibt, gilt für den Erwartungswert von X

$$\mathbb{E}[X] = \sum_{k=1}^n \binom{n}{k} \Pr[F \text{ mit } |F| = k \text{ ist eine Clique}].$$

Darauf kann nun mit Hilfe von Lemma 3.1 der Term nach oben abgeschätzt werden durch

$$\begin{aligned} & \sum_{k=1}^n \binom{n}{k} (p_{uv})^{k-1} \cdot 1 \\ &= \frac{1}{p_{uv}} \sum_{k=1}^n \binom{n}{k} (p_{uv})^k \cdot 1^{n-k} \\ &\leq \frac{1}{p_{uv}} \sum_{k=0}^n \binom{n}{k} (p_{uv})^k \cdot 1^{n-k}. \end{aligned}$$

Auf diesen Term kann der allgemeine Binomische Lehrsatz angewendet werden und der Term vereinfacht sich zu

$$\frac{(1 + p_{uv})^n}{p_{uv}}$$

Da die Kantenwahrscheinlichkeit p_{uv} in $\Theta(\frac{1}{n})$ liegt, gilt, dass der Nenner des Bruchs in $\Theta(\frac{1}{n})$ liegt und der Term vereinfacht sich zu

$$\Theta(n)(1 + p_{uv})^n.$$

Da die Kantenwahrscheinlichkeit p_{uv} in $\Theta(\frac{1}{n})$ liegt, ist $(1 + p_{uv})^n$ eine Folge der Art $(1 + O(\frac{1}{n}))^n$ und liegt somit in $O(1)$. Insgesamt kann also der Erwartungswert, der Anzahl aller inklusionsmaximalen Cliques in einem geometrischen Zufallsgraphen, nach oben durch $\Theta(n)$ abgeschätzt werden. Damit liegt der Erwartungswert in $O(n)$. □

Lemma 3.3. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit n Knoten und Kantenwahrscheinlichkeit p_{uv} . Sei x ein beliebiger Knoten aus V . Dann ist die Wahrscheinlichkeit, dass x isoliert ist, $(1 - p_{uv})^{n-1}$.*

Proof. Wenn x isoliert ist, darf kein anderer Punkt mit x verbunden sein. Die Wahrscheinlichkeit, dass zwei Knoten nicht miteinander verbunden sind beträgt $1 - p_{uv}$. Da die Wahrscheinlichkeiten, ob die Knoten nicht miteinander verbunden sind, unabhängig voneinander sind, beträgt die Wahrscheinlichkeit, dass x isoliert ist $(1 - p_{uv})^{n-1}$. □

Theorem 3.4. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit n Knoten und einem Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei X die Anzahl inklusionsmaximaler Cliques des Graphens G . Dann gilt dass der Erwartungswert von X in $\Omega(n)$ liegt.*

Proof. Der Erwartungswert, der Anzahl aller inklusionsmaximalen Cliques in dem geometrischen Zufallsgraphen, ist größer als der Erwartungswert der Anzahl inklusionsmaximale Cliques, der Größe eins. Die Anzahl der inklusionsmaximalen Cliques der Größe eins entspricht der Anzahl isolierter Knoten. Aufgrund der Linearität des Erwartungswertes und da es n mögliche Knoten gibt, gilt

$$\mathbb{E}[\text{Anzahl isolierter Knoten}] = n \Pr[x \in V \text{ ist isoliert}].$$

Mit Lemma 3.3 ist dieser Term gleich

$$n(1 - p_{uv})^{n-1}.$$

Da der Radius p_{uv} in $\Theta(\frac{1}{n})$ liegt, ist $(1 - p_{uv})^{n-1}$ eine Folge der Art $(1 - O(\frac{1}{n}))^{\Theta(n)}$ und liegt somit in $\Theta(1)$. Somit kann der Erwartungswert, der Anzahl aller inklusionsmaximalen Cliques in einem geometrischen Zufallsgraphen, nach unten durch $\Theta(n)$ abgeschätzt werden. Damit liegt der Erwartungswert in $\Omega(n)$. \square

Theorem 3.5. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Dann liegt der Erwartungswert der Anzahl inklusionsmaximaler Cliques des Graphens und der Erwartungswert aller Cliques in $\Theta(n)$.*

Proof. Theorem 3.4 besagt, dass der Erwartungswert aller Cliques in $O(n)$ liegt. Theorem 3.2 besagt, dass der Erwartungswert aller inklusionsmaximaler Cliques in $\Omega(n)$ liegt. Da die Anzahl inklusionsmaximaler Cliques von der Anzahl aller Cliques dominiert wird, liegt sowohl der Erwartungswert der Anzahl aller Cliques als auch der Erwartungswert der Anzahl inklusionsmaximaler Cliques in $\Theta(n)$. \square

3.2 Konzentrationsschranke

In diesem Abschnitt zeigen wir, dass die Anzahl Cliques mit hoher Wahrscheinlichkeit kleiner ist als αn für ausreichend große Konstanten α .

Zu Beginn betrachten wir eine Motivation des Beweises. Die Idee ist ein Gitter auf den Torus zu legen und dann für jede Zelle des Gitters die Anzahl an Cliques mit Knoten in dieser Zelle zu betrachten. Es zeigt sich, dass die Anzahl Cliques pro Zelle für die meisten Zellen quasi unabhängig voneinander ist. Die Summe über alle Zellen ist also quasi eine Summe unabhängiger gleicher Zufallsvariablen. Nach dem Gesetz der großen Zahlen folgt, dass die Summe mit hoher Wahrscheinlichkeit relativ nah an ihrem Erwartungswert liegt. Entlang dieser Idee orientiert sich der Beweis.

Technisch umgesetzt wird der Beweis mithilfe der tschebyscheffschen Ungleichung, welche mithilfe der Varianz eine Abschätzung für die Abweichung einer Zufallsvariable von ihren Erwartungswert gibt.

$$\Pr[|X - \mathbb{E}[X]| < \alpha] \geq 1 - \frac{\mathbb{V}[X]}{\alpha^2}$$

Da sich die Varianz der Anzahl Cliques nur schwer berechnen lässt, verwenden wir stattdessen eine nur leicht abweichende Zufallsvariable, welche die Anzahl Cliques dominiert.

3.2.1 Gittergröße

Dazu betrachten wir ein Gitter über die Oberfläche des Torus. Wir bezeichnen die Menge aller Zellen mit Z . Die Idee ist, die Größe des Gitters so zu wählen, dass wir n gleiche quadratische Zellen haben. Da die Einteilung in n quadratische Zellen nur möglich ist wenn n eine Quadratzahl ist, wird n auf die nächste mindestens so große Quadratzahl d aufgerundet und statt einem $\sqrt{n} \times \sqrt{n}$ - Gitter ein $\sqrt{d} \times \sqrt{d}$ - Gitter über die Oberfläche des Torus gelegt. Da die nächst größere Quadratzahl einer Zahl um maximal einen konstanten Faktor größer ist, liegt d in $\Theta(n)$. Somit ist das $\sqrt{d} \times \sqrt{d}$ - Gitter auch ein $\Theta(\sqrt{n}) \times \Theta(\sqrt{n})$ - Gitter. Die Fläche einer Zelle z ist somit auch $\Theta(\frac{1}{n})$ groß.

3.2.2 Nachbarschaft einer Zelle

Sei $z \in Z$ eine beliebige Zelle. Knoten aus weit entfernten Zellen können nicht mit Knoten aus z verbunden sein. Wir bezeichnen die Zellen, die nah genug an z liegen, sodass sie Knoten enthalten können, die mit Knoten aus Zelle z verbunden sind als *Nachbarschaft* der Zelle z

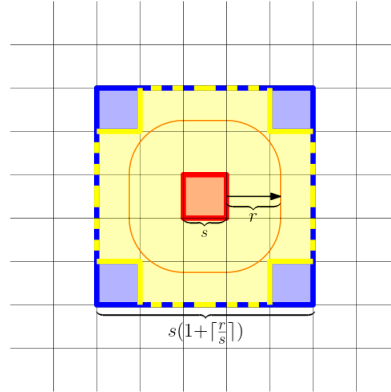


Figure 3.1: Die dünne orange Linie um das rote Zelle umrandet das Gebiet welches Knoten enthält, die mit Knoten aus der roten Zelle verbunden sein können. Die gelbe Linie umrandet alle Zellen, die in der Nachbarschaft von der roten Zelle enthalten sind. Die blaue Linie umrandet $(1 + 2\lceil \frac{r}{s} \rceil)^2$ viele Zellen. Also die Abschätzung nach Maximumsnorm für die Anzahl Zellen in der Nachbarschaft einer Zelle.

Lemma 3.6. Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei Z die Menge aller Zellen eines $\Theta(\sqrt{n}) \times \Theta(\sqrt{n})$ - Gitters. Dann besteht die Nachbarschaft einer Zelle $z \in Z$ aus konstant vielen Zellen.

