Geometric Inhomogeneous Random Graphs
and Graph Coloring Games

A thesis submitted to attain the degree of
Doctor of Sciences of ETH Zurich
(Dr. sc. ETH Zurich)

presented by

Ralph Keusch
MSc ETH Math.
born on 27.06.1986
citizen of Boswil AG

accepted on the recommendation of

Prof. Dr. Angelika Steger, examiner
Prof. Dr. Tobias Friedrich, co-examiner
Dr. Johannes Lengler, co-examiner

2018
Abstract

Real-world networks, like social networks or the internet infrastructure, are consistently found to possess short paths between nodes and to be scale-free, i.e., their degree sequences resemble power laws. These properties have been theoretically explained by classic models like Chung-Lu random graphs or preferential attachment graphs. However, real-world networks exhibit additional structural properties such as strong clustering that can best be described by using geometry. For instance, in social networks two people are more likely to know each other if they live in the same region and share hobbies, both of which can be encoded as spatial information. This is why the focus of the literature on random graph models for complex networks shifted towards modern, geometry-based models. In fact, in 2010, Boguñá et al. observed that the internet graph can be embedded into hyperbolic geometry [BPK10] and therefore proposed the model of hyperbolic random graphs [Kri+10]. As many desirable features of real-world networks naturally emerge in hyperbolic random graphs, they gained a lot of attention during the last years and serve as benchmarks for network algorithms like routing or link prediction.

In the first part of this thesis, we propose a class of scale-free random graphs models with underlying geometry that we call geometric inhomogeneous random graphs (GIRGs). GIRGs generalize hyperbolic random graphs to higher dimensions and are technically simpler while preserving their qualitative behavior. In GIRGs, every vertex draws independently a random position in a geometric space and a positive weight according to a power law. For each pair of nodes, we insert an edge with a certain probability that is higher if the nodes have larger weights and if they are geometrically close.
We prove that GIRGs reproduce the main structural properties that are associated with social and technological networks such as a power-law degree sequence, a large clustering coefficient, the presence of a unique giant connected component, poly-logarithmic diameter, and small separating sets. In addition, we show that the average distance between two vertices is only of order $\Theta(\log \log n)$ and present a sampling algorithm with expected linear running-time, significantly improving upon the best previous algorithm for hyperbolic random graphs [LMP15]. The results on the degree sequence, on the connectivity structure, and on distances also extend to a much more general class of random graphs: a generic augmented version of Chung-Lu random graphs.

Since Milgram's letter forwarding experiment from the 60s it is not only known that short paths exist in social networks, but also that human beings are actually able to find them without global knowledge by using only local information. This phenomenon was first reproduced in theory by Kleinberg [Kle00a]. However, from today's perspective his model has several shortcomings, limiting the applicability to real-world networks. To give a more cogent explanation of Milgram's experiments, we study greedy routing in GIRGs. We prove that the process succeeds with constant probability and that in case of success it finds asymptotically shortest paths of length $O(\log \log n)$. Hence, in GIRGs greedy routing is considerably faster than in previously analyzed models, explaining the algorithmic small-world phenomenon much more accurately. Moreover, we investigate local backtracking mechanisms and verify that such patching methods can ensure success probability 1 within an asymptotically tight number of steps. We also provide formal proofs of experimental findings on hyperbolic random graphs and furthermore address the question of finding efficient local routing protocols for the internet graph.

In the second part of the thesis we study combinatorial two-player games, thereby focusing on graph coloring games. Over the last decades, the area of positional games has been established as a central field of combinatorics, with many beautiful and surprising results and tight links to other branches of discrete mathematics. Vice-versa, coloring a given graph with few colors is one of the archetypical problems in graph theory. Graph coloring games connect
these two topics of graph theory and arose in the context of coloring planar graphs with 4 colors. The most classic game is played as follows: Maker and Breaker take turns in alternately coloring the vertices of a given graph $G$ with $k$ colors such that the partial coloring stays proper. Maker wins if the process ends with a proper coloring of the complete vertex set whereas Breaker wins if at some point, every proper coloring is forbidden. The *game chromatic number* $\chi_g(G)$ then denotes the smallest $k$ for which Maker has a winning strategy. Bohman, Frieze, and Sudakov initiated the analysis of the game chromatic number of binomial random graphs $G_{n,p}$ [BFS08] and conjectured that asymptotically almost surely, it is twice as large as the ordinary chromatic number of random graphs. We verify their conjecture for sufficiently dense random graphs, more precisely, for the regime $p \geq e^{-o(\log n)}$.

A natural variant of the game is the setting where the players color the edges instead of the vertices. For this variation, the *game chromatic index* $\chi'_g(G)$ denotes the smallest number of colors for which Maker has a winning strategy. We show that there exist global constants $c, C > 0$ such that for every graph $G$ with maximum degree $\Delta(G) \geq C \log n$ it holds $\chi'_g(G) \leq (2 - c)\Delta(G)$. This result partially solves a conjecture of Beveridge, Bohman, Frieze, and Pikhurko [Bev+08] and yields the first non-trivial bound on $\chi'_g(G)$ for many graph classes.

At last, we study *colorability saturation games* where the game process is not about coloring a given graph but about creating a graph with certain colorability properties. Concretely, Mini and Maxi start with the empty graph on $n$ vertices and alternately insert edges such that the graph remains $k$-colorable. At some point, the graph is saturated and the process stops. Mini wants that this happens as soon as possible whereas Maxi aims at prolonging the game as much as possible. The score of the game denotes the number of edges in the final graph, given that both players followed an optimal strategy. In this generality, the game was first analyzed recently by Hefetz, Krivelevich, Naor, and Stojaković [Hef+16]. We improve their estimates on the score of the game and prove almost matching lower and upper bounds.
ZUSAMMENFASSUNG


Im ersten Teil dieser Arbeit stellen wir mit *geometrischen inhomogenen Zufallsgraphen* (GIRGs) eine neue Klasse von mathematischen Modellen für komplexe Netzwerke vor, die skalenfrei ist und zugrundeliegende Geometrie verwendet. In GIRGs beziehen die Knoten unabhängig voneinander zufällige Positionen in einem geometrischen Raum und zufällige Gewichte gemäss einem Potenzgesetz. Für jedes Paar von Knoten fügen wir danach mit einer gewissen Wahrscheinlichkeit eine Kante ein, wobei dies wahrscheinlicher ist, wenn die beiden Knoten stärkere Gewichte aufweisen und näher beieinander liegen. Damit
sind GIRGs technisch einfacher als hyperbolische Zufallsgraphen, besitzen aber die gleichen qualitativen Merkmale und verallgemeinern jenes Modell in höhere Dimensionen.

Wir weisen nach, dass GIRGs tatsächlich die wesentlichen strukturellen Eigenschaften von komplexen Netzwerken besitzen: Gradverteilung gemäß einem Potenzgesetz, grosser Clusteringkoeffizient, eine Zusammenhangskomponente von linearer Größe, Durchschnittsdistanz der Größe $\Theta(\log \log n)$ zwischen den Knoten, polylogarithmischer Durchmesser und kleine Separatoren. Andererseits beschreiben wir einen Algorithmus, welcher GIRGs in erwartet linearer Zeit generieren kann; eine deutliche Verbesserung gegenüber dem bisher schnellsten Algorithmus für hyperbolische Zufallsgraphen [LMP15].


Im zweiten Teil der Arbeit beschäftigen wir uns mit kombinatorischen Spielen. Solche Spiele haben sich während den letzten Jahrzehnten als zentraler Bestandteil der Kombinatorik etabliert. Hier fokussieren wir uns auf Spiele, welche in Relation zu Graphfärbungsproblemen stehen. Das klassische

Bohman, Frieze und Sudakov untersuchten dieses Spiel als Erste im klassischen Zufallsgraphen $G_{n,p}$ [BFS08] und vermuteten, dass $\chi_g(G_{n,p})$ asymptotisch fast sicher doppelt so groß ist wie die normale chromatische Zahl von $G_{n,p}$. Wir beweisen diese Vermutung für dichte Zufallsgraphen; genauer gesagt für den Bereich $p \geq e^{-o(\log n)}$.

Eine natürliche Variante dieses Spiels ergibt sich, wenn die beiden Spieler die Kanten anstelle der Knoten färben. Bei dieser Variation bezeichnet der *spielchromatische Index* $\chi'_g(G)$ die kleinste Anzahl von Farben, bei der eine Gewinnstrategie für Maker existiert. Wir zeigen, dass globale Konstanten $c, C > 0$ existieren, so dass für alle Graphen mit Maximalgrad $\Delta(G) \geq C \log n$ die Ungleichung $\chi'_g(G) \leq (2 - c)\Delta(G)$ gilt. Dieses Resultat stellt für viele Graphklassen die erste nicht-triviale Schranke an $\chi'_g(G)$ dar und löst teilweise eine Vermutung von Beveridge, Bohman, Frieze und Pikhurko [Bev+08].

Schließlich analysieren wir Spiele, bei welchen nicht Graphen gefärbt werden, sondern Graphen mit bestimmten Färbungseigenschaften erzeugt werden: Mini und Maxi starten mit dem leeren Graphen auf $n$ Knoten und fügen abwechslungsweise eine neue Kante ein so dass der Graph $k$-färbbar bleibt. An einem bestimmten Punkt ist dies nicht mehr möglich, da der Graph gesättigt ist. Das Ziel von Mini ist, dass dies so früh wie möglich geschieht, während Maxi diesen Zeitpunkt so weit als möglich hinauszögern will. Das Ergebnis des Spiels ist definiert als die Anzahl Kanten im resultierenden Graphen, unter der Annahme dass beide Spieler eine optimale Strategie verfolgen. Das allgemeine Spiel für eine beliebige Wahl von $k$ wurde kürzlich von Hefetz, Krivelevich, Naor und Stojaković eingeführt [Hef+16]. Wir verbessern ihre Ergebnisse und zeigen untere und obere Schranken an das Spielergebnis, die fast übereinstimmen.
First and foremost, I want to thank my advisor Angelika Steger: for increasing my interest on combinatorial structures with your lectures, for co-supervising my Master thesis, for making me the unexpected offer of starting a PhD in her group, for all advices during the last four years, for being patient, for supporting flexible working hours, for covering a rich variety of topics in your group, and for your efforts to enhance the competitive environment of academia with a friendly group atmosphere. I wish you all the best for your personal future and the future of your research group.

I’m very thankful to Tobias Friedrich for co-refereeing my thesis, for the helpful feedback, for many interesting inputs, and for earlier suggestions on the topic of GIRGs and related models. Next, I thank Johannes Lengler for all the energy you put in our joint papers, for many insightful meetings in particular during the middle part of my PhD, for fighting together with deadlines, and in general for your steady optimism. I’m grateful to Karl Bringmann for initiating the GIRG project and for letting me join it as a rather novel PhD Candidate. I also give many thanks to Yannic Maus and Anisur Molla for the intensive and fruitful collaboration and for your continuous and motivating enthusiasm.

Next, I would like to thank Ueli Peter and Henning Thomas for being great supervisors of my Bachelor and Master theses, for providing so interesting topics, and thus introducing me to the two subjects of this dissertation. I also thank my further co-authors Andreas Noever, Konstantinos Panagiotou, and Pascal Pfister for the enriching collaboration.

Furthermore, I am very thankful to Florian Meier for running Algorithms and Complexity together with me during the last two years. It was much more
pleasant and instructive to organize exercises, meetings, and exams together with you than doing everything alone. Many thanks also go to Pascal Su for joining the AuK squad and for sharing the interest on saturation games, and to Hafsteinn Einarsson and Christoph Koch for valuable discussions on GIRGs. Moreover, I thank Anders, Asaf, Asier, Felix, Frank, Frederik, Katarzyna, Marcelo, Miloš, Nemanja, Rajko, and Xun for the enjoyable atmosphere in our retreats, at coffee breaks, or when playing board games.

Apart from my group mates, I am grateful to Manuela Fischer for being an extremely motivating and supportive room-neighbor, to all Gremos for the great company in Mittagsseminar and the stimulating talks on a great variety of topics, and to Mohsen Ghaffari for bringing new drive on G-floor. I thank Thomas Holenstein for many interesting and valuable discussions when preparing the Algorithms and Complexity lecture. I was so often in your office that people started assuming we are writing complexity papers together! Furthermore, I would like to thank Andrea Salow for constantly supplying us with fresh fruits and for running all the administration so perfectly, and Tomas Hruz for the smooth collaboration.

Persuading a PhD in our institute comes together with many teaching opportunities, in particular during semesters with triple-assistance. First, I thank everybody who helped me organizing the Algorithms and Complexity lecture during the past four years. Next, I’m thankful to Alessandro, Samuel, Daniel B., Steve, Seijin, Lierni, and Daniel K. for the enjoyable meetings while I was supervising your theses or talks. Along these lines, I thank as well all further AuK TAs and all students of the 9 exercise groups that I was holding during my PhD, for hundreds of creative and smart ideas. It was a pleasure to spend time together with so many clever and motivated people.

However, writing a doctoral thesis is not possible without a lot of support from outside academia. At this point, I want to thank my family for your love and support during good and bad times, for the way you raised me, for accepting all my decisions, for being so open-minded, and for not making too many jokes about the strange formulas in my work. I can not be grateful enough to my girlfriend Manuela for the years we spent together, for sharing interests, for
believing in me, for always being understanding and helpful during busy times before deadlines, and for many more reasons.

I would like to thank all friends for distracting me from working life. And last but not least, I’m very thankful to my long-term piano teacher Franz Martin for, to date, almost 16 years of collaboration, for awakening my passion for classical music and thus enriching my life so heavily, for passing on everything you learned from your teachers, and for showing me the grandeur of Liszt’s sonata, the disruption of Schumann's Kreisleriana, the pain of late Chopin, the colors of Scriabin, and, maybe standing over everything else, the sublimity of Beethoven's latest sonatas.
# Table of Contents

Abstract i

Zusammenfassung v

Acknowledgment ix

List of Figures xvii

1 Introduction 1
   1.1 Complex Networks and Random Graph Models . . . . . . . . . . 2
   1.2 Geometric Inhomogeneous Random Graphs . . . . . . . . . . . . 10
   1.3 Positional Games . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
   1.4 Graph Coloring Games . . . . . . . . . . . . . . . . . . . . . . . . 18
   1.5 Organization of the Thesis . . . . . . . . . . . . . . . . . . . . . 23

2 Preliminaries 25
   2.1 Notations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25
   2.2 Concentration Inequalities . . . . . . . . . . . . . . . . . . . . . . 27

Part I Geometric Inhomogeneous Random Graphs 31

3 A General Class of Scale-Free Networks 33
   3.1 The General Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Basic Properties of the Model</td>
<td>38</td>
</tr>
<tr>
<td>3.3</td>
<td>First Examples</td>
<td>43</td>
</tr>
<tr>
<td>3.4</td>
<td>Example: Geometric Inhomogeneous Random Graphs</td>
<td>47</td>
</tr>
<tr>
<td>3.5</td>
<td>Example: Hyperbolic Random Graphs</td>
<td>55</td>
</tr>
<tr>
<td>3.6</td>
<td>Further Related Models</td>
<td>62</td>
</tr>
<tr>
<td>3.7</td>
<td>Summary</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>GIRGs: Structural and Algorithmic Properties</td>
<td>65</td>
</tr>
<tr>
<td>4.1</td>
<td>Results</td>
<td>66</td>
</tr>
<tr>
<td>4.2</td>
<td>Degree Sequence</td>
<td>70</td>
</tr>
<tr>
<td>4.3</td>
<td>Giant Component, Diameter, and Average Distance</td>
<td>74</td>
</tr>
<tr>
<td>4.4</td>
<td>Clustering</td>
<td>90</td>
</tr>
<tr>
<td>4.5</td>
<td>Sampling Algorithm</td>
<td>93</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Cells</td>
<td>93</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Building Blocks</td>
<td>94</td>
</tr>
<tr>
<td>4.5.3</td>
<td>The Algorithm</td>
<td>97</td>
</tr>
<tr>
<td>4.5.4</td>
<td>Analysis</td>
<td>99</td>
</tr>
<tr>
<td>4.5.5</td>
<td>Sampling Threshold GIRGs</td>
<td>100</td>
</tr>
<tr>
<td>4.6</td>
<td>Separators, Entropy, and Compression Algorithm</td>
<td>101</td>
</tr>
<tr>
<td>5</td>
<td>Greedy Routing in GIRGs</td>
<td>109</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>110</td>
</tr>
<tr>
<td>5.2</td>
<td>Random Graph Model &amp; Greedy Routing</td>
<td>116</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>118</td>
</tr>
<tr>
<td>5.4</td>
<td>Comparison with Experimental Results</td>
<td>123</td>
</tr>
<tr>
<td>5.5</td>
<td>Patching Algorithms: Discussion and Examples</td>
<td>125</td>
</tr>
<tr>
<td>5.6</td>
<td>Proof Sketches and the Typical Trajectory</td>
<td>128</td>
</tr>
<tr>
<td>5.7</td>
<td>Preparations for the Proofs</td>
<td>134</td>
</tr>
<tr>
<td>5.7.1</td>
<td>Auxiliary Properties of GIRGs</td>
<td>134</td>
</tr>
<tr>
<td>5.7.2</td>
<td>Where to expect Neighbors &amp; Unlikely Jumps</td>
<td>139</td>
</tr>
<tr>
<td>5.8</td>
<td>Success Probability and Length of Greedy Routing</td>
<td>142</td>
</tr>
<tr>
<td>5.8.1</td>
<td>Main Lemma</td>
<td>142</td>
</tr>
<tr>
<td>5.8.2</td>
<td>Proof of Theorem 5.1</td>
<td>152</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.8.3</td>
<td>Proof of Theorem 5.2</td>
<td>155</td>
</tr>
<tr>
<td>5.8.4</td>
<td>Proof of Theorem 5.3</td>
<td>161</td>
</tr>
<tr>
<td>5.9</td>
<td>Patching Algorithms: Proof of Theorem 5.4</td>
<td>165</td>
</tr>
<tr>
<td>5.10</td>
<td>Relactions: Proof of Theorem 5.5</td>
<td>170</td>
</tr>
<tr>
<td>5.11</td>
<td>Greedy Routing in Hyperbolic Random Graphs</td>
<td>179</td>
</tr>
<tr>
<td>5.12</td>
<td>Proofs of Auxiliary Lemmas of Section 5.7.2</td>
<td>182</td>
</tr>
</tbody>
</table>

**Part II  Graph Coloring Games**  

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>The Game Chromatic Number of Random Graphs</td>
<td>193</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>193</td>
</tr>
<tr>
<td>6.2</td>
<td>Box Games</td>
<td>198</td>
</tr>
<tr>
<td>6.3</td>
<td>Outline of Proof Strategy</td>
<td>202</td>
</tr>
<tr>
<td>6.4</td>
<td>Proofs</td>
<td>207</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Properties of Random Graphs</td>
<td>207</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Proof of Lemma 6.9</td>
<td>212</td>
</tr>
<tr>
<td>6.4.3</td>
<td>Proof of Lemma 6.10</td>
<td>215</td>
</tr>
<tr>
<td>6.4.4</td>
<td>Proof of Theorem 6.4</td>
<td>217</td>
</tr>
<tr>
<td>6.5</td>
<td>Concluding Remarks</td>
<td>220</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>The Game Chromatic Index of Graphs</td>
<td>221</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>222</td>
</tr>
<tr>
<td>7.2</td>
<td>Upper Bounds</td>
<td>225</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Notations</td>
<td>225</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Maker's Strategy</td>
<td>225</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Main Proof</td>
<td>228</td>
</tr>
<tr>
<td>7.2.4</td>
<td>Missing Proofs</td>
<td>234</td>
</tr>
<tr>
<td>7.3</td>
<td>Lower Bound for the Biased Game</td>
<td>239</td>
</tr>
<tr>
<td>7.4</td>
<td>Open Problems</td>
<td>241</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Colorability Saturation Games</td>
<td>243</td>
</tr>
<tr>
<td>8.1</td>
<td>Introduction</td>
<td>243</td>
</tr>
<tr>
<td>8.2</td>
<td>Preliminaries</td>
<td>248</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>8.3 Lower Bound</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>8.4 Upper Bound</td>
<td>258</td>
<td></td>
</tr>
<tr>
<td>8.5 The Four Color Game</td>
<td>267</td>
<td></td>
</tr>
<tr>
<td>8.6 Concluding Remarks</td>
<td>275</td>
<td></td>
</tr>
<tr>
<td>Bibliography</td>
<td>277</td>
<td></td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td>299</td>
<td></td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Degree Sequence of the Internet Infrastructure</td>
<td>5</td>
</tr>
<tr>
<td>3.1</td>
<td>Hierarchy of GIRGs and Various Scale-Free Random Graph Models</td>
<td>34</td>
</tr>
<tr>
<td>3.2</td>
<td>Instance of a GIRG with Parameters $d = 2$, $\alpha = 8$, and $\beta = 2.8$</td>
<td>50</td>
</tr>
<tr>
<td>3.3</td>
<td>Instance of a GIRG with Parameters $d = 2$, $\alpha = 8$, and $\beta = 2.3$</td>
<td>51</td>
</tr>
<tr>
<td>3.4</td>
<td>Instance of a GIRG with Parameters $d = 2$, $\alpha = 2$, and $\beta = 2.5$</td>
<td>52</td>
</tr>
<tr>
<td>3.5</td>
<td>Instance of a GIRG with Parameters $d = 2$, $\alpha = 40$, and $\beta = 2.5$</td>
<td>53</td>
</tr>
<tr>
<td>5.1</td>
<td>Trajectory of a Typical Greedy Path in GIRGs</td>
<td>132</td>
</tr>
<tr>
<td>6.1</td>
<td>A Planar Graph on 12 Vertices with Game Chromatic Number 6</td>
<td>195</td>
</tr>
</tbody>
</table>
Graph theory is a enriching, manifold, and very active field in the intersection of mathematics and computer science that is strongly related to combinatorics, probability, algorithm design and has applications in many disciplines such as sociology, physics, or biology. On the one hand, during the last 300 years graphs have been studied intensively as abstract combinatorial objects, leading to thousands of beautiful theoretical results. On the other hand, graphs serve as a mathematical model for many types of binary relations and processes appearing in social, economical, biological, or physical networks. These applications provide motivation and value for researching graphs extensively on an abstract level. Vice-versa, from theoretical studies of graphs we gain intuition, tools, methods, and algorithms that are necessary for employing graph models when solving practical problems or explaining real-world phenomena.

This thesis is devoted to two aspects of graph theory and located in the interface between theory and applications. In the first part of the thesis, we propose geometric inhomogeneous random graphs (GIRGs) as a new model for complex networks and study structural and algorithmic properties of the
model. In the second part, we analyze several families of combinatorial games by investigating graph coloring games on various classes of graphs.

In this first chapter, we introduce the two subjects, summarize our results, and present an overview of the thesis.

1.1 Complex Networks and Random Graph Models

Complex networks appear everywhere in today’s world where pairwise relations are present, from business-relations between companies in economy and social relations between personal acquaintances to co-authorship-relations between scientists. The internet infrastructure interconnects computer networks. Transmission lines of the power grid carry power between sources, transformers, and customers. Or movie stars are “connected” if they co-starred at least once in the same movie.

Network science is a highly active, interdisciplinary field between mathematics, physics, computer science, biology, and social sciences, and aims at describing, understanding, modeling, and predicting the topology, organization, and evolution of such real-world networks. In the last two decades, network research moved on to study complex networks on a large scale. This is because most complex networks contain thousands or millions of nodes, especially new and modern networks such as the internet, and the computer age made it possible to gather and investigate these large data sets. Although a network is defined via local rules, from looking only at single nodes and their neighbors we don’t gain any insight about the global structure of the network. Consequently, the focus shifted towards a macroscopic perspective, statistical properties, and stylized facts of complex networks, aiming to understand networks on large scale.

Many real-world networks of different areas have been experimentally analyzed over the last decades. Studies on technological networks include the power-grid infrastructure [Ama+00; WS98], the internet infrastructure [FFF99; Mah+06], or airline route maps [Lor+16]. Examples for biological networks are protein-protein interactions in cells [Jeo+01] or gene regulatory networks
[Jeo+00], whereas the web graph [AJB99; BAJ00] or citation networks [Red98] are two famous examples for information networks. When considering social networks, the vertices are typically individuals and the edges are given by some sort of social tie (e.g., friendship [RH61], co-authorships [New01]). Finally, an example of economic systems are R&D alliances where official collaborations between companies are depicted [Tom+16]. There exist several excellent reviews on recent developments in network analysis, e.g., the book of Dorogovtsev and Mendes [DM02], the review of Newman [New03], the book of Watts [Wat04a] from a sociologic perspective, the popular book of Barabási [Bar02], or the book of Newman [New10] with a focus on measuring methods for empirical data.

Almost all these examples share two crucial structural properties. The first fundamental fact is the small-world phenomenon: despite being sparse networks (meaning that the average degree is very small compared to the size of the network), their structure is such that the typical graph-theoretical distance between two nodes is surprisingly small. More precisely, when representing networks as graphs in the language of mathematics and computer science, real-world networks have a unique largest connected component that contains a significant fraction of all nodes, and within this giant component, both the average distance and the maximum distance (i.e., diameter) between vertices are small. For social networks, this effect was established in the 1960s by the seminal experiments of Milgram [Mil67; TM69]: several random people were given a target person’s name and address and asked to send a letter to the target person by forwarding it to a personal acquaintance who is more likely to know the target person, iterating this process until the letter reached the target (or was lost). Among the more than 20% successful trials, Milgram reported an average distance of about 6, nowadays known as the phenomenon that everyone is connected to everybody through “6 degrees of separation”. Since then, this result inspired researchers in various disciplines to study networks with small-world properties. Recently, it was demonstrated that people might be separated by even fewer steps [Bac+12]. Apart from social networks, other examples exhibit exactly the same phenomenon, e.g., for the internet infrastructure Mahadevan et al. verified that the number of Autonomous Systems (AS) traversed by a data
set is average between 3 and 4, and at most 10 [Mah+06].

The second shared feature is the shape of the degree distribution: it is almost ubiquitously a power law for real-world networks. That is, the fraction of vertices with degree $k$ is proportional to $k^{-\beta}$ for some power-law parameter $\beta$, and thus decays slowly in $k$. In such a case, the network is sparse and most vertices have small degree, but there exist also strong nodes (“hubs”) whose degree is substantially larger than the average degree. Such networks are usually called scale-free [AB02; DM02; Mit03]. It turns out that many real-world networks are scale-free with parameter $2 < \beta < 3$. In this regime, the variance of the degree distribution is large and a heavy tail is present. In 1999, Faloutsos, Faloutsos, and Faloutsos [FFF99] first observed that the degree distribution of the internet follows a power law with parameter $2.15 \leq \beta \leq 2.2$. In Figure 1.1 on page 5, the degree sequence of the internet infrastructure is depicted. For protein networks, Jeong et al. found a power-law coefficient of 2.4 [Jeo+01]. Further examples of scale-free networks are listed in Table 1.1 on page 8 below.

Often, obtaining data of real-world networks is rather difficult and we suffer from a lack of knowledge about microscopic details. For instance, the internet infrastructure is maintained by separated organizations and its topology can only be reconstructed. In such situations, appropriate models enable us to nevertheless design algorithms and run simulations on complex networks. Mathematical models also fulfill the purpose of gaining an ensemble perspective, understanding the mechanisms that form real-world networks, and defining abstract but simple rules that reproduce the key features of complex networks.

It turned out that real-world networks can be best modeled by using random graphs. On the one hand, random graph models provide simple rules for generating large networks, on the other hand, with probabilistic rules we avoid over-fitting and presuming too much structure. Perhaps the simplest model are Erdős-Rényi random graphs as introduced in 1959 [ER59; Gil59], where, in the binomial model, we have $n$ vertices and between every pair of vertices we add an edge independently with probability $p$. We then denote by $G_{n,p}$ a random graph that is sampled according to this procedure. Note that the choice $p = \frac{1}{2}$ yields the uniform distribution over all labeled graphs on $n$ vertices. Random graphs
are one of the most central objects in combinatorics, allowing us to answer questions about typical graphs of certain densities. Over the last decades, hundreds of beautiful and surprising phenomena on random graphs have been discovered, see, e.g., the books of Bollobás [Bol01], Frieze and Karoński [FK15c], or Janson, Luczak, and Rucinski [JLR11]. However, one drawback of the model $G_{n,p}$ is that it is not scale-free. Vertices are homogeneous and the degree sequence follows a
Poisson distribution, leaving us with the necessity of developing more advanced random graph models.

Let us mention three celebrated, early models for scale-free networks. The configuration model provides a natural method of incorporating almost arbitrary degree sequence into a random graph model. We first insert for every node \( v \) a fixed number \( n_v \) of half-edges ("stubs"). In a second step, we connect all half-edges by sampling a random matching. For a fixed degree sequence, this procedure generates a uniform distribution over all possible graphs. In 1995, Molloy and Reed determined precise conditions for the existence of a giant connected component in the configuration model and thus explained the presence of a giant component in scale-free networks.

Barabási and Albert (1999) proposed the following preferential-attachment mechanism \([BA99]\): we start with a complete graph on \( m_0 \) vertices. Then, in every round a new vertex is born and chooses \( m \leq m_0 \) neighbors at random, but proportional to the degrees that the existing nodes already have. This mechanism reproduces the "rich get richer" effect and serves, for instance, as a very accurate model for citation networks. Bollobás, Riordan, Spencer, and Tusnády rigorously verified a power-law degree sequence with parameter \( \beta = 3 \) \([Bol+01]\), and later Bollobás and Riordan proved that for \( m \geq 2 \), after \( n \) steps the diameter is asymptotically \( \frac{\log n}{\log \log n} \) \([BR04]\). The model was also generalized in order to obtain power laws for arbitrary choices of \( \beta > 2 \) \([Bol+01]\). Moreover, in the regime \( 2 < \beta < 3 \), the average distance is of order \( \log \log n \) \([DHH09]\).

The by now classic Chung-Lu model uses a different approach \([CL02b; CL02a; CL04]\). Here, we have \( n \) vertices, each vertex \( v \) coming with a positive weight \( w_v \) that essentially is its expected degree. Then two vertices \( u \neq v \) are connected by an edge independently with probability

\[
p_{uv} = \min\{1, \frac{w_u w_v}{W}\},
\]

where \( W = \sum_{v \in V} w_v \) denotes the total weight. If the weight sequence follows a power law with parameter \( 2 < \beta < 3 \), then the average distance of the random graph is \( \frac{2+o(1)}{|\log(\beta-2)|} \log \log n \). Compared to the preferential attachment mechanism, Chung-Lu random graphs are only a static model and do not evolve over time,
but have the clear advantage of being technically much easier, because once the weights are fixed all edges are independently present, similarly as in the $G_{n,p}$ model. Due to their technical simplicity, Chung-Lu random graphs are often used as underlying model for the theoretical analysis of algorithms on scale-free networks (e.g., de-anonymization [BFK14] or information dissemination [FPS12]).