Proof. Sei s die Seitenlänge einer Zelle. Dann enthält die Nachbarschaft von z sicher weniger als $(1 + 2\lceil \frac{r}{s} \rceil)^2$ Zellen, da dies die Zellen sind, welche nach Maximumnorm Knoten enthalten können, die weniger als r von Knoten aus z entfernt sein können (siehe Abbildung 1). Die Anzahl $(1 + 2\lceil \frac{r}{s} \rceil)^2$ liegt in $O(1)$, da sowohl der Radius r als auch die Seitenlänge s in $\Theta(\frac{1}{\sqrt{n}})$ liegt. \square

3.2.3 Abschätzung der Anzahl Cliques

Um die Anzahl Cliques abzuschätzen ordnen wir jeder Zelle alle Cliques zu, die Knoten in dieser Zelle haben. Beachte, dass jede Clique immer in mindestens einer Zelle liegt, ist die Summe über alle Zellen größer als die Anzahl Cliques. einer oder mehreren Zellen zugewiesen wird. Die Anzahl zugewiesener Zellen summiert über alle Zellen ist also größer als die Anzahl aller Cliques. Die Anzahl Cliques mit Knoten in Zelle z kann also durch $2^{|N_z|}$ abgeschätzt werden, wobei N_z die Menge Knoten in der Nachbarschaft von Zelle z bezeichnet. Insgesamt können wir die Anzahl Cliques nach oben abschätzen durch $C = \sum_{z \in Z} 2^{|N_z|}$. Diese Zufallsvariable C domiert also die Anzahl aller Cliques X

3.2.4 Erwartungswert einer Zelle

Um später die Tschebyscheffsche Ungleichung auf C anzuwenden benötigen wir den Erwartungswert von C . Dazu betrachten wir zunächst den Erwartungswert $\mathbb{E}[2^{|N_z|}]$ einer

einzelnen Zelle z . Dafür brauchen wir noch eine weitere Zufallsvariable als Hilfe.

Sei N_z^v die Indikatorvariable dafür ob Knoten v in Nachbarschaft von Zelle z liegt. Sie ist also 1, wenn v in der Nachbarschaft von z liegt und 0 wenn v außerhalb der Nachbarschaft von z liegt und $\Pr[v \in N_z] = \Pr[N_z^v = 1]$. Damit gilt für die Anzahl der Knoten in der Nachbarschaft einer Zelle $|N_z| = \sum_{z \in Z} N_z^v$.

Es sei angemerkt, dass das folgende Lemma etwas allgemeiner formuliert ist, da wir es für den Beweis der Varianz brauchen.

Lemma 3.7. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei Z die Menge der Zellen des $\Theta(\sqrt{n}) \times \Theta(\sqrt{n})$ -Gitters. Sei N_z die Menge aller Knoten in der Nachbarschaft einer Zelle z . Dann gilt für eine beliebige Konstante $a \in \mathbb{N}$ und eine beliebige Zelle $z \in Z$, dass $\mathbb{E}[a^{|N_z|}] = (1 + (a - 1)\Pr[v \in N_z])^n \in O(1)$ (für $v \in V$).*

Proof. Aus der Definition für die Indikatorvariable N_z^v wissen wir, dass die Anzahl der Knoten in der Nachbarschaft von z als Summe über die einzelnen Indikatorvariablen der Knoten formuliert werden kann. Wir erhalten

$$\mathbb{E}[a^{|N_z|}] = \mathbb{E}\left[a^{\sum_{z \in Z} N_z^v}\right] = \mathbb{E}\left[\prod_{z \in Z} a^{N_z^v}\right].$$

Da die Wahrscheinlichkeit ob ein Knoten innerhalb der Nachbarschaft einer Zelle liegt, für alle Knoten unabhängig voneinander ist, kann der Erwartungswert in das Produkt hereingezogen werden. Damit gilt

$$\begin{aligned} \mathbb{E}\left[\prod_{z \in Z} a^{N_z^v}\right] &= \prod_{z \in Z} \mathbb{E}[a^{N_z^v}] \\ &= \prod_{z \in Z} a^1 \Pr[N_z^v = 1] + a^0 \Pr[N_z^v = 0] \end{aligned}$$

Da N_z^v nur entweder 1 oder 0 sein kann, ist $N_z^v = 0$ das Gegenereignis zu $N_z^v = 1$ und $\Pr[N_z^v = 0] = 1 - \Pr[N_z^v = 1]$. Damit vereinfacht sich der obige Term zu

$$\begin{aligned} &\prod_{z \in Z} (a \Pr[N_z^v = 1] + (1 - \Pr[N_z^v = 1])) \\ &= \prod_{z \in Z} (1 + (a - 1)\Pr[N_z^v = 1]). \end{aligned}$$

Da die Wahrscheinlichkeit, dass ein Knoten innerhalb der Nachbarschaft einer Zelle liegt, für alle Knoten und Zellen gleich ist, kann das Produkt als Potenz mit der Basis eines beliebigen Knotens v und einer beliebigen Zelle z

$$\begin{aligned} &(1 + (a - 1)\Pr[N_z^v = 1])^n \\ &= (1 + (a - 1)\Pr[v \in N_z])^n \end{aligned}$$

formuliert werden. Nach Lemma 3.6 besteht die Nachbarschaft einer Zelle aus konstant vielen Zellen. Da die Fläche einer Zelle $O\left(\frac{1}{n}\right)$ groß ist, ist auch die Nachbarschaft einer Zelle $O\left(\frac{1}{n}\right)$ groß. Damit liegt auch die Wahrscheinlichkeit, dass ein Knoten in der Nachbarschaft liegt, in $O\left(\frac{1}{n}\right)$. Der obige Term kann also umgeformt werden zu

$$\left(1 + O\left(\frac{1}{n}\right)\right)^n$$

Aus den Grundlagen wissen wir, dass diese Folge in $O(1)$ liegt. □

3.2.5 Erwartungswert aller Zellen

Nun kann der Erwartungswert von C , also der Summe über alle Zellen, berechnet werden. Wir zeigen, dass der Erwartungswert von C , ebenso wie der Erwartungswert der Anzahl Cliques in $O(n)$ liegt.

Lemma 3.8. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei Z die Menge der Zellen eines $\Theta(\sqrt{n}) \times \Theta(\sqrt{n})$ -Gitters. Sei N_z die Menge aller Knoten in der Nachbarschaft einer Zelle $z \in Z$. Sei $C = \sum_{z \in Z} 2^{|N_z|}$. Dann ist $\mathbb{E}[C] \in O(n)$.*

Proof. Durch Einsetzen der Definition von C folgt

$$\mathbb{E}[C] = \mathbb{E} \left[\sum_{k \in Z} 2^{|N_k|} \right].$$

Aufgrund der Linearität des Erwartungswert kann der Erwartungswert in die Summe hereingezogen werden. Damit gilt

$$\mathbb{E}[C] = \sum_{k \in Z} \mathbb{E} \left[2^{|N_k|} \right]$$

Da die Anzahl Zellen $|Z|$ in $\Theta(n)$ liegt (siehe 3.2.1) und der Erwartungswert von $2^{|N_z|}$ nach Lemma 3.7 in $O(1)$ liegt, liegt der Erwartungswert von C somit in $O(n)$. \square

Da C eine Abschätzung nach oben für die Anzahl an Cliques ist, folgt aus diesem Lemma automatisch, dass die Anzahl Cliques auch in $O(n)$ liegt. Es ist also ein alternativer aber etwas komplizierterer Beweis als der Beweis aus dem ersten Kapitel. Der Vorteil von C ist jedoch, dass wir nun leichter die Varianz berechnen können.

3.3 Varianz der Abschätzung

Um die Varianz von $C = \sum_{z \in Z} 2^{|N_z|}$ abschätzen zu können, reicht es die Kovarianzen, der Zufallsvariablen $2^{|N_z|}$ zu betrachten (siehe 2.1). Zuerst machen wir eine Abschätzung für die Kovarianz von $2^{|N_{z_1}|}$ und $2^{|N_{z_2}|}$ für beliebige Zellen z_1 und z_2 . Es kann also insbesondere auch sein, dass die Nachbarschaften von z_2 und z_1 sich überschneiden oder sogar $z_2 = z_1$ gilt und damit $2^{|N_{z_1}|}$ und $2^{|N_{z_2}|}$ eine relativ hohe Kovarianz haben. Danach machen wir eine Abschätzung für die Kovarianz von $2^{|N_{z_1}|}$ und $2^{|N_{z_2}|}$ für Zellen z_2 und z_1 , deren Nachbarschaften sich nicht überschneiden. Für diese Zellen zeigen wir, dass $2^{|N_{z_1}|}$ und $2^{|N_{z_2}|}$ sogar eine negative Kovarianz haben.

3.3.1 Kovarianz benachbarter Zellen

Lemma 3.9. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei Z die Menge der Zellen eines $\Theta(\sqrt{n}) \times \Theta(\sqrt{n})$ -Gitters. Sei N_z die Menge aller Knoten in der Nachbarschaft einer Zelle $z \in Z$. Dann gilt für zwei beliebige Zellen $i, k \in Z$, dass $\text{Cov}(2^{|N_k|}, 2^{|N_i|}) \in O(1)$.*

Proof. Nach der Definition der Kovarianz gilt die folgende Gleichung.

$$\text{Cov}(2^{|N_k|}, 2^{|N_i|}) = \mathbb{E} \left[2^{|N_k|} 2^{|N_i|} \right] - \mathbb{E} \left[2^{|N_k|} \right] \mathbb{E} \left[2^{|N_i|} \right]$$

Da der Erwartungswert von $2^{|N_k|}$ und $2^{|N_i|}$ größer als Null ist kann der Term nach oben abgeschätzt werden durch

$$\mathbb{E} \left[2^{|N_k|} 2^{|N_i|} \right]$$

Aufgrund der Cauchy-Schwarz Ungleichung (siehe 2.3) kann dieser Term noch weiter nach oben abgeschätzt werden durch

$$\sqrt{\mathbb{E} \left[2^{|N_k|} 2^{|N_k|} \right] \mathbb{E} \left[2^{|N_i|} 2^{|N_i|} \right]}$$

Da sich die Zellen k und k nicht unterscheiden ist auch der Erwartungswert von $2^{|N_i|} 2^{|N_i|}$ und $2^{|N_k|} 2^{|N_k|}$ für alle Zellen gleich. Somit gilt

$$\begin{aligned} \sqrt{\mathbb{E} \left[2^{|N_k|} 2^{|N_k|} \right] \mathbb{E} \left[2^{|N_i|} 2^{|N_i|} \right]} &= \sqrt{\mathbb{E} \left[2^{|N_k|} 2^{|N_k|} \right] \mathbb{E} \left[2^{|N_k|} 2^{|N_k|} \right]} \\ &= \mathbb{E} \left[2^{|N_k|} 2^{|N_k|} \right] \\ &= \mathbb{E} \left[4^{|N_k|} \right] \end{aligned}$$

Mit Lemma 3.7 folgt, dass dieser Erwartungswert in $O(1)$ liegt. Damit liegt auch die Kovarianz von $2^{|N_k|}$ und $2^{|N_i|}$ in $O(1)$. \square

3.3.2 Kovarianz nicht benachbarter Zellen

Lemma 3.10. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei Z die Menge der Zellen eines $\Theta(\sqrt{n}) \times \Theta(\sqrt{n})$ -Gitters. Sei N_z die Menge aller Knoten in der Nachbarschaft einer Zelle $z \in Z$. Dann gilt für zwei beliebige Zellen $i, k \in Z$ deren Nachbarschaften sich nicht überschneiden, dass $\text{Cov}(2^{|N_k|}, 2^{|N_i|}) < 0$.*

Proof. Nach der Definition der Kovarianz gilt

$$\text{Cov}(2^{|N_k|}, 2^{|N_i|}) = \mathbb{E} \left[2^{|N_k|} 2^{|N_i|} \right] - \mathbb{E} \left[2^{|N_k|} \right] \mathbb{E} \left[2^{|N_i|} \right]$$

Da die Zellen i und k gleich sind ist auch der Erwartungswert der Zufallsvariablen $2^{|N_k|}$ und $2^{|N_i|}$ gleich. Somit vereinfacht sich der obige Term zu

$$\mathbb{E} \left[2^{|N_k|} 2^{|N_i|} \right] - \mathbb{E} \left[2^{|N_k|} \right]^2.$$

Wie im Beweis des Lemma 3.7 kann $|N_k|$ auch als Summe über die Indikatorvariablen N_j^k geschrieben werden. Damit kann der obige Term umgeformt werden zu

$$\begin{aligned} &\mathbb{E} \left[2^{\sum_{j=1}^n N_j^k} 2^{\sum_{j=1}^n N_j^i} \right] - \mathbb{E} \left[2^{|N_k|} \right]^2 \\ &= \mathbb{E} \left[2^{\sum_{j=1}^n N_j^k + N_j^i} \right] - \mathbb{E} \left[2^{|N_k|} \right]^2. \end{aligned}$$

Nach Lemma 3.7 ist $\mathbb{E} \left[2^{|N_k|} \right] = (1 + \Pr[v \in N_z])^n$ für ein beliebigen Knoten v und eine beliebige Zelle z . Damit kann der obige Term umgeformt werden zu

$$\begin{aligned} &\mathbb{E} \left[2^{\sum_{j=1}^n N_j^k + N_j^i} \right] - (1 + \Pr[v \in N_z])^{2n} \\ &= \mathbb{E} \left[\prod_{j=1}^n 2^{N_j^k + N_j^i} \right] - (1 + \Pr[v \in N_z])^{2n} \end{aligned}$$

Da die Positionen der Knoten unabhängig voneinander gewählt werden, sind auch die Zufallsvariablen $2^{N_k^j + N_i^j}$ für alle Knoten j unabhängig voneinander. Der Erwartungswert kann somit in das Produkt rein gezogen werden. Damit kann der obige Term umgeformt werden zu

$$\left(\prod_{j=1}^n \mathbb{E} \left[2^{N_k^j + N_i^j} \right] \right) - (1 + \Pr[v \in N_z])^{2n}$$

Der Erwartungswert von $2^{N_k^j + N_i^j}$ ist für alle Knoten gleich. Das Produkt kann also als Potenz mit der Basis eines beliebigen Knotens j geschrieben werden. Damit kann der obige Term umgeformt werden zu

$$= \left(\mathbb{E} \left[2^{N_k^j + N_i^j} \right] \right)^n - (1 + \Pr[v \in N_z])^{2n}$$

Die Nachbarschaften von i und k überschneiden sich nicht. Damit kann ein Knoten j nicht in beiden Nachbarschaften gleichzeitig sein. Die Zufallsvariable $N_k^j + N_i^j$ kann also nur entweder 1 oder 0 sein. Damit kann der obige Term umgeformt werden zu

$$\left(2^1 \Pr[N_k^j + N_i^j = 1] + 2^0 \Pr[N_k^j + N_i^j = 0] \right)^n - (1 + \Pr[v \in N_z])^{2n}$$

Da die Zufallsvariable $N_k^j + N_i^j$ entweder 1 oder 0 sein kann ist $N_k^j + N_i^j = 0$ das Gegenereignis zu $N_k^j + N_i^j = 1$ und damit ist $\Pr[N_k^j + N_i^j = 0] = 1 - \Pr[N_k^j + N_i^j = 1]$. Der obige Term kann damit umgeformt werden zu

$$= \left(2 \Pr[N_k^j + N_i^j = 1] + \left(1 - \Pr[N_k^j + N_i^j = 1] \right) \right)^n - (1 + \Pr[v \in N_z])^{2n}$$

Da die Nachbarschaften von Zelle k und Zelle i sich nicht überschneiden, ist die Wahrscheinlichkeit, dass ein Knoten innerhalb einer der beiden Nachbarschaften liegt, die Summe der Flächen der beiden Nachbarschaften und damit doppelt so groß wie die Wahrscheinlichkeit, dass ein Knoten z innerhalb der Nachbarschaft einer Zelle k liegt. Der obige Term kann damit umgeformt werden zu

$$\begin{aligned} &= (4 \Pr[v \in N_z] + (1 - 2 \Pr[v \in N_z]))^n - (1 + \Pr[v \in N_z])^{2n} \\ &= (1 + 2 \Pr[v \in N_z])^n - (1 + 2 \Pr[v \in N_z] + \Pr[v \in N_z]^2)^n \\ &< 0 \end{aligned}$$

□

3.3.3 Abschätzung der Varianz

Um nun den Beweis für die Varianz von C zusammenzusetzen, ist es noch notwendig zu wissen, für wie viele Zellen wir die zweite bessere Abschätzung der Kovarianz verwenden können. Wir hatten bereits gezeigt, dass die Nachbarschaft einer Zelle nur aus konstant vielen Zellen besteht. Nun wollen wir zeigen, dass es für jede Zelle z nur konstant viele Zellen gibt deren Nachbarschaft sich mit der Nachbarschaft von z überschneidet.

Lemma 3.11. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei Z die Menge aller Zellen eines $\Theta(\sqrt{n}) \times \Theta(\sqrt{n})$ - Gitters. Sei $z \in Z$ eine beliebige Zelle. Dann gibt es nur konstant viele Zellen deren Nachbarschaften sich mit der Nachbarschaft von z überschneiden.*

Proof. Sei i eine weitere beliebige Zelle deren Nachbarschaft sich mit der Nachbarschaft von z überschneidet. Dann gibt es eine Zelle k die sowohl in der Nachbarschaft von z als auch in der Nachbarschaft i liegt. Da i in der Nachbarschaft von k liegt, liegt auch k in der Nachbarschaft von i . Zelle i muss somit in der Nachbarschaft einer Zelle liegen, die in der Nachbarschaft von z liegt. Da die Nachbarschaft einer Zelle aus $O(1)$ Zellen besteht kann es nur $O(1) \cdot O(1) = O(1)$ Zellen k geben die in Nachbarschaft einer Zelle liegen, die in der Nachbarschaft von z liegt. \square

Mit diesen Lemmata können wir nun endlich zeigen, dass die Varianz in $O(n)$ liegt.