In 2006, Norros and Reittu defined a related model [NR06] which can be interpreted as a Poissonized variant of Chung-Lu random graphs. Bollobás, Janson, and Riordan [BJR07] introduced another, fairly general class of inhomogeneous random graphs. For a rigorous and deep mathematical discourse on scale-free random graph models we refer to the substantial book of van der Hofstad [Hof16]. Further, the book of Newman, Barabási, and Watts [NWB06] comprises a collection of significant papers.

Beyond the small-world and scale-free properties, most real-world networks have additional important features that are unfortunately not reproduced by the aforementioned models. For instance, in social networks two of one’s friends are likely to also know each other. This effect is called transitivity or clustering, and measured in terms of the clustering coefficient. For a vertex $v$, the local clustering coefficient $CC(v)$ is defined as the number of triangles containing $v$, divided by the number of paths $\{u, v, w\}$. Then the clustering coefficient of the graph is the average of the local clustering coefficients, for a formal definition see Definition 4.21. Almost consistently, real-world networks are found to possess strong clustering [DM02], meaning that the clustering coefficient is dramatically larger than the one of a random graph with the same degree distribution. Below in Table 1.1, we also list the clustering coefficient of several complex networks.

It is quite surprising that many real-world networks are found to be sparse, to possess clustering, and to have short distances at the same time. The aforementioned models don’t have clustering, being a major drawback and limiting their applicability. The simultaneous presence of these three features has been first explained by Watts and Strogatz [WS98] with the following model: we start with a ring lattice on $n$ vertices, each vertex being connected to its $k$ closest neighbors, where $k$ is a small integer. Then for each edge, with probability $p$
we randomly rewire one of the two endpoints. For a suitable choice of $p$, we obtain constant average degree, logarithmic average distance, and a constant clustering coefficient. However, this “small-world model” has the disadvantage of being homogeneous, disagreeing with most real-world data.

Apart from power-law degrees, short distances, and strong clustering, let us briefly mention three additional relevant features of complex networks: Milgram’s letter forwarding experiment not only demonstrated that there exist short paths through social networks, but also that the network is navigable: human beings are actually able to find such short paths without global knowledge of the network, by greedily routing the letter to local acquaintances. Next, for many real-world networks one of the crucial properties is the stability under attacks. Often, complex networks have relatively small separators [BBK03], that is, it is possible to cut the giant component into two halves by only removing a small

<table>
<thead>
<tr>
<th>Network</th>
<th>$n$</th>
<th>$m$</th>
<th>$\langle k \rangle$</th>
<th>$\langle \ell \rangle$</th>
<th>$\beta$</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS topology</td>
<td>17 446</td>
<td>40 805</td>
<td>4.68</td>
<td>3.69</td>
<td>2.15</td>
<td>0.29</td>
</tr>
<tr>
<td>WWW nd.edu</td>
<td>269 504</td>
<td>1 497 135</td>
<td>5.55</td>
<td>11.27</td>
<td>2.1/2.4</td>
<td>0.29</td>
</tr>
<tr>
<td>Bank transactions</td>
<td>5 086</td>
<td>76 614</td>
<td>15.06</td>
<td>2.62</td>
<td>2.1/2.2</td>
<td>0.53</td>
</tr>
<tr>
<td>Co-authorships</td>
<td>337 454</td>
<td>496 489</td>
<td>2.94</td>
<td>7.73</td>
<td>2.81</td>
<td>0.15</td>
</tr>
<tr>
<td>Movie actors</td>
<td>449 913</td>
<td>25 516 482</td>
<td>113.43</td>
<td>3.48</td>
<td>2.3</td>
<td>0.78</td>
</tr>
<tr>
<td>Metabolic network</td>
<td>765</td>
<td>3 686</td>
<td>9.64</td>
<td>2.56</td>
<td>2.2</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table 1.1. Key properties of some scale-free networks. Depicted are the number of vertices $n$ and edges $m$, the average degree $\langle k \rangle$, the average shortest path length $\langle \ell \rangle$, the empirical power-law exponent $\beta$, and the clustering coefficient $CC$. The first row describes the “BGP tables topology” of the internet infrastructure on the Autonomous Systems level (AS) [Mah+06], the second row the WWW of the nd.edu domain [AJB99; BAJ00], the third row payment transfers between commercial banks [Sor+06], the fourth row collaborations in mathematical research papers [Gro02], the fifth row co-stardom of movie actors [Ama+00; WS98], and the last row the metabolic network of biological cells [Jeo+00]. For the directed networks of the second and third example, $\langle k \rangle$ denotes the average in-degree and the two values of $\beta$ result from the in- and out-degrees.
number of nodes. For example, the internet infrastructure can be disconnected by removing about 3% of the nodes [AJB01]. Finally, real-world networks mostly exhibit community structures, that is, some subsets of vertices are forming relatively dense subgraphs [Moo01].

The reason behind all these effects is that real-world networks often emerge over underlying spatial information. Two people are more likely to know each other if they share hobbies and live in the same region. Two scientists are more likely to have joint publications if they work in the same field. Or two companies are more likely to invest in a joint project if they have similar interests. Classic scale-free models such as Chung-Lu or preferential attachment random graphs are not able to capture these phenomena, and over time, people realized that these effects can be best reproduced by using models with underlying geometry which directly encodes spatial information. This is why the focus of the literature on models for complex networks shifted from the classic models towards modern, geometry-based random graphs.

In recent years, dozens of new, realistic models have been proposed which also incorporate spatial information, often by using underlying geometry: spatial preferential attachment [JM15], scale-free percolation [DHH13], the forest fire model [LKF07; Kan+16], and many others [Aie+08; BHP08; BJP10; SKB08; DHW15]. One groundbreaking result that is based on underlying geometry is the seminal work of Kleinberg: he gave the first theoretical explanation of Milgram's experiment by introducing a grid-based random graph model and verifying that in his model, geometric greedy routing (the process of sending a packet forward to the neighbor who is geometrically closest to a target destination) succeeds within $O(\log^2(n))$ steps [Kle00a; Kle00b].

Among the proposed models, hyperbolic random graphs [Pap+10; Kri+10] stand out because they have been extensively experimentally validated. E.g., Boguña et al. [BPK10] computed and studied a (heuristic) maximum likelihood fit of the internet graph into the hyperbolic geometry and demonstrated its quality by showing that greedy routing in the underlying geometry of the fit finds near-optimal shortest paths. The simplest variant of the model is defined as follows: we place $n$ vertices independently and uniformly at random on the
hyperbolic disk of radius $R = 2\log n$ and vertices are connected via an edge if and only if their hyperbolic distance is smaller than $R$. Theoretical results on hyperbolic random graphs cover the degree sequence [GPP12] and small-world properties [ABM17; KM15; FK15b; BFM15b; TS17], but in contrast to most other models, also algorithmic aspects such as treewidth [BFK16], embedding algorithms [Bl16], or sampling [LMP15] have been investigated.

1.2 Geometric Inhomogeneous Random Graphs

One of our goals is to improve algorithmic and structural results on the promising model of hyperbolic random graphs. However, it turns out to be beneficial to work with a more general model that we call geometric inhomogeneous random graphs (GIRGs). Similarly as in Chung-Lu random graphs, we start with a sequence $(w_1, \ldots, w_n)$ of weights that we require to follow a power law, usually with power-law parameter $\beta > 2$. Again, we denote by $W$ the total weight sum. In addition, for a fixed parameter $d \in \mathbb{N}$, every node $v$ picks uniformly at random position $x_v$ in the $d$-dimensional cube $[0, 1]^d$ or in the $d$-dimensional torus $T^d = \mathbb{R}^d / \mathbb{Z}^d$. Then for every pair of vertices $u, v$, the edge $\{u, v\}$ is independently present with probability $p_{uv}$, where we require for some $\alpha > 1$ the following edge probability condition:

$$p_{uv} = \Theta\left( \min\left\{ 1, \left( \frac{w_u w_v}{\|x_u - x_v\|^\alpha W} \right)^\alpha \right\} \right).$$

Note that we can use an arbitrary norm as distance function. We also allow the special case $\alpha = \infty$ which yields a threshold behavior. We argued that underlying geometry is beneficial for reproducing structural properties that are caused by spatial information. In this sense, the geometric positions $x_v$ not only capture geographic coordinates but also additional information, for example occupation, hobbies, or music taste in social networks.

One of our first results shows that all models of hyperbolic random graphs are special cases of 1-dimensional GIRGs. Hence, GIRGs generalize hyperbolic random graphs to higher dimensions. But a difference between the two models remains: we ignore constant factors in the edge probabilities $p_{uv}$, allowing us to
simplify the expressions for $p_{uv}$ and reducing the technical overhead by avoiding hyperbolic cosines while preserving the qualitative behavior of the model. We also believe that GIRGs are technically easier than other recent models [BJR07; JM15].

We prove the following fundamental results on GIRGs. We provide a sampling algorithm that generates a random instance from our model in expected linear time, improving the best-known sampling algorithm for hyperbolic random graphs by a substantial factor $O(\sqrt{n})$ [LMP15]. Next, we show that GIRGs have small separators, i.e., it suffices to delete a sublinear number of edges to break the giant component into two large pieces. For hyperbolic random graphs, this was already proven by Bläsius, Friedrich, and Krohmer [BFK16]. We also verify that GIRGs can be stored by using $O(n)$ bits in expectation, and establish that GIRGs have a large clustering coefficient of $\Omega(1)$.

The expression for $p_{uv}$ ensures that the marginal probability of vertices $u, v$ with weights $w_u, w_v$ forming an edge is $\Pr[u \sim v] = \Theta(\min\{1, \frac{w_u w_v}{W}\})$, as in the Chung-Lu model. This probability does not change by more than a constant factor if we fix either $x_u$ or $x_v$. Furthermore, the expected degree of a vertex coincides with its weight. Hence, GIRGs can be interpreted as a variant of Chung-Lu random graphs with underlying geometry. Along these lines, we create a generic augmented and very general version of Chung-Lu random graphs. More concretely, we replace the torus by an arbitrary ground space $\mathcal{X}$ and let the dependence of $p_{uv}$ on $x_u, x_v, w_u, w_v$ be arbitrary as long as for random $x_v$ the marginal edge probability satisfies

$$\mathbb{E}_{x_v}[p_{uv} | x_u, w_u, w_v] = \Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right).$$

Note that we strip off any geometric specifics and in fact, the ground space is not even required to be metric. We only retain the most important features, namely power-law weights and the appropriate marginal edge probabilities.

As one of our main results we prove that all geometric variants of Chung-Lu random graphs that we obtain in this way have the same average distance $\frac{2+o(1)}{\log(\beta-2)} \log \log n$ for $2 < \beta < 3$, showing universality of the ultra-small property and extending it from Chung-Lu, Norros-Reittu, and hyperbolic random graphs,
where it was deduced before [CL02b; NR06; ABM17], to a much broader class of random graph models, including examples with non-metric underlying geometry. This result shows that specific choices, such as the underlying geometry being Euclidean or the dependence on the distance being inversely polynomial, do not significantly influence the average distance.

Beyond the average distance, we establish that this general model (together with its special case GIRGs) is scale-free and that in the regime $2 < \beta < 3$, we have a unique giant connected component and polylogarithmic diameter. Consequently, all classes of augmented Chung-Lu random graphs are reasonable models for real-world networks. Apart from providing new random graph models for complex networks, our general framework also reveals how previously existing scale-free models such as hyperbolic random graphs or Chung-Lu random graphs are related to each other.

All these results on GIRGs and their generalizations are joint work with Karl Bringmann and Johannes Lengler [BKL15; BKL16; BKL17].

After establishing that GIRGs reproduce the key structural properties of real-world networks and thus serve as a reasonable and flexible model, we demonstrate the utility of GIRGs by studying greedy routing in our model. As described above, Milgram’s experiment and the algorithmic small-world phenomenon have been first rigorously explained by Kleinberg in 2000 [Kle00a]. But from today’s perspective, his model has several severe shortcomings that have not been resolved simultaneously in subsequent work [MNW04; MN04; FGP06; FM06; CFL08]: routing succeeds only if the network is based on a perfect lattice structure, the result is fragile with respect to small changes in the model, and the graph model is unrealistically homogeneous. We address these drawbacks by studying greedy routing in GIRGs, as they overcome all three shortcomings. We therefore believe that studying greedy routing in GIRGs yields a more cogent theoretical explanation for Milgram’s findings. Moreover, there is a very natural interpretation of Milgram’s instruction to route to an “acquaintance who is more likely than you to know the target person” [TM69]: pick the neighbor most likely being adjacent to the target.

Indeed, greedy routing in GIRGs with power-law parameter $2 < \beta < 3$ per-
forms very well. For a random source and a random target node, we obtain that greedy routing succeeds with probability $\Omega(1)$. This is essentially best possible, as the target vertex is isolated with constant probability. Next, we show that in case of success, the stretch (the ration of the lengths of the greedy path and the shortest path) is $1 + o(1)$. Hence, greedy routing is asymptotically optimal, and the lengths of the greedy paths are $\frac{2 + o(1)}{\log(\beta - 2)} \log \log n$. Arguably, this is closer to the “six degrees of separation” than Kleinberg’s $O(\log^2(n))$. All previous analyses of greedy routing in small-world models yielded at least $\Omega(\log n)$ steps, e.g. [FG10]. It is quite surprising that greedy routing is much faster in GIRGs than in previously analyzed models, as we use GIRGs as a preexisting model whereas in previous studies of greedy routing, models were often specifically designed for the sole purpose of accelerating routing processes. Since hyperbolic random graphs are a special case of GIRGs, our results explain at the same time many findings of the extensive experimental work on hyperbolic random graphs [Pap+10; Kri+10], on the embedding of the internet graph [BPK10], and on related models [BKC09; Kri+09].

Since constant success probability is too low for practical applications, it is natural to ask whether there are patching methods that use backtracking to enforce success. We investigate this question and deduce that every method that satisfies three natural criteria has success probability 1, conditional on the source and the target being in the same component, and still has stretch $1 + o(1)$. These results also address the question of Krioukov et al. whether there are local routing protocols for the internet graph [Kri+07]. This question remained unsolved theoretically. Our results give, in GIRGs, for the first time a theoretical, affirmative answer to Krioukov’s question.

From a technical point of view, analyzing greedy routing in GIRGs turns out to be quite involved because dependencies between different hops of the routing process rapidly emerge, massively exacerbating the calculations. To overcome this issue, we develop a layer-based method for uncovering the random graph, on which our proofs heavily rely. In this way, we get rid of many obstructive dependencies and are only left with those which actually support the routing process and thus our analysis.
The work on greedy routing in GIRGs is joint work with Karl Bringmann, Johannes Lengler, Yannic Maus, and Anisur Molla [Bri+16; Bri+17].

1.3 Positional Games

Throughout history, in most cultures and societies games have been played. Humans play games for entertainment, achievement, reward, and for educational reasons. Even 5000 years ago in prehistoric Egypt, a board game called Senet was played [Tyl08]. Although defined via simple rules, games are usually not predictable and hard to analyze which is the reason why traditional games such as chess or Go are still exciting after being played for hundreds of years. Consider for example Tic-Tac-Toe played in three dimensions: while the first player easily wins the $3 \times 3 \times 3$ game by starting with the middle element, we only know by a computer-assisted proof that he also has a winning strategy for the $4 \times 4 \times 4$ game. According to Patashnik, one of the solvers, the complete winning strategy would fill a phone book [Pat80]. Moving one step higher, the $5 \times 5 \times 5$ game is already claimed to be a hopeless open problem.

Based on the seminal works of van Neumann and Morgenstern [Neu28; NM44] and Nash [Nas50], classical or algorithmic game theory evolved to a popular interdisciplinary field over the past centuries. It involves games with imperfect information: when deciding about the next action, players don't have all information available. For example, players have to decide simultaneously, they don't see the cards of the opponent, or some chance moves are included. Nowadays, game theory is applied in a wide range between economics, social sciences, psychology, and biology, whenever interactions include logical decisions [Osb04].

Vice-versa, in the discipline of positional games we consider games with perfect information. Its aim is to discover mathematical foundations on two-player games without chance moves, where during the whole game process, all information is available to the players. Examples include popular games such as chess, Go, Hex, or Tic-Tac-Toe just as more advanced, abstract games. Despite the perfect-information assumption, their high complexity makes it is usually
impossible to solve positional games by exhaustive search, leaving us with the necessity of developing smart combinatorial ideas.

The origin of the systematic studies of combinatorial games can be attributed to three pioneering papers. Hales and Jewett (1963) studied the $n$-dimensional $k \times k$ Tic-Tac-Toe [HJ63]. Using the strategy-stealing argument, they proved that either the first player has a winning strategy or the second player can force a draw. Moreover, for fixed $k$ they verified that the starting player wins if $n$ is sufficiently large, and on the other hand for fixed $n$ the game is a draw if $k$ is large enough. In 1973, Erdős and Selfridge generalized the framework to what is called a positional game today [ES73]: we have a set $X$ of elements, the board of the game, and the winning sets are given by an arbitrary family $\mathcal{A}$ of subsets of $X$. Using a beautiful and short argument, they showed that the second player can force a tie whenever $\sum_{A \in \mathcal{A}} 2^{-|A|} < \frac{1}{2}$, no matter how the concrete structure of the family $\mathcal{A}$ looks like. Furthermore, a simple construction shows that this criterion is tight.

Whenever the setting is symmetric in the sense that both players have the same goal, the game is called a strong game. By strategy-stealing, the first player can guarantee at least a draw in every strong game. So we are left with the question whether the first player actually has a winning strategy or not. Unfortunately, even simple strong games such as the $5 \times 5 \times 5$ Tic-Tac-Toe are hard to analyze and the current knowledge is rather disappointing. Turning back to the second player, when observing that he has no winning strategy at all, it is natural to slightly change the rules and modify his goal: now, he only wants to prevent the first player from winning, by “breaking” into every winning set $A \in \mathcal{A}$. This leads to the notion of Maker-Breaker games where the first player (“Maker”) wins if he claims an entire winning set $A \in \mathcal{A}$ and the second player (“Breaker”) wins in all other cases. The difference between strong games and Maker-Breaker games can be easily observed in the toy example $3 \times 3$ Tic-Tac-Toe: the outcome of its usual form as a strong game is a draw, whereas in the Maker-Breaker variant Maker has a winning strategy. For Maker-Breaker games, it turns out that the aforementioned criterion of Erdős and Selfridge again provides a winning condition for Breaker.
Very often, the board of Maker-Breaker games is given by a graph $G = (V, E)$ whose edges can be claimed by the players. Then the winning sets are for example spanning trees, Hamilton cycles, cliques, or $s$-$t$-paths. One other specific example of Maker-Breaker games are so-called Box games as introduced by Chvátal and Erdős in 1978 [CE78]. In Box games, the winning sets of the family $\mathcal{A}$ are disjoint. As soon as every set $A$ has size at least two, Breaker has a simple winning strategy for the Box game. This is not surprising since many Maker-Breaker are characterized by giving extreme advantage to one of the two players. To counterbalance this advantage, Chvátal and Erdős established the notion of biased positional games, where Maker is allowed to claim $m$ and Breaker is allowed to claim $b$ elements per turn. Usually, the parameter for the player favored by the game is set to 1, and we are interested to determine the threshold bias of the game: the smallest bias (i.e., choice of the second parameter) such that the originally disadvantaged player has a winning strategy for the biased game.

Based on these underlying papers and the methods introduced therein, the area of positional games has turned into a vivid and central field of combinatorics. In particular, József Beck developed a systematic theory behind combinatorial two player games, e.g., [Bec81; Bec82; Bec02]. One of the most fascinating phenomenons in positional games is the so-called random player intuition, being first observed of course by Erdős. This notion refers to the fact that for many positional games, the outcome between two optimal strategies is the same as if both players play randomly. One captivating example is the biased Maker-Breaker connectivity game on the complete graph on $n$ vertices. By the early result of Lehman [Leh64], the unbiased version is an easy win for Maker. Here, the random player intuition suggests a threshold bias of $\frac{n}{\log n}$ (cf. [Bol01]), and indeed, it turned out that this is the correct value [CE78; GS09]. This phenomenon appears quite often (cf., [Bec93; Bec94; Bec96]) and reveals a suprising but deep connection between deterministic games and random structures. Consequently, nowadays games are often analyzed via probabilistic tools even though they are completely deterministic. Amazingly, there exist even games where random strategies are almost optimal [BL00].
One other important aspect of positional games is that they have tight connections to other branches of discrete mathematics, which we are going to illustrate with two examples. First, the theorem of Hales and Jewett demonstrates that for $n$ large enough, every 2-coloring of $k^n$ contains a monochromatic line. This result not only built a link between the outcome of games and Ramsey-type questions, but also initiated the studies of combinatorial lines. For the second example, we turn back to the Erdős-Selfridge-criterion. Its correctness can be verified by using the probabilistic intuition and advising Breaker to always take the element $x \in X$ that minimizes the expected number of winning sets claimed entirely by Maker, assuming that the remaining elements are distributed independently and uniformly at random between Maker and Breaker. This method of conditional expectation is not at all restricted to games and was transferred to other fields: for example, it can be used to prove that every $k$-CNF with less than $2^k$ clauses is satisfiable. Furthermore, inspired by his research on positional games, Beck achieved the first efficient derandomization of the Lovász-Local-Lemma by designing a polynomial-time algorithm for finding a satisfying assignment for a $k$-CNF, under the assumption that every clause intersects at most $2^{k/48}$ other clauses [Bec91]. In this sense, positional games often open a new perspective and new methods for other combinatorial problems. For an overview, details, and further examples of positional games we refer to the recent books of Beck [Bec08] and Hefetz et al. [Hef+14].

Apart from Maker-Breaker games, strong games, and closely related families of positional games, in the last quarter-century a new class of positional games, related to extremal graph theory, gained attention in combinatorics. Let $\mathcal{P}$ be a monotone graph property. Then the saturation game is played as follows: two players, called Maxi and Mini, start with the empty graph on $n$ vertices and take turns in which they extend the current graph $G$ by adding one additional edge $e$ such that $G \cup \{e\}$ does not satisfy $\mathcal{P}$. At some point, the process yields a graph $G_{\text{end}}$ which is saturated with respect to $\mathcal{P}$, that is, no additional edge can be inserted without violating the rule. There the game stops. Mini wants to minimize the number of edges in $G_{\text{end}}$ whereas Maxi’s goal is to maximize this number. The score of the game, denoted by $s(n, \mathcal{P})$, is the total number of edges
in \( G_{\text{end}} \), given that both players followed an optimal strategy. When analyzing saturation games, we attempt to determine the score of the game as precisely as possible.

Even though being defined via simple combinatorial rules, saturation games define surprisingly involved problems as they combine two issues: (i) both players want to follow their own strategy and simultaneously “break” into the strategy of the adversary, and (ii) the two players have different goals. Therefore, a priori it is not clear at all which type of strategy a player should apply and for how many edges one should aim at. If a player is too ambitious, the adversary perhaps easily breaks into her strategy. But if a player only creates a weak structure, it is likely that she is far away from playing optimally. The challenge of finding a suitable trade-off between the two extremes is omnipresent in the triangle-free game. Here, Mini and Maxi are forced to keep the graph triangle-free, i.e., \( \mathcal{P} = K_3 = \{ G \text{ contains a copy of } K_3 \} \). Clearly we have the two trivial bounds \( n - 1 \leq s(n, K_3) \leq n^2/4 \) (see [EHM64], e.g.). Füredi, Reimer, and Seress proved \( s(n, K_3) \geq (1/2 + o(1)) n \log_2 n \) [FRS91; Ser92], only a small improvement compared to the trivial lower bound but still the only lower bound that is known to date. An upper bound of \( n^2/5 \) is attributed to Erdős, but unfortunately the proof is lost, and so far this bound couldn’t be reconstructed. We see that the remaining gap is tremendous. Nevertheless, the random player intuition allows us to make a plausible guess on the score: transferring Bohman’s remarkable analysis of the triangle-free process [Boh09], it is reasonable that the score of the triangle-free game would be \( \tilde{\Theta}(n^{3/2}) \). In general, knowledge on saturation games is still rather scarce and many challenging open problems are waiting for being solved.

### 1.4 Graph Coloring Games

Finding a proper vertex coloring is one of the most fundamental questions of graph theory. For a given graph \( G = (V, E) \) and a set \( I \) of \( k \) colors, we want to find an assignment \( c : V \to I \) such that for every edge \( \{u, v\} \in E \) we have \( c(u) \neq c(v) \). The chromatic number \( \chi(G) \) is defined as the smallest integer \( k \) for which such a
coloring \( c \) exists. Undoubtedly the most famous problem in the area of graph colorings is to prove that every planar graph can be colored with only four colors. After decades of frustrating research on this particular problem, finally in 1976 Appel and Haken managed to find a computer-assisted proof of the theorem [AH77a; AH77b], but unfortunately, still no traditional proof of the statement exists.

In 1981, in his column in the Scientific American Martin Gardner suggested the following Maker-Breaker game [Gar81]: given a planar graph \( G \) and a set of \( k \) colors, Maker and Breaker take turns in which they alternately color a vertex of \( G \) such that two neighboring vertices never get the same color. Maker wins if the process terminates with a proper coloring of \( G \). Breaker wins as soon as the partial coloring cannot be extended to a proper coloring of \( G \) anymore. If Maker could win the game on every planar graph with only 4 colors, this would immediately imply a proof of the four color theorem. However, as demonstrated by Kierstead and Trotter (1994), there exist counter-examples where Maker has no chance in winning the game [KT94]. Anyhow, this game became one of the classic Maker-Breaker games, being analyzed for many different graph families and extended by several variations. What makes the game especially interesting is that colorability properties are *global* graph properties, and thus one bad move of a player can hurt for the remainder of the game. For a graph \( G \) the *game chromatic number* \( \chi_g(G) \) is defined as the smallest integer \( k \) for which Maker has a winning strategy in the coloring game. Clearly, for every graph \( G \) we have the two trivial bounds \( \chi(G) \leq \chi_g(G) \leq \Delta(G) + 1 \), where \( \Delta(G) \) denotes the maximum degree of \( G \). Hence, this graph parameter is well-defined.

In recent years it has become popular to study positional games on random graphs, cf. [SS05]. For games, random graphs are not only interesting as board because they yield a distribution over all possible graphs or because they are one of the central topics of combinatorics in general, but essentially because they provide a sparser board than complete graphs and implicitly add a bias to the game, making the game more balanced while preserving at the same time certain global graph structures. In terms of the game chromatic number, we obviously benefit from the sparser graph since the vertex coloring game is
trivial on the complete graph. Hence, in 2008, Bohman, Frieze, and Sudakov initiated the analysis of the game chromatic number of random graphs $G(n, p)$ [BFS08]. Let $\varepsilon > 0$ and let $(\log n)^{C(\varepsilon)} / n < p < 1$, where $C(\varepsilon) > 0$ is a sufficiently large constant. Bohman, Frieze, and Sudakov proved that asymptotically almost surely (a.a.s.), the random graph $G_{n, p}$ is such that

$$(1 - \varepsilon) \frac{n}{\log_b(np)} \leq \chi_g(G_{n, p}) \leq (2 + \varepsilon) \frac{n}{\log_b(np)},$$

where $b = \frac{1}{1-p}$. For the remaining gap, they conjectured that as long as $np \to \infty$, a.a.s. it holds $\chi_g(G_{n, p}) = (1 + o(1)) \frac{n}{\log_b(np)}$.

We investigate this coloring game on dense random graphs and prove their conjecture for dense random graphs. More precisely, we show that for every $p \geq e^{-o(\log n)}$ and every $\alpha > 1$, a.a.s. it holds

$$\chi_g(G_{n, p}) \leq \alpha \frac{n}{\log_b(np)}.$$ 

In particular, our result applies to random graphs with constant parameter $p$. Our proof is based on a greedy strategy: we advise Maker to either color the vertex with fewest available colors or to use the least frequent color at an arbitrary vertex (if possible). Then, one of the key ingredients of the proof is to construct a reduction from the graph coloring game to Box games.

This result on the game chromatic number of random graphs is joint work with Angelika Steger [KS14].

A very natural variant of the vertex coloring game is the variation where Maker and Breaker color the edges instead of the vertices, as first considered by Lam, Shiu, and Xu [LSX99]. Here, the game chromatic index $\chi'_g(G)$ denotes the smallest number of colors $k$ for which Maker has a winning strategy. Clearly, for every graph $G$ we have the trivial bounds $\Delta(G) \leq \chi'_g(G) \leq 2\Delta(G) - 1$. Unfortunately, it is believed that this game is hard to analyze and in general, no powerful strategies for the players are known [Bev08], not even when the game is played on complete graphs. To date, accurate bounds on the game chromatic index are only known for special and sparse graph classes.

Therefore, Beveridge, Bohman, Frieze, and Pikhurko (2008) proposed that it is reasonable to first decide whether $\chi'_g(G)$ is bounded away from the trivial
lower bound, from the trivial upper bound, or from both [Bev+08]. In this spirit, they demonstrated that for every $\delta > 0$ there exists $\varepsilon > 0$ such that for every graph $G$ on $n$ vertices with maximum degree $\Delta(G) \geq (\frac{1}{2} + \delta)n$ it holds

$$\chi'_G(G) \leq (2 - \varepsilon)\Delta(G).$$

Furthermore, they conjectured that there exists a universal constant $c > 0$ such that $\chi'_G(G) \leq (2 - c)\Delta(G)$ holds for any graph [Bev+08]. Here, we make a first step towards a proof of their conjecture. We show that there exist $C > 0$ and $\varepsilon > 0$ such that for every graph $G$ with $\Delta(G) \geq C \log n$, it holds

$$\chi'_G(G) \leq (2 - \varepsilon)\Delta(G).$$

This gives the first non-trivial bound on $\chi'_G(G)$ for a large portion of graphs (e.g., bipartite graphs or random graphs with $C \log n / n \leq p < 1/2$). We obtain the result by proposing a strategy where in each of his turns, Maker chooses both the edge and the color at random. In addition, we investigate a biased version of the game where Breaker is allowed to color $b$ edges per turn and give bounds on the number of colors that Maker needs to win this biased variant of the game.