Lemma 3.12. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei Z die Menge der Zellen des $\Theta(\sqrt{n}) \times \Theta(\sqrt{n})$ -Gitters. Sei N_z die Menge aller Knoten in der Nachbarschaft einer Zelle $z \in Z$. Sei $C = \sum_{z \in Z} 2^{|N_z|}$. Dann ist $\mathbb{V}[C] \in O(n)$.*

Proof. Da C eine Summe von Zufallsvariablen ist, kann die Varianz von C als Summe über alle paarweisen Kovarianzen der Summanden geschrieben werden (siehe 2.1). Damit gilt

$$\mathbb{V}[C] = \sum_{k \in Z} \sum_{i \in Z} \text{Cov}(2^{|N_k|}, 2^{|N_i|})$$

Da alle Zellen gleich sind, ist auch die Summe über die Kovarianzen mit allen Zellen für alle Zellen gleich. Somit entspricht der Term genau $|Z|$ -mal der Summe über die Kovarianzen mit allen Zellen für eine beliebige Zelle z . Somit gilt

$$\mathbb{V}[C] = |Z| \sum_{i \in Z} \text{Cov}(2^{|N_z|}, 2^{|N_i|})$$

Wir wissen aus Lemma 3.11, dass es nur konstant viele Zellen i gibt deren Nachbarschaft sich mit der Nachbarschaft von z schneidet. Auf diese Kovarianzen müssen wir die $O(1)$ -Abschätzung von Lemma 3.9 verwenden. Für die restlichen Zellen können wir allerdings die kleiner 0 Abschätzung aus Lemma 3.10 verwenden. Der obige Term kann somit nach oben abgeschätzt werden durch

$$|Z|(O(1) * O(1) + O(n) * 0)$$

Da die Anzahl an Zellen $|Z|$ in $O(n)$ liegt (siehe 3.2.1), gilt, dass die Varianz von C in $O(n)$ liegt. \square

3.3.4 Konzentrationsschranke

Mithilfe der Erkenntnis, dass sowohl der Erwartungswert als auch die Varianz der Zufallsvariable C in $O(n)$ liegt, kann nun mit der tschebyscheffschen Ungleichung, die eigentliche Konzentrationsschranke für die Anzahl Cliques gezeigt werden. Der nun folgende Beweis ist im Prinzip eine Formalisierung des in Abbildung 3.3.4 dargestellten Beweises.

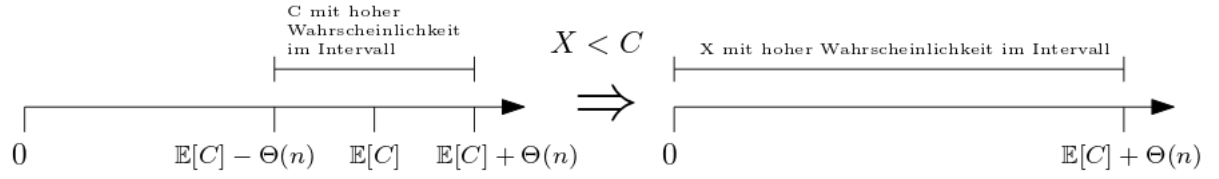


Figure 3.2: Mithilfe der tschebyscheffschen Ungleichung ergibt sich die linke Seite der Skizze. Durch die im vorherigen Abschnitt gezeigten Eigenschaften ergibt sich die rechte Seite.

Theorem 3.13. Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$. Sei X die Anzahl an Cliques in G . Dann existiert eine Konstante α , sodass gilt $\Pr[X < \alpha n] \geq 1 - O\left(\frac{1}{n}\right)$.

Proof. Sei α eine Konstante für die gilt, dass $\alpha n - \mathbb{E}[C] \in \Theta(n)$. Da $\mathbb{E}[C] \in O(n)$ (Lemma 3), existiert diese Konstante auch wirklich.

Die Zufallsvariable C dominiert die Anzahl Cliques X (siehe 3.2.3). Somit gilt

$$\begin{aligned} \Pr[X < \alpha n] &\geq \Pr[C < \alpha n] \\ &= \Pr[C - \mathbb{E}[C] < \alpha n - \mathbb{E}[C]] \end{aligned}$$

An dieser Stelle kann die Voraussetzung für α eingesetzt werden und es gilt

$$\Pr[C - \mathbb{E}[C] < \alpha n - \mathbb{E}[C]] = \Pr[|C - \mathbb{E}[C]| < \Theta(n)]$$

Mithilfe der tschebyscheffschen Ungleichung kann die Wahrscheinlichkeit nach unten abgeschätzt werden durch

$$1 - \frac{\mathbb{V}[C]}{\Theta(n^2)}$$

Da die Varianz nach Lemma 13 in $O(n)$ liegt kann dieser Term nach unten abgeschätzt werden durch

$$1 - O\left(\frac{1}{n}\right)$$

□

Die eben gezeigte Konzentrationsschranke gilt ebenfalls für die Anzahl an inklusionsmaximaler Cliques, da die Anzahl inklusionsmaximaler Cliques auf jeden Fall kleiner ist als die Anzahl Cliques.

4. Geometrische Zufallsgraphen (RGG) mit Temperatur

4.1 Erwartungswert der Anzahl Cliques

In diesem Abschnitt wird für $T \in (0, 1)$ gezeigt, dass die Kantenwahrscheinlichkeit p_{uv} in $\Theta(\frac{1}{n})$ liegt. Damit folgt mit Lemma 3.4, dass dann auch der Erwartungswert der Anzahl Cliques in $\Theta(n)$ liegt.

4.1.1 Wahrscheinlichkeitsverteilung des Abstands

Die Wahrscheinlichkeit p_{uv} , dass zwei beliebige Knoten u und v miteinander verbunden sind, ist abhängig von dem Abstand der Knoten $\text{dist}(u, v)$. Da der Abstand der Knoten $\text{dist}(u, v)$ eine Zufallsvariable ist, müssen zunächst dessen Wahrscheinlichkeitsverteilung F_{dist} betrachten um p_{uv} berechnen zu können.

Für einen beliebigen Knoten u gilt, dass alle Knoten innerhalb eines Kreises mit Radius r um u ein Abstand weniger als r von u haben. Die Wahrscheinlichkeit, dass beliebiger Knoten $v \neq u$ einen Abstand kleiner gleich r zu u hat, entspricht also genau der Fläche des Kreises πr^2 . Dies gilt allerdings nur für einen Abstand kleiner als $\frac{1}{2}$, da sich ein Kreis mit einem größerem Radius auf dem \mathbb{T}^2 mit sich selbst schneidet und damit nicht mehr genau eine Fläche von πr^2 abdeckt. Für den Beweis ist allerdings nicht notwendig den genauen Verlauf von F_{dist} ab diesem Bereich zu wissen sondern es reicht zu wissen, dass der Abstand zwischen zwei Knoten nicht größer werden kann als $\frac{1}{\sqrt{2}}$ auf dem Torus.

Aus der Wahrscheinlichkeitsfunktion lässt sich auch die Dichtefunktion f_{dist} herleiten. Für Dichtefunktion f_{dist} gilt im Intervall $[0, \frac{1}{2}]$

$$f_{\text{dist}}(x) = F'_{\text{dist}}(x) = 2\pi r. \quad (4.1)$$

4.1.2 Kantenwahrscheinlichkeit

Um die Kantenwahrscheinlichkeit $p_{uv} = \Pr[uv \in E]$ aus der Kantenwahrscheinlichkeit in Abhängigkeit der Distanz $\Pr[uv \in E \mid \text{dist}(u, v) = x]$ zu berechnen, brauchen wir das Gesetz der totalen Wahrscheinlichkeit 2.3. Wir können Es verwenden, da wir formal gesehen

$\Pr[uv \in E]$ auch als Wahrscheinlichkeit betrachten können, dass die Indikatorvariable ob die Kante uv existiert gleich eins ist.

$$\Pr[uv \in E] = \int_{x=0}^{\frac{1}{\sqrt{2}}} \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx.$$

In den Folgenden drei Lemmata wird diese Integral in unterschiedlichen Intervallen betrachtet.

Das erste Intervall ist $[0, r]$. Der daraus resultierende Anteil an der Wahrscheinlichkeit entspricht der Wahrscheinlichkeit im einfachen Modell ohne Temperatur.

Lemma 4.1. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$ und Temperatur $T \in (0, 1)$. Dann gilt für zwei beliebige Knoten $u, v \in V$ dass*

$$\int_{x=0}^r \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx \in \Theta\left(\frac{1}{n}\right).$$

Proof. Da die Wahrscheinlichkeit, dass die Kante uv existiert eins ist wenn die Knotendistanz kleiner als r ist gilt

$$\begin{aligned} & \int_{x=0}^r \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx \\ &= \int_{x=0}^r f_{\text{dist}}(x) dx \\ &= [F_{\text{dist}}(x)]_0^r \\ &= \pi r^2. \end{aligned}$$

Da $r \in \Theta\left(\frac{1}{\sqrt{n}}\right)$ gilt, liegt dieser teil des Integrals in $\Theta\left(\frac{1}{n}\right)$ □

Das nächste Intervall $[r, \frac{1}{2}]$ ist der Bereich wo die Wahrscheinlichkeit mit der Distanz geringer wird, aber die Distanz noch nicht so groß wird, dass ein Kreis mit Radius aus $[r, \frac{1}{2}]$ sich selbst schneidet.