The last class of games that we consider are colorability saturation games. For a given integer $k$, we investigate the saturation game with respect to the monotone property

$$\chi_{>k} = \{G \text{ is not } (k + 1)-\text{colorable}\}.$$

That is, Maxi and Mini draw edges such that the graph remains $k$-colorable. In contrast to the aforementioned coloring games, in colorability saturation games the game process is not about coloring a given graph but about creating a graph with certain colorability properties. Note that by the Turán number, every $k$-colorable graph contains at most $\binom{n}{2}(1 - 1/k + o(1))$ edges, yielding a first yet trivial upper bound on the score of these games. For the bipartite game with $k = 2$ it is straightforward to see that the score is actually $\lfloor n^2 / 4 \rfloor$ [Car+17], hence Mini has no power in this specific game. But as soon as $k > 2$, the game becomes much more interesting. The analysis of the general colorability saturation game has been initiated recently by Hefetz, Krivelevich, Naor, and Stojaković [Hef+16].
Specifically, they proved that there exists a constant $C > 0$ such that for all $k$ and all $n$ sufficiently large compared to $k$, it holds

$$s(n, \chi > k) \geq \left( \binom{n}{2} \left(1 - \frac{C \log k}{k}\right) \right).$$

We improve these bounds on the score of the colorability saturation game and provide almost matching lower and upper bounds. Concretely, for all integers $k$ and all $n > k$ we prove

$$\left( \binom{n}{2} \left(1 - \frac{1}{\lceil k/2 \rceil}\right) \right) \leq s(n, \chi > k) \leq \left( \binom{n}{2} \left(1 - \frac{1}{k - \lfloor (k - 1)/3 \rfloor}\right) \right) + n.$$

In contrast to the previous lower bound (whose proof uses $C > 3000$) our bounds are also non-trivial for small (but intuitive!) choices of $k$. While previous results on saturation games often base on rather long case distinctions, we use a different approach: we work with carefully chosen potential functions that are closely related to the density of induced subgraphs, constituting a novel method for the analysis of saturation games. In particular, this allows us to propose and analyze a very adaptive strategy for Mini. We thus believe that our upper bound on the score is tight or almost tight. This guess is supported by a detailed analysis of the specific game with parameter $k = 4$, where we argue that the score indeed is $s(n, \chi > 4) = \frac{n^2}{3} + O(n)$.

All two-player games studied in this thesis are closely related to each other, having graph colorings as common subject: during the games, either given graphs are colored or graphs with colorability properties are constructed. Nevertheless, the results are based on very different strategies for the players, covering almost all frequent methods in the area of positional games: a greedy approach for the game chromatic number, a random strategy for the game chromatic index, and potential functions and case distinctions for the saturation games. This variety of employed techniques indicates once again the richness and diversity that is inherently present in the field of combinatorial games.
1.5 Organization of the Thesis

We start by introducing general notations and collecting some basic inequalities in Chapter 2. In particular, we prove an Azuma-type inequality which bounds large deviations, taking into account some bad events. This inequality is especially useful for obtaining concentration results on GIRGs.

In Chapter 3, we formally define GIRGs, variations, and generalizations, list various examples and study basic properties of the models. Furthermore, we compare GIRGs to other, preexisting models. Figure 3.1 on page 34 summarizes the hierarchy of the different random graph models that we study in this thesis. In Section 3.5, we describe hyperbolic random graphs in more details, summarize recent advances in the theoretical analysis of the model, and formally prove that they are a subclass of GIRGs. In Section 3.4 on pages 50-53 we present plots of GIRGs with different parameter choices (Figures 3.2-3.5). Next, in Chapter 4 we establish GIRGs as a reasonable model for large real-world networks by proving fundamental structural and algorithmic properties: power-law degree sequence, existence of a giant connected component, ultra-small average distance, poly-logarithmic diameter, clustering, small separators, and compression and sampling algorithms that run in expected linear time. We refer to Section 4.1 for a condensed list of these results. Afterwards, Chapter 5 contains an in-depth analysis of greedy routing in GIRGs and additional related work. The formal statements of the various results on greedy routing in GIRGs are presented in Section 5.3. This chapter concludes the first part of the thesis.

Afterwards, we proceed with the second part where we study three classes of graph coloring games. In Chapter 6, we determine the asymptotic value of the game chromatic number of dense random graphs. In Chapter 7 we analyze the game chromatic index of arbitrary graphs with sufficiently high maximum degree. Finally, we study colorability saturation games in Chapter 8. In each of these chapters, we present further background on the respective games and discuss existing results on related positional games.
In this short chapter we introduce basic notations and collect several concentration inequalities that will be used throughout the thesis. In particular, we prove an Azuma-type inequality that takes into account some bad events. This inequality will be useful when proving concentration results on geometric inhomogeneous random graphs.

### 2.1 Notations

We use the standard Landau symbols $O$, $\Omega$, $\Theta$, $o$, $\omega$ for describing the asymptotic behavior of functions $f(n)$ w.r.t. $n \to \infty$. Note that the Landau symbols may hide positive or negative functions. We say that an event holds with high probability (w.h.p.) if it holds with probability $1 - n^{-\omega(1)}$ and that it holds asymptotically almost surely (a.a.s.) if it holds with probability $1 - o(1)$. We use $\log n$ for the natural logarithm and $\log_b n$ for the logarithm to base $b$. In general, we denote by $G = (V, E)$ a graph $G$ with vertex set $V$ and edge set $E$, where $E$ is a subset of the two-element subsets of $V$. For graphs, we use the following notation.
<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u \sim v$</td>
<td>${u, v} \in E$</td>
</tr>
<tr>
<td>$u \sim A$</td>
<td>there exists $v \in A$ such that $u \sim v$</td>
</tr>
<tr>
<td>$A \sim B$</td>
<td>there exist $u \in A, v \in B$ such that $u \sim v$</td>
</tr>
<tr>
<td>$\Gamma_G(v)$</td>
<td>the neighborhood of a vertex $v$ in $G$, i.e., $\Gamma_G(v) = {u \in V</td>
</tr>
<tr>
<td>$\deg_G(v)$</td>
<td>the degree of a vertex $v$ in $G$, i.e., $\deg_G(v) =</td>
</tr>
<tr>
<td>$\Delta(G)$</td>
<td>the maximum degree of $G$, i.e., $\Delta(G) = \max{\deg_G(v)</td>
</tr>
<tr>
<td>$\delta(G)$</td>
<td>the minimum degree of $G$, i.e., $\delta(G) = \min{\deg_G(v)</td>
</tr>
<tr>
<td>$\dist_G(u,v)$</td>
<td>the number of edges in a shortest $u$-$v$-path</td>
</tr>
<tr>
<td>$\chi(G)$</td>
<td>the chromatic number of $G$: the smallest $k \in \mathbb{N}$ such that the vertices of $G$ can be colored with colors ${1, \ldots, k}$ without creating a monochromatic edge</td>
</tr>
<tr>
<td>$G[A]$</td>
<td>the subgraph of $G$ induced by the vertex set $A \subseteq V$</td>
</tr>
<tr>
<td>$E(A)$</td>
<td>the set of edges in the induced subgraph $G[A]$</td>
</tr>
<tr>
<td>$E(A,B)$</td>
<td>the set of edges ${u,v}$ such that $u \in A$ and $v \in B$, where we assume that $A$ and $B$ are disjoint subsets of $V$</td>
</tr>
</tbody>
</table>

Whenever the notation is clear, we drop subscripts and simply write $\Gamma(v)$, $\deg(v)$, and $\dist(u,v)$. For geometric inhomogeneous random graphs, we need some additional notation, which is collected in the next table. In the following definitions, we always assume that $w$ is a strictly positive real number.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_v$</td>
<td>the geometric position of a vertex $v$</td>
</tr>
<tr>
<td>$\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}$</td>
<td>the $d$-dimensional torus</td>
</tr>
<tr>
<td>$\text{VOL}(A)$</td>
<td>the Lebesgue-measure of a measurable set $A \subseteq [0,1]^d$</td>
</tr>
<tr>
<td>$w_v$</td>
<td>the weight of a vertex $v$</td>
</tr>
<tr>
<td>$V_{\geq w}$ resp. $V_{\leq w}$</td>
<td>${v \in V</td>
</tr>
<tr>
<td>$W$</td>
<td>the total weight $\sum_{v \in V} w_v$</td>
</tr>
<tr>
<td>$W_{\geq w}$, resp. $W_{\leq w}$</td>
<td>$\sum_{v \in V_{\geq w}} w_v$ resp. $\sum_{v \in V_{\leq w}} w_v$</td>
</tr>
<tr>
<td>$w_{\text{min}}$, resp. $w_{\text{max}}$</td>
<td>$\min{w_v</td>
</tr>
</tbody>
</table>
2.2 Concentration Inequalities

Let us start with Markov’s inequality.

**Theorem 2.1** (Markov’s inequality). *Let $X$ be a non-negative random variable. Then for all $t > 0$, it holds*

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$  

We repeatedly use the following well-known concentration bounds for the tails of sums of Bernoulli random variables or for the tails of Poisson random variables.

**Theorem 2.2** (Chernoff-Hoeffding bound, Theorem 1.1 in [DP09]). *For $1 \leq i \leq n$, let $X_i$ be random variables that are independently distributed in $[0,1]$, and let $X := \sum_{i=1}^n X_i$. Then*

(i) \[ \Pr \left[ X > (1 + \varepsilon)\mathbb{E}[X] \right] \leq \exp \left( -\frac{\varepsilon^2}{3} \mathbb{E}[X] \right), \text{ for all } 0 < \varepsilon < 1, \]

(ii) \[ \Pr \left[ X < (1 - \varepsilon)\mathbb{E}[X] \right] \leq \exp \left( -\frac{\varepsilon^2}{2} \mathbb{E}[X] \right), \text{ for all } 0 < \varepsilon < 1, \text{ and} \]

(iii) \[ \Pr[X > t] \leq 2^{-t}, \text{ for all } t > 2e\mathbb{E}[X]. \]

*Moreover, if $X$ is a Poisson random variable, then (i)-(iii) hold as well (e.g., Theorem 5.4 in [MU05]).*

We also need Le Cam’s theorem which allows us to bound the total variation distance of a binomial distribution to a Poisson distribution with the same mean.

**Theorem 2.3** (Le Cam, Proposition 1 in [LC60]). *For $1 \leq i \leq n$, let $X_i$ be independent Bernoulli random variables such that $\Pr[X_i = 1] = p_i$. Let $\lambda_n = \sum_{i=1}^n p_i$ and $S_n = \sum_{i=1}^n X_i$. Then*

\[ \sum_{k=0}^{\infty} \left| \Pr[S_n = k] - \frac{\lambda_n^k e^{-\lambda_n}}{k!} \right| < 2 \sum_{i=1}^n p_i^2. \]

*In particular, if $\lambda_n = \Theta(1)$ and $\max_{i \in [n]} p_i = o(1)$, then $\Pr[S_n = k] = \Theta(1)$, for $k = O(1)$.***
Next, we formulate a concentration inequality which bounds large deviations taking into account some bad event $B$. We start with the following variant of McDiarmid's inequality.

**Theorem 2.4** (Theorem 3.6 in [Kut02], slightly simplified). Let $X_1, \ldots, X_m$ be independent random variables over $\Omega_1, \ldots, \Omega_m$. Let $X = (X_1, \ldots, X_m)$, $\Omega = \prod_{k=1}^m \Omega_k$ and let $f : \Omega \to \mathbb{R}$ be measurable with $0 \leq f(\omega) \leq M$, for all $\omega \in \Omega$ and for some $M > 0$. Let $B \subseteq \Omega$ such that for some $c > 0$ and for all $\omega \in B$, $\omega' \in \Omega$ that differ in only one component $X_i$ we have

$$|f(\omega) - f(\omega')| \leq c.$$

Then for all $t > 0$

$$\Pr \left[ |f(X) - \mathbb{E}[f(X)]| \geq t \right] \leq 2e^{-\frac{t^2}{8mc^2}} + 2 \frac{mM}{c} \Pr[B].$$

Our improved version of this theorem is the following, where in the Lipschitz condition both $\omega$ and $\omega'$ come from the good set $\overline{B}$, but we have to consider changes of two components at once. Recently, a similar inequality has been proven by Combes [Com15], see also [War16]. This inequality was obtained in joint work with Karl Bringmann and Johannes Lengler, e.g., [BKL16].

**Theorem 2.5.** Let $X_1, \ldots, X_m$ be independent random variables over $\Omega_1, \ldots, \Omega_m$. Let $X = (X_1, \ldots, X_m)$, $\Omega = \prod_{k=1}^m \Omega_k$ and let $f : \Omega \to \mathbb{R}$ be measurable such that $0 \leq f(\omega) \leq M$, for all $\omega \in \Omega$ and for some $M > 0$. Let $B \subseteq \Omega$ such that for some $c > 0$ and for all $\omega \in \overline{B}, \omega' \in \overline{B}$ that differ in at most two components $X_i, X_j$ we have

$$|f(\omega) - f(\omega')| \leq c.$$

Then for all $t \geq 2M\Pr[B]$

$$\Pr \left[ |f(X) - \mathbb{E}[f(X)]| \geq t \right] \leq 2e^{-\frac{t^2}{32mc^2}} + \left(2 \frac{mM}{c} + 1\right) \Pr[B].$$

**Proof.** We say that $\omega, \omega' \in \Omega$ are *neighbors* if they differ in exactly one component $\Omega_k$. Given a function $f$ as in the statement, we define a function $f'$ as
follows. On $\overline{\mathcal{B}}$ the functions $f$ and $f'$ coincide. Let $\omega \in \mathcal{B}$. If $\omega$ has a neighbor $\omega' \in \overline{\mathcal{B}}$, then choose any such $\omega'$ and set $f'(\omega) := f(\omega')$. Otherwise set $f'(\omega) := f(\omega)$.

The constructed function $f'$ satisfies the precondition of Theorem 2.4. Indeed, let $\omega \in \overline{\mathcal{B}}$ and $\omega' \in \Omega$ differ in only one position. If $\omega' \in \overline{\mathcal{B}}$, then clearly $f'(\omega) = f(\omega)$ and $f'(\omega') = f(\omega')$, and we obtain $|f'(\omega) - f'(\omega')| \leq c$ by the assumption on $f$. Otherwise we have $\omega' \in \mathcal{B}$, and since $\omega'$ has at least one neighbor in $\overline{\mathcal{B}}$, namely $\omega$, we have $f'(\omega') = f(\omega'')$ for some neighbor $\omega'' \in \overline{\mathcal{B}}$ of $\omega'$. Note that both $\omega$ and $\omega''$ are in $\overline{\mathcal{B}}$, and as they are both neighbors of $\omega'$ they differ in at most two components. Thus, by the assumption on $f$ we have

$$|f'(\omega) - f'(\omega')| = |f(\omega) - f(\omega'')| \leq c.$$ 

Hence, we can use Theorem 2.4 on $f'$ and obtain concentration of $f'(X)$. Specifically, since $\Pr[f(X) \neq f'(X)] \leq \Pr[\mathcal{B}]$, and thus $|\mathbb{E}[f(X)] - \mathbb{E}[f'(X)]| \leq M\Pr[\mathcal{B}]$, we obtain

$$\Pr[|f(X) - \mathbb{E}[f(X)]| \geq t] \leq \Pr[\mathcal{B}] + \Pr[|f'(X) - \mathbb{E}[f'(X)]| \geq t - M\Pr[\mathcal{B}]]$$

$$\leq \Pr[\mathcal{B}] + \Pr[|f'(X) - \mathbb{E}[f'(X)]| \geq t/2],$$

since $t \geq 2M\Pr[\mathcal{B}]$, which together with Theorem 2.4 proves the claim. ■

Finally, we need Jensen’s inequality for concave and convex functions.

**Theorem 2.6 (Jensen’s inequality).** Let $X$ be a random variable with finite expectation. Then for every convex function $f$ it holds

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$$

and for every concave function $g$ it holds

$$\mathbb{E}[g(X)] \leq g(\mathbb{E}[X]).$$
PART I

GEOMETRIC INHOMOGENEOUS RANDOM GRAPHS
In this chapter we formally introduce geometric inhomogeneous random graphs (GIRGs) as a new mathematical model for real-world networks. We also define several variants, modifications, and generalizations of GIRGs. Thereby, our goal is not only to provide a new class of realistic and flexible random graph models, but also to discover how previously existing models are related to each other. We refer to Figure 3.1 on page 34 for an overview of the different models that we study in this thesis.

For the sake of clarity, in Section 3.1 we start with the most general model where we generically augment Chung-Lu random graphs by underlying geometry. We verify some elementary properties of this model in Section 3.2. Afterwards, in Section 3.3 we proceed by defining more and more concrete subclasses of the model and discussing several examples. In particular, we regain the Chung-Lu model as a special case of our general class.

In Section 3.4 we then define geometric inhomogeneous random graphs. In order to illustrate the model and give additional intuition, with Figures 3.2-3.5 on pages 50-53 we present four plots of GIRGs with different parameter
choices. Afterwards, in Section 3.5 we study hyperbolic random graphs, present a brief overview of recent results on this model, and rigorously prove that hyperbolic random graphs are a special case of one-dimensional GIRGs. Finally, in Section 3.6 we compare GIRGs with further related models.

Most parts of this chapter are based on joint work with Karl Bringmann and Johannes Lengler [BKL15; BKL16; BKL17].

3.1 The General Model

We introduce a general random graph model with underlying geometry which can be seen as a generic augmented version of Chung-Lu random graphs. In our general class of scale-free networks, randomness is involved in both the vertex set $V$ and the edge set $E$. In a nutshell, each vertex $v$ comes with a weight $w_v$, corresponding to its expected degree, and with a random position $x_v$ in a geometric space $\mathcal{X}$. Afterwards, for every pair of vertices we toss a coin whether the edge is present in the graph. We now present the full definition by starting with the weight sequence.
**Power-law weights:** For $n \in \mathbb{N}$ let $w = (w_1, \ldots, w_n)$ be a non-increasing sequence of positive weights. We call $W := \sum_{v=1}^{n} w_v$ the **total weight**. We assume that the weights follow a **power law**: the fraction of vertices with weight at least $w$ is $\approx w^{1-\beta}$ for some parameter $\beta > 2$ (the **power-law exponent** of $w$). More precisely, we assume that for some $\bar{w} = \bar{w}(n)$ with $n^{\omega(1/\log \log n)} \leq \bar{w} \leq n^{(1-\Omega(1))/(\beta-1)}$, the sequence $w$ satisfies the following conditions:

(PL1) the minimum weight is constant, i.e., $w_{\text{min}} := \min\{w_v \mid 1 \leq v \leq n\} = \Omega(1)$;

(PL2) for all $\eta > 0$ there exist constants $c_1, c_2 > 0$ such that

$$c_1 \frac{n}{w^{\beta-1+\eta}} \leq |\{1 \leq v \leq n \mid w_v \geq w\}| \leq c_2 \frac{n}{w^{\beta-1-\eta}},$$

where the first inequality holds for all $w_{\text{min}} \leq w \leq \bar{w}$ and the second holds for all $w \geq w_{\text{min}}$.

We remark that these are standard assumptions for power-law graphs with constant average degree. A concrete example is the widely used weight function $w_v := w_{\text{min}} \cdot (n/v)^{1/(\beta-1)}$ with parameter $w_{\text{min}} = \Theta(1)$.

Note that since $\bar{w} \leq n^{(1-\Omega(1))/(\beta-1)}$, there are $n^{\Omega(1)}$ vertices with weight at least $\bar{w}$. On the other hand, no vertex has weight larger than $(c_2 n)^{1/(\beta-1-\eta)}$.

It is natural to state the power-law assumptions in terms of the **cumulative distribution function**, i.e., by considering the number of vertices of weight at least $w$, and not in terms of the probability density function. On the one hand, this makes the definition more general and flexible. On the other hand, in empirical analysis of real-world data it is often tricky to measure the tail of a degree distribution as data sets might be noisy. Therefore, data is mostly analyzed in terms of the cumulative distribution function [New03].

In some complex networks, the degree distribution closely follows a power law for small and intermediate values, but rather vaguely for large values. In such situations, the lack of sufficiently many vertices with ultra-large degree can be explained by large degrees requiring some sort of additional costs. Then, the degree distribution rather follows a power law with **exponential cutoff**, that is, the fraction of vertices with degree $k$ is proportional to $k^{-\beta} e^{-\lambda k}$ for some small parameter $\lambda$ [CSN09; Hof16]. Note that our definition of (PL2) takes this
observation into account as its lower bound is only required for values \( w \leq \bar{w} \).

In this way, we include a model with exponential cutoff as long as \( \lambda \leq \beta \frac{\log \bar{w}}{\bar{w}} \).

**Random graph model:** Let \( \mathcal{X} \) be a non-empty set, equipped with a probability measure \( \mu \) that allows to sample elements from \( \mathcal{X} \). We call \( \mathcal{X} \) the *ground space* of the model and the elements in \( \mathcal{X} \) *positions*. The random graph \( \mathcal{G}(n, \mathcal{X}, w, p) \) has vertex set \( V = [n] = \{1, \ldots, n\} \). For any vertex \( v \) we independently draw a position \( x_v \in \mathcal{X} \) according to measure \( \mu \). Conditional on \( x_1, \ldots, x_n \), we connect any two vertices \( u \neq v \) independently with probability

\[
p_{uv} := p_{uv}(x_u, x_v) := p_{uv}(x_u, x_v; n, \mathcal{X}, w),
\]

where \( p \) is a (symmetric in \( u, v \) and measurable) function mapping to \([0, 1]\) that satisfies the following condition:

**(EP1)** for any \( u, v \), if we fix position \( x_u \in \mathcal{X} \) and draw position \( x_v \) from \( \mathcal{X} \) according to \( \mu \), then the marginal edge probability is

\[
\mathbb{E}_{x_v} [p_{uv}(x_u, x_v) \mid x_u] = \Theta \left( \min \left\{ 1, \frac{w_u w_v}{W} \right\} \right).
\]

For some results we also need an additional condition, to ensure the existence of a unique giant component:

**(EP2)** for all \( \eta > 0 \), any \( u, v \) with \( w_u, w_v \geq \bar{w} \), and any fixed positions \( x_u, x_v \in \mathcal{X} \) we have

\[
p_{uv}(x_u, x_v) \geq \left( \frac{n}{w \beta - 1 + \eta} \right)^{-1+\omega(1/\log \log n)}.
\]

**Discussion of the model:** Let us first argue why condition (EP2) is necessary to obtain a unique giant component. Suppose we have an instantiation \( \mathcal{G} \) of our model on a space \( \mathcal{X} \). We will see in Section 4.3 that with high probability \( \mathcal{G} \) has a giant component that contains all high-degree vertices. Now make a copy \( \mathcal{X}' \) of \( \mathcal{X} \), and consider a graph where all vertices draw geometric positions from \( \mathcal{X} \cup \mathcal{X}' \). Vertices in \( \mathcal{X} \) are never connected to vertices in \( \mathcal{X}' \), but within \( \mathcal{X} \) and \( \mathcal{X}' \) we use the same connection probabilities as for \( \mathcal{G} \). Then the resulting graph satisfies all properties of our model except for (EP2), but it has two giant
components, one in $\mathcal{X}$ and one in $\mathcal{X}'.$ As we will see, (EP2) ensures that the high-weight vertices form a single dense network, so that the graph indeed has a unique giant component. However, for instance for our results on the degree sequence (EP2) is not necessary.

Since the right hand side of (EP1) is the edge probability of Chung-Lu graphs, it is a natural condition for any augmented version of Chung-Lu graphs. In particular, (EP1) ensures that the expected degree of a vertex $v$ with weight $w_v$ is indeed $\Theta(w_v)$. For similar reasons as discussed for (EP2), we cannot further relax (EP1) to a condition on the marginal probability over random positions $x_u$ and $x_v$, i.e., a condition like $E_{x_u,x_v}[p_{uv}(x_u,x_v)] = \Theta(\min\{1, \frac{w_u w_v}{W}\})$. Indeed, consider the same setup as above, with $\mathcal{G}$, $\mathcal{X}$, and copy $\mathcal{X}'$. For two vertices of weight at most $\bar{w}$, connect them only if they are in the same copy of $\mathcal{X}$. For two vertices of weight larger than $\bar{w}$, always treat them as if they came from the same copy (then condition (EP2) is satisfied). For a vertex $u$ of weight at most $\bar{w}$ and $v$ of weight larger than $\bar{w}$, connect them only if $u$ is in $\mathcal{X}'$. Then the high-weight vertices form a unique component, but it is only connected to vertices in $\mathcal{X}'$, while the low-weight vertices in $\mathcal{X}$ may form a second giant component. Thus, in (EP1) it is necessary to allow any fixed $x_u$.

**Sampling the weights:** In the definition we assume that the weight sequence $w$ is fixed, as it is the case in Chung-Lu random graphs. Some other models like Norros-Reittu random graphs (Example 3.7 on page 43) use a different approach and assume that the weights are chosen at random from an appropriate distribution. If we sample the weights from a power-law distribution, then the sampled weight sequence itself follows a power law with probability $1 - n^{-\Omega(1)}$, so that a model with sampled weights is almost surely included in our model. For the precise statement, see Lemma 3.5 below.

**Outlook:** In the remainder of this chapter, we see that our general framework includes preexisting models such as Chung-Lu random graphs and hyperbolic random graphs as well as geometric inhomogeneous random graphs that we study in more details. Hence, we encompass graphs with very different local
structure as, for example, the clustering coefficient varies drastically between Chung-Lu random graphs where it is $n^{-\Omega(1)}$ and hyperbolic random graphs where it is constant [GPP12].

In Chapter 4 we prove that in the range $2 < \beta < 3$, all instantiations of our model have w.h.p. a unique giant connected component of linear size with polylogarithmic diameter. Furthermore, a.a.s. the same average distance is $\frac{2+o(1)}{\log(\beta-2)} \log \log n$, showing universality of the ultra-small world property. Beyond these properties, we establish for the regime $\beta > 2$ that our general model has a power-law degree sequence. Hence, all variants that are included in our generic class are reasonable models for real-world networks.

The generality of this result is rather surprising. Note that by the scale-free property, we always have $\Theta(n)$ edges w.h.p. If an instance has high clustering, many edges are local edges inside well-connected subgraphs, and therefore futile for finding short paths between distant vertices. Still, we demonstrate that in such graphs, the average distance is asymptotically the same as in Chung-Lu random graphs, where we have no clustering and every edge is potentially helpful when searching for short paths. We also remark that in the regime $\beta > 3$ our model is too general to obtain any meaningful results about the component structure or distances as in this range, there exist examples with a giant component but also examples without giant connected component.

### 3.2 Basic Properties of the Model

In this section we prove some basic properties of the considered random graph model. In particular we calculate the expected degree of a vertex and the marginal probability that an edge between two vertices with given weights is present. Let us start by calculating the partial weight sums $W_{\geq w} := \sum_{v \in V_{\geq w}} w_v$ and $W_{\leq w} := \sum_{v \in V_{\geq w}} w_v$. The values of these sums follow from the assumptions on power-law weights in Section 3.1.

**Lemma 3.1 (Weight sums).** The total weight satisfies $W = \Theta(n)$. Moreover, for all sufficiently small $\eta > 0$,

(i) $W_{\geq w} = O(nw^{2-\beta+\eta})$ for all $w \geq w_{\text{min}}$. 


(ii) \( W \geq w = \Omega(nw^{2-\beta-\eta}) \) for all \( w_{\text{min}} \leq w \leq \overline{w} \),

(iii) \( W \leq w = O(n) \) for all \( w \), and

(iv) \( W \leq w = \Omega(n) \) for all \( w = \omega(1) \).

**Proof.** Let \( w_1 \geq w_0 \geq 0 \) be two fixed weights. We start by summing up all vertex-weights between \( w_0 \) and \( w_1 \). By Fubini’s theorem (see [Hal74], e.g.), we can rewrite this sum as

\[
\sum_{v \in V, w_0 \leq w_v \leq w_1} w_v = \int_0^\infty |V_{\geq w_0} \setminus V_{> w_1}| dx
\]

(3.1)

We start with (i) and apply (3.1) with \( w_0 = w \) and \( w_1 = w_{\text{max}} \). Then, the set \( V_{> w_1} \) is empty, and we have

\[
W_{\geq w} = w \cdot |V_{\geq w}| + \int_{w}^{w_{\text{max}}} |V_{\geq x}| dx.
\]

Then the assumption (PL2) implies

\[
W_{\geq w} = |V_{\geq w}| \cdot w + \int_w^{\infty} |V_{\geq x}| dx = O\left(nw^{2-\beta+\eta} + \int_w^{\infty} nx^{1-\beta+\eta} dx\right) = O(nw^{2-\beta+\eta}).
\]

For (ii) we similarly obtain

\[
W_{\geq w} = \Omega\left(nw^{2-\beta-\eta} + \int_w^{w_{\text{max}}} nx^{1-\beta-\eta} dx\right) = \Omega(nw^{2-\beta-\eta}).
\]

For (iii), we see that if \( w < w_{\text{min}} \), then clearly \( W_{\leq w} = 0 \). Otherwise, Equation (3.1) with \( w_0 = w_{\text{min}} \) and \( w_1 = w \) implies

\[
W_{\leq w} = |V_{\geq w_{\text{min}}}| \cdot w_{\text{min}} + \int_{w_{\text{min}}}^{w} |V_{\geq x}| dx - |V_{> w}| \cdot w
\]

\[
\leq nw_{\text{min}} + O\left(\int_{w_{\text{min}}}^{w} nx^{1-\beta+\eta} dx\right) = O(n),
\]

and for (iv) we obtain

\[
W_{\leq w} \geq \int_{w_{\text{min}}}^{w} |V_{\geq x}| dx - |V_{> w}| \cdot w = \Omega\left(\int_{w_{\text{min}}}^{w} nx^{1-\beta-\eta} dx\right) - O(nw^{2-\beta+\eta})
\]

\[
= \Omega(n) - o(n) = \Omega(n).
\]

In particular, (iii) and (iv) yield \( W = \Theta(n) \) by choosing \( w = w_{\text{max}} \).
Next we consider the marginal edge probability $\Pr[u \sim v]$ of two vertices $u, v$ with weights $w_u, w_v$. For a fixed position $x_u \in \mathcal{X}$, we already know this probability by (EP1).

**Lemma 3.2.** Let $u \in [n]$ and let $x_u \in \mathcal{X}$ be any fixed position. Then in $\mathcal{G}(n, \mathcal{X}, w, p)$ all edges $\{u, v\}, u \neq v$, are independently present with probability

$$\Pr[u \sim v | x_u] = \Theta\left( \Pr[u \sim v] \right) = \Theta\left( \min\left\{1, \frac{w_u w_v}{W}\right\}\right).$$

**Proof.** Let $u, v \in [n]$. Then by (EP1), it follows directly

$$\Pr[u \sim v] = \mathbb{E}_{x_u} \left[ \Pr_{x_v}[u \sim v | x_u] \right] = \mathbb{E}_{x_u} \left[ \Theta\left( \min\left\{1, \frac{w_u w_v}{W}\right\}\right) \right] = \Theta\left( \min\left\{1, \frac{w_u w_v}{W}\right\}\right).$$

Furthermore, for every fixed $x_u \in \mathcal{X}$ the edges incident to $u$ are independently present with probability $\Pr_{x_v}[u \sim v | x_u], as the event “$u \sim v$” only depends on $x_v$, and an independent random choice for the edge $\{u, v\}$ (after fixing $x_u$). $\square$

The following lemma shows that the expected degree of a vertex is of the same order as the weight of the vertex, thus we can interpret a given weight sequence $w$ as a sequence of expected degrees.

**Lemma 3.3 (Expected degrees).** in $\mathcal{G}(n, \mathcal{X}, w, p)$, for any $v \in [n]$ we have

$$\mathbb{E}[\deg(v)] = \Theta(w_v).$$

**Proof.** Let $v$ be any vertex. We estimate the expected degree both from below and above. By Lemma 3.2, the expected degree of $v$ is at most

$$\sum_{u \neq v} \Pr[u \sim v] = \Theta\left( \sum_{u \neq v} \min\left\{1, \frac{w_u w_v}{W}\right\}\right) = O\left( \sum_{u \in V} \frac{w_u w_v}{W}\right) = O\left( \frac{W}{W} \sum_{u \in V} w_u\right) = O(w_v).$$

For the lower bound, $\Pr[u \sim v] = \Theta\left( \frac{w_u w_v}{W}\right)$ holds for all $w_u \leq \frac{W}{W_v}$. We set $w' := \frac{W}{W_v}$ and observe that $w' = \omega(1)$. Using Lemma 3.1, we obtain

$$\mathbb{E}[\deg(v)] \geq \sum_{u \neq v, u \in V_{\leq w'}} \Pr[u \sim v] = \Omega\left( \frac{W}{W} \sum_{u \leq w'} w\right) = \Omega(w_v).$$

$\square$
As the expected degree of a vertex is roughly the same as its weight, it is no surprise that w.h.p. the degrees of all vertices with sufficiently large weight are concentrated around the expected value. The following lemma gives a precise statement.

**Lemma 3.4.** The following properties hold w.h.p. in $G(n, X, w, p)$.

(i) $\deg(v) = O(w_v + \log^2 n)$ for all $v \in [n]$.

(ii) $\deg(v) = (1 + o(1))E[\deg(v)] = \Theta(w_v)$ for all $v \in V_{\geq \omega(\log^2 n)}$.

(iii) $\sum_{v \in V_{\geq w}} \deg(v) = \Theta(W_{\geq w})$ for all $w = \omega(\log^2 n)$.

**Proof.** Let $v \in V$ with fixed position $x_v \in X$ and let $\mu := E[\deg(v) | x_v] = \Theta(w_v)$. By definition of the model, conditioned on the position $x_v$ the degree of $v$ is a sum of independent Bernoulli random variables. By Lemma 3.3 there exists a constant $C$ such that $2e\mu < C \log^2 n$ holds for all vertices $v \in V_{\leq \log^2 n}$ and all positions $x_v \in X$. Thus, if $v \in V_{\leq \log^2 n}$, we apply a Chernoff bound (Theorem 2.2 (iii)), and obtain $\Pr[\deg(v) > C \log^2 n] \leq 2^{-C \log^2 n} = n^{-\omega(1)}$. If $v \in V_{\log^2 n}$, we similarly obtain $\Pr[\deg(v) > 3\mu/2] \leq e^{-\Theta(\mu)} = n^{-\omega(1)}$ and $\mu = \Theta(w_v)$ by Lemma 3.3. Then (i) follows by applying a union bound over all vertices.

For (ii), let $v \in V$ such that $w_v = \omega(\log^2 n)$, let $\mu$ be as defined above and put $\epsilon = \frac{\log n}{\sqrt{\mu}} = o(1)$. Thus by the Chernoff bound,

$$\Pr[|\deg(v) - \mu| > \epsilon \cdot \mu] \leq e^{-\Theta(\epsilon^2 \cdot \mu)} = n^{-\omega(1)},$$

and we obtain (ii) by applying Lemma 3.3 and a union bound over all such vertices. Finally, from (ii) we infer $\sum_{v \in V_{\geq w}} \deg(v) = \sum_{v \in V_{\leq w}} \Theta(w_v) = \Theta(W_{\geq w})$ for all $w = \omega(\log^2 n)$, which shows (iii). 

We conclude this section by proving that if we sample the weights randomly from an appropriate distribution, then almost surely the resulting weights satisfy our conditions on power-law weights.

**Lemma 3.5 (Weight sampling model).** Let $w_{\min}$ be a strictly positive constant, let $F = F_n : [0, 1] \rightarrow [0, 1]$ be non-decreasing such that $F(z) = 0$ for all $z \leq w_{\min}$, and
$F(z) = 1 - \Theta(z^{1-\beta})$ for all $z \geq w_{\text{min}}$. Suppose that for every vertex $v \in [n]$, we choose the weight $w_v$ independently according to the cumulative probability distribution $F$. Then with $\bar{w} = (n/\log^2 n)^{1/(\beta-1)}$, the resulting weight vector $w$ satisfies deterministically (PL1), with high probability the lower bound of (PL2), and for all $\eta = \eta(n) = \omega(\log \log n / \log n)$ with probability $1 - n^{-\Omega(\eta)}$ the upper bound of (PL2).

In particular, this lemma proves that for any fixed $\eta > 0$, with probability $1 - n^{-\Omega(1)}$ (PL1) and (PL2) are fulfilled for weights sampled according to $F(\cdot)$. Moreover, it follows that any property which holds with probability $1 - q$ for weights satisfying (PL1) and (PL2) also holds in a model of sampled weights with probability at least $1 - q - n^{-\Omega(1)}$. However, we claim without proof that all our main results hold with the original probability in a model of sampled weights.

**Proof of Lemma 3.5.** Condition (PL1) is fulfilled by definition of $F$, and we only need to prove (PL2). For all $z > w_{\text{min}}$, denote by $Y_z$ the number of vertices with weight at least $z$ and observe that

$$E[Y_z] = n(1 - F(z)) = \Theta(nz^{1-\beta}).$$

Let us first consider the case $z \in [w_{\text{min}}, \bar{w}]$. For all $z$ in this range we have $E[Y_z] = \Omega(\log^2 n)$, so for any $z \in [w_{\text{min}}, \bar{w}]$ the Chernoff bound (Theorem 2.2 (i) and (ii)) yields

$$\Pr[|Y_z - E[Y_z]| \geq 0.5E[Y_z]] \leq \exp(-\Omega(E[Y_z])) = n^{-\Omega(\log n)}.$$

Note that $Y_z$ is always an integer and at most $n$. Clearly, $E[Y_z]$ is decreasing. Hence, we can assume without loss of generality that either $z \in [w_{\text{min}}, \bar{w}]$ or $0.5E[Y_z]$ or $1.5E[Y_z]$ is an integer, because if $0.5E[Y_z] < Y_z < 1.5E[Y_z]$ holds for these values of $z$, then it must hold for all other $z \in [w_{\text{min}}, \bar{w}]$ as well. Thus, we can restrict $z$ to a set of size $O(n)$. This allows us to take a union bound, and it follows that with probability $1 - n^{-\omega(1)}$, $Y_z = \Theta(nz^{1-\beta})$ holds for all $z \in [w_{\text{min}}, \bar{w}]$. In this case, all $z$ in our range satisfy both the lower and upper bound of (PL2) even for $\eta = 0$. In particular, this proves that with probability $1 - n^{-\omega(1)}$, the lower bound of (PL2) holds for all $\eta \geq 0$. 

It only remains the upper bound of (PL2) for \( z \geq \bar{w} \). Let \( z \geq \bar{w} \) and choose \( \eta = \eta(n) = \omega(\log \log n/ \log n) \). By Markov’s inequality and (3.2),

\[
\Pr[Y \geq n z^{1-\beta+\eta}] \leq \Pr[Y \geq \Omega(z^\eta)\mathbb{E}[Y_z]] \leq O(z^{-\eta}) \leq n^{-\Omega(\eta)}.
\]

By the same argument as above, we can restrict \( z \) to \( \bar{w} \) and values where the intended bound \( \Omega(z^\eta)\mathbb{E}[Y_z] \) is integral, which happens only for \( O(\log^2 n) \) values of \( z \) above \( \bar{w} \). Hence, we can use the union bound to obtain error probability

\[
O(n^{-\Omega(\eta)} \log^2 n) = n^{-\Omega(\eta)},
\]

since \( \eta(n) = \omega(\log \log n/ \log n) \). In particular it also follows that with probability \( 1 - n^{-\Omega(\eta)} \), the maximum weight satisfies \( w_{\max} \leq n^{1/(\beta-1-\eta)} \).

### 3.3 First Examples

In this section we discuss a few first examples of the general model. First, we observe that Chung-Lu random graphs and Norros-Reittu random graphs are special cases of the model defined in Section 3.1. We then study a class which is still fairly general, the so-called distance model. In Section 3.4, we will see that geometric inhomogeneous random graphs are a special case of the distance model. Finally, we discuss a non-metric example of the distance model.

**Example 3.6** (Chung-Lu random graphs). We regain the Chung-Lu model as a special case by setting \( \mathcal{X} = \{x\} \) (the trivial ground space) and \( p_{uv} = \min\{1, \frac{w_u w_v}{\mathbb{W}}\} \).

(EP2) is trivially satisfied, and (EP1) is satisfied by the lower bound on \( \bar{w} \), since \( \mathbb{W} = \Theta(n) \) by Lemma 3.1. It was proven that a.a.s. this model has a unique giant connected component [CL02a]. Furthermore, in the range \( 2 < \beta < 3 \) the diameter is \( \Theta(\log n) \) and the average distance is \( \frac{2+o(1)}{|\log(\beta-2)|} \log \log n \) [CL02b; CL04].

**Example 3.7** (Norros-Reittu random graphs). Suppose every vertex \( v \in V \) draws a weight \( w_v \) independently at random according to a power-law distribution. Afterwards, each pair \( \{u, v\} \) is connected independently with \( E_{u,v} \) edges, where the random variable \( E_{u,v} \) is Poisson distributed with parameter \( \frac{w_u w_v}{\mathbb{W}} \). Norros and
Reittu proved that the obtained random graph has a.a.s. a giant connected component and that in the regime $2 < \beta < 3$, the average distance 
\[ \frac{2+o(1)}{\log(\beta-2)} \log \log n \] 
[NR06]. We claim that when considering every multiedge as a single edge, this model is a.a.s. included in our general framework. Indeed, we can take again the trivial geometry. By Lemma 3.5 the random weights satisfy (PL1) and (PL2) a.a.s. Furthermore, for any pair \( \{u, v\} \) we have

\[
\Pr[E_{u,v} > 0 | w_u, w_v] = 1 - \exp\left(- \frac{w_u w_v}{W}\right) = \Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right).
\]

**The distance model:** We now consider the following situation, which will cover both geometric inhomogeneous random graphs and the non-metric example. As our underlying geometry we specify the ground space \( X = [0,1]^d \), where \( d \geq 1 \) is a (constant) parameter of the model. We sample from this set according to the standard (Lebesgue) measure. This is in the spirit of the classical random geometric graphs [Pen03].

To describe the distance of two points \( x, y \in X \), assume we have some measurable function \( \|\cdot\|: [-1/2,1/2)^d \rightarrow \mathbb{R}_{\geq 0} \) such that \( \|0\| = 0 \) and \( \|-x\| = \|x\| \) for all \( x \in [-1/2,1/2)^d \). Note that \( \|\cdot\| \) does not need to be a norm or seminorm. We extend \( \|\cdot\| \) to \( \mathbb{R}^d \) via \( \|z\| := \|z - u\| \), where \( u \in \mathbb{Z}^d \) is the unique lattice point such that \( z - u \in [-1/2,1/2)^d \). For \( r \geq 0 \) and \( x \in X \), we define the \( r \)-ball around \( x \) to be \( B_r(x) := \{x \in X | \|x - y\| \leq r\} \), and we denote by \( V(r) \) the volume of the \( r \)-ball around 0. Intuitively, \( B_r(x) \) is the ball around \( x \) in \([0,1]^d\) with the torus geometry, i.e., with 0 and 1 identified in each coordinate. Assume that \( V : \mathbb{R}_{\geq 0} \rightarrow [0,1] \) is surjective, i.e., for each \( V_0 \in [0,1] \) there exists \( r \) such that \( V(r) = V_0 \). Finally, let \( \alpha \in \mathbb{R}_{>0} \) be a long-range parameter. Since the case \( \alpha = 1 \) deviates slightly from the general case, we assume \( \alpha \neq 1 \).

**Theorem 3.8.** Let \( p \) be any edge probability function such that for all \( u, v \) and all positions \( x_u, x_v \in X = [0,1]^d \),

\[
(3.4) \quad p_{uv}(x_u, x_v) = \Theta\left(\min\left\{1, V(\|x_u - x_v\|)^{-\alpha} \cdot \left(\frac{w_u w_v}{W}\right)^{\max\{\alpha, 1\}}\right\}\right).
\]

Then conditions (EP1) and (EP2) are satisfied, and we obtain an instance of the general model \( G(n, X, w, p) \).
Proof. Fix $u, v$, and also $x_u$. Note that $V(r)$ is the cumulative probability distribution $\Pr_{x_u}(\|x_u - x_v\| \leq r)$. The marginal edge probability is given by the Riemann-Stieltjes integral over $r$ (see, e.g., [CB00]),

$$E := \mathbb{E}_{x_v}[p_{uv}(x_u, x_v) \mid x_u] = \Theta\left(\int_0^\infty \Lambda_{u,v}(r)dV(r)\right),$$

where

$$\Lambda_{u,v}(r) := \min\left\{1, V(r)^{-\alpha} \cdot \left(\frac{w_u w_v}{W}\right)^{\max\{a, 1\}}\right\}.$$ 

In particular, for every sequence of partitions $r^{(t)} = \{0 = r_0^{(t)} < \ldots < r_{\ell(t)}^{(t)}\}$ with meshes tending to zero, the upper Darboux sum with respect to $r^{(t)}$ converges to the expectation,

$$E = \Theta\left(\lim_{t \to \infty} \sum_{s=1}^{\ell(t)} \sup_{r_{s-1}^{(t)} \leq r_s^{(t)} \leq r_{s}^{(t)}} \Lambda_{u,v}(r) \left(V(r_s^{(t)}) - V(r_{s-1}^{(t)})\right)\right).$$

Since $V$ is surjective, we may refine the meshes $r^{(t)}$ if necessary such that the meshes of the partitions $V^{(t)} = \{V(r_0^{(t)}), \ldots, V(r_{\ell(t)}^{(t)})\} = : \{V_0^{(t)}, \ldots, V_{\ell(t)}^{(t)}\}$ also tend to zero. Hence,

$$E = \Theta\left(\lim_{t \to \infty} \sum_{s=1}^{\ell(t)} \Lambda_{u,v}(r_s^{(t)}) \left(V(r_s^{(t)}) - V(r_{s-1}^{(t)})\right)\right)$$

$$= \Theta\left(\lim_{t \to \infty} \sum_{s=1}^{\ell(t)} \min\left\{1, (V_s^{(t)})^{-\alpha} \cdot \left(\frac{w_u w_v}{W}\right)^{\max\{a, 1\}}\right\} \left(V_s^{(t)} - V_{s-1}^{(t)}\right)\right)$$

$$= \Theta\left(\int_0^1 \min\left\{1, V^{-\alpha} \cdot \left(\frac{w_u w_v}{W}\right)^{\max\{a, 1\}}\right\} dV\right),$$

where the latter integral is an ordinary Riemann integral. If $w_u w_v / W \geq 1$, the integrand is 1 and we obtain $E = \Theta(1) = \Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right)$. On the other hand, if $w_u w_v / W < 1$ then let $r_0 := \left(\frac{w_u w_v}{W}\right)^{\max\{a, 1\}/\alpha} < 1$. Note that if $r_0 = \Theta(1)$, then also $r_0 = \Theta(w_u w_v / W)$. Therefore,

$$E = \Theta\left(\int_0^{r_0} 1 dV + \left(\frac{w_u w_v}{W}\right)^{\max\{a, 1\}} \int_{r_0}^1 V^{-\alpha} dV\right)$$

$$= \begin{cases} \Theta\left(r_0 + \frac{w_u w_v}{W} (1 - r_0^{1-\alpha})\right) = \Theta\left(\frac{w_u w_v}{W}\right), & \text{if } \alpha < 1, \text{ and} \\ \Theta\left(r_0 + \left(\frac{w_u w_v}{W}\right)^{\alpha} (r_0^{1-\alpha} - 1)\right) = \Theta\left(\frac{w_u w_v}{W}\right), & \text{if } \alpha > 1, \end{cases}$$
as required.

It remains to show that $p$ satisfies (EP2), i.e., that for all vertices $u, v$ with $w_u, w_v \geq \bar{w}$, all $x_u, x_v \in \mathcal{X}$, and all $\eta > 0$ it holds $p_{uv} \geq (\frac{n}{\bar{w}^{\beta-1-\eta}})^{-1+\omega(1/\log\log n)}$. Since $V(\|x_u - x_v\|) \leq 1$, we may use Equation (3.4) to bound

$$p_{uv} \geq \Omega\left(\min\left\{1, \left(\frac{w_u w_v}{\bar{W}}\right)^{\max\{\alpha, 1\}}\right\}\right).$$

If $w_u w_v / \bar{W} \geq 1$ then there is nothing to show (since the right hand side of (EP2) is $o(1)$ by the upper bound on $\bar{w}$). Otherwise, if $w_u w_v / \bar{W} < 1$, then

$$p_{uv} \geq \Omega\left(\left(\frac{w_u w_v}{\bar{W}}\right)^{\max\{\alpha, 1\}}\right) \geq \Omega\left(\frac{\bar{w}^2}{n}\right) \geq \left(\frac{n}{\bar{w}^{\beta-1+\eta}}\right)^{-1+\omega(1/\log\log n)},$$

where the last step follows from the lower bound $n^{\omega(1/\log\log n)} \leq \bar{w}$. This concludes the proof. □

**The minimum component distance model:** We consider a distance measure which is particularly useful to model social networks: assume that two individuals share one feature (e.g., they are in the same sports club), but are very different in many other features (work, music, ...). Then they are still likely to know each other, which is captured by the minimum component distance.

**Example 3.9.** Let the **minimum component distance** be defined by

$$\|x\|_{\min} := \min\{x_i \mid 1 \leq i \leq d\}$$

for $x = (x_1, \ldots, x_d) \in [-1/2, 1/2]^d$.

Note that the minimum component distance is not a metric for $d \geq 2$, since there are $x, y, z \in \mathcal{X}$ such that $x$ and $y$ are close in one component, $y$ and $z$ are close in one (different) component, but $x$ and $z$ are not close in any component. Thus the triangle inequality is not satisfied. However, $\|\| \,$ still satisfies the requirements specified above, so we get a non-metric example of our model.

The minimum component distance model has been studied in more details in [LT17]. It has been proven that the clustering coefficient is still $\Omega(1)$, but in contrast to geometric inhomogeneous random graphs, a.a.s. the removal of any sublinear number of edges only decreases the size of the largest component by a sublinear number of vertices.
3.4 Example: Geometric Inhomogeneous Random Graphs

The Standard GIRG model: We consider again the distance model as introduced above. Then, we obtain geometric inhomogeneous random graphs (GIRGs) by choosing \( \| \cdot \| \) to be any norm on \( \mathbb{R}^d \), for instance \( \| \cdot \|_2 \) or \( \| \cdot \|_\infty \). A similar model was presented in [SKB08] in the physics community.

Example 3.10. Let \( \alpha > 1 \) and let \( \| \cdot \| \) be the Euclidean distance \( \| \cdot \|_2 \) or any equivalent norm such as \( \| \cdot \|_\infty \). If we choose \( p \) to be any edge probability function that satisfies (3.4) for all \( u, v \) and \( x_u, x_v \in \mathcal{X} \), we obtain the so-called GIRG model (or Standard GIRG model) as a special case of the distance model.

Note that for any norm \( \| \cdot \| \) on \( \mathbb{R}^d \) and any \( r \leq 1 \) we have \( V(r) = \Theta(r^d) \). Hence, GIRGs are defined via (1) a weight sequence \( w \) that follows a power law, (2) the \( d \)-dimensional torus \( \mathbb{T}^d \) equipped with any norm as geometric ground space, and (3) the edge probability condition

\[
(3.5) \quad p_{uv} = \Theta\left(\min\left\{ 1, \frac{1}{\| x_u - x_v \|^{\alpha d}} \left( \frac{w_u w_v}{W} \right)^{\alpha d} \right\} \right).
\]

Indeed, the exact choice of the norm is irrelevant as we hide constant factors in the Landau notation and it is well-known that all norms on \( \mathbb{R}^d \) are equivalent.

For a complex network model with underlying geometry, it is reasonable to require that adjacent vertices are likely to have small distance, therefore \( p_{uv} \) should decrease with increasing distance \( \| x_u - x_v \| \). The constraint \( \alpha > 1 \) is necessary to cancel the growth of the volume of the ball of radius \( r \) proportional to \( r^d \), so that we expect most neighbors of a vertex to lie close to it. GIRGs are technically easy as the weight of every vertex coincides with its expected degree (Lemma 3.3) and because for any vertex \( u \) with fixed position \( x_u \), the incident edges are independent (Lemma 3.2).

Although we defined GIRGs as a “special case”, it is still a rather general model. The free parameters of the model are the dimension \( d \in \mathbb{N} \), the long-range parameter \( \alpha > 1 \), the concrete weights \( w \) with power-law exponent \( \beta > 2 \) and average weight \( W/n \), and the concrete function \( f_{uv}(x_u, x_v) \) replacing the \( \Theta \) in \( p_{uv} \). For example, when the minimum in expression (3.5) for \( p_{uv} \) is obtained
by the first term, we can require that the edge \( \{u, v\} \) is present \emph{deterministically}, but we are also allowed to include such an edge with probability, say, \( \frac{1}{2} \). We emphasize that when embedding a real-world network into the GIRG model, the real world coordinates do not necessarily resemble the positions in the geometric space since additional spatial information might be included, e.g., hobbies in social networks [LN+05].

As choice for the ground space, we prefer the torus \( \mathbb{T}^d \) to the hyper-cube for technical simplicity as it yields symmetry. However, one could easily replace \( \mathbb{T}^d \) by \([0, 1]^d\). If we set \( \mathcal{X} := [0, 1]^d \), then for any \( r \geq 0 \) and \( x \in \mathcal{X} \) the volume of the ball \( B_r(x) \) decreases at most by a constant factor \( 2^d \) compared to the torus geometry. Thus all structural results hold verbatim for both ground spaces. One advantage of the torus geometry is that it allows us to prove that hyperbolic random graphs are a special case of GIRGs, as we will see below in Section 3.5.

Figures 3.2-3.5 on pages 50-53 illustrate four instances of GIRGs on 3000 vertices with underlying ground space \([0,1]^2\). With Figure 3.2 and Figure 3.3 we first demonstrate the effect of the power-law parameter \( \beta \). Figure 3.4 and Figure 3.5 then indicate the influence of the long-range parameter \( \alpha \).

**The Threshold model:** Finally, we discuss a variation of Example 3.10 where we consider again the distance model, but let \( \alpha \to \infty \) and thus obtain a threshold function for \( p \).

**Example 3.11.** Let \( \| \cdot \| \) be the Euclidean distance \( \| \cdot \|_2 \) or any equivalent norm such as \( \| \cdot \|_\infty \). Let \( p \) again satisfy (3.4), but this time we assume that \( \alpha = \infty \). More precisely, we require

\[
p_{uv}(x_u, x_v) = \begin{cases} 
\Theta(1), & \text{if } \|x_u - x_v\| \leq O\left(\frac{w_u w_v}{W}\right)^{1/d}, \text{ and} \\
0, & \text{if } \|x_u - x_v\| \geq \Omega\left(\frac{w_u w_v}{W}\right)^{1/d},
\end{cases}
\]

where the constants hidden by \( O \) and \( \Omega \) do not have to match, i.e., there can be an interval \([c\left(\frac{w_u w_v}{W}\right)^{1/d}, C\left(\frac{w_u w_v}{W}\right)^{1/d}]\) for \( \|x_u - x_v\| \) where the behavior of \( p_{uv}(x_u, x_v) \) is arbitrary.
Notice that the volume of a ball with radius \( r_0 = \Theta((\frac{w_u w_v}{W})^{1/d}) \) around any fixed \( x \in \mathcal{X} \) is \( \Theta(\min\{1, \frac{w_u w_v}{W}\}) \). Thus, by (3.6), for fixed \( x_u \) it follows directly that

\[
\mathbb{E}_{x_v}[p_{uv}(x_u, x_v) \mid x_u] = \Theta\left( \Pr_{x_v}[\|x_u - x_v\| \leq r \mid x_u]\right) = \Theta\left( \min\left\{1, \frac{w_u w_v}{W}\right\}\right).
\]

Hence, (EP1) is satisfied, which will be sufficient for most structural and algorithmic results that we are going to prove. In order to also fulfill (EP2), we put \( w = \omega\left(\frac{n}{\sqrt{d}}\right) \) and require that \( 2 < \beta < 3 \), which is fine as (EP2) is only needed for theorems that have this precondition too. Our choice of \( w \) ensures that for all \( w_u, w_v \geq w \) we have \( \frac{w_u w_v}{W} = \omega(1) \). For all positions \( x_u, x_v \in \mathcal{X} \) we thus obtain \( p_{uv}(x_u, x_v) = \Theta(1) \) by (3.6).

**SubGIRGs:** Let \( \mathcal{X} \in \{\mathbb{T}^d, [0,1]^d\} \) and let \( G = (V,E) \) be an instance of a GIRG on ground space \( \mathcal{X} \). We consider the subgraph of \( G \) induced by a small subspace \( A \in \mathcal{X} \). As the vertices pick their positions independently and uniformly at random, the vertex set \( V[A] \) of \( G[A] \) is a random subset of \( V \) whose size is proportional to \( |V| \cdot \text{VOL}(A) \), where \( \text{VOL}(A) \) denotes the geometric volume of \( A \), and we can expect that the weights of the vertex set \( V[A] \) follow a power law. It turns out that the edge probabilities scale properly, and thus \( G[A] \) itself is a GIRG. For simplicity, we only study the case where \( A \) is a sub-cuboid of \( \mathbb{T}^d \) or \([0,1]^d\).

**Lemma 3.12.** Consider a GIRG on ground space \( \mathcal{X} \in \{\mathbb{T}^d, [0,1]^d\} \) where the weights are sampled from the power-law distribution given by

\[
f(w) = \begin{cases} (\beta - 1) w_{\min}^{\beta - 1} w^{-\beta}, & \text{if } w \geq w_{\min}, \\ 0, & \text{else.} \end{cases}
\]

Let \( r \leq 1 \) such that \( r^d = \omega(n^{-1}) \) and let \( A \) be a sub-cuboid of \( \mathcal{X} \) where each side has length \( \Theta(r) \). Then a.a.s., the induced subgraph \( G[A] \) is itself a GIRG, with ground space \([0,1]^d\) after appropriate re-scaling.

**Proof.** Let \( n' := \text{VOL}(A) \cdot n \) and observe that \( n' = n'(n) = \omega(1) \) as \( \text{VOL}(A) = \Theta(r^d) \). Since weights and coordinates of GIRGs are sampled independently, we can assume that in the sampling process of the original GIRG on \( \mathcal{X} \in \{\mathbb{T}^d, [0,1]^d\} \),
Figure 3.2. Instance of a GIRG on ground space $[0, 1]^2$ with parameters $n = 3000$, $\alpha = 8$, $\beta = 2.8$, and $w_{\text{min}} = 3$. The graph has clustering coefficient 0.12, average degree 7.13 and maximum degree 741, the giant component contains 2891 vertices, and within the giant the average distance is 3.90 while the diameter is 14. The graph possesses few vertices of high degree because we use a rather large $\beta$, stressing local community structure.
Figure 3.3. Instance of a GIRG on ground space $[0, 1]^2$ with parameters $n = 3000$, $\alpha = 8$, $\beta = 2.3$, and $w_{\min} = 3$. The graph has clustering coefficient 0.09, average degree 8.79 and maximum degree 921, the giant component contains 2903 vertices, and within the giant the average distance is 3.15 while the diameter is 11. As the parameter $\beta$ is smaller than in Figure 3.2, the number of high degree nodes is considerably larger.
Figure 3.4. Instance of a GIRG on ground space $[0, 1]^2$ with parameters $n = 3000$, $\alpha = 2$, $\beta = 2.5$, and $w_{\text{min}} = 3$. The graph has clustering coefficient 0.13, average degree 9.25 and maximum degree 564, the giant component contains 2970 vertices, and within the giant the average distance is 3.51 while the diameter is 10. Due to the small value of $\alpha$, many long-range edges are present, linking vertices that are far away from each other.
Figure 3.5. Instance of a GIRG on ground space $[0, 1]^2$ with parameters $n = 3000$, $\alpha = 40$, $\beta = 2.5$, and $w_{\text{min}} = 3$. The graph has clustering coefficient 0.13, average degree 7.82 and maximum degree 623, the giant component contains 2899 vertices, and within the giant the average distance is 3.76 while the diameter is 16. Compared to Figure 3.4, the larger value of $\alpha$ ensures that long-range edges appear almost exclusively between vertices of high degree.
the vertices first pick their random positions, where \( \Pr[x_v \in A] = \text{vol}(A) \) holds for each vertex \( v \in V \). Let \( V' \) be the set of vertices that fall into \( A \). We have \( \mathbb{E}[|V'|] = n' \), and by a Chernoff bound (Theorem 2.2), \( 0.5n' \leq |V'| \leq 1.5n' \) a.a.s. Considering the weights of the vertices in \( V' \), by Lemma 3.5 a.a.s. the weight sequence induced by \( V' \) follows a power law.

It remains to study the edge probabilities. We stretch \( A \) by a factor \( \Theta(r^{-1}) \) along each of the \( d \) coordinates and transform it into \( A' = [0, 1]^d \). Assume first \( \alpha < \infty \). For two vertices \( u, v \in V' \), by (3.5) the edge probability is given by

\[
p_{uv} = \Theta\left( \min \left\{ 1, \left( \frac{W_u W_v}{\|x_u - x_v\|_{A'}^d} \right)^{1/\alpha} \right\} \right),
\]

where \( W \) denotes the total weight of the original GIRG. Note that by Lemma 3.1 we have \( W = \Theta(n) \) and \( W' = \Theta(n') \), where \( W' \) denotes the total weight of the vertices in \( V' \). Hence,

\[
\|x_u - x_v\|_{A'}^d \cdot W = \Theta\left( \|x_u - x_v\|_{A'}^d r^{-d} n \right) = \Theta\left( \|x_u - x_v\|_{A'}^d \cdot n' \right).
\]

Therefore, after the re-scaling \( p_{uv} \) is given by (3.5) for the distances in \( A' \) and a total number of \( \Theta(n') \) vertices, proving that indeed, a.a.s. \( G[A] \) is itself a GIRG.