Lemma 4.2. *Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$ und*

Temperatur $T \in (0, 1)$. Dann gilt für zwei beliebige Knoten $u, v \in E$, dass $\int_{x=r}^{\frac{1}{2}} \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx \in O\left(\frac{1}{n}\right)$

Proof. Durch einsetzen der Dichtefunktion und der Kantenwahrscheinlichkeit in abhängigkeit der Distanz gilt

$$\begin{aligned} & \int_{x=r}^{\frac{1}{2}} \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx \\ &= \int_{x=r}^{\frac{1}{2}} \left(\frac{r}{x}\right)^{\frac{2}{T}} 2\pi x dx \\ &= 2\pi(r^2)^{\frac{1}{T}} \int_{x=r}^{\frac{1}{2}} x^{1-\frac{2}{T}} dx. \end{aligned}$$

Da $T \in (0, 1)$ und $r \in \Theta\left(\frac{1}{\sqrt{n}}\right)$ liegt der Faktor vor dem Integral in $O\left(\frac{1}{n}\right)$ und der Term vereinfacht sich zu

$$O\left(\frac{1}{n}\right) \int_{x=r}^{\frac{1}{2}} x^{1-\frac{2}{T}} dx.$$

Da $T \neq 1$ kann mit der normalen Potenzregel integriert werden. Damit gilt für das Integral

$$\int_{x=r}^{\frac{1}{2}} x^{1-\frac{2}{T}} dx = \left[\left(\frac{1}{2-\frac{2}{T}} \right) x^{2-\frac{2}{T}} \right]_r^{\frac{1}{2}} = \left[x^{2-\frac{2}{T}} \right]_r^{\frac{1}{2}} = \left(\left(\frac{1}{2} \right)^{2-\frac{2}{T}} - r^{2-\frac{2}{T}} \right) \leq \left(\frac{1}{2} \right)^{2-\frac{2}{T}}$$

Da T eine konstante ist, liegt dieser Term in $O\left(\frac{1}{n}\right)$

□

Im letzten Intervall $[\frac{1}{2}, \frac{1}{\sqrt{2}}]$ können wir für $f_{dist}(x)$ nicht mehr den einfachen Term $f_{dist}(x)$ verwenden.

Lemma 4.3. Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta\left(\frac{1}{\sqrt{n}}\right)$ und Temperatur $T \in (0, 1)$. Dann gilt für zwei beliebige Knoten $u, v \in V$ dass

$$\int_{x=\frac{1}{2}}^{\frac{1}{\sqrt{2}}} \Pr[uv \in E | dist(u, v) = x] f_{dist}(x) dx \in \Theta\left(\frac{1}{n}\right)$$

Proof. Da die Kantenwahrscheinlichkeit monoton sinkend in der Knotendistanz ist, ist das Maximum der bedingten Wahrscheinlichkeit $\Pr[uv \in E | dist(u, v) = x]$ auf dem Intervall $[\frac{1}{2}, \frac{1}{\sqrt{2}}]$ an der Stelle $\frac{1}{2}$. Somit gilt für das Integral

$$\begin{aligned} \int_{x=\frac{1}{2}}^{\frac{1}{\sqrt{2}}} \Pr[uv \in E | dist(u, v) = x] f_{dist}(x) dx &\leq \int_{x=\frac{1}{2}}^{\frac{1}{\sqrt{2}}} \Pr\left[uv \in E | dist(u, v) = \frac{1}{2}\right] f_{dist}(x) dx \\ &= \Pr\left[uv \in E | dist(u, v) = \frac{1}{2}\right] \int_{x=\frac{1}{2}}^{\frac{1}{\sqrt{2}}} f_{dist}(x) dx \\ &= \left(\frac{r^2}{\frac{1}{2}}\right)^{\frac{1}{T}} \int_{x=\frac{1}{2}}^{\frac{1}{\sqrt{2}}} f_{dist}(x) dx \end{aligned}$$

Da $T \in (0, 1)$ und $r \in \Theta\left(\frac{1}{\sqrt{n}}\right)$ liegt der Faktor vor dem Integral in $O\left(\frac{1}{n}\right)$ und der obige Term vereinfacht sich zu

$$O\left(\frac{1}{n}\right) \int_{x=\frac{1}{2}}^{\frac{1}{\sqrt{2}}} f_{dist}(x) dx = O\left(\frac{1}{n}\right) [F_{dist}(x)]_{\frac{1}{2}}^{\frac{1}{\sqrt{2}}} = O\left(\frac{1}{n}\right) \left(1 - \frac{\pi}{4}\right) = O\left(\frac{1}{n}\right)$$

□

Lemma 4.4. Sei $G = (V, E)$ ein geometrischer Zufallsgraph mit Radius $r \in \Theta(\frac{1}{\sqrt{n}})$ und Temperatur $T \in (0, 1)$. Dann gilt für zwei beliebige Knoten $u, v \in E$ dass $\Pr[uv \in E] \in O(\frac{1}{n})$

Proof. Nach dem Gesetz der totalen Wahrscheinlichkeit kann wie in 4.1 die folgende Umformung gemacht werden.

$$\begin{aligned} \Pr[uv \in E] &= \int_{x=0}^{\sqrt{2}} \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx \\ &= \int_{x=0}^r \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx \\ &\quad + \int_{x=r}^{\frac{1}{2}} \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx \\ &\quad + \int_{x=\frac{1}{2}}^{\frac{1}{\sqrt{2}}} \Pr[uv \in E | \text{dist}(u, v) = x] f_{\text{dist}}(x) dx \end{aligned}$$

Auf diese Integrale lassen sich Lemma 4.1 bis 4.3 anwenden. Somit gilt, dass $\Pr[uv \in E]$ in $\Theta(\frac{1}{n})$ liegt. \square

4.1.3 Erwartungswert

Nun kann mit dem selben Beweis wie in Lemma 3.5 gezeigt werden, dass der Erwartungswert der Anzahl Cliques und auch der Erwartungswert der Anzahl inklusionsmaximaler Cliques in $\Theta(\frac{1}{n})$ liegen.

5. GIRG

In diesem Kapitel zeigen wir für $\beta \geq 3$, dass die Anzahl inklusionsmaximaler Cliques mit hoher Wahrscheinlichkeit in $O(n)$ liegt.

Dies machen wir mit der selben Methodik wie in dem Beweis der Konzentrationsschranke für RGGs. Auch Mathematisch ist dieser Beweis recht ähnlich. Wir zeigen also zunächst, dass sowohl Erwartungswert als auch die Varianz einer Abschätzung in $O(n)$ liegen und können dann mit Tschebyscheff eine Konzentrationsschranke zeigen. Wir zeigen auch in diesem Kapitel, dass der Erwartungswert der Anzahl aller Cliques und damit auch der Erwartungswert der Anzahl inklusionsmaximaler Cliques in $O(n)$ liegen. Für den Beweis der Varianz werden wir dann allerdings die inklusionsmaximalität brauchen und somit die Konzentrationsschranke nur für die Anzahl inklusionsmaximaler Cliques zeigen.

5.0.1 Abschätzung der Anzahl Cliques

Beim Beweis für RGGs haben wir die Cliques Zellen zugeordnet. In diesem Beweis ordnen wir jede Clique nun den Knoten zu die Teil der Clique sind. Wir können dann die Anzahl zugeordneter Cliques eines Knotens dann durch zwei hoch die Anzahl Knoten in der Nachbarschaft des Knotens abschätzen. Diese Abschätzung reicht ist allerdings noch zu ungenau um zu zeigen, dass der Erwartungswert der Anzahl Cliques in $O(n)$ liegt, da z.b bei $\beta = 3$ allein Knoten $v = 1$ im Erwartungswert $\Theta(\sqrt{n})$ viele Knoten in seiner Nachbarschaft hat.

Eine Möglichkeit die Abschätzung genauer zu machen ist die Cliques nicht mehr jedem ihrer Knoten zuzuordnen, sondern nur noch dem Knoten mit geringstem Gewicht. Das erlaubt uns die Anzahl zugeordneter Cliques pro Knoten nun über $2^{|N_v|}$ abzuschätzen, wobei $|N_v|$ in diesem Beweis die innere Nachbarschaft von v ist. Die *innere Nachbarschaft* von v definieren wir als die Teilmenge Nachbarschaft, deren Knoten ein höheres Gewicht als v haben, inklusive dem Knoten v selbst. Die Idee hinter dieser Zuordnung ist, dass jede Clique nur noch genau einem Knoten zugeordnet wird und dass Knoten mit hohem Gewicht zwar relativ viele Nachbarn haben, aber es aber nur wenige Knoten mit noch höheren Gewicht gibt.

Insgesamt lässt sich die also die Anzahl aller Cliques durch $C = \sum_{v \in V} 2^{|N_v|}$ nach oben abschätzen. Wobei N_v die Teilmenge der Nachbarschaft von v oder v selbst sind und deren Gewicht mindestens so groß ist, wie das Gewicht von v .