Finally, for the threshold case \( \alpha = \infty \) the argument can be applied analogously.

\[\blacksquare\]

**Outlook:** As GIRGs are an example of our general class of scale-free random graphs, all results that we obtain for \( G(n, X, w, p) \) also hold for GIRGs: average distance \( \frac{2+o(1)}{\log(n - 2)} \log \log n \) in the regime \( 2 < \beta < 3 \), power-law degree sequence, a giant connected component, and polylogarithmic diameter.

Moreover, in Chapter 4 we prove additional structural and algorithmic results that fundamentally depend on an underlying geometry, and thus do not hold in the general model. We will see that in the regime \( \beta > 2 \), both the Standard and Threshold GIRG can be sampled in expected linear time, have w.h.p. clustering coefficient \( \Theta(1) \), and can be compressed in expected linear time. Further, in the regime \( 2 < \beta < 3 \) a.a.s. they have small separating sets of sublinear size.

In Chapter 5 we analyze greedy routing in GIRGs, again on both variants. But for technical reasons, we slightly modify the definition of the random graph model, e.g., we will use a Poisson point process.
Koch and Lengler studied bootstrap percolation on GIRGs [KL16]. In particular, it was demonstrated that the speed of infection in the process crucially depends on the underlying geometry. Further properties on so-called explosion processes have been conjectured in [HK17] by Komjáthy and van der Hofstad. In a very recent paper, Komjáthy and Lodewijks [KL18] indeed solved most of these problems. Thereby, they also presented further variations of GIRGs. For instance, the edge probability condition (3.5) was generalized by replacing the first argument in the minimum. Concretely, it was only required

\[ p_{uv} = \Theta \left( \min \left\{ g(u, v), \frac{1}{\|x_u - x_v\|^\alpha d(w_u, w_v)^\alpha} \right\} \right), \]

where \( g(u, v) \leq 1, g(u, v) = \Omega(\exp(-\max\{w_u, w_v\}^\gamma)), \) and \( 0 < \gamma < 1 \) is an additional parameter of the model.

### 3.5 Example: Hyperbolic Random Graphs

The goal of this section is to prove that the celebrated model of hyperbolic random graphs is a special case of one-dimensional GIRGs. More precisely, we obtain hyperbolic random graphs from GIRGs by setting the weights to a specific power law and the \( \Theta \) in the edge probability function \( p_{uv} \) to a specific, complicated function. A priori, it is not obvious that there actually exists such an embedding between hyperbolic random graphs, based on hyperbolic geometry, and GIRGs, based on Euclidean geometry, exists. Our proof is grounded on Taylor expansions and the appropriate application of preexisting technical lemmas on hyperbolic random graphs.

With this result we gain an additional, non-obvious example that falls into the general class of scale-free random graphs. In particular, all results on GIRGs hold for hyperbolic random graphs, too. Moreover, as our proofs are much less technical than many proofs for hyperbolic random graphs, we think that it is considerable to switch from hyperbolic random graphs to GIRGs in future theoretical studies.

As already discussed in Section 1.1, the model of hyperbolic random graphs has been proposed 2010 in the physics community by Papadopoulos, Krioukov,
et al. [Pap+10; Kri+10]. Since then, it gained a lot of attention in both experimental and theoretical studies. Here, we give a formal definition, summarize recent theoretical results on hyperbolic random graphs, and present our embedding.

There exist several different representations of hyperbolic geometry, all with advantages and disadvantages. For the random graph model it is convenient to use the native representation. It can be described by a disk $H$ of radius $R$ around the origin $0$, where the position of every point $x$ is given by its polar coordinates $(r_x, \nu_x)$. The model is isotropic around the origin. The hyperbolic distance between two points $x$ and $y$ is given by the non-negative solution $d_H = d_H(x, y)$ of the equation

$$\cosh(d_H) = \cosh(r_x) \cosh(r_y) - \sinh(r_x) \sinh(r_y) \cos(\nu_x - \nu_y).$$

In the following definition, we follow the notation introduced by Gugelmann et al. [GPP12].

**Definition 3.13.** Let $\alpha_H > 0, C_H \in \mathbb{R}, T_H > 0, n \in \mathbb{N}$, and set $R = 2 \log n + C_H$. Then the random hyperbolic graph $\mathcal{G}(n, \alpha_H, C_H, T_H)$ is a graph on $n$ vertices with the following properties:

- Every vertex $v \in V$ independently draws random coordinates $(r_v, \nu_v)$. The angle $\pi_v$ is chosen uniformly at random in $[0, 2\pi)$ and the radius $r_v \in [0, R]$ is chosen at random with density $f(r) := \frac{\alpha_H \sinh(\alpha_H r)}{\cosh(\alpha_H R) - 1}$.

- Every potential edge $e = \{u, v\}$, $u, v \in V$, is independently present with probability

  $$p_H(d_H(u, v)) = \left(1 + e^{\frac{1}{2 T_H} (d_H(u, v) - R)}\right)^{-1}.$$

In the limit $T_H \to 0$, we obtain the threshold hyperbolic random graph $\mathcal{G}(n, \alpha_H, C_H)$, where every edge $e = \{u, v\}$ is present if and only if $d_H(u, v) \leq R$.

Unfortunately, it is quite technical to analyze the general model. This is why most theory papers focus on the threshold model where $T_H$ is set to zero. The rigorous study of hyperbolic random graphs was initiated by Gugelmann, Panagiotou, and Peter [GPP12]. They veryfied a power-law degree sequence of
the threshold model with exponent $2\alpha_H + 1$ and that the clustering coefficient is constant. In [CF16b], Candellero and Fountoulakis calculated the clustering coefficient precisely for both models. Furthermore, Bode, Fountoulakis, and Müller determined the degree distribution of arbitrary vertices [BFM15a].

Regarding the component structure, the same authors demonstrated that in the regime $\frac{1}{2} < \alpha_H < 1$ there exists a giant connected component of linear size, and vice-versa, there exists no giant if $\alpha_H > 1$ [BFM15b]. After a sequence of papers, we know that in the regime $\frac{1}{2} < \alpha_H < 1$, the second largest component has size $\Theta((\log n)^{1/(1-\alpha_H)})$ [BFM16; KM15; KM17]. For the largest component, Friedrich and Krohmer showed that the diameter is $\Omega(\log n)$ [FK15b] and very recently, Müller and Staps proved a matching upper bound [TS17]. Finally, Abdullah, Fountoulakis, and Fountoulakis demonstrated that the average distance is $\frac{2+o(1)}{|\log(2\alpha_H-1)|} \log \log n$ [ABM17].

Bläsius, Friedrich, and Krohmer (2016) analyzed the separator hierarchy and the treewidth for both models of hyperbolic random graphs [BFK16]. Furthermore, they demonstrated how their results yield a polynomial-time approximation scheme for finding the largest independent set and a polynomial-time algorithm for finding maximum matchings. In the same year, Kiwi and Mitsche analyzed the spectral gap and the conductance of hyperbolic random graphs [KM16]. Further results on the structure of hyperbolic random graphs include the number of $k$-cliques [FK15a].

Moreover, Candellero and Fountoulakis studied bootstrap percolation and determined the threshold for the initial infection such that infection spreads over a linear number of vertices [CF16a]. Von Looz et al. gave a sampling algorithm of expected time $O(n^{3/2})$ [LMP15]. Finally, Bläsius, Friedrich, Krohmer, and Laue provided a quasi-linear embedding algorithm [Bl16] with good performance results.

After this short tour on recent results on hyperbolic random graphs, we come back to our original goal and show that an instance of hyperbolic random graphs is a.a.s. an instance of 1-dimensional GIRGs. Thus, we reduce the geometry of the hyperbolic disk to the geometry of a circle, but gain additional freedom as we can choose the weights of vertices. Notice that a single point on the hyperbolic
disk has measure zero, so we can assume that no vertex has radius $r_v = 0$. For the parameters, we put

$$d := 1, \quad \beta := 2\alpha_H + 1, \quad \alpha := 1/T_H, \quad w_{\text{min}} := e^{-C_H/2}.$$ 

In the threshold case, we consequently map $T_H = 0$ to $\alpha = \infty$. In Chapter 5, when considering greedy routing in hyperbolic random graphs, we also study how the results depend on $w_{\text{min}}$. This is why we don’t hide the constants $w_{\text{min}}$ respectively $C_H$ in the Landau symbols but make the mapping explicit. Next, we define the mapping

$$w_v := ne^{-r_v/2} \quad \text{and} \quad x_v := \frac{\nu_v}{2\pi}.$$ 

Since this is a bijection between $H \setminus \{0\}$ and $[1, e^{R/2}) \times \mathbb{T}^1$, there exists as well an inverse function $g(w_u, x_u) = (r_u, \nu_u)$. Finally for any two vertices $u \neq v$ on the torus, we set

$$p_{uv} := p_H(d_H(g(w_u, x_u), g(w_v, x_v))).$$

This finishes our embedding. The following lemma, which we prove later, demonstrates that under this mapping almost surely the weights follows a power law.

**Lemma 3.14.** Let $\alpha_H > \frac{1}{2}$. Then for all $\eta = \eta(n) = \omega\left(\frac{\log \log n}{\log n}\right)$, with probability $1 - n^{-\Omega(\min\{\eta, 1\})}$ the induced weight sequence $w$ satisfies (PL1) and (PL2) and thus follows a power law with parameter $\beta = 2\alpha_H + 1$.

Now we come to the main statement of this section. In the following we assume that if we sample an instance of the hyperbolic random graph model, we first sample the radii, next the angles, and at last the edges.

**Theorem 3.15.** Let $\alpha_H > \frac{1}{2}, n \in \mathbb{N}$ and fix a set of radii $(r_1, \ldots, r_n) \in [0, R]^n$, inducing a power-law weight sequence $w$ with parameter $\beta = 2\alpha_H + 1$. Then the random positions $x_u$ and the edge probabilities $p_{uv}(x_u, x_v)$ induced by our mapping satisfy the properties of the GIRG model. I.e., for fixed radii inducing power-law weights, hyperbolic random graphs are a special case of GIRGs.
Note that the precondition of Theorem 3.15 on the weight sequence \( w \) holds for any \( \eta = \eta(n) = \omega(\log \log n / \log n) \) with probability \( 1 - n^{-\Omega(\min\{\eta, 1\})} \) by Lemma 3.14. Therefore an instance of random hyperbolic graphs is almost surely included in our GIRG model with parameters as set above. In particular, any property that holds with probability \( 1 - q \) for GIRGs also holds for hyperbolic random graphs with probability at least, say, \( 1 - q - n^{-o(1)} \).

Before proving Lemma 3.14 and Theorem 3.15, we consider the following basic property of hyperbolic random graphs.

**Lemma 3.16.** Let \( \alpha_H > \frac{1}{2} \). Then with probability \( 1 - n^{-\Omega(1)} \) every vertex has radius at least \( r_0 := (1 - \frac{1}{2 \alpha_H}) \log n \). Furthermore, for all \( r = \omega(1), r \leq R \) and \( v \in V \), we have

\[
\Pr[r_v \leq r] = e^{-\alpha_H(r-R)}(1 + o(1)).
\]

**Proof.** Let \( v \in V \). By the given density \( f \) it follows immediately that

\[
\Pr[r_v \leq r] = \int_0^r f(x) \, dx = \alpha_H \int_0^r \frac{\sinh(\alpha_H x)}{\cosh(\alpha_H R) - 1} \, dx = \frac{\cosh(\alpha_H r) - 1}{\cosh(\alpha_H R) - 1},
\]

where we used \( \cosh(x) = \frac{e^x + e^{-x}}{2} = \frac{e^x}{2} (1 + o(1)) \) whenever \( x = \omega(1) \). Now let \( X_{r_0} \) be the random variable counting the vertices of radius at most \( r_0 \). We observe that the above expression for \( \Pr[r_v \leq r] \) implies

\[
\mathbb{E}[X_{r_0}] = n e^{-\alpha_H(R-r_0)}(1 + o(1)) = e^{\alpha_HC_H n^{1/2 - \alpha_H}} (1 + o(1)) = n^{-\Omega(1)}.
\]

By Markov’s inequality (Theorem 2.1), with probability \( 1 - n^{-\Omega(1)} \) we have \( X_{r_0} = 0 \).

**Proof of Lemma 3.14.** For every vertex of the random hyperbolic graph, the radius is chosen independently and uniformly according to \( f(r) \). Hence under our mapping, we sample the weights independently. We now prove that this sampling procedure fulfills the prerequisites of Lemma 3.5. Let \( \varepsilon > 0 \) be sufficiently small. By Lemma 3.16, the probability that a vertex \( v \) has radius at most \( r \geq \varepsilon \log n \) is \( e^{-\alpha_H(R-r)}(1 + o(1)) \). Let \( w_{\min} \leq z \leq o(n^{1-\varepsilon/2}) \). Then
\[ R - C_H - 2 \log z \geq \epsilon \log n, \] 

and

\[ F(z) := \Pr[w_v \leq z] = 1 - \Pr[r_v \leq R - C_H - 2 \log z] = 1 - e^{-2\alpha_H \log z - \alpha_H C_H} (1 + o(1)) \]

\[ = 1 - z^{-2\alpha_H w_{\min}^2} (1 + o(1)) = 1 - z^{1-\beta} w_{\min}^{\beta-1} (1 + o(1)). \]

Furthermore, for \( z < w_{\min} \) we get

\[ F(z) := \Pr[w_v \leq z] = \Pr[r_v \geq 2 \log n - 2 \log z] = \Pr[r_v \geq R] = 0. \]

By Definition 3.13, \( F(\cdot) \) is non-decreasing and therefore satisfies the preconditions of Lemma 3.5 together with our choice of \( w_{\min} \), at least in the range \( z \leq o(n^{1-\epsilon/2}) \) which suffices to apply the lemma for \( \eta = O(1) \). It follows that with sufficiently high probability, the weight sequence \( w \) follows a power law with parameter \( \beta \).

\[ \square \]

**Proof of Theorem 3.15.** Let us start by considering the sampling process of a random hyperbolic graph. First we sample the radii of the vertices, for which the precondition of the theorem assumes that they induce a power-law weight sequence under our mapping. Next we sample the angles. This corresponds to coordinates chosen independently and uniformly at random on \( \mathbb{T}^1 \). It remains to prove that \( p_{uv} \) as defined above satisfies conditions (3.5) and (3.6).

Let \( u \neq v \) be two vertices of the random hyperbolic graph with coordinates \( (r_u, \nu_u) \) and \( (r_v, \nu_v) \) and consider their mappings \( (w_u, x_u) \) and \( (w_v, x_v) \). Since the hyperbolic model is isotropic around the origin, we can assume without loss of generality that \( r_u \geq r_v \), \( \nu_v = 0 \) and \( \nu_u \leq \pi \).

Let us first consider the threshold model, corresponding to the case \( \alpha = \infty \) and \( T_H = 0 \). We claim that there exist constants \( M > m > 0 \) such that whenever \( \|x_u - x_v\| \geq M \frac{w_u w_v}{W} \), then \( p_{uv} = 0 \), and whenever \( \|x_u - x_v\| \leq m \frac{w_u w_v}{W} \), we have \( p_{uv} = 1 \). This will imply (3.6), as we set \( d = 1 \). Recall that in the threshold model, two vertices \( u \) and \( v \) are connected if and only if \( d_H(u, v) \leq R \). When \( r_u + r_v \leq R \), this is the case for all angles \( \nu_u \) and \( \nu_v \). Otherwise, for \( \nu_v = 0 \) and \( \nu_u \leq \pi \), the distance between \( u \) and \( v \) is increasing in \( \nu_u \) and there exists a critical value \( \nu_0 \) such that \( d_H((r_u, \nu_u), (r_v, 0)) \leq R \) if and only if \( \nu_u \leq \nu_0 \). The following lemma estimates \( \nu_0 \).
Lemma 3.17 (Lemma 3.1 in [GPP12]). Let \(0 \leq r_u \leq R\), \(r_u + r_v \geq R\) and assume \(\nu_v = 0\). Then
\[
\nu_0 = 2e^{-\frac{R-r_u-r_v}{2}}(1 + \Theta(e^{R-r_u-r_v})).
\]

Suppose \(\|x_u - x_v\| \geq M\frac{w_u w_v}{W}\). Notice that by our transformation we have \(\|x_u - x_v\| = \frac{v_u}{2\pi}\) and
\[
(3.8) \quad \frac{w_u w_v}{W} = \Theta\left(\frac{w_u w_v}{n}\right) = \Theta\left(ne^{-(r_u+r_v)/2}\right) = \Theta\left(e^{(R-r_u-r_v)/2}\right),
\]
where we used \(W = \Theta(n)\) by Lemma 3.1. Hence, we have \(\frac{v_u}{2\pi M} = \Omega(e^{(R-r_u-r_v)/2})\).

If we choose the constant \(M\) sufficiently large this implies \(r_u + r_v > R\), since \(v_u \leq 1\), and \(v_u > \nu_0\), for sufficiently large \(n\). Thus, by Lemma 3.17 the two vertices \(u\) and \(v\) are not connected and indeed \(p_{uv} = 0\).

On the other hand, assume \(\|x_u - x_v\| \leq m\frac{w_u w_v}{W}\). Then either \(r_u + r_v < R\) and thus \(\{u, v\} \in E\) follows directly, or \(r_u + r_v \geq R\) and \(v_u < \nu_0\), if \(m\) is sufficiently small. In the second case, Lemma 3.17 implies \(p_{uv} = 1\).

We now turn to the case \(\alpha < \infty\). By our assumptions on \(v_u\) and \(v_v\) and by the identity \(\cosh(x \pm y) = \cosh(x) \cosh(y) \pm \sinh(x) \sinh(y)\), we can rewrite (3.7) as
\[
(3.9) \quad \cosh(d_H) = \cosh(r_u - r_v) + (1 - \cos(v_u)) \sinh(r_u) \sinh(r_v).
\]

Next, note that \(\cosh(x) = \Theta(e^{x^2})\) for all \(x\) and \(\sinh(x) = \Theta(e^x)\) for all \(x = \omega(1)\). Observe that (PL2) and \(w_v = ne^{-r_v/2}\) imply \(r_v = \Theta(\log n)\) for all vertices \(v\). Furthermore, we perform a Taylor approximation of \(1 - \cos(v_u)\) around 0 and get
\[
1 - \cos(v_u) = \frac{v_u^2}{2} - \frac{v_u^4}{24} + \ldots = \Theta(v_u^2), \text{ as } v_u \text{ is at most a constant.}
\]
Combining these observations with (3.9) and the assumption \(r_u \geq r_v\), we deduce
\[
(3.10) \quad e^{d_H-R} = \Theta(\cosh(d_H)e^{-R}) = \Theta\left(e^{r_u-r_v-R} + v_u^2 e^{r_u+r_v-R}\right).
\]

In the condition (3.5) on \(p_{uv}\) the minimum is obtained by the second term whenever \(\|x_u - x_v\| \leq \frac{w_u w_v}{W}\). Mapping \(u\) and \(v\) to the hyperbolic disk, this condition corresponds to \(v_u = O(e^{(R-r_u-r_v)/2})\). We claim that whenever it holds \(v_u = O(e^{(R-r_u-r_v)/2})\), then the two vertices \(u\) and \(v\) are connected with constant probability and therefore \(p_{uv} = \Theta(1)\). Indeed, in this case by (3.10) we have \(e^{d_H-R} = O(1)\), and using Definition 3.13 we deduce
\[
p_{uv} = p_H(d_H(u, v)) = \left(1 + (e^{d_H-R})^{1/(2\Theta)}\right)^{-1} = \Theta(1).
\]
On the other hand, suppose \( \|x_u - x_v\| \geq \frac{w_u w_v}{W} \). Then \( v_u = \Omega(e^{(R-r_u-r_v)/2}) \). In this case by (3.10) we have \( e^{d_H-R} = \Theta(v_u^2 e^{r_u+r_v-R}) = \Omega(1) \). However, if \( e^{d_H-R} = \Omega(1) \), we can use Definition 3.13 and (3.8) to obtain

\[
p_{uv} = \left(1 + e^{\frac{1}{2\tau} (d_H-R)}\right)^{-1} = \Theta\left(\frac{1}{\|x_u - x_v\| \cdot \frac{w_u w_v}{W}}\right)^{1/T_H} = \Theta\left(\left(\frac{1}{\|x_u - x_v\| \cdot \frac{w_u w_v}{W}}\right)^{\alpha}\right).
\]

This finishes the case \( \alpha < \infty \) and thus the proof.

### 3.6 Further Related Models

Here, we discuss two random graph models that are not examples of our general model, but still related to GIRGs.

**Scale-free percolation:** This is an infinite random graph model, introduced by Deijfen et al. [DHH13], where the vertex set is given by the \( d \)-dimensional grid \( \mathbb{Z}^d \). Each vertex \( x \in \mathbb{Z}^d \) chooses independently a weight \( W_x \) at random from a distribution \( W \), where we assume that \( W \) follows a power law with exponent \( \tau_S \). Let \( \alpha_S > 0 \) be a long-range parameter and \( \lambda_S \) be a percolation parameter. Then, for every two vertices \( x, y \in \mathbb{Z}^d \) the edge \( \{x, y\} \) is independently present in the graph with probability

\[
\Pr[\{x, y\} \in E] = 1 - \exp\left(-\lambda_S \frac{W_x W_y}{\|x - y\|^{\alpha_S}}\right).
\]

Sometimes, it is assumed in addition that geometric neighbors (i.e., pairs \( \{x, y\} \) with \( \|x - y\| = 1 \)) deterministically share an edge so that an underlying grid structure is present. It is shown in [DHH13] that the degree sequence has again a power-law distribution, with parameter \( \beta = (\tau_S - 2)\alpha_S / d \), and that in the regime \( 2 < \beta < 3 \) we have a connected component of infinitely many vertices. Moreover, in this regime, typical graph-theoretic distances are doubly logarithmic in the geometric distance. More precisely, for every \( \eta > 0 \) it holds

\[
\lim_{|x| \to \infty} \Pr\left[d(0, x) \leq \frac{2 + \eta}{\log(\beta - 2)} \log|\log|x|| |d(0, x) < \infty\right] = 1.
\]
Further structural results have been proven in [DHW15; HHJ17].

By setting $\alpha := \alpha_S/d$ and $\beta := (\tau_S - 2)\alpha_S/d$, scale-free percolation model and GIRGs are very similar models. More concretely, scale-free percolation can be seen as an infinite variant of GIRGs where a box of $\mathbb{Z}^d$ of volume $n$ induces a GIRG on $n$ vertices. In scale-free percolation, the positions of the vertices are fixed on the grid instead of being random and the precise edge probability function is fixed as well. The tight connection between the two models was described in [HK17], where explosion processes on the scale-free percolation model have been analyzed, and has been further strengthened in [KL18]. We believe that GIRGs are slightly more natural as the weight of a vertex coincides with its expected degree, up to constant factors, which is not the case in scale-free percolation. It is also plausible that a setting with random coordinates is more realistic than a perfect lattice structure. Finally, GIRGs include hyperbolic random graphs as a special case (Theorem 3.15), which we believe to be beneficial since all the experimental work on hyperbolic random graphs further validates the model.

**Kleinberg’s model:** For the purpose of explaining Milgram’s letter forwarding experiment, Kleinberg defined the following random graph model [Kle00b; Kle00a]. We start with an $n \times n$ lattice graph, that is, node $v = (i, j)$ has position $x_v = (i/n, j/n)$ in the unit square. Next, any two nodes with $||x_u - x_v|| \leq 1/n$ are connected by an edge. We then add for each node $u$ a constant number of directed long-range edges, where the other endpoint $v$ is chosen proportional to $||x_u - x_v||^{-2\alpha}$ with decay parameter $\alpha$. Note that this graph is homogeneous, i.e., all vertices have the same out-degree, which is not realistic for social networks.

In subsequent work, Kleinberg’s model was modified and amended. In particular, some variants have a power-law degree sequence, e.g., [FG14]. These models are related to GIRGs as they use the concept of long-range edges too and the geometric space is somewhat similar. However, Kleinberg’s model and its variants are mainly constructed for the purpose of analyzing and optimizing routing processes, limiting the applicability as models for complex networks. In general, there is much more evidence for GIRGs to resemble real-world networks.
than for Kleinberg’s model and its variations, both from a theoretical and from an empirical perspective, the latter due to the extensive experimental work on hyperbolic random graphs.

### 3.7 Summary

We introduced a new class of random graphs that generically augment Chung-Lu random graphs with underlying geometry, i.e., every vertex has a random position in the ground space and edge probabilities may arbitrarily depend on the vertex positions, as long as marginal edge probabilities are preserved. Since our model is very general, it relates pre-existing models like hyperbolic random graphs and Chung-Lu random graphs to each other as they are contained in our framework as special cases. One specific example of the general class are geometric inhomogeneous random graphs (GIRGs) which we study in more details in Chapter 4. Beyond these well-studied models, our class also includes non-metric ground spaces, which are motivated by social networks, where two persons are likely to know each other if they share a hobby, regardless of their other hobbies.

These examples also show that our model is incomparable to the (also very general) model of inhomogeneous random graphs studied by Bollobás, Janson, and Riordan [BJR07]. Their model requires many long-range edges, so that setting $\alpha > 1$ in (3.5) yields an edge probability that is not supported by their model, and metric distances, so that the minimum component distance is also not supported by their model.

We have seen that in the regime $\beta > 3$, the general model contains both examples with a giant connected component such as Chung-Lu random graphs and examples without giant connected component such as hyperbolic random graphs. Therefore, when analyzing the component structure and distances in Chapter 4, we restrict our analysis to the regime $2 < \beta < 3$. 
GIRGs: Structural and Algorithmic Properties

We start with the general class of random graph models that we introduced in Section 3.1. We prove that $G(n, X, w, p)$ always has the same average distance as Chung-Lu random graphs, up to a factor $1 + o(1)$. This shows in particular that specific choices, such as the underlying geometry being Euclidean or the dependence on the distance being inversely polynomial, do not significantly influence the average distance. The proof also yields that our model has a giant component and polylogarithmic diameter with high probability. Furthermore, we verify that every random graph model from our general class has a power-law degree sequence.

After these results on the most general model, we specifically study geometric inhomogeneous random graphs as defined in Section 3.4. As main contribution for GIRGs we provide a sampling algorithm that generates a random graph from our model in expected linear time, improving the best-known sampling algorithm for hyperbolic random graphs by a substantial factor $O(\sqrt{n})$ [LMP15]. Further, we establish that GIRGs have clustering coefficients in $\Omega(1)$ and small separators, i.e., it suffices to delete a sublinear number of edges to break the
giant component into two large pieces. Finally, we show how to compress GIRGs in expected linear time. We prove all these results for both Standard GIRGs \((\alpha < \infty)\) and Threshold GIRGs \((\alpha = \infty)\).

Everything of this chapter is joint work with Karl Bringmann and Johannes Lengler [BKL16; BKL15; BKL17].

4.1 Results

We start by presenting a condensed list of our structural and algorithmic results. Recall from Theorem 3.15 that hyperbolic random graphs are a special case of one-dimensional GIRGs, and thus all results on GIRGs hold for hyperbolic random graphs too. Moreover, by Theorem 3.8 GIRGs are a special case of our general model \(G(n, X, w, p)\), hence all findings for \(G(n, X, w, p)\) transfer to GIRGs. We first analyze the general model as introduced in Section 3.1 and then turn to GIRGs, where we consider Standard GIRGs \((\alpha < \infty)\) and Threshold GIRGs \((\alpha = \infty)\) together. We study the following fundamental structural questions.

Scale-freeness: As we always use a weight sequence \(w\) that satisfies the power-law conditions (PL1) and (PL2), we expect our model to be scale-free.

**Theorem 4.1** (Section 4.2). With high probability, for all \(\beta > 2\) the degree sequence of \(G(n, X, w, p)\), where \(w\) not necessarily fulfills (EP2), follows a power law with exponent \(\beta\) and average degree \(\Theta(1)\).

Giant component and diameter: The connectivity properties of the model for \(\beta > 3\) are not very well-behaved. In this regime, the model is too general to deduce significant results, in particular because in this case even threshold hyperbolic random graphs do not possess a giant component of linear size [BFM15b]. Hence, for connectivity properties we restrict our attention to the regime \(2 < \beta < 3\), which holds for most real-world networks [DM02].

**Theorem 4.2** (Section 4.3). Let \(2 < \beta < 3\). W.h.p. the largest component of \(G(n, X, w, p)\) has linear size, while all other components have size at most \(\log^{O(1)} n\). Moreover, w.h.p. the diameter is at most \(\log^{O(1)} n\).
4.1. RESULTS

A better bound of \(\Theta(\log n)\) holds for the diameter of Chung-Lu graphs \([CL02b; CL04]\) and for hyperbolic random graphs \([FK15b; TS17]\), as discussed in Section 3.5. It remains an open problem whether the upper bound \(O(\log n)\) holds in general for our model.

**Average distance:** As main result for our general random graph model, we determine the average distance between two randomly chosen nodes in the giant component to be the same as in Chung-Lu random graphs up to a factor \(1 + o(1)\), showing that the underlying geometry is irrelevant for the value of the average distance.

**Theorem 4.3** (Section 4.3). Let \(2 < \beta < 3\). Then the average distance of the random graph \(\mathcal{G}(n, \mathcal{X}, w, p)\) is \(\frac{2 + o(1)}{|\log(\beta - 2)|} \log \log n\) in expectation and a.a.s.

It is quite surprising that the average distance can be computed so precisely in this generality. For instance, our general class of scale-free random graphs includes both models with clustering and without clustering. The fact that \(\mathcal{G}(n, \mathcal{X}, w, p)\) is too general to obtain any meaningful results in the regime \(\beta > 3\) makes it even more surprising that in the regime \(2 < \beta < 3\) the average distance can be determined precisely for all instances at once.

A common property of all models in our general class is that for a set \(S\) of vertices whose weights sum to \(W_S\) (often called *volume* in the literature), the expected number of half-edges going out from \(S\) is \(\Theta(W_S)\). For the classic Chung-Lu random graphs without geometry, the targets of these half-edges are independent of each other. Thus the quantity \(W_S\) is essentially sufficient to determine the size and the volume of the neighborhood \(\Gamma(S)\) of \(S\) and the analyses of Chung-Lu random graphs are based on this property. However, for non-trivial geometries the size of the neighborhood crucially depends on the geometric position of the vertices in \(S\). For example, if the clustering coefficient is constant, then even if \(S\) consists of only two adjacent vertices there is already a non-negligible probability that they share some neighbors. Thus the proofs for classic Chung-Lu random graphs do not carry over to the general setting. On the
other hand, existing proofs for geometric scale-free networks [ABM17; DHH13] rely rather heavily on the specifics of the underlying geometry.

In the general setting, we can therefore only borrow one step from previous proofs, namely the “greedy path” argument (Lemma 4.13). We use this idea to prove that for all vertices of at least poly-logarithmic weight there exists an ultra-short path to the “heavy core”, which is well-connected and contains the vertices of highest weight. From a technical point of view, the most important contribution is the “bulk lemma” (Lemma 4.16). It contains a delicate and subtle analysis of the neighborhood of a vertex restricted to small-weight vertices. The lemma studies the probability that the \( k \)-neighborhood of a random vertex \( v \) of low-weight contains a node \( v' \) that is connected to a high-weight vertex, from which we can then apply the “greedy path” argument. We emphasize that both the size and the shape of the \( k \)-neighborhood of such a vertex crucially depend on the underlying geometry. Therefore, we are forced to use the geometry implicitly to develop a general argument that is valid universally for \( \mathcal{G}(n, x, w, p) \). Finally, we obtain the bound on the average distance by applying the bulk lemma repeatedly for different values of \( k \) and carefully summing up the resulting terms.

**Clustering:** In social networks, two friends of the same person are likely to also be friends with each other. This property of having many triangles is captured by the clustering coefficient, defined as the probability when choosing a random vertex \( v \) and two random neighbors \( v_1 \neq v_2 \) of \( v \) that \( v_1 \) and \( v_2 \) are adjacent (if \( v \) does not have two neighbors then its contribution to the clustering coefficient is 0). Chung-Lu random graphs have a very small clustering coefficient of \( n^{-\Omega(1)} \) which is why we now turn to the more concrete model of GIRGs. We show that the clustering coefficient of GIRGs is \( \Theta(1) \). This is consistent with empirical data of real-world networks [DM02; New03; AB02] (see as well Table 1.1) and with the constant clustering coefficient of hyperbolic random graphs, determined in [CF16b].

**Theorem 4.4** (Section 4.4). Let \( \beta > 2 \). Then w.h.p. the clustering coefficient of a GIRG is \( \Theta(1) \).
4.1. RESULTS

**Sampling:** Regarding applications and simulations, it is crucial that a model can be generated fast and produces sufficiently large samples. Sampling algorithms that generate a random graph from a fixed distribution are known for Chung-Lu random graphs and others, running in expected linear time [BB05; MH11]. As our main result, we present such an algorithm for GIRGs. This greatly improves the trivial $O(n^2)$ sampling algorithm (throwing a biased coin for each possible edge), as well as the best previous algorithm for threshold hyperbolic random graphs with expected time $O(n^{3/2})$ [LMP15]. It allows to run experiments on much larger graphs than the ones with $\approx 10^4$ vertices in [BPK10]. In addition to our model assumptions, here we assume that the $\Theta$ in our requirement on $p_{uv}$ is sufficiently explicit, i.e., we can compute $p_{uv}$ exactly and we know a constant $c > 0$ such that replacing $\Theta$ by $c$ yields an upper bound on $p_{uv}$, see Section 4.5 for details.

**Theorem 4.5 (Section 4.5).** Let $\beta > 2$. Then GIRGs can be sampled in expected time $O(n)$.

**Stability:** For real-world networks, a key property to analyze is their stability under attacks. It has been empirically observed that many real-world networks have small separators of size $n^c$, $c < 1$ [BBK03]. In contrast, Chung-Lu random graphs are unrealistically stable, since any deletion of $o(n)$ nodes or edges reduces the size of the giant component by at most $o(n)$ [BJR07]. We show that GIRGs agree with the empirical results much better. Specifically, if we cut the ground space $T^d$ into two halves along one of the axes then we roughly split the giant component into two halves, but the number of edges passing this cut is quite small, namely $n^{1-\Omega(1)}$. Thus, GIRGs are prone to (quite strong) adversarial attacks, just as many real-world networks. Furthermore, their small separators are useful for many algorithms, e.g., the compression scheme of the next paragraph.

**Theorem 4.6 (Section 4.6).** Let $2 < \beta < 3$. Then almost surely it suffices to delete

$$O\left(n^{\max\{2-a,3-\beta,1-1/d\}+o(1)}\right)$$

edges of a GIRG to split its giant component into two parts of linear size each.
Since we assume $\alpha > 1$, $\beta > 2$, and $d = \Theta(1)$, the number of deleted edges is indeed $n^{1-\Omega(1)}$. Recently, Bläsius, Friedrich, and Krohmer [BFK16] proved that threshold hyperbolic random graphs (corresponding to GIRGs with parameters $d = 1$ and $\alpha = \infty$) can be separated by removing $O(n^{(3-\beta)/2})$ vertices. Moreover, they obtained $O(\log n)$ for the case $\beta = 3$, $O(1)$ for the regime $\beta > 3$, and $n^{1-\Omega(1)}$ for the general hyperbolic random graph model (recall Definition 3.13). In addition, they provided a number of algorithmic applications.

**Entropy:** The internet graph has empirically been shown to be well compressible, using only 2-3 bits per edge [BBK03; BV04]. This is not the case for the Chung-Lu model, as its entropy is $\Theta(n \log n)$ [Chi+09a]. We show that GIRGs have linear entropy.

**Theorem 4.7** (Section 4.6). Let $\beta > 2$. Then we can store a GIRG using $O(n)$ bits in expectation. The resulting data structure allows to query the degree of any vertex and its $i$-th neighbor in time $O(1)$. The compression algorithm runs in time $O(n)$.

### 4.2 Degree Sequence

We assume that the weight sequence $w$ satisfies (PL1) and (PL2) and thus follows a power law. Since the expected degree of a vertex with weight $w_v$ is $\Theta(w_v)$ by Lemma 3.3, it is not surprising that the resulting random graph is scale-free w.h.p. In this section, we give details and prove Theorem 4.1, where we use Theorem 2.5 for showing concentration. Some ideas of our proof are based on [GPP12]. We start with the maximum degree $\Delta(G)$, which is a simple corollary of Lemma 3.4.

**Corollary 4.8** (Maximum degree). With high probability, $\Delta(G) = \Theta(w_{\text{max}})$, where $w_{\text{max}} = \max\{w_v \mid v \in V\}$. In particular, for all $\eta > 0$, w.h.p., $\Delta(G) = \Omega(\overline{w})$ and $\Delta(G) = O(n^{1/(\beta-1-\eta)})$.

**Proof.** We deduce from (PL2) that $\omega(\log^2 n) \leq \overline{w} \leq w_{\text{max}} = O(n^{1/(\beta-1-\eta)})$. Then Lemma 3.4 directly implies the statement.
Next, we calculate the expected number of vertices with degree at least \( k \).

**Lemma 4.9.** Let \( \eta > 0 \) be sufficiently small. Then for all \( k \geq 1, k = k(n) = o(\overline{w}) \), we have

\[
\Omega(nk^{1-\beta-\eta}) \leq \mathbb{E}[|\{v \in V \mid \deg(v) \geq k\}|] \leq O(nk^{1-\beta+\eta}).
\]

**Proof.** Let \( \eta \) be sufficiently small. Recall that for every vertex \( v \in V \) it holds \( \mathbb{E}[\deg(v)] = \Theta(w_v) \) by Lemma 3.3. Let \( k \geq 1, k = k(n) = o(\overline{w}) \), and let \( v \) be any vertex with weight \( w_v \geq \Omega(k) \) large enough such that \( \mathbb{E}[\deg(v)] \geq 2k \). Then by a Chernoff bound

\[
\Pr[\deg(v) < k] \leq \Pr[\deg(v) < 0.5\mathbb{E}[\deg(v)]] \leq e^{-\mathbb{E}[\deg(v)]/8} \leq e^{-k/4} \leq e^{-1/4}.
\]

By the power-law assumption (PL2) there are \( \Omega(nk^{1-\beta-\eta}) \) vertices with weight \( \Omega(k) \), and a single vertex of this set has degree at least \( k \) with probability at least \( 1 - e^{-1/4} \). By linearity of expectation,

\[
\mathbb{E}[|\{v \in V \mid \deg(v) \geq k\}|] = \sum_{v \in [n]} \Pr[\deg(v) \geq k] = \Omega(nk^{1-\beta-\eta}).
\]

Next let \( v \) be a vertex with weight \( w_v \leq O(k) \) small enough such that \( 2e\mathbb{E}[\deg(v)] \leq 3k/4 \). By a Chernoff bound (Theorem 2.2 (iii)) we obtain

\[
\Pr[\deg(v) \geq k] \leq \Pr[\deg(v) > 3k/4] \leq 2^{-3k/4}.
\]

Thus, for the upper bound it follows

\[
\mathbb{E}[|\{v \in V \mid \deg(v) \geq k\}|] = \sum_{v \in [n]} \Pr[\deg(v) \geq k] \leq |V_{\geq O(k)}| + \sum_{v \in V_{< O(k)}} \Pr[\deg(v) \geq k]
\leq O(nk^{1-\beta+\eta}) + n \cdot 2^{-3k/4}.
\]

Note that \( k^2 \leq 3 \cdot 2^{3k/4} \) holds for all \( k \geq 1 \). Hence \( n \cdot 2^{-3k/4} \leq 3nk^{-2} < 3nk^{1-\beta+\eta} \) and indeed it holds \( \mathbb{E}[|\{v \in V \mid \deg(v) \geq k\}|] = O(nk^{1-\beta+\eta}). \)

After these preparations we come to the main theorem of this section which is a more precise formulation of Theorem 4.1 and states that the degree sequence follows a power law with the same exponent \( \beta \) as the weight sequence.
Theorem 4.10 (Degree sequence). In $\mathcal{G}(n, X, w, p)$, for all $\eta > 0$ w.h.p. we have
\[
\Omega(n k^{1-\beta-\eta}) \leq |\{v \in V \mid \deg(v) \geq k\}| \leq O(n k^{1-\beta+\eta}),
\]
where the first inequality holds for all $1 \leq k \leq \overline{w}$ and the second inequality holds for all $k \geq 1$.

Before we prove Theorem 4.10, we note that together with the standard calculations from Lemma 3.1 we immediately obtain the average degree in the graph.

Corollary 4.11 (Average degree). W.h.p., it holds $\frac{1}{n} \sum_{v \in V} \deg(v) = \Theta(1)$ and thus $|E| = \Theta(n)$.

Proof of Theorem 4.10. We first consider the case $k \geq \log^3 n = o(\overline{w})$. From Condition (PL2) on the vertex weights and Lemma 3.3 it follows that
\[
|\{v \in V \mid \mathbb{E}\deg(v) \geq 1.5k\}| = \Omega(n k^{1-\beta-\eta})
\]
holds for all $\log^3 n \leq k \leq \overline{w}$. Then by Lemma 3.4, w.h.p. every vertex $v$ with $\mathbb{E}\deg(v) \geq 1.5k$ has degree at least $(1 - o(1))1.5k \geq k$ for $n$ large enough. Hence w.h.p. there exist at least $\Omega(n k^{1-\beta-\eta})$ vertices with degree at least $k$. Vice-versa, by Lemma 3.3 we have
\[
|\{v \in V \mid \mathbb{E}\deg(v) \geq 0.5k\}| = O(n k^{1-\beta+\eta}).
\]
By the same arguments as above, w.h.p. every vertex $v$ with $\mathbb{E}\deg(v) < 0.5k$ has degree at most $(1 + o(1))0.5k < k$. Thus in total there can be at most $O(n k^{1-\beta+\eta})$ vertices with degree at least $k$. This proves the theorem for $k \geq \log^3 n$.

Let $1 \leq k \leq \log^3 n$, $\varepsilon > 0$ be sufficiently small, $V' := V \leq n^\varepsilon$ be the set of small-weight vertices, and $G' := G[V']$. First, we introduce some notation and define the two random variables
\[
g_k := |\{v \in V \mid \deg(v) \geq k\}| \quad \text{and} \quad f_k := |\{v \in V' \mid \deg_{G'}(v) \geq k\}|.
\]
Note that by Lemma 4.9, we already have $\Omega(n k^{1-\beta-\eta}) \leq \mathbb{E}[g_k] \leq O(n k^{1-\beta+\eta})$, and it remains to prove concentration. Clearly,
\[
f_k \leq g_k \leq f_k + 2 \sum_{v \in V \setminus V'} \deg(v). \tag{4.1}
\]
Next we apply Lemma 3.4 together with Lemma 3.1 and see that w.h.p.
\[ \sum_{v \in V \setminus V'} \deg(v) = \Theta(W_{\geq n^\varepsilon}) = O(n^{1+(2-\beta+\eta)\varepsilon}) = n^{1-\Omega(1)}. \]

Recall that we assume \( k \leq \log^3 n \), so in particular \( \mathbb{E}[g_k] = \Omega(n/(\log n)^{3(\beta-1+\eta)}) \).
Thus \( \mathbb{E}\left[ \sum_{v \in V \setminus V'} \deg(v) \right] = o(\mathbb{E}[g_k]) \), and \((4.1)\) implies \( \mathbb{E}[f_k] = (1+o(1))\mathbb{E}[g_k]. \)
Hence, it is sufficient to prove that the random variable \( f_k \) is concentrated around its expectation, because this will transfer immediately to \( g_k \).

We aim to show this concentration result via Theorem 2.5. We need to argue that the probability space of our random graph model is a product of independent random variables. Recall that we apply two different randomized processes to create the geometric graph. First, for every vertex \( v \) we chose \( x_v \in X \) independently at random. Afterwards, every edge is present with some probability \( p_{uv} \). So far, only the \( n \) random variables \( x_1, \ldots, x_n \) defining the vertex set are independent. We introduce a second set of \( n-1 \) independent random variables. For every \( u \in \{2, \ldots, n\} \) we let \( Y_u := (Y_u^1, \ldots, Y_u^{u-1}) \), where every \( Y_u^v \) is a real number chosen independently and uniformly at random from the interval \([0, 1]\). Then for \( v < u \), we include the edge \( \{u, v\} \) in the graph if and only if
\[ p_{uv} > Y_u^v. \]

We observe that indeed this implies \( \Pr[u \sim v \mid x_u, x_v] = p_{uv}(x_u, x_v) \) as desired. Furthermore, the \( 2n-1 \) random variables \( x_1, \ldots, x_n, Y_2, \ldots, Y_n \) are independent and define a product probability space \( \Omega \) which is equivalent to our random graph model. Formally, every \( \omega \in \Omega \) defines a graph \( G(\omega) \). Similarly we use \( G' = G'(\omega) \) and \( f_k = f_k(\omega) \). We now consider the bad event
\[ \mathcal{B} := \{ \omega \in \Omega \mid \text{the maximum degree in } G'(\omega) \text{ is at least } n^{2\varepsilon} \}. \]
We observe that \( \Pr[\mathcal{B}] = n^{-\omega(1)} \), since by Lemma 3.4 w.h.p. every vertex \( v \in V' \) has degree at most \( O(w_v + \log^2 n) = o(n^{2\varepsilon}) \). Let \( \omega, \omega' \in \overline{\mathcal{B}} \) such that they differ in at most two coordinates. We observe that changing one coordinate \( x_u \) or \( Y_u \) can only influence the degrees of \( u \) itself and of the vertices which are neighbors of \( u \) before or after the coordinate change. It follows that \( |f_k(\omega) - f_k(\omega')| \leq 4n^{2\varepsilon} =: c. \)
Therefore, $f_k$ satisfies the Lipschitz condition of Theorem 2.5 with bad event $\mathcal{B}$. Let $t = n^{1-\varepsilon} = o(\mathbb{E}[f_k])$. Then since $n \Pr[\mathcal{B}] = n^{-\omega(1)}$, Theorem 2.5 implies

$$\Pr[|f_k - \mathbb{E}[f_k]| \geq t] \leq 2e^{-\frac{t^2}{64c^2n}} + \left(\frac{4n^2}{c} + 1\right) \Pr[\mathcal{B}] = e^{-\Omega(n^{1-4\varepsilon})} + n^{-\omega(1)} = n^{-\omega(1)},$$

which proves the concentration and concludes the proof.

\section{4.3 Giant Component, Diameter, and Average Distance}

Throughout this section we assume $2 < \beta < 3$. Under this assumption we prove that w.h.p. $\mathcal{G}(n, \mathcal{X}, w, p)$ has a giant component with diameter at most $(\log n)^{O(1)}$, and that all other components are only of polylogarithmic size. We further show that the average distance of any two vertices in the giant is $(2 + o(1)) \log \log n/|\log(\beta - 2)|$ in expectation and with probability $1 - o(1)$. The same formula has been known to hold for various graph models, including Chung-Lu [CL04] and hyperbolic random graphs [ABM17]. The lower bound follows from the first moment method on the number of paths of different types. Note that the probability that a fixed path $P = (v_1, \ldots, v_k)$ exists in our model is the same as in Chung-Lu random graphs, since the marginal probability of the event $v_i \sim v_{i+1}$ conditioned on the positions of $v_1, \ldots, v_i$ is $\Theta(\min\{1, w_{v_i}w_{v_{i+1}}/W\})$, as in the Chung-Lu model. In particular, the expected number of paths coincides for both models (save the factors coming from the $\Theta(\cdot)$-notation). Not surprisingly, the lower bound for the expected average distance follows from general statements on power-law graphs, bounding the expected number of too short paths by $o(1)$, cf. [DMM12, Theorem 2]. The main contribution of this section is to prove a matching upper bound for the average distance.

The proof-strategy is as follows. We first prove that w.h.p. for every vertex of weight at least $(\log n)^C$ there exists an ultra-short path to the “heavy core”, which has diameter $o(\log \log n)$ and contains the vertices of highest weight. Afterwards, we show that a random low-weight vertex has a large probability to connect to a vertex of weight at least $(\log n)^C$ within a small number of steps. The statement is formalized below as the “bulk lemma” (Lemma 4.16). This
lemma is the crucial step of the main proof and new compared to previous studies of Chung-Lu random graphs and similar models. It contains a delicate analysis of the $k$-hop-neighborhood of a random vertex, restricted to small weights. Thereby, the underlying geometry is used implicitly, in order to make the argument applicable for the fairly general model that we study.

In the whole section let $G$ be a graph sampled from $G(n, X, w, p)$. We start by considering the subgraph induced by the heavy vertices $\bar{V} := V_{\geq \bar{w}}$, where $\bar{w}$ is given by the definition of power-law weights, see condition (PL2) in Section 3.1. We call the induced subgraph $\bar{G} := G[\bar{V}]$ the heavy core.

**Lemma 4.12** (Heavy core). *With high probability $\bar{G}$ is connected and has diameter $o(\log \log n)$.*

**Proof.** Let $\bar{n}$ be the number of vertices in the heavy core, and let $\eta > 0$ be small enough. Since $\bar{w} \leq n^{(1-\Omega(1))/(\beta-1)}$, we may bound $\bar{n} = \Omega(n^{w^{-1-\beta-\eta}}) = n^{\Omega(1)}$. By (EP2), the connection probability for any heavy vertices $u, v$, regardless of their position, is at least

$$p_{uv}(x_u, x_v) \geq \left(\frac{n}{\bar{w}^{\beta-1+\eta}}\right)^{-1+\omega(1/\log \log n)} \geq \bar{n}^{-1+\omega(1/\log \log n)}.$$

Therefore, the diameter of the heavy core is at most the diameter of an Erdős-Rényi random graph $G(\bar{n}, p)$, with $p = \bar{n}^{-1+\omega(1/\log \log n)}$. However, with probability $1 - \bar{n}^{-\omega(1)}$ this diameter is $\Theta(\log \bar{n} / \log (p \bar{n})) = o(\log \log n)$ [DM10]. Since $\bar{n} = n^{\Omega(1)}$, this proves the lemma. 

Next we show that if we start at a vertex of weight $w$, going greedily to neighbors of largest weight yields a short path to the heavy core with a probability that approaches 1 as $w$ increases.

**Lemma 4.13** (Greedy path).

(i) Let $0 < \varepsilon < 1$ and let $v$ be a vertex of weight $2 \leq w < \bar{w}$. Then with probability at least $1 - O\left(\exp\left(-w^{\Omega(\varepsilon)}\right)\right)$ there exists a weight-increasing path of length at most $(1 + \varepsilon) \frac{\log \log n}{\log(\beta-2)}$ from $v$ to the heavy core.
(ii) For every $\epsilon > 0$ there exists a constant $C = C(\epsilon) > 0$ such that w.h.p. for all $v \in V_{\geq (\log n)^C}$ there exists a weight-increasing path of length at most $(1 + \epsilon) \frac{\log \log n}{|\log(\beta - 2)|}$ from $v$ to the heavy core.

(iii) W.h.p., there are $\Omega(n)$ vertices in the same component as the heavy core.

Proof. Let $0 < \epsilon < 1$, let $v$ be a vertex of weight $2 \leq w \leq \overline{w}$, and let

$$\tau = \tau(\epsilon) := (\beta - 2)^{-1/(1+\epsilon/2)}.$$  

Note that $1 < \tau < 1/(\beta - 2)$, and that $1/ \log \tau = (1 + \epsilon/2)/|\log(\beta - 2)|$. Moreover, we define an increasing weight sequence $w_0, w_1, \ldots, w_{i_{\text{max}}} := \overline{w}$ such that for all $1 \leq i \leq i_{\text{max}}$ it holds $w_i := w_{i-1}^\tau$, and such that $w_0 \leq w_v < w_1$. For all $i < i_{\text{max}}$ we put $V_i := V_{\geq w_i \setminus V_{\geq w_{i+1}}}$. Furthermore, we put $V_{i_{\text{max}}} := \overline{V}$ and $v_0 := v$. We will show that with sufficiently high probability, for all $0 \leq i < i_{\text{max}}$ the vertex $v_i$ has at least one neighbor $v_{i+1} \in V_{i+1}$. Note that $i_{\text{max}} = \lceil \log \left( \frac{\log \overline{w}}{\log w} \right) \rceil$, so this implies that there is a path from $v$ to the heavy core of length at most

$$i_{\text{max}} \leq (1 + \epsilon/2) \frac{\log \log n}{|\log(\beta - 2)|} + 1 \leq (1 + \epsilon) \frac{\log \log n}{|\log(\beta - 2)|},$$

for sufficiently large $n$, and thus proves statement (i).

Let $0 \leq i < i_{\text{max}}$ and assume by induction that there exists a weight-increasing path from $v_0$ to $v_i$ where $v_i \in V_i$. Note that this event only depends on the random graph induced by the vertex set $V_{<w_{i+1}}$. We want to verify that $v_i$ connects to at least one vertex $v_{i+1} \in V_{i+1}$. First, observe that by condition (PL2), each layer $V_i$ contains at least $\Omega(nw_i^{1-\beta-\eta})$ and at most $O(nw_i^{1-\beta+\eta})$ vertices. Next, by condition (EP1) the edges from $v_i$ to vertices $v$ with $v \in V_{i+1}$, are independently present with probability $\Omega(\min\{w_v w_i / \mathcal{W}, 1\})$, respectively. If $w_i w_{i+1} \geq \mathcal{W}$, this probability is $\Omega(1)$. However, then $w_i \geq n^{1/(1+\tau)}$ and we deduce $|V_{i+1}| = n^{\Omega(1)}$.

In this case, the probability that $v_i$ connects to at least one vertex of the next weight layer is $1 - \exp(-n^{\Omega(1)}) = 1 - \exp(-w_i^{\Omega(1)})$. So assume $w_i w_{i+1} < \mathcal{W}$, where we can lower-bound the edge probability by $\Omega(w_i w_{i+1} / \mathcal{W})$. Thus, for any $\eta > 0$
the probability that \( v_{i-1} \) does not connect to a vertex in \( V_{i+1} \) is at most
\[
p_i := \prod_{v \in V, w_v \geq w_{i+1}} \left( 1 - \Omega \left( \frac{w_v w_{i+1}}{W} \right) \right) \leq \exp \left( -\Omega \left( \frac{w_i w_{i+1}}{W} \cdot |V_{i+1}| \right) \right) \leq \exp \left( -\Omega \left( w_i w_{i+1}^{2-\beta-\eta} \right) \right),
\]
where we used Lemma 3.1 in the last step. Since \( w_{i+1} \leq w_{i}^\tau \), we obtain
\[
p_i \leq \exp \left( -\Omega \left( w_i^{1-\tau(\beta-2+\eta)} \right) \right).
\]
Note that as \( \tau < 1/(\beta - 2) \), the exponent of \( w_i \) in this expression is positive for sufficiently small \( \eta > 0 \). More precisely, we have
\[
1 - \tau(\beta - 2) = 1 - (\beta - 2)^{\varepsilon/(2+\varepsilon)} = \Omega(\varepsilon),
\]
and thus for \( \eta > 0 \) sufficiently small compared to \( \varepsilon \) we have
\[
(4.2) \quad p_i \leq \exp \left( -w_i^{\Omega(\varepsilon)} \right).
\]
By the union bound, the probability that for every \( 0 \leq i < i_{\text{max}} \) the vertex \( v_i \) has a neighbor in the next weight layer is at least
\[
1 - \sum_i \exp(-w_i^{\Omega(\varepsilon)}) = 1 - \exp(-w_i^{\Omega(\varepsilon)}),
\]
which proves the first claim.

For the second statement, let \( C = C(\varepsilon) = \Omega(1/\varepsilon) \) with sufficiently large hidden constant. If a vertex \( v \) has weight at least \((\log n)^C \) then the error probability estimated above is at least \( 1 - e^{-\Omega(\varepsilon)} = 1 - n^{-\omega(1)} \). The claim now follows from a union bound over all vertices of weight at least \((\log n)^C \).

For the size of the giant component, we apply the same arguments as before in the proof of (i) for \( w = 2 \). Let \( \varepsilon > 0 \) be sufficiently small, let \( \eta > 0 \) sufficiently small compared to \( \varepsilon \), and consider the same system of weight layers \( V_i \) as before. Let \( i \geq 0 \) such that \( w_i \leq (\log n)^C \), where \( C \) is the constant \( C(\varepsilon) \) given by the proof of statement (ii). For every \( v \in V_i \), let \( \Gamma_i(v) := \{ u \in V_{i+1} \mid v \sim u \} \), and let \( E_i(v) := \mathbb{E}[\Gamma_i(v)] \). Moreover, let \( \gamma := \tau(2 - \beta - \eta) + 1 > 0 \). Then for every \( v \in V_i \), by (EP1),
\[
E_i(v) \geq \Omega \left( n w_i^{1-\beta-\eta} \cdot \frac{w_i w_{i+1}}{W} \right) \geq \Omega(w_i^{2-\beta-\eta} w_i) \geq \Omega(w_i^{\tau(2-\beta-\eta)+1}) \geq \Omega(w_i^{\gamma}).
\]
As this lower bound is independent of $v$, we also have

$$E_i := \min_{v \in V_i} E_i(v) = \Omega(w_i^\gamma).$$

Let $\lambda := \min\{\gamma, \frac{1}{2C\tau}\}$. Furthermore put $B_i := \{v \in V_i \mid |\Gamma_i(v)| \leq E_i/2\}$. This set will play the role of "bad" vertices.

**Claim 4.14.** There is a constant $c > 0$ such that for all $i \geq 0$ with $w_i \leq (\log n)^C$, w.h.p. it holds $|B_i| \leq 2 \exp(-c w_i^\lambda) \cdot |V_i|$.

We postpone the proof of Claim 4.14 (and Claim 4.15 below) until we have finished the main argument. We uncover the sets $V_i$ one by one, starting with the largest weights. Let $\delta > 0$ be so small that $\tau(\lambda - \delta) > \lambda$. Note that when applying Claim 4.14, we may replace the factor 2 by any other factor $D_1 \geq 2$ without violating the statement of the claim. We will show by induction that if $D_1 = O(1)$ is sufficiently large, then w.h.p. the fraction of vertices in $V_i$ with a weight-increasing path to the inner core is at least $1 - D_1 \exp(-c w_i^{\lambda - \delta})$. Note that for any $i_0 = i_0(c) = O(1)$ the statement is trivial for all $i \leq i_0$, if we choose $D_1 = D_1(i_0, c)$ sufficiently large. Also, if $w_i \geq (\log n)^C$ then we already know that w.h.p. all vertices in $V_i$ are connected to the inner core with weight-increasing paths. For the remaining values of $i$, denote by $V'_i$ the set of vertices in $V_i$ for which there is no weight-increasing path to the inner core that uses exactly one vertex per layer. Furthermore, let $\Lambda_i$ be the set those $D_1 \exp(-c w_i^{\lambda - \delta}) |V_i|$ vertices in $V_i$ with the smallest neighborhood in $V_{i+1} \setminus V'_{i+1}$ (where we break ties according to some previously fixed order). We then use the following claim.

**Claim 4.15.** There exists $D_2 > 0$ such that w.h.p., for all $i \geq i_0$ with $w_i \leq (\log n)^C$, it holds

$$|E(V_i, \Lambda_{i+1})| \leq D_1 \exp(-c w_i^{\lambda - \delta}) \cdot |V_i| \cdot E_i \cdot w_i^{D_2}.$$  

Consider some $i_0 < i \leq i_{\text{max}}$ such that $w_i \leq (\log n)^C$, and assume by induction that for sufficiently many vertices of $V_{i+1}$ there is a weight-increasing path to the inner core, that is, $|V'_{i+1}| \leq D_1 \exp(-c w_{i+1}^{\lambda - \delta}) \cdot |V_{i+1}|$. By construction, this implies $V'_{i+1} \subseteq \Lambda_{i+1}$. Now we consider $B'_i := \{v \in V_i \mid |E(\{v\}, V'_{i+1})| \geq E_i/2 \}$. If
the low-probability event of Claim 4.15 does not occur, using $w_i^{\lambda-\delta} = w_i^{\lambda+\Omega(1)}$ it follows

\begin{equation}
|B_i'| \leq \frac{2|E(V_i, V_{i+1}')|}{E_i} \leq 2D_1 \exp(-cw_i^{\lambda-\delta}) \cdot |V_i| \cdot w_i^{D_2} \leq D_1 \exp(-cw_i^{\lambda}) \cdot |V_i|,
\end{equation}

provided that $i_0 = i_0(c)$ (and thus, $w_{i_0}$) is a sufficiently large constant. It remains to observe that every vertex in $V_i \setminus (B_i \cup B_i')$ has at least one edge into $V_{i+1} \setminus V_{i+1}'$. Since the latter vertices are all connected to the inner core, we have at least $|V_i| - |B_i| - |B_i'|$ vertices in $V_i$ that are connected to the inner core. By Claim 4.14 and Equation (4.3), w.h.p. both $B_i$ and $B_i'$ have size at most $D_1 \exp(-cw_i^{\lambda}) |V_i|$, so together they have size at most $D_1 \exp(-cw_i^{\lambda-\delta}) |V_i|$, for all $i \geq i_0$ where $i_0$ is sufficiently large. This concludes the induction modulo Claims 4.14 and 4.15. The existence of the giant component now follows because w.h.p. a constant fraction of $V_{i_0}$ is connected to the inner core, and $V_{i_0}$ has linear size by (PL2).