5.0.2 Erwartungswert

Zunächst zeigen wollen wir zeigen, dass der Erwartungswert von C in $O(n)$ liegt. Dazu zeigen wir zunächst, dass der Erwartungswert $\mathbb{E}[2^{|N_v|}]$ für alle Knoten v in $O(1)$ liegt. Das Folgende Lemma ist wieder etwas allgemeiner formuliert, da es später noch für den Beweis der Varianz benötigt wird. Für $\mathbb{E}[2^{|N_v|}]$ gilt der Beweis mit den Parameter $a = 2$ und $b = 1$, da dann N_v der inneren Nachbarschaft entspricht.

Lemma 5.1. *Sei $G = (V, E)$ ein GIRG mit Gewichtsfunktion $w_v = (\frac{v}{n})^{\frac{1}{\beta-1}}$. Sei N_v^u die Indikatorvariable dafür ob Knoten u in der Nachbarschaft von v liegt. Sei $|N_v| = \sum_{u=b}^v N_v^u$.*

Dann gilt für eine beliebige Konstante $a \in \mathbb{N}$, $\mathbb{E}[a^{|N_v|}] = 2 \prod_{u=1}^{v-1} (1 + (a-1)\Pr[uv \in E]) \in O(1)$

Proof. Per Definition gilt

$$\mathbb{E}[a^{|N_v|}] = \mathbb{E}\left[a^{\sum_{u=b}^v N_v^u}\right] = \mathbb{E}\left[\prod_{u=b}^v a^{N_v^u}\right].$$

Da alle Knoten unabhängig voneinander platziert werden kann der Erwartungswert in das Produkt hereingezogen werden. Damit gilt

$$\mathbb{E}\left[\prod_{u=b}^v a^{N_v^u}\right] = \prod_{u=b}^v \mathbb{E}[a^{N_v^u}] = \prod_{u=b}^v (a^1 \Pr[N_v^u = 1] + a^0 \Pr[N_v^u = 0]).$$

$N_u^v = 0$ ist das Gegenereignis zu $N_u^v = 1$. Somit ist $\Pr[N_u^v = 0] = 1 - \Pr[N_u^v = 1]$. Der Term vereinfacht sich also zu

$$\prod_{u=b}^v (1 + (a-1)\Pr[N_u^v = 1])$$

Wenn $u = v$ gilt ist N_u^v per Definition immer eins und ist der obige Term gleich

$$a \prod_{u=b}^{v-1} (1 + (a-1)\Pr[N_u^v = 1])$$

Für $u \neq v$ ist die Wahrscheinlichkeit $\Pr[N_u^v = 1]$ die Wahrscheinlichkeit, dass uv eine Kante ist also $\Pr[uv \in E]$. Da die Knoten gleichverteilt auf dem Intervall $[0, 1]$ sind und Knoten maximal $\frac{1}{2}$ voneinander entfernt sein können Kantenwahrscheinlichkeit kann durch die Funktion $\frac{2w_u w_v}{W}$ nach oben abgeschätzt werden. Der Term kann also nach oben abgeschätzt werden durch

$$a \prod_{u=1}^{v-1} \left(1 + \frac{2(a-1)w_u w_v}{W}\right)$$

Nach Lemma 2.4 liegt das Produkt in $O(1)$ womit auch der gesamte Term auch in $O(1)$ liegt. \square

Lemma 5.2. *Sei $G = (V, E)$ ein GIRG mit Gewichtsfunktion $w_v = (\frac{v}{n})^{\frac{1}{\beta-1}}$. Sei $\beta > 3$. Sei N_v die innere Nachbarschaft von v , inklusive v . Sei $C = \sum_{v \in V} 2^{|N_v|}$. Dann gilt $\mathbb{E}[C] \in O(n)$.*

Proof. Mit der Linearität des Erwartungswertes und Lemma 5.1 folgt

$$\mathbb{E}[C] = \mathbb{E}\left[\sum_{v=1}^n 2^{|N_v|}\right] = \sum_{v=1}^n \mathbb{E}\left[2^{|N_v|}\right] \in O(n).$$

□

Da die Anzahl inklusionsmaximaler Cliques kleiner ist als die Anzahl aller Cliques, gilt dieser Beweis also auch für Erwartungswert der Anzahl inklusionsmaximaler Cliques.

5.1 Konzentrationsschranke

5.1.1 Abschätzung der Anzahl inklusionsmaximaler Cliques

Die Konzentrationsschranke zeigen wir nur für die Anzahl inklusionsmaximaler Cliques, da uns dies erlaubt eine genauere Abschätzung zu verwenden, von der wir leichter eine Schranke für die Varianz zeigen können. Die Idee ist, dass alle Knoten aus der Nachbarschaft eines Knoten ab einem gewissen Gewicht sicher eine Clique bilden und wir diese Knoten dann in unserer Abschätzung ignorieren können. Die folgenden zwei Lemmata zeigen dies.

Lemma 5.3. *Sei $G = (V, E)$ ein GIRG mit Gewichtsfunktion $w_v = \left(\frac{v}{n}\right)^{\frac{1}{\beta-1}}$. Sei Knoten $v \in V$ beliebig. Sei A_v die Teilmenge der Nachbarschaft von v für die gilt, dass für alle $u \in A$ gilt $u < 2^{1-\beta}v$. Dann ist A_v eine Clique.*

Proof. Um zu zeigen, dass A_v eine Clique ist, zeigen wir dass zwei beliebigen Knoten aus A_v eine Kante existiert.

Seien u_1 und u_2 beliebige Knoten aus A_v . Dann sind u_1 und u_2 am weitesten voneinander entfernt wenn sie auf gegenüberliegenden Seiten von v liegen und genau so weit, dass sie jeweils gerade noch mit v verbunden sind. Sie können also maximal $\frac{w_{u_1}w_v}{W} + \frac{w_{u_2}w_v}{W}$ voneinander entfernt sein. Genau genommen gilt dies allerdings nur wenn $\frac{w_{u_1}w_v}{W} + \frac{w_{u_2}w_v}{W}$ kleiner ist als $\frac{1}{2}$, da Knoten auf dem Einheitskreis maximal $\frac{1}{2}$ voneinander entfernt sein können. Diesen Fall decken wir jedoch auch wenn wir die Entfernung mit $\frac{w_{u_1}w_v}{W} + \frac{w_{u_2}w_v}{W}$ nach oben abschätzen. Damit u_1 und u_2 miteinander verbunden sind müssen u_1 und u_2 auf jeden Fall weniger als $\frac{w_{u_1}w_{u_2}}{W}$ voneinander entfernt sein. Für die Knoten u_1 und u_2 muss also die Bedingung

$$\frac{w_{u_1}w_v}{W} + \frac{w_{u_2}w_v}{W} < \frac{w_{u_1}w_{u_2}}{W}$$

gelten. Durch einsetzen der Gewichtsfunktion folgt mit einfachen Gleichungsumformungen, dass dies äquivalent ist zu

$$u_1^{\frac{1}{\beta-1}} + u_2^{\frac{1}{\beta-1}} < v^{\frac{1}{\beta-1}}$$

Wenn nun $i < 2^{1-\beta}v$ und $u < 2^{1-\beta}v$ gilt ist diese Ungleichung erfüllt. □

Den Faktor $2^{1-\beta}$ bezeichnen wir ab sofort mit α . Diese Eigenschaft hilft uns, da wir nun die Anzahl inklusionsmaximaler Cliques, deren Knoten mit dem geringsten Gewicht v ist durch $2^{|N'_v|}$ abschätzen können, wobei N'_v die Teilmenge der Nachbarschaft von v , inklusive v , aber ohne die Clique A_v , ist. Das folgende Lemma zeigt Dies.

Lemma 5.4. Sei $G = (V, E)$ ein GIRG mit Gewichtsfunktion $w_v = (\frac{v}{n})^{\frac{1}{\beta-1}}$. Sei Knoten $v \in V$ beliebig. Sei A_v die Teilmenge der Nachbarschaft von v , für die gilt dass für alle $u \in A_v$ gilt $u < 2^{1-\beta}v$. Sei N'_v die Teilmenge der Nachbarschaft von v inklusive v aber ohne A_v . Dann kann die Anzahl inklusionsmaximaler Cliques, deren Knoten mit dem geringsten Gewicht v ist durch $2^{|N'_v|}$ abgeschätzt werden.

Proof. Wir beweisen dies indem wir zeigen, dass jede inklusionsmaximale Clique deren Knoten mit geringstem Gewicht v ist eindeutig durch ihre Teilmenge der Knoten in N'_v identifiziert werden kann. Da es nur $2^{|N'_v|}$ verschiedene Teilmengen von N'_v gibt folgt damit, dass es maximal $2^{|N'_v|}$ viele inklusionsmaximale Cliques geben kann.

Angenommen es gäbe zwei inklusionsmaximale Cliques $C_1 = B \cup D_1$ und $C_2 = B \cup D_2$ mit $D_1 \neq D_2$ deren Knoten mit geringsten Gewicht v ist, wobei B jeweils Teilmenge der Knoten aus N'_v und D_1 bzw. D_2 jeweils die Teilmenge der Knoten aus A_v sind. Dann ist auch $C_3 = B \cup D_1 \cup D_2$ eine Clique, da nach Lemma 1 alle Knoten aus D_1 und D_2 paarweise miteinander verbunden sind und nach der Definition von C_1 und C_2 dasselbe für B_1 und D_1 bzw. B_1 und D_1 gilt. Die Clique C_3 widerspricht allerdings der inklusionsmaximalität von C_1 und C_2 . Es kann also keine zwei inklusionsmaximalen Cliques mit derselben Teilmenge der Knoten aus N'_v geben.