**Proof of Claim 4.14** Let $i \geq 0$ such that $w_i \leq (\log n)^C$. For a single $v \in V_{i+1}$, the events “$v \sim u$” are independent for all $u \in V_{i+1}$. So by the Chernoff bound (Theorem 2.2), there is a constant $c > 0$ such that $\Pr[v \in B_i] \leq \exp(-cw_i^{\gamma})$ and $\mathbb{E}[|B_i|] \leq \exp(-cw_i^{\gamma})|V_i|$. Let $G_i$ be the subgraph induced by $V_i$ and $V_{i+1}$ and observe that the size of $B_i$ only depends on $G_i$. We aim to use Theorem 2.5 for proving the concentration of $|B_i|$. Similarly as in the proof of Theorem 4.10, we can assume that the considered probability space $\Omega$ is a product space of independent coordinates: for each $u \in V_i \cup V_{i+1}$ we have a random variable $x_u$ for its position and a random variable $Y_u$ for the incident edges in $G_i$. Again, for every $\omega \in \Omega$, we denote by $G_i(\omega)$ the resulting graph. We observe that changing one coordinate $x_u$ or $Y_u$ can only influence the degrees of $u$ itself and of the vertices that are neighbors of $u$ before or after the coordinate change.

We study the bad event $\mathcal{B}$ that there exists a vertex $v \in V_i \cup V_{i+1}$ with degree larger than $(\log n)^{2Cr^2}$ in $G_i$. Since $w_i \leq (\log n)^C$ and therefore $w_v \leq (\log n)^{Cr^2}$ for all $v \in V_i \cup V_{i+1}$, by Lemma 3.4 we have $\Pr[\mathcal{B}] = n^{-\omega(1)}$. Let $\omega, \omega' \in \overline{\mathcal{B}}$ such that they differ in at most two coordinates of our product probability space $\Omega$. We observe that changing one coordinate $x_u$ or $Y_u$ can only influence the degrees of
$u$ itself and of the vertices that are neighbors of $u$ before or after the coordinate change. Therefore, $|B_i(\omega) - |B_i(\omega')|| \leq (\log n)^{O(1)}$. We pick $t = \exp(-cw_i^4) \cdot |V_i|$ and observe that $w_i^4 \leq (\log n)^{1/2}$ by our choice of $\lambda$. Then Theorem 2.5 implies

$$
\Pr[|B_i| - \mathbb{E}[|B_i|] \geq t] \leq 2 \exp\left(-\frac{t^2}{128|V_i|(\log n)^{O(1)}}\right) + n^{O(1)} \Pr[\mathcal{B}]
$$

$$
\leq 2 \exp\left(-\frac{e^{-cw_i^4}|V_i|}{128(\log n)^{O(1)}}\right) + n^{O(1)} \Pr[\mathcal{B}] = n^{-\omega(1)}.
$$

Hence, w.h.p. we have $|B_i| \leq \mathbb{E}[|B_i|] + t \leq (\exp(-cw_i^4) + \exp(-cw_i^4)) \cdot |V_i|$. For $i$, the statement now follows since $\lambda < \gamma$ and $w_i > 1$, and then the proof of the claim is finished by a union bound over all $O(\log \log n)$ choices of $i$.  

**Proof of Claim 4.15** Let $i \geq i_0$ such that $w_i \leq (\log n)^C$. We assume that the subgraph induced by $V_{\geq w_{i+2}}$ is given, and now we uncover $V_i$ and $V_{i+1}$ to obtain the subgraph induced by $V_{\geq w_i}$. Similarly as in the proof of Claim 4.14 we can assume that the considered probability space $\Omega$ is a product probability space with $2(|V_i| + |V_{i+1}|)$ coordinates. Recall that $G_i$ denotes the subgraph induced by $V_i \cup V_{i+1}$. We consider the same bad event $\mathcal{B}$ as in the proof of Claim 4.14, i.e., $\mathcal{B}$ denotes the event that the maximum degree in $G_i$ is larger than $(\log n)^{2C^2}$. Note that $\mathcal{B}$ is independent of $V_{\geq w_{i+2}}$, so indeed Lemma 3.4 can be again applied to deduce $\Pr[\mathcal{B}] = n^{-\omega(1)}$.

Let $Z_i := |E(V_i, \Lambda_{i+1})|$, and let $\omega, \omega' \in \overline{\mathcal{B}}$ such that they differ in at most two coordinates of $\Omega$. If we change a coordinate of $\Omega$ that stems from a vertex $v \in V_i$, under $\overline{\mathcal{B}}$ the influence on $Z_i$ is at most $(\log n)^{O(1)}$. If a coordinate belonging to a vertex of $V_{i+1}$ is changed, this may result in a different set $\Lambda_{i+1}$. However, the symmetric difference between the old $\Lambda_{i+1}$ and the new $\Lambda_{i+1}$ is at most two (as $\Lambda_{i+1}$ is defined via a fixed ordering), and under $\overline{\mathcal{B}}$ the influence on $Z_i$ is again upper-bounded by $(\log n)^{O(1)}$. Finally, the same is true if the set $\Lambda_{i+1}$ does not change. Repeating the argument for the second coordinate change, we conclude that $|Z_i(\omega) - Z_i(\omega')| \leq (\log n)^{O(1)}$.

Next, we want to upper-bound $\mathbb{E}[Z_i]$. First, we uncover $V_{i+1}$ to obtain the subgraph induced by $V_{\geq w_{i+1}}$. Then the set $\Lambda_{i+1}$ is determined. In a second step,
we uncover $V_i$. By (EP1) and linearity of expectation, we deduce

$$
\mathbb{E}[Z_i] \leq |\Lambda_{i+1}| \cdot |V_i| \cdot O\left(\frac{w_i^{1+\eta} w_{i+1}^{1+\eta}}{W}ight)
$$

$$
= O\left(D_1 \exp(-cw_i^{\lambda-\delta}) n w_i^{1-\beta+\eta} \cdot |V_i| \cdot \frac{w_i^{1+\eta} w_{i+1}^{1+\eta}}{W}\right)
$$

$$
\leq D_1 \exp(-cw_i^{\lambda-\delta}) \cdot |V_i| \cdot O\left(w_i^{2-\beta+2\eta} w_i^{1+\eta}\right)
$$

$$
\leq D_1 \exp(-cw_i^{\lambda-\delta}) \cdot |V_i| \cdot O\left(w_i^{(2-\beta)+1+\eta(2\tau+1)}\right)
$$

$$
\leq D_1 \exp(-cw_i^{\lambda-\delta}) \cdot |V_i| \cdot E_i \cdot O\left(w_i^{\eta(3\tau+1)}\right).
$$

Since we assumed $w_i \geq 2$, we may upper-bound the $O(\cdot)$-term by $0.5 w_i^{D_2}$ for a sufficiently large $D_2 > 0$.

Now we can apply Theorem 2.5 with $t = 0.5 D_1 \exp(- cw_i^{\lambda-\delta}) \cdot |V_i| \cdot E_i \cdot w_i^{D_2}$. Using $\lambda \tau C \leq \frac{1}{2}$ by our choice of $\lambda$, it follows similarly as in the proof of Claim 4.14 that $\Pr\left[|Z_i - \mathbb{E}[Z_i]| \geq t\right] = n^{-\omega(1)}$, and we conclude that with probability $1 - n^{-\omega(1)}$ it holds

$$
|Z_i| \leq \mathbb{E}[|Z_i|] + t \leq D_1 \exp(-cw_i^{\lambda-\delta}) \cdot |V_i| \cdot E_i \cdot w_i^{D_2}.
$$

Now the claim follows by a union bound over all $O(\log \log n)$ choices of $i$. ■

By Lemma 4.13 (ii), w.h.p. every vertex of weight at least $(\log n)^C$ has small distance from the heavy core. It remains to show that every vertex in the giant component has a large probability to connect to such a high-weight vertex in a small number of steps. The next lemma shows that the more vertices of small weight we have in the neighborhood of a vertex, the more likely it is that there is an edge from the neighborhood to a vertex of large weight.

**Lemma 4.16 (Bulk lemma).** Let $\epsilon > 0$. Let $w_{\min} \leq w \leq \overline{w}$ be a weight, and let $k \geq \max\{2, w^{\beta+\epsilon}\}$ be an integer. For a vertex $v \in V_{<w}$, let $N_v$ be the set of all vertices within distance at most $k$ of $v$ in the graph $G_{<w}$. Then for a random vertex $v \in V_{<w}$,

$$
\Pr\left[\text{dist}(v, V_{\geq w}) > k \text{ and } |N_v| \geq k\right] \leq O\left(e^{-w^{\Omega(1)}}\right).
$$

**Proof.** Before starting with the formal proof, let us sketch some of the main ideas. We first uncover the graph $G_{<w}$ induced by vertices of weights less than...
\[ w. \] For a fixed vertex \( v \in V_{<w} \) and a vertex \( u \in V_{\geq w} \), we know (a lower bound on) \( \mathbb{E}[|N_v \cap \Gamma(u)||G_{<w}] \) by Lemma 3.2. Then two cases can occur: either \( \Pr[u \sim N_v] \) is large, or we have a relatively large probability that \( u \) connects to many vertices of \( N_v \) at the same time. If \( N_v \) is such that we are in the first case, it is not difficult to conclude that it is very unlikely that there is no edge between \( N_v \) and \( V_{\geq w} \). Now suppose that the second case occurs for a large fraction of the vertices in \( V_{<w} \). In this case we carefully choose a set \( V_u \subseteq V_{<w} \) that has in particular the property that the sets \( N_v \) for \( v \in V_u \) are disjoint. Such a situation would significantly increase the probability that \( \deg(u) \) is large, which would contradict Lemma 3.3. We can deduce that it is unlikely that the second case happens for a large fraction of \( V_{<w} \), from which the statement follows.

We may assume \( w \leq n^{1/2} \), since otherwise \( k > n \), and the statement is trivial. Let \( c > 0 \) be such that for all vertices \( u \) of weight at least \( w \), all vertices \( u' \in V \), and every fixed position \( x_{u'} \in \mathcal{X} \) we have \( \Pr[u \sim u' | x_{u'}] \geq cw/n \), i.e., \( c \) is the hidden constant of condition (EP1). Finally by the power-law assumption (PL2), for any sufficiently small \( 0 < \eta < 1 \) we may choose \( \bar{w} = O(w^{1+\eta}) \) such that there are at least \( \Omega(n/w^{\beta-1+\eta}) \) vertices with weights between \( w \) and \( \bar{w} \).

We first uncover the graph \( G_{<w} \) induced by vertices of weight less than \( w \), i.e., we uncover the positions of these vertices and the edges in the induced subgraph. Let \( v \in V_{<w} \) and let \( N_v := N_v(k, w) \) be the \( k \)-neighborhood of \( v \) in \( G_{<w} \). Once \( G_{<w} \) is fixed, consider a random vertex \( u \), conditioned on \( w_u \in [w, \bar{w}] \). Let \( R := R(v) := \Pr_u[u \sim N_v | G_{<w}] \).

**Claim 4.17.** \( Q := \Pr_u[|N_v \cap \Gamma(u)| \geq cw|N_v|/(2nR) | G_{<w}] \geq \frac{cw}{2n} \).

**Proof.** Let \( x := cw|N_v|/(2nR) \). We first use \( |N_v \cap \Gamma(u)| \leq |N_v| \) to bound

\[
\mathbb{E}[|N_v \cap \Gamma(u)||G_{<w}] \leq Q|N_v| + (R - Q)x \leq Q|N_v| + Rx.
\]

On the other hand, the left hand side is at least \( cw|N_v|/n \) by our choice of \( c \). Together, it follows \( Q \geq cw/n - Rx|N_v| = cw/(2n) \), proving the claim. \( \square \)

Now we distinguish three cases for the vertex \( v \). (1) If \( |N_v| < k \) then obviously
there is nothing to show. (2) If \( R \geq w^\beta / n \), then

\[
\mathbb{E}[|\{u \mid w_u \in [w, \tilde{w}] \text{ and } u \sim N_v| \rangle G_{<w}, R \geq w^\beta / n ] \geq \Omega\left( \frac{w^\beta}{n} \cdot \frac{n}{w^\beta - 1 + \eta} \right) = \Omega(w^{\Omega(1)}),
\]

because the number of vertices of weight in \([w, \tilde{w}]\) is at least \( \Omega(\frac{n}{w^\beta - 1 + \eta}) \) by (PL2). Since every \( u \) draws its position and its edges to \( V_{\leq w} \) independently from each other, we may apply the Chernoff bounds and obtain

\[
\Pr[\exists u \mid w_u \in [w, \tilde{w}] \text{ and } u \sim N_v| \rangle G_{<w}, R \geq w^\beta / n ] \geq 1 - O(e^{-\Omega(w)}),
\]

as desired.

(3) For the last case, \( |N_v| \geq k \) and \( R < w^\beta / n \), we will show that it is very unlikely that this case occurs for a random \( v \) (over a random choice in \( V_{<w} \)). More precisely, let \( V_R \subseteq V_{<w} \) be the set of vertices \( v \) of weight less than \( w \) for which \( |N_v| \geq k \) and \( R(v) < w^\beta / n \). Further, let \( \mathcal{E} \) be the event that \( |V_R| \geq ne^{-c'w} \), where \( c' \) is a constant to be fixed later. Then we will show that \( \Pr[\mathcal{E}] = e^{-\Omega(w)} \).

Note that with this statement, we can conclude the proof as follows. Let \( v \) be a random vertex of weight less than \( w \). When we uncover \( G_{<w} \), then \( \mathcal{E} \) occurs only with probability \( e^{-\Omega(w)} \). On the other hand, if \( \mathcal{E} \) does not occur, then there at most \( ne^{-c'w} \) vertices \( v' \in V_{<w} \) for which \( |N_{v'}| \geq k \) and \( R(v') < w^\beta / n \), and the probability that \( v \) is among them is at most

\[
\frac{ne^{-c'w}}{|V_{<w}|} = O\left( \frac{ne^{-c'w}}{n(1 - w^{1-\beta + \eta})} \right) = O(e^{-\Omega(w)})
\]

for any \( \eta > 0 \). Finally, if \( v \) is not among these vertices, then either \( |N_v| < k \), and we are done, or \( R(v) \geq w^\beta / n \), and then \( N_v \not\sim V_{\geq w} \) with probability at most \( O(e^{-w^{\Omega(1)}}) \) by (4.4). Thus the theorem follows by a union bound, and it remains to show the following claim.

**Claim 4.18.** Let \( V_R := \{ v \in V_{<w} \mid |N_v| \geq k \text{ and } R(v) < w^\beta / n \} \) and denote by \( \mathcal{E} \) the event that \( |V_R| \geq ne^{-c'w} \). Then

\[
\Pr[\mathcal{E}] = O(e^{-\Omega(w)}).
\]
Before we prove Claim 4.18, we need some preparation. Sort the vertices \( v \in V_R \) decreasingly by \(|N_v|\). We go through the list one by one, and pick greedily a set \( V_{Gr} \subseteq V_R \) such that the \( N_v, v \in V_{Gr} \) are pairwise disjoint. Then after this procedure, the following holds.

**Claim 4.19.** \( \sum_{v \in V_{Gr}} 2|N_v|^5 \geq |V_R| \).

**Proof of Claim 4.19.** We prove Claim 4.19 by the following charging argument. Whenever we pick a vertex \( v \) to be included into \( V_{Gr} \), we inductively define levels \( L_s(v) \subseteq V_R \), \( s \geq 0 \) by \( L_0(v) := \{ u \in N_v \mid |N_u| \leq |N_v|^2 \} \) and

\[
L_{s+1}(v) := \bigcup_{v' \in L_s(v)} \{ u \in N_{v'} \mid |N_u| \leq |N_{v'}|^{2^{-s}} \}.
\]

The vertex \( v \) pays one coin to each vertex in \( \bigcup_{s \geq 0} L_s(v) \). We claim that (i) every vertex \( v \) that we pick pays at most \( 2|N_v|^5 \) coins, and (ii) every vertex in \( V_R \) is paid at least one coin. Note that (i) and (ii) together will imply Claim 4.19.

To prove (i), we observe that \(|L_0(v)| \leq |N_v|\) and \(|L_1(v)| \leq |L_0(v)| \leq |N_v|^2\) by definition of \( L_0(v) \), and \(|L_{s+1}(v)| \leq |L_s(v)|^{2^{-s}} \) for all \( s \geq 1 \) by definition of \( L_s(v) \). Therefore, \(|L_s(v)| \leq |N_v|^{1+2+\sum_{j=1}^{s} 2^{1-j}}\). Moreover, for all \( s > s_0 := \lfloor \log_2 \log_k |N_v| \rfloor \) we have \(|N_v|^{2^{-s}} < k\), so \( L_{s+1} = \emptyset \) by definition of \( V_R \). On the other hand, for all \( s \leq s_0 \) we have \(|N_v|^{2^{-s}} \geq k \geq 2\), thus the terms \(|N_v|^{3+\sum_{j=1}^{s} 2^{1-j}}\) increase at least geometrically fast for \( s \leq s_0 \). Hence,

\[
\sum_{s=0}^{s_0} |L_s(v)| \leq \sum_{s=0}^{s_0} |N_v|^{3+\sum_{j=1}^{s} 2^{1-j}} \leq 2|N_v|^{3+\sum_{j=1}^{\infty} 2^{1-j}} \leq 2|N_v|^5,
\]

proving (i).

For (ii), we show the following statement inductively for all vertices \( v \). After \( v \) has paid its coins, every vertex \( u \) which comes after \( v \) in the ordering, and for which \( N_u \cap N_v \neq \emptyset \) holds, has received at least one coin. Note that it will follow that each vertex that we consider and that we do not pick has been paid by an earlier vertex. So assume that \( u \) comes after \( v \) in the ordering, and that \( N_u \cap N_v \neq \emptyset \). Since we go through the vertices in descending order with respect to \(|N_v|\), we have \(|N_u| \leq |N_v|\). Let \( v' \in N_u \cap N_v \). If \(|N_{v'}| \leq |N_v|^2\), then \( v' \in L_0 \) and \( u \in L_1 \), so \( v \) pays to \( u \). If \(|N_{v'}| > |N_v|^2\), then we have considered \( v' \) before \( v \).
However, since we picked $v$, and since $v' \in N_v$ (and thus, $v \in N_{v'}$), $v'$ was not picked. Therefore, by induction hypothesis $v'$ had been paid by some earlier vertex $v''$, so $v' \in L_s(v'')$ for some $s \geq 0$. Since $|N_u| \leq |N_{v'}| < |N_{v''}|^{2^{-s}}$, we obtain $u \in L_{s+1}(v'')$, so $u$ has been paid by $v''$ as well. This proves (ii), and thus concludes the proof of Claim 4.19. Note that $2\sum_{v \in V_{Gr}} |N_v|^5 \geq |V_R| \geq ne^{-c'}w$ if $E$ holds.

\[ \square \]

**Proof of Claim 4.18.** With Claim 4.19, we can finally prove Claim 4.18 as follows. Fix a vertex $u$ such that $w_u \leq \tilde{w}$. Then for each position $x_u$ of $u$, the expected degree of $u$ conditioned on $x_u$ is in $O(\tilde{w})$, and it is the sum of independent random variables by Lemmas 3.2 and 3.3. Note that the hidden constant in the $O(\cdot)$-notation is independent of $w_u$ and of $x_u$. Therefore, by the Chernoff-Hoeffding bound (Theorem 2.2 (iii)), there are constants $c', C > 0$ independent of $w_u$ and $x_u$ such that $\Pr[\deg(u) \geq i] \leq e^{-2c'i}$ for all $i \geq C\tilde{w}$, and this also holds if $u$ is a random vertex with weight in $[w, \tilde{w}]$. So let $u$ be a random vertex with weight in $[w, \tilde{w}]$, and let $V_u := \{v \in V_{Gr} | |N_v \cap \Gamma(u)| \geq |N_v|cw^{1-\beta}/2\}$.

Consider the random variables

$$S_1 := 2 \sum_{v \in V_u} |N_v|^5 \quad \text{and} \quad S_2 := \frac{cw^{1-\beta}}{2} \sum_{v \in V_u} |N_v|.$$ 

Note that $S_2 \leq \deg(u)$ by definition of $V_u$, and since all $v \in V_u \subseteq V_{Gr}$ have disjoint $N_v$ by construction. Hence, $\Pr[S_2 \geq i] \leq e^{-2c'i}$ for all $i \geq C\tilde{w}$. Now consider the expectation of $S_1$ conditioned on $E$. On the one hand, since we are in the case $R < w^{\beta}/n$, we have $|N_v|cw^{1-\beta}/2 < |N_v|cw/(2nR)$, and thus $\Pr[v \in V_u | v \in V_{Gr}] \geq cw/(2nR)$. By Claim 4.17, we have

$$\mathbb{E}[S_1 | E] \geq \frac{cw}{n} \sum_{v \in V_{Gr}} |N_v|^5 \geq \frac{cw}{2} e^{-c'w}.$$ 

On the other hand, since $\sum_{v \in V_u} |N_v|^5 \leq (\sum_{v \in V_u} |N_v|)^5$, we may lower-bound
$S_1 \leq 2 \cdot (2w^\beta - 1) S_2/2^5$. Both inequalities together yield

$$\frac{cw e^{-c'w}}{2} \leq \mathbb{E}[S_1 \mid \mathcal{E}] \leq 2 \cdot \left(\frac{2w^\beta - 1}{c} \right)^5 \cdot \mathbb{E}[S_2^5 \mid \mathcal{E}] = 2 \cdot \left(\frac{2w^\beta - 1}{c} \right)^5 \sum_{i \geq 1} i^5 \Pr[S_2 = i \mid \mathcal{E}]$$

$$\leq 2 \cdot \left(\frac{2w^\beta - 1}{c} \right)^5 \cdot \frac{\sum_{i \geq 1} i^5 \Pr[S_2 = i]}{\Pr[\mathcal{E}]}.$$ 

Solving for $\Pr[\mathcal{E}]$ yields $\Pr[\mathcal{E}] \leq w^{O(1)} e^{c'} w \sum_{i \geq 1} i^5 \Pr[S_2 = i]$. Observe that $S_2 > 0$ already implies $S_2 > cw^{1-\beta}k/2 = w(\tilde{w})$, since $|N_v| \geq k$ for all $v \in V_{Gr}$. So if $w$ is sufficiently large then the first $C\tilde{w}$ terms of $\sum_{i \geq 1} i^5 \Pr[S_2 = i]$ vanish. On the other hand, recall that $\Pr[S_2 \geq i] \leq e^{-2c'i}$ for all $i \geq C\tilde{w}$. Hence, if $w$ is sufficiently large,

$$\Pr[\mathcal{E}] \leq w^{O(1)} e^{c'} w \sum_{i \geq C\tilde{w}} i^5 \Pr[S_2 \geq i] \leq w^{O(1)} e^{c'} w \sum_{i \geq C\tilde{w}} i^5 e^{-2c'i}$$

$$= \tilde{w}^{O(1)} e^{-\Omega(\tilde{w})} = O(e^{-\Omega(w)}).$$

This concludes the proof of Claim 4.18, and thus of the lemma. □

The upper bounds on the diameter and the average distance now follow easily from the lemmas we proved so far. We collect the results for $\mathcal{G}(n, \mathcal{X}, w, p)$ in the following theorem, which reformulates and specifies Theorem 4.2 and Theorem 4.3.

**Theorem 4.20 (Components and Distances).**

(i) W.h.p., there is a giant component, i.e., a connected component which contains $\Omega(n)$ vertices.

(ii) W.h.p., all other components have at most polylogarithmic size.

(iii) W.h.p., the giant component has polylogarithmic diameter.

(iv) In expectation and a.a.s., the average distance (i.e., the expected distance of two uniformly random vertices in the largest component) is $\frac{2 + o(1)}{\log(\beta - 2)} \log \log n$. 

(v) A.a.s., a \((1 - o(1))\)-fraction of all pairs of vertices in the giant component have distance at most \(\frac{2 + o(1)}{\log \beta - 2} \log \log n\).

**Proof.** (i) has been proven with Lemma 4.13 (iii). For (ii) and (iii) we fix a sufficiently small constant \(\epsilon > 0\) and conclude from the same lemma that w.h.p. the giant contains all vertices of weight at least \(w := (\log n)^C\), for a suitable constant \(C > 0\), and that w.h.p. all such vertices have distance at most \(\frac{1 + \epsilon}{\log \beta} \log \log n\) from the heavy core \(\tilde{V}\). We apply Lemma 4.16 with \(\ell = w^{\beta + \epsilon}\). Then a random vertex in \(V_{< w}\) has probability at least \(1 - e^{-w^{O(1)}}\) to either be at distance at most \(\ell \log \log n\) from \(\tilde{V}\), or to be in a component of size less than \(\ell\). Note that for sufficiently large \(C\) this probability is at least \(1 - n^{-o(1)}\). By the union bound, w.h.p. one of the two options happens for all vertices in \(V_{< w}\). This already shows that w.h.p. all non-giant components are of size less than \(\ell = (\log n)^{O(1)}\). For the diameter of the giant, recall that w.h.p. the heavy core has diameter \(o(\log \log n)\) by Lemma 4.12. Therefore, w.h.p. the diameter of the giant component is \(O(\ell + \log \log n) = (\log n)^{O(1)}\).

For the average distance, let \(\epsilon > 0\), and let \(v \in V\) be a vertex chosen uniformly at random. Fix \(\ell \geq 3\), \(\ell = n^{o(1)}\), and let \(w := w(\ell) = \ell^{1/(\beta + 1)}\). We sort the vertices by weight and uncover the graph vertex by vertex in increasing order, until either (1) we see for the first time a vertex \(v' \in V_{\geq w}\) such that in the subgraph induced by \(V_{\leq w, v'}\) there exists a path of length at most \(\ell\) from \(v\) to \(v'\), or (2) we have uncovered the full graph and (1) never happened. If \(w_v \geq w\), then (1) trivially occurs. Otherwise, by Lemma 4.16, with probability \(1 - O(\exp(-w^{O(1)})\)) either case (1) happens or the connected component of \(v\) in \(G\) has size less than \(\ell\). In the latter case, \(v\) is not connected to the core and there is nothing to show. Otherwise, we have uncovered only the vertices of weight at most \(w_{v'}\), which allows us to apply Lemma 4.13 (i) since its statement only depends on vertices of higher weight. By Lemma 4.13 (i), with probability \(1 - O(\exp(-w^{O(\epsilon)})\)) there is a weight-increasing path from \(v'\) to the heavy core of length at most \(\lambda_{\epsilon} := (1 + \epsilon)^{\frac{\log \log n}{\log(\beta - 2)}}\). Summarizing, we have shown that for a random vertex \(v\) and every \(\ell \geq 3\) with \(\ell = n^{o(1)}\)

\[
(4.6) \quad \Pr[\infty > \text{dist}(v, V_{\text{core}}) \geq \ell + \lambda_{\epsilon}] \leq e^{-\Omega(w(\ell)^{O(\epsilon)})} = O(e^{-\ell^{O(\epsilon)}}).
\]
Let us first consider the expectation of the average distance, i.e., if $u, u'$ denote random vertices in the largest component of a random graph $G$ then we consider $\mathbb{E}_G[\mathbb{E}_{u, u'}[\text{dist}(u, u')]]$. Since $\text{dist}(u, u') \leq n$ we can condition on any event happening with probability $1 - n^{-o(1)}$, in particular we can condition on the event $\mathcal{E}$ that $G$ has a giant component containing $V_{\text{core}}$, all other components have size $(\log n)^O(1)$, $G$ has diameter $(\log n)^O(1)$, and finally the core has diameter $d_{\text{core}} = o(\log \log n)$. Moreover, by bounding

$$\text{dist}(u, u') \leq \text{dist}(u, V_{\text{core}}) + \text{dist}(u', V_{\text{core}}) + d_{\text{core}}$$

it suffices to bound $2 \cdot \mathbb{E}_G[\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \mid \mathcal{E}] + d_{\text{core}}$. Now we use the fact $\mathbb{E}[X] = \sum_{\ell > 0} \Pr[X \geq \ell]$ for a random variable $X$ taking values in $\mathbb{N}_{\geq 0}$ to bound

$$\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \leq \lambda_{\epsilon} + \sum_{\ell = 1}^{(\log n)^O(1)} \Pr_u[\text{dist}(u, V_{\text{core}}) \geq \ell + \lambda_{\epsilon}].$$

Note that conditioned on $\mathcal{E}$, since $u$ is chosen uniformly at random from the giant component, $\text{dist}(u, V_{\text{core}}) < \infty$. Taking expectation over $G$, conditioned on $\mathcal{E}$, we may use (4.6) to bound the probability that $\text{dist}(v, V_{\text{core}})$ is too large for a vertex chosen uniformly at random from $V$. Since the giant has size $\Omega(n)$, this probability increases at most by a constant factor if we instead choose $v$ uniformly at random from the giant. Hence, for every constant $\epsilon > 0$ we obtain

$$\mathbb{E}_G[\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \mid \mathcal{E}] \leq \lambda_{\epsilon} + \sum_{\ell = 1}^{(\log n)^O(1)} O(e^{-\ell \Omega(\epsilon)}) + n^{-o(1)}.$$  

We now use the inequality

$$\sum_{\ell = 1}^{\infty} e^{-\ell \kappa} \leq \int_{x = 0}^{\infty} e^{-x\kappa} \, dx = \Gamma(1 + 1/\kappa),$$

where $\Gamma$ is Euler’s Gamma function. Since $\Gamma(x)$ is monotonically increasing on the real axis for $x \geq 2$ and $\Gamma(1 + n) = n!$, we have $\Gamma(1 + 1/\kappa) \leq [1/\kappa]! \leq (1/\kappa)^{O(1/\kappa)}$ for $\kappa \leq 1$. Plugging this into equation (4.7) yields

$$\mathbb{E}_G[\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \mid \mathcal{E}] \leq \lambda_{\epsilon} + O(1/\epsilon)^{O(1/\epsilon)} + n^{-o(1)}.$$
Note that for sufficiently slowly falling $\varepsilon = \varepsilon(n) = o(1)$ we have

$$O(1/\varepsilon)^{O(1/\varepsilon)} = o(\log\log n).$$

This yields the desired bound on the expected average distance of

$$2\lambda o(1) + o(\log\log n) = \frac{2 + o(1)}{\log|\beta - 2|} \log\log n.$$

For the concentration, we want to show $Pr_G[\mathbb{E}_{u, u'}[\text{dist}(u, u')] \geq 2\lambda \varepsilon] = o(1)$, where we choose the same $\varepsilon(n) = o(1)$ as before. Similarly as before, we may bound

$$Pr_G[\mathbb{E}_{u, u'}[\text{dist}(u, u')] \geq 2\lambda \varepsilon] \leq n^{-\omega(1)} + Pr_G[2 \cdot \mathbb{E}_u[\text{dist}(u, V_{\text{core}})] + d_{\text{core}} \geq 2\lambda \varepsilon | \mathcal{E}].$$

Let $\gamma > 0$ be a sufficiently small constant, and let

$$\rho = \rho(n) = \frac{\varepsilon}{3|\log(\beta - 2)| \log\log n} = o(\log\log n).$$

We claim that for sufficiently large $n$, $2 \cdot \mathbb{E}_u[\text{dist}(u, V_{\text{core}})] + d_{\text{core}} \geq 2\lambda \varepsilon$ can only happen if for some $\ell > \rho$ we have

$$(4.8) \quad Pr_u[\text{dist}(u, V_{\text{core}}) \geq \ell + \lambda \varepsilon/3] \geq e^{-2\ell \gamma \varepsilon}.$$

Indeed, otherwise we have (conditioned on $\mathcal{E}$), similarly as before

$$\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \leq \lambda \varepsilon/3 + \rho + \sum_{\ell = \rho}^{(\log n)^{O(1)}} Pr_u[\text{dist}(u, V_{\text{core}}) \geq \ell + \lambda \varepsilon/3] \leq \lambda_{2\varepsilon/3} + O(1/\varepsilon)^{O(1/\varepsilon)},$$

and thus, indeed $\mathbb{E}_{u, u'}[\text{dist}(u, u')]$ is at most

$$2 \cdot \mathbb{E}_u[\text{dist}(u, V_{\text{core}})] + d_{\text{core}} \leq 2\lambda_{2\varepsilon/3} + o(\log\log n) < 2\lambda \varepsilon = 2\lambda o(1),$$

if $\varepsilon = \varepsilon(n) = o(1)$ decreases sufficiently slowly compared to $(\rho + d_{\text{core}}) / \log\log n$.