□

Wir können also nun für jeden Knoten v die Anzahl inklusionsmaximaler Cliques, deren Knoten mit geringstem Gewicht v ist durch $2^{|N'_v|}$ abschätzen und damit die Anzahl inklusionsmaximaler Cliques im gesamten Graphen durch $C' = \sum_{v \in V} 2^{|N'_v|}$ nach oben abschätzen.

Der Erwartungswert von C' liegt auch in $O(n)$. Dies kann man auch mit Lemma 3.19 durch Ersetzen von N_v mit N'_v und Wahl der entsprechenden Parameter zeigen.

5.1.2 Varianz der Abschätzung

Lemma 5.5. Sei $G = (V, E)$ ein GIRG mit Gewichtsfunktion $w_v = (\frac{v}{n})^{\frac{1}{\beta-1}}$. Sei Knoten $v \in V$ beliebig. Sei N'_v die Teilmenge der Nachbarschaft von v , inklusive v , für die gilt dass für alle $u \in N'_v$ gilt $2^{1-\beta}v < u \leq v$. Sei $C' = \sum_{v \in V} 2^{|N'_v|}$. Dann gilt $\mathbb{V}[C'] \in O(n)$

Proof. Da C' eine Summe von Zufallsvariablen ist, kann die Varianz von C' über die paarweisen Kovarianzen auszurechnen 2.1.

Für jeden Knoten v gilt, dass es nur Knoten u aus dem Intervall $(\lceil \alpha v \rceil, \min(n, \lfloor \frac{v}{\alpha} \rfloor))$, dass es Knoten geben kann die sowohl in N'_v als auch N'_u enthalten sind. Für alle anderen Knoten sind die Zufallsvariablen $2^{|N'_v|}$ und $2^{|N'_u|}$ somit unabhängig voneinander und deren Kovarianz damit null. Insgesamt gilt also für die Varianz

$$\mathbb{V}[C'] = \sum_{v=1}^n \sum_{u=\lceil \alpha v \rceil}^{\min(n, \lfloor \frac{v}{\alpha} \rfloor)} \text{Cov}(2^{|N'_v|}, 2^{|N'_u|}) = \sum_{v=1}^n \sum_{u=\lceil \alpha v \rceil}^{\min(n, \lfloor \frac{v}{\alpha} \rfloor)} \mathbb{E}[2^{|N'_v|} 2^{|N'_u|}] - \mathbb{E}[2^{|N'_v|}] \mathbb{E}[2^{|N'_u|}].$$

Zunächst betrachten wir separat den Sonderfall $u = v$. Hier entspricht die Kovarianz der Varianz der Zufallsvariable und kann durch den Erwartungswert $\mathbb{E}[2^{|N'_v|} 2^{|N'_v|}] = \mathbb{E}[4^{|N'_v|}]$

abgeschätzt werden. Mit Lemma 5.1 folgt, dass auch dieser Erwartungswert in $O(1)$ liegt. Damit gilt

$$\mathbb{V}[C'] = \sum_{v=1}^n (O(1) + \sum_{u=\lceil \alpha v \rceil \neq v}^{\min(n, \lfloor \frac{v}{\alpha} \rfloor)} (\mathbb{E}[2^{|N'_v|} 2^{|N'_u|}] - \mathbb{E}[2^{|N'_v|}] \mathbb{E}[2^{|N'_u|}])).$$

Bei Beweis zu RGGs haben wir zwischen Zellen, deren Nachbarschaften sich überschneiden, und Zellen, deren Nachbarschaft sich nicht überschneiden, unterschieden. Analog dazu unterscheiden wir hier zwischen Knotenpaaren u und v die nah genug zusammen liegen, sodass es möglich ist, dass ein Knoten sowohl in N'_v als auch in N'_u liegt. Um dies machen zu können verwenden wir die Formel des totalen Erwartungswertes

$$\mathbb{E}[2^{|N'_v|} 2^{|N'_u|}] = \mathbb{E}[2^{|N'_v|} 2^{|N'_u|} \mid B] \Pr[B] + \mathbb{E}[2^{|N'_v|} 2^{|N'_u|} \mid \neg B] \Pr[\neg B]$$

und wählen als Bedingung B , dass u und v nah genug zusammen liegen, sodass es möglich ist, dass es ein Knoten k gibt der sowohl in N'_v als auch in N'_u liegt.

Damit dies der Fall ist dürfen die Knoten maximal so weit voneinander entfernt sein, dass es einen dritten Knoten k gibt, der sowohl mit u als auch mit v verbunden ist. Die Knoten u und v sind somit auf jeden Fall weniger als $\frac{w_u w_k}{W} + \frac{w_k w_v}{W}$ voneinander entfernt. Da u in $O(v)$ liegt, muss auch k in $O(v)$ liegen, damit es möglich ist, dass k sowohl in N'_v als auch in N'_u liegt. Mit $v \leq n$ und $\beta \geq 3$ folgt, dass $\frac{w_u w_k}{W} + \frac{w_k w_v}{W}$ in $O(\frac{1}{v})$ liegt und somit die Wahrscheinlichkeit $\Pr[B]$ in $O(\frac{1}{v})$ liegt. Die Wahrscheinlichkeit $\Pr[\neg B]$ können wir durch eins nach oben abschätzen.

Den Erwartungswert $\mathbb{E}[2^{|N'_v|} 2^{|N'_u|} \mid B]$ können wir mithilfe der Cauchy-Schwarz-Ungleichung nach oben abschätzen durch $\sqrt{\mathbb{E}[4^{|N'_u|} \mid B] \mathbb{E}[4^{|N'_v|} \mid B]}$. Der Erwartungswert $\mathbb{E}[4^{|N'_v|} \mid B]$ ist kleiner als $4\mathbb{E}[4^{|N'_v|}]$, da die Bedingung B nur Punkt u beeinflusst und die $|N'_v|$ am größten wird wenn u in der Nachbarschaft von v enthalten ist und somit den Wert der Zufallsvariable um maximal einen Faktor von vier erhöht. Der Erwartungswert $\mathbb{E}[4^{|N'_v|}]$ liegt nach Lemma 5.1 in $O(1)$. Analog folgt, dass auch $\mathbb{E}[4^{|N'_u|} \mid B]$ in $O(1)$ liegt. Somit liegt auch $\mathbb{E}[2^{|N'_v|} 2^{|N'_u|} \mid B]$ in $O(1)$ und der Term für die Varianz vereinfacht sich zu

$$\sum_{v=1}^n (O(1) + \sum_{u=\lceil \alpha v \rceil \neq v}^{\min(n, \lfloor \frac{v}{\alpha} \rfloor)} (O(\frac{1}{v}) + \mathbb{E}[2^{|N'_v|} 2^{|N'_u|} \mid \neg B] - \mathbb{E}[2^{|N'_v|}] \mathbb{E}[2^{|N'_u|}])).$$

Es bleibt zu zeigen, dass die Differenz $\mathbb{E}[2^{|N'_v|} 2^{|N'_u|} \mid \neg B] - \mathbb{E}[2^{|N'_v|}] \mathbb{E}[2^{|N'_u|}]$ kleiner als 0 ist. Diese Differenz entspricht quasi den Kovarianzen von Zellen deren Nachbarschaften sich nicht überschneiden im Beweis der Varianz bei RGGs.

Auch in diesem Beweis können wir $|N'_v|$ und $|N'_u|$ als Summe von Indikatorvariablen formulieren. Es gilt also $|N'_v| = \sum_{k=\lceil \alpha v \rceil}^v N'_v{}^k$ bzw. $|N'_u| = \sum_{k=\lceil \alpha v \rceil}^v N'_u{}^k$, wobei $N'_v{}^k$ die

Indikatorvariable dafür ist ob Knoten k in N'_v liegt. Wir nehmen im folgenden o.B.d.A $u < v$ an. Da die Knoten unabhängig voneinander auf dem Kreis platziert werden und da Knoten v sicher in N'_v enthalten ist und u sicher in N'_u aber nach Bedingung $\neg B$ Knoten u nicht in N'_v und v nicht in N'_u enthalten ist ergibt sich

$$\mathbb{E}[2^{|N_u|} 2^{|N_v|} \mid \neg B] = 4 \prod_{k=\lceil \alpha u \rceil}^{\lceil \alpha v \rceil - 1} \mathbb{E}[2^{N_u{}^k} \mid \neg B] \prod_{k=\lceil \alpha v \rceil}^{u-1} \mathbb{E}[2^{N_u{}^k + N_v{}^k} \mid \neg B] \prod_{k=u+1}^{v-1} \mathbb{E}[2^{N_v{}^k} \mid \neg B].$$

Der Erwartungswert von $\mathbb{E} \left[2^{N_u^k + N_v^k} \mid \neg B \right]$ ist $1 + \Pr[uk \in E] + \Pr[vk \in E]$. Dies liegt daran, dass $N_u^k + N_v^k$ nur eins oder Null sein kann nach Bedingung $\neg B$ und somit $\Pr[N_u^k + N_v^k = 0] = 1 - \Pr[N_u^k + N_v^k = 1]$ gilt. Die Wahrscheinlichkeit, dass $N_u^k + N_v^k$ eins ist die Summe der Wahrscheinlichkeiten von $N_u^k = 1$ und $N_v^k = 1$, da nach Bedingung $\neg B$, der Knoten k nicht in N'_u und N'_v gleichzeitig sein kann.

Der Erwartungswert $\mathbb{E} \left[2^{N'_u} \mid \neg B \right]$ entspricht $\mathbb{E} \left[2^{N'_u} \right]$, wenn $k \neq u$ gilt. Der Erwartungswert $\mathbb{E} \left[2^{N'_v} \mid \neg B \right]$ entspricht $\mathbb{E} \left[2^{N'_v} \right]$, wenn $k \neq v$ gilt. Damit gilt

$$\mathbb{E} \left[2^{|N_u|} 2^{|N_v|} \mid \neg B \right] = 4 \prod_{k=\lceil \alpha u \rceil}^{\lceil \alpha v \rceil - 1} \mathbb{E} \left[2^{N'_u} \right] \prod_{k=\lceil \alpha v \rceil}^{u-1} (1 + \Pr[uk \in E] + \Pr[vk \in E]) \prod_{k=u+1}^{v-1} \mathbb{E} \left[2^{N'_v} \right].$$

Nun schauen wir uns den zweiten Teil der Differenz $\mathbb{E} \left[2^{|N'_v|} 2^{|N'_u|} \mid \neg B \right] - \mathbb{E} \left[2^{|N'_v|} \right] \mathbb{E} \left[2^{|N'_u|} \right]$ an. Für das Produkt der Erwartungswerte gilt nach Lemma 5.1

$$\begin{aligned} & \mathbb{E} \left[2^{|N'_v|} \right] \mathbb{E} \left[2^{|N'_u|} \right] \\ &= 4 \prod_{k=\lceil \alpha u \rceil}^{\lceil \alpha v \rceil - 1} \mathbb{E} \left[2^{N'_v} \right] \prod_{k=\lceil \alpha v \rceil}^{u-1} (1 + \Pr[uk \in E])(1 + \Pr[vk \in E]) \prod_{k=u+1}^{v-1} \mathbb{E} \left[2^{N'_v} \right] \\ &= 4 \prod_{k=\lceil \alpha u \rceil}^{\lceil \alpha v \rceil - 1} \mathbb{E} \left[2^{N'_v} \right] \prod_{k=\lceil \alpha v \rceil}^{u-1} (1 + \Pr[uk \in E] + \Pr[vk \in E] + \Pr[uk \in E] \Pr[vk \in E]) \prod_{k=u+1}^{v-1} \mathbb{E} \left[2^{N'_v} \right]. \end{aligned}$$

Die Differenz $\mathbb{E} \left[2^{|N'_v|} 2^{|N'_u|} \mid \neg B \right] - \mathbb{E} \left[2^{|N'_v|} \right] \mathbb{E} \left[2^{|N'_u|} \right]$ ist also aufgrund des zusätzlichen Summanden $\Pr[vk \in E] \Pr[uk \in E]$ im mittleren Produkt also kleiner als Null. Die Varianz kann somit nach oben abgeschätzt werden durch

$$\sum_{v=1}^n (O(1) + \sum_{u=\lceil \alpha v \rceil \neq v}^{\min(n, \lfloor \frac{v}{\alpha} \rfloor)} O(\frac{1}{v})).$$

Dieser Term liegt in $O(n)$ □

5.1.3 Konzentrationsschranke

Wir haben also nun für unsere Abschätzung C' der Inklusionsmaximaler Cliques gezeigt, dass sowohl der Erwartungswert als auch die Varianz von C' in $O(n)$ liegen. Es lässt sich nun also mit einem Beweis wie in 3.13 eine Konzentrationsschranke die Anzahl inklusionsmaximaler Cliques zeigen.

6. Zusammenfassung und Ausblick

Wir haben in dieser Arbeit gezeigt, dass für RGGs mit $r \in \Theta(\frac{1}{\sqrt{n}})$ sowohl der Erwartungswert der Anzahl aller Cliques in $\Theta(n)$ liegt als auch der Erwartungswert der Anzahl der inklusionsmaximaler Cliques in $\Theta(n)$ liegt.

Desweiteren haben wir gezeigt, dass die Ergebnisse für den Erwartungswert auch gelten, wenn wir das RGG-Modell um das Konzept der Temperatur erweitern und die Temperatur T im Intervall $(0, 1)$ liegt.

Außerdem haben wir gezeigt, dass die Anzahl aller Cliques und damit auch die Anzahl aller inklusionsmaximaler Cliques mit hoher Wahrscheinlichkeit in $O(n)$ liegt.

In letztem Kapitel haben wir für eine einfache Version des GIRG-Modell ohne Temperatur und mit konkreter Gewichtsfunktion $w_v = (\frac{n}{v})^{\frac{1}{\beta-1}}$ gezeigt, dass für Power-Law Exponent $\beta \geq 3$ der Erwartungswert, der Anzahl Cliques und damit auch der Anzahl Inklusionsmaximaler Cliques, in $O(n)$ liegt.

Ganz am Ende haben wir gezeigt, dass die Anzahl aller inklusionsmaximalen Cliques mit hoher Wahrscheinlichkeit in $O(n)$ liegt. Hier besteht natürlich auch die Frage ob diese Konzentrationsschranke auch für die Anzahl aller Cliques gilt.

Es bleiben allerdings noch viele weitere Fragen für zukünftige Arbeit offen. Interessant wäre z.B. ob es möglich ist genauere Konzentrationsschranken mit anderen Methoden wie z.B. mit der „bounded differences inequality“ zu zeigen. Außerdem wäre noch interessant was in den Parameterbereichen passiert, die wir nicht in dieser Arbeit untersucht haben. Insbesondere für GIRGs mit Power-Law Exponent $\beta \in (2, 3)$, wären Aussagen besonders interessant, da dies der Parameterbereich ist, welcher viele echte Netzwerke besser modelliert. Aber auch für RGGs mit anderem r bleibt die Anzahl Inklusionsmaximaler Cliques eine spannende Frage. Da in dieser Arbeit eigentlich alle Aussagen mit der O -Notation gemacht wurden, ist die Frage nach den genauen Konstanten eine weitere mögliche Fragestellung für zukünftige Arbeit.

Bibliography

- [BFG⁺] Thomas Bläsius, Tobias Friedrich, Andreas Göbel, Jordi Levy und Ralf Rothenberger: *The Impact of Heterogeneity and Geometry on the Proof Complexity of Random Satisfiability*. <https://arxiv.org/abs/2004.07319>.
- [BFK18] Thomas Bläsius, Tobias Friedrich und Anton Krohmer: *Cliques in Hyperbolic Random Graphs*. *Algorithmica*, 80, 2018. <https://doi.org/10.1007/s00453-017-0323-3>.
- [BKL15] Karl Bringmann, Ralph Keusch und Johannes Lengler: *Sampling Geometric Inhomogeneous Random Graphs in Linear Time*, 2015. <https://arxiv.org/abs/1511.00576>.
- [BPK10] M. Boguñá, F. Papadopoulos und D. Krioukov: *Sustaining the Internet with hyperbolic mapping*. *Nature Communications*, 62, 2010. <https://doi.org/10.1038/ncomms1063>.
- [Dal02] Michael Dall, Jesper und Christensen: *Random geometric graphs*. *Phys. Rev. E*, 66:016121, Jul 2002. <https://link.aps.org/doi/10.1103/PhysRevE.66.016121>.
- [DMV13] Laurent Decreusefond, Philippe Martins und Anaïs Vergne: *Clique number of random geometric graphs*. <https://hal.archives-ouvertes.fr/hal-00864303>, working paper or preprint, 2013.
- [ER59] Paul Erdős und Alfred Rényi: *On Random Graphs I*. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.
- [GWG05] Rajarshi Gupta, Jean Walrand und Olivier Goldschmidt: *Maximal Cliques in Unit Disk Graphs: Polynomial Approximation*. In: *IN PROCEEDINGS INOC 2005*, 2005.
- [HS95] M. L. Huson und A. Sen: *Broadcast scheduling algorithms for radio networks*. In: *Proceedings of MILCOM '95*, Seiten 649–650. IEEE, 1995.
- [KPK⁺10] Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat und Marián Boguñá: *Hyperbolic geometry of complex networks*. *Phys. Rev. E*, 82:036106, Sep 2010. <https://link.aps.org/doi/10.1103/PhysRevE.82.036106>.
- [MU17] Michael Mitzenmacher und Eli Upfal: *Probability and Computing*. Cambridge University Press, 2017.
- [Pen03] Mathew Penrose: *Random Geometric Graphs*. Oxford University Press, 2003. <https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780198506263.001.0001/acprof-9780198506263>.
- [XLN03] Yuan Xue, Baochun Li und Klara Nahrstedt: *Price-Based Resource Allocation in Wireless Ad Hoc Networks*. Band 2707, Seiten 155–155, Mai 2003, ISBN 978-3-540-40281-7.