However, using the union bound over all $\rho \leq \ell \leq (\log n)^{O(1)}$, the probability that $G$ is such that $(4.8)$ holds for some $\ell > \rho$ is bounded from above by

$$\sum_{\ell = \rho}^{(\log n)^{O(1)}} Pr_G[Pr_u[\text{dist}(u, V_{\text{core}}) \geq \ell + \lambda \varepsilon/3] \geq e^{-\ell \gamma \varepsilon} | \mathcal{E}].$$
By (4.6) it follows that $\mathbb{E}_G[\Pr_{u}[\infty > \text{dist}(u, V_{\text{core}}) \geq \ell + \lambda_{\varepsilon/3}]] \leq O(\exp(-2\ell^{\gamma}\varepsilon))$, for $\gamma > 0$ sufficiently small. We apply Markov’s inequality (Theorem 2.1) and deduce that

$$\Pr_G[\Pr_{u}[\infty > \text{dist}(u, V_{\text{core}}) \geq \ell + \lambda_{\varepsilon/3} > e^{-\ell_{\gamma}\varepsilon}] \leq O(e^{-\ell_{\Omega}(\varepsilon)}).$$

Because the giant has linear size, this probability increases at most by a constant factor if we instead draw $v$ from the giant component (conditioned on $\mathcal{E}$). Thus, the desired probability is bounded by

$$\sum_{\ell = \rho}^\infty \frac{(\log n)^{O(1)}}{\rho} O(e^{-\ell_{\Omega}(\varepsilon)}),$$

which is $o(1)$, since $\rho = \varepsilon \cdot \log \log n$ grows sufficiently quickly compared to a sufficiently slowly falling $\varepsilon = o(1)$. This shows the concentration of the average distance and proves statement (iv).

Regarding the last statement (v), (4.9) shows that a.a.s. $G$ is such that (4.8) does not hold for any $\ell > \rho$. However, in this case the fraction of pairs $\{u, u'\}$ of vertices in the giant that have distance at least $2\ell + 2\lambda_{\varepsilon/3}$ is at most $e^{-4\ell_{\gamma}\varepsilon}$. By taking $\ell = 2\rho$ and assuming that $\rho = \varepsilon \log \log n$ grows sufficiently quickly compared to $\varepsilon = o(1)$, we see that a $(1 - o(1))$-fraction of pairs $\{u, u'\}$ has distance at most $2\lambda_{\varepsilon/3}$, given that $\rho = \varepsilon \cdot \log \log n$ grows sufficiently fast compared to $\varepsilon = o(1)$. This finishes the proof of Theorem 4.20.

\[\Box\]

### 4.4 Clustering

In this section we study the clustering coefficient of Standard GIRGs and Threshold GIRGs and give a proof of Theorem 4.4. We start by stating the formal definition of the clustering coefficient.

**Definition 4.21.** In a graph $G = (V, E)$ the clustering coefficient of a vertex $v \in V$ is defined as

$$\text{CC}(v) := \text{CC}_G(v) := \begin{cases} \frac{|\{\{u, u'\} \subseteq \Gamma(v) | u \sim u'\}|}{\binom{\deg(v)}{2}}, & \text{if } \deg(v) \geq 2, \\ 0, & \text{otherwise,} \end{cases}$$
and the (mean) clustering coefficient of $G$ is

$$
\text{CC}(G) := \frac{\sum_{v \in V} \text{CC}(v)}{|V|}.
$$

**Proof of Theorem 4.4.** Let $V' := V_{\leq n^{1/8}}$ and $G' = G[V']$. We first show that for the subgraph $G'$ we have $\mathbb{E}[\text{CC}(G')] = \Omega(1)$. Let $w_0 = \Theta(1)$ be a weight such that there are linearly many vertices with weight at most $w_0$ and denote by $\bar{V} := V_{\leq w_0}$ the set of vertices of weight at most $w_0$.

Since $\text{CC}(G') = \frac{1}{|V'|} \sum_{v \in V'} \text{CC}_{G'}(v) = \Theta\left(\frac{1}{n} \sum_{v \in \bar{V}} \text{CC}_{G'}(v)\right)$, it suffices to show that a vertex $v$ of weight at most $w_0$ fulfills $\mathbb{E}[\text{CC}_{G'}(v)] = \Omega(1)$. Fix a vertex $v \in \bar{V}$ at position $x_v \in \mathbb{T}^d$, and let $U(v)$ be the ball around $x_v$ with radius $cn^{-1/d}$ for a sufficiently small constant $c > 0$. Clearly the volume of this ball is $\Theta(n^{-1})$. Thus, the expected number of vertices in $\bar{V}$ with position in $U(v)$ is $\Theta(1)$. We now consider the event $\mathcal{E} = \mathcal{E}(v)$ that the following three properties hold.

(i) $v$ has at least two neighbors in $\bar{V}$ with positions in $U(v)$.

(ii) $v$ does not have neighbors in $\bar{V}$ with positions in $\mathbb{T}^d \setminus U(v)$.

(iii) $v$ does not have neighbors in $V \setminus \bar{V}$.

We claim that $\Pr[\mathcal{E}] = \Theta(1)$. For (i), note that the expected number of vertices in $\bar{V}$ with position in $U(v)$ is $\Theta(1)$. Since the position of every vertex is independent, by Le Cam’s theorem (Theorem 2.3) the probability that there are at least two vertices in $\bar{V}$ with position in $U(v)$ is $\Theta(1)$. Moreover, by (3.5) and (3.6), for each such vertex the probability to connect to $v$ is $\Theta(1)$ (in the Threshold GIRG with $\alpha = \infty$ this follows because we have chosen $c$ small enough), so (i) holds with probability $\Theta(1)$. For (ii), for any vertex $u \in \bar{V} \setminus \{v\}$, we can bound

$$
\Pr[v \sim u \mid x_u \in \mathbb{T}^d \setminus U(v)] \leq \Pr[v \sim u] = \Theta(1/n).
$$

Hence, by Le Cam’s theorem, (ii) holds with probability $\Theta(1)$, and this probability can only increase if we condition on (i). Finally, for every fixed position $x$, (iii) holds independently of (i) or (ii) with probability $\Theta(1)$, again by Le Cam’s theorem. This proves the claim that $\Pr[\mathcal{E}] = \Theta(1)$. 
Conditioned on $\mathcal{E}$, let $v_1$ and $v_2$ be two random neighbors of $v$. Then $x_{v_1}, x_{v_2} \in U(v)$, and $w_{v_1}, w_{v_2} \leq w_0$. Furthermore, by the triangle inequality we obtain $\|x_{v_1} - x_{v_2}\| \leq 2cn^{-1/d}$. For $c$ sufficiently small, we deduce from (3.5) and (3.6) that $v_1 \sim v_2$ holds with probability $\Theta(1)$. Thus we have shown that $\mathbb{E}[CC_{G'}(v) \mid \mathcal{E}(v)] = \Omega(1)$ for all $v \in \bar{V}$. Since $\mathbb{P}[\mathcal{E}(v)] = \Theta(1)$, this verifies $\mathbb{E}[CC_{G'}(v)] = \Omega(1)$ for all $v \in \bar{V}$, implying $\mathbb{E}[cc(G')]=\Omega(1)$.

Next we show that $cc(G')$ is concentrated around its expected value. We will use again Theorem 2.5. Similar as in our previous applications of Theorem 2.5 in Section 4.2 and Section 4.3, we can assume that the probability space $\Omega$ is a product space of independent random variables. We consider the bad event

$$\mathcal{B} := \{\omega \in \Omega \mid \text{the maximum degree in } G'(\omega) \text{ is at least } n^{1/4}\}.$$ 

We observe that $\mathbb{P}[\mathcal{B}] = n^{-o(1)}$, since by Lemma 3.4, w.h.p. every vertex $v \in V'$ has degree at most $O(w_v \log^2 n) = o(n^{1/4})$. Let $\omega, \omega' \in \bar{\mathcal{B}}$ such that they differ in at most two coordinates. We observe that changing one coordinate $x_i$ or $Y_i$ can influence only the local clustering coefficients of $i$ itself and of the vertices which are neighbors of $i$ either before or after the coordinate change. Unless $\mathcal{B}$ holds, every vertex in $G'$ has degree at most $n^{1/4}$ and therefore every coordinate of the probability space has effect at most $2n^{1/4}/n$ onto $cc(G')$. Thus, we obtain $|cc(G'(\omega)) - cc(G'(\omega'))| \leq 4n^{-3/4}$. We apply Theorem 2.5 with $t = n^{-1/8}$ and $c := 4n^{-3/4}$ and deduce

$$\mathbb{P} \left[ |cc(G') - \mathbb{E}[cc(G')]| \geq t \right] \leq 2e^{-\Omega(t^2/n^{1-3/2})} + n^{O(1)} \mathbb{P}[\mathcal{B}] = n^{-o(1)},$$

where we used $\frac{t^2}{n^{1-3/2}} = n^{1/4}$. Hence, we have $cc(G') = (1 + o(1))\mathbb{E}[cc(G')] = \Omega(1)$ w.h.p.

In order to compare $cc(G)$ with $cc(G')$, we observe that every additional edge $e = \{u, v\}$ which we add to $G'$ can decrease only $cc(u)$ and $cc(v)$, both by at most one. Thus,

$$CC(G) \geq \frac{|V'|}{n} CC(G') - \frac{2}{n} \sum_{v \in V \setminus V'} \deg(v).$$

By Lemma 3.4, $\frac{2}{n} \sum_{v \in V \setminus V'} \deg(v) = \Theta(n^{-1}W_{>n^{1/8}}) = o(1)$ w.h.p. Together with $|V'| = \Theta(n)$, this concludes the argument and proves that $cc(G) = \Theta(1)$ w.h.p. \blacksquare
4.5 Sampling Algorithm

In this section we show that GIRGs can be sampled in expected time $O(n)$. The running time depends exponentially on the fixed dimension $d$. In addition to the definition of GIRGs (Section 3.4), in this section we require that (1) edge probabilities $p_{uv}$ can be computed in constant time (given any vertices $u, v$ and positions $x_u, x_v$) and (2) we know an explicit constant $c > 0$ such that in the case $\alpha < \infty$ of Standard GIRGs we have

$$p_{uv} \leq \min\left\{c \frac{1}{\|x_u - x_v\|^\alpha d \cdot \left(\frac{w_u w_v}{W}\right)^\alpha}, 1\right\},$$

and in the case $\alpha = \infty$ of the Threshold GIRG we have

$$p_{uv} \leq \begin{cases} 1, & \text{if } \|x_u - x_v\| < c \left(\frac{w_u w_v}{W}\right)^{1/d}, \text{ and} \\ 0, & \text{otherwise.} \end{cases}$$

Note that existence of $c$ follows from our model assumptions. In the remainder of this section we introduce a geometric ordering of the vertices (Section 4.5.1) which will be useful as well for the compression algorithm that we study later in Section 4.6. We then define building blocks of our algorithm (Section 4.5.2) and present our algorithm (Section 4.5.3) and its analysis (Section 4.5.4). Thereby, we always assume $\alpha < \infty$. In the last part of this chapter (Section 4.5.5), we show how the sampling algorithm can be adapted to Threshold GIRGs with $\alpha = \infty$.

4.5.1 Cells

We introduce a geometric ordering of the vertices. Consider the ground space $\mathbb{T}^d$, split it into $2^d$ equal cubes, and repeat this process with each created cube; we call the resulting cubes cells. Cells are cubes of the form

$$C = [x_1 2^{-\ell}, (x_1 + 1)2^{-\ell}) \times \ldots \times [x_d 2^{-\ell}, (x_d + 1)2^{-\ell})$$

with $\ell \geq 0$ and $0 \leq x_i < 2^\ell$. We represent cell $C$ by the tuple $(\ell, x_1, \ldots, x_d)$. The volume of $C$ is $\text{VOL}(C) = 2^{-\ell \cdot d}$. For $0 < x \leq 1$ we let $[x]_{2^d}$ be the smallest number larger or equal to $x$ that is realized as the volume of a cell, or in other words $x$
rounded up to a power of $2^d$, $[x]_{2^d} = \min\{2^{-\ell \cdot d} \mid \ell \in \mathbb{N}_0 : 2^{-\ell \cdot d} \geq x\}$. Note that the cells of a fixed level $\ell$ partition the ground space. We obtain a geometric ordering of these cells by following the recursive construction of cells in a breadth-first-search manner, yielding the following lemma.

**Lemma 4.22** (Geometric ordering). *There is an enumeration of the cells $C_1, \ldots, C_{2^\ell d}$ of level $\ell$ such that for every cell $C$ of level $\ell' < \ell$ the cells of level $\ell$ contained in $C$ form a consecutive block $C_i, \ldots, C_j$ in the enumeration.*

**Proof.** We construct the geometric ordering by induction on the level $\ell$. For $\ell = 0$ there is only one cell to enumerate, so let $\ell > 0$. Given an enumeration $C_1, \ldots, C_{2^{(\ell-1)d}}$ of the cells of level $\ell - 1$, we first enumerate all cells of level $\ell$ contained in $C_1$, starting with the cell which is smallest in all $d$ coordinates, and ending with the cell which is largest in all $d$ coordinates. Then we enumerate all cells of level $\ell$ contained in $C_2$ (starting with smallest coordinates, and ending with largest coordinates), and so on. Evidently this gives us a geometric ordering of the cells of level $\ell$. ■

### 4.5.2 Building Blocks

**Data structures:** Recall the definition of cells from Section 4.5.1. We first build a basic data structure on a set of points $P$ that allows to access the points in a given cell $C$ (of volume at least $\nu$) in constant time.

**Lemma 4.23.** *Given a set of points $P$ and $0 < \nu \leq 1$, in time $O(|P| + 1/\nu)$ we can construct a data structure $\mathcal{D}_\nu(P)$ supporting the following queries in time $O(1)$:*

1. *given a cell $C$ of volume at least $\nu$, return $|C \cap P|$,*

2. *given a cell $C$ of volume at least $\nu$ and a number $k$, return the $k$-th point in $C \cap P$ (in a fixed ordering of $C \cap P$ depending only on $P$ and $\nu$).*

**Proof.** Let $\mu = [\nu]_{2^d} = 2^{-\ell \cdot d}$, so that $\nu \leq \mu \leq O(\nu)$. Following the recursive construction of cells, we can determine a geometric ordering of the cells of volume $\mu$ as in Lemma 4.22 in time $O(1/\mu) = O(1/\nu)$; say $C_1, \ldots, C_{1/\mu}$ are the
cells of volume $\mu$ in the geometric ordering. We store this ordering by storing a pointer from each cell $C_i = (\ell, x_1, \ldots, x_d)$ to its successor $C_{i+1} = (\ell, x'_1, \ldots, x'_d)$, which allows to scan the cells $C_1, \ldots, C_{1/\mu}$ in linear time. For any point $x \in P$, using the floor function we can determine in time $O(1)$ the cell $(\ell, x_1, \ldots, x_d)$ of volume $\mu$ that $x$ belongs to. This allows to determine the numbers $|C_i \cap P|$ for all $i$ in time $O(|P| + 1/\nu)$. We also compute each prefix sum $s_i := \sum_{j<i} |C_j \cap P|$ and store it at cell $C_i = (\ell, x_1, \ldots, x_d)$. Using an array $A[.]$ of size $|P|$, we store (a pointer to) the $k$-th point in $C_i \cap P$ at position $A[s_i + k]$. Note that this preprocessing can be performed in time $O(|P| + 1/\nu)$.

A given cell $C$ of volume at least $\nu$ may consist of several cells of volume $\mu$. By Lemma 4.22, these cells form a contiguous subsequence $C_i, C_{i+1}, \ldots, C_{j-1}, C_j$ of $C_1, \ldots, C_{1/\mu}$, so that the points $C \cap P$ form a contiguous subsequence of $A$. For constant access time, we store for each cell $C$ of volume at least $\nu$ the indices $s_C, e_C$ of the first and last point of $C \cap P$ in $A$. Then $|C \cap P| = e_C - s_C + 1$ and the $k$-th point in $C \cap P$ is stored at $A[s_C + k]$. Thus, both queries can be answered in constant time. Note that the ordering $A[.]$ of the points in $C \cap P$ is a mix of the geometric ordering of cells of volume $\mu$ and the given ordering of $P$ within a cell of volume $\mu$, in particular this ordering indeed only depends on $P$ and $\nu$. ■

Next we construct a partitioning of $\mathbb{T}^d \times \mathbb{T}^d$ into products of cells $A_i \times B_i$. This partitioning allows to split the problem of sampling the edges of a GIRG into one problem for each $A_i \times B_i$, which is beneficial, since each product $A_i \times B_i$ has one of two easy types. For any $A, B \subseteq \mathbb{T}^d$ we denote the distance of $A$ and $B$ by $d(A, B) = \inf_{a \in A, b \in B} \|a - b\|$.

**Lemma 4.24.** Let $0 < \nu \leq 1$. We can construct a set $\mathcal{P}_\nu = \{(A_1, B_1), \ldots, (A_s, B_s)\}$ in time $O(1/\nu)$ such that

1. $A_i, B_i$ are cells with $\text{VOL}(A_i) = \text{VOL}(B_i) \geq \nu$,

2. for all $i$, either $d(A_i, B_i) = 0$ and $\text{VOL}(A_i) = [\nu]_{2^d}$ (type I) or $d(A_i, B_i) \geq \text{VOL}(A_i)^{1/d}$ (type II),

3. the sets $A_i \times B_i$ partition $\mathbb{T}^d \times \mathbb{T}^d$. 
\[(4) \ s = O(1/\nu).\]

**Proof.** Note that for cells \(A, B\) of equal volume we have \(d(A, B) = 0\) if and only if either \(A = B\) or (the boundaries of) \(A\) and \(B\) touch. For a cell \(C\) of level \(\ell\) we let \(\text{par}(C)\) be its *parent*, i.e., the unique cell of level \(\ell - 1\) that \(C\) is contained in. Let \(\mu = [\nu]_{2^d}\). We define \(\mathcal{P}_\nu\) as follows. For any pair of cells \((A, B)\) with \(\text{VOL}(A) = \text{VOL}(B) \geq \nu\), we add \((A, B)\) to \(\mathcal{P}_\nu\) if either (i) \(\text{VOL}(A) = \text{VOL}(B) = \mu\) and \(d(A, B) = 0\), or (ii) \(d(A, B) > 0\) and \(d(\text{par}(A), \text{par}(B)) = 0\).

Property (1) follows by definition. Regarding property (2), the pairs \((A, B)\) added in case (i) are clearly of type I. Observe that two cells \(A, B\) of equal volume that are not equal or touching have distance at least the sidelength of \(A\), which is \(\text{VOL}(A)^{1/d}\). Thus, in case (ii) \(d(A, B) > 0\) implies \(d(A, B) \geq \text{VOL}(A)^{1/d}\), so that \((A, B)\) is of type II.

For property (3), consider \((x, y) \in \mathbb{T}^d \times \mathbb{T}^d\) and let \(A, B\) be the cells of volume \(\mu\) containing \(x, y\). Let \(A^{(0)} := A\) and \(A^{(i)} := \text{par}(A^{(i-1)})\) for any \(i \geq 1\), until \(A^{(k)} = \mathbb{T}^d\). Similarly, define \(B = B^{(0)} \subset \ldots \subset B^{(k)} = \mathbb{T}^d\) and note that \(\text{VOL}(A^{(i)}) = \text{VOL}(B^{(i)})\). Observe that each set \(A^{(i)} \times B^{(i)}\) contains \((x, y)\). Moreover, any set \(A' \times B'\), where \(A', B'\) are cells with \(\text{VOL}(A') = \text{VOL}(B')\) and \((x, y) \in A' \times B'\), is of the form \(A^{(i)} \times B^{(i)}\). Thus, to show that \(\mathcal{P}_\nu\) partitions \(\mathbb{T}^d \times \mathbb{T}^d\) we need to show that it contains exactly one of the pairs \((A^{(i)}, B^{(i)})\) (for any \(x, y)\). To show this, we use the monotonicity \(d(A^{(i)}, B^{(i)}) \geq d(A^{(i+1)}, B^{(i+1)})\) and consider two cases. If \(d(A, B) = 0\) then we add \((A, B)\) to \(\mathcal{P}_\nu\) in case (i), and we add no further \((A^{(i)}, B^{(i)})\), since \(d(A^{(i)}, B^{(i)}) = 0\) for all \(i\). If \(d(A, B) > 0\) then since \(d(A^{(k)}, B^{(k)}) = d(\mathbb{T}^d, \mathbb{T}^d) = 0\) there is a unique index \(0 \leq i < k\) with \(d(A^{(i)}, B^{(i)}) > 0\) and \(d(A^{(i+1)}, B^{(i+1)}) = 0\). Then we add \((A^{(i)}, B^{(i)})\) in case (ii) and no further \((A^{(j)}, B^{(j)})\). This proves property (3).

Property (4) follows from the running time bound of \(O(1/\nu)\), which we show in the following. Note that we can enumerate all \(1/\mu = O(1/\nu)\) cells of volume \(\mu\), and all of the at most \(3^d = O(1)\) touching cells of the same volume, in time \(O(1/\nu)\), proving the running time bound for case (i). Moreover, we can enumerate all \(2^{\ell - d}\) cells \(C\) in level \(\ell\), together with all of the at most \(3^d = O(1)\) touching cells \(C'\) in the same level. Then we can enumerate all \(2^d = O(1)\) cells \(A\) that have \(C\) as parent as well as all \(O(1)\) cells \(B\) that have \(C'\) as parent. This enumerates (a
superset of) all possibilities of case (ii). Summing the running time \( O(2^\ell d) \) over all levels \( \ell \) with volume \( 2^{-\ell} d \geq \nu \) yields a total running time of \( O(1/\nu) \). ■

**Weight layers:** We set
\[
w_0 := w_{\text{min}} \quad \text{and} \quad w_i := 2w_{i-1} \quad \text{for} \quad i \geq 1.\]
This splits the vertex set \( V = [n] \) into weight layers \( V_i := \{v \in V \mid w_{i-1} \leq v < w_i\} \) for \( 1 \leq i \leq L \) with \( L = O(\log n) \). We write \( V_i^C \) for the restriction of weight layer \( V_i \) to cell \( C \),
\[
V_i^C := \{v \in V_i \mid x_v \in C\}.
\]

**Geometric random variates:** For \( 0 < p \leq 1 \) we write Geo\((p)\) for a geometric random variable, taking value \( i \geq 1 \) with probability \( p (1-p)^{i-1} \). Geo\((p)\) can be sampled in constant time using the simple formula
\[
\left\lceil \frac{\log(R)}{\log(1-p)} \right\rceil,
\]
where \( R \) is chosen uniformly at random in \((0,1)\), see [Dev86]. To evaluate this formula exactly in time \( O(1) \) we need to assume the RealRAM model of computation. However, also on a bounded precision machine like the WordRAM Geo\((p)\) can be sampled in expected time \( O(1) \) [BF13].

### 4.5.3 The Algorithm

Given the model parameters, our Algorithm 4.1 samples the edge set \( E \) of a GIRG. To this end, we first sample all vertex positions \( x_v \) uniformly at random in \( T^d \). Given weights \( w_1, \ldots, w_n \) we can determine the weight layers \( V_i \) in linear time (we may use counting sort or bucket sort since there are only \( L = O(\log n) \) layers). Then we build the data structure from Lemma 4.23 for the points in \( V_i \) setting \( \nu = \nu(i) = \frac{w_i w_0}{W} \), i.e., we build \( \mathcal{D}_{v(i)}(\{x_v \mid v \in V_i\}) \) for each \( i \). In the following, for each pair of weight layers \( V_i, V_j \) we sample the edges between \( V_i \) and \( V_j \). To this end, we construct the partitioning \( \mathcal{P}_{v(i,j)} \) from Lemma 4.24 with \( \nu(i,j) = \frac{w_i w_j}{W} \). Since \( \mathcal{P}_{v(i,j)} \) partitions \( T^d \times T^d \), every pair of vertices \( u \in V_i, v \in V_j \) satisfies \( x_u \in A, x_v \in B \) for exactly one \( (A, B) \in \mathcal{P}_{v(i,j)} \). Thus, we can iterate over all \( (A, B) \in \mathcal{P}_{v(i,j)} \) and sample the edges between \( V_i^A \) and \( V_j^B \).

If \( (A, B) \) is of type I, then we simply iterate over all vertices \( u \in V_i^A \) and \( v \in V_j^B \) and add the edge \( \{u, v\} \) with probability \( p_{uv} \); this is the trivial sampling algorithm. Note that we can efficiently enumerate \( V_i^A \) and \( V_j^B \) using the data structure \( \mathcal{D}_{v(i)}(\{x_v \mid v \in V_i\}) \) that we constructed above.
Algorithm 4.1 Sampling algorithm for GIRGs in expected time $O(n)$

1: $E := \emptyset$

2: sample the positions $x_v, v \in V$, and determine the weight layers $V_i$

3: for all $1 \leq i \leq L$ do build data structure $\mathcal{D}_{V(i)}(\{x_v \mid v \in V_i\})$ with $\nu(i) := \frac{w_i w_j}{W}$

4: for all $1 \leq i \leq L$ do

5: construct partitioning $\mathcal{D}_{V(i,j)}$ with $\nu(i,j) := \frac{w_i w_j}{W}$

6: for all $(A, B) \in \mathcal{D}_{V(i,j)}$ of type I do

7: for all $u \in V_i$ and $v \in V_j$ do with probability $p_{uv}$ add edge $\{u, v\}$ to $E$

8: for all $(A, B) \in \mathcal{D}_{V(i,j)}$ of type II do

9: $\bar{p} := \min \{c \cdot \frac{1}{d(A,B)^{\alpha d}} \cdot \left(\frac{w_i w_j}{W}\right)^{\alpha}, 1\}$

10: $r := \text{Geo}(\bar{p})$

11: while $r \leq |V_i^A| \cdot |V_j^B|$ do

12: determine the $r$-th pair $(u, v)$ in $V_i^A \times V_j^B$

13: with probability $p_{uv}/\bar{p}$ add edge $\{u, v\}$ to $E$

14: $r := r + \text{Geo}(\bar{p})$

15: if $i = j$ then remove all edges with $u > v$ sampled in this iteration

If $(A, B)$ is of type II, then the distance $\|x - y\|$ of any two points $x \in A, y \in B$ satisfies $d(A, B) \leq \|x - y\| \leq d(A, B) + \text{VOL}(A)^{1/d} + \text{VOL}(B)^{1/d} \leq 3d(A, B)$, by the definition of type II. Thus, $\bar{p} = \min \{c \cdot \frac{1}{d(A,B)^{\alpha d}} \cdot \left(\frac{w_i w_j}{W}\right)^{\alpha}, 1\}$ is an upper bound on the edge probability $p_{uv}$ for any $u \in V_i^A, v \in V_j^B$, and it is a good upper bound since $d(A, B)$ is within a constant factor of $\|x_u - x_v\|$ and $w_i, w_j$ are within constant factors of $w_u, w_v$. Now we first sample the set of edges $\bar{E}$ between $V_i^A$ and $V_j^B$ that we would obtain if all edge probabilities were equal to $\bar{p}$, i.e., any $(u, v) \in V_i^A \times V_j^B$ is in $\bar{E}$ independently with probability $\bar{p}$. From this set $\bar{E}$, we can then generate the set of edges with respect to the true edge probabilities $p_{uv}$ by throwing a coin for each $\{u, v\} \in \bar{E}$ and letting it survive with probability $p_{uv}/\bar{p}$. Then in total we choose a pair $(u, v)$ as an edge in $E$ with probability

$$\bar{p} \cdot \frac{p_{uv}}{\bar{p}} = p_{uv},$$

proving that we sample from the correct distribution. Note that here we used $p_{uv} \leq \bar{p}$. It is left to show how to sample the “approximate” edge set $\bar{E}$. First note that the data structure $\mathcal{D}_{V}(\{x_v \mid v \in V_i\})$ defines an ordering on $V_i^A$, and we can determine the $\ell$-th element in this ordering in constant time, similarly for $V_j^B$. 
Using the lexicographic ordering, we obtain an ordering on \( V_i^A \times V_j^B \) for which we can again determine the \( \ell \)-th element in constant time. In this ordering, the first pair \((u, v) \in V_i^A \times V_j^B\) that is in \( \tilde{E} \) is geometrically distributed, according to \( \text{Geo}(\tilde{p}) \). Since geometric random variates can be generated in constant time, we can efficiently generate \( \tilde{E} \), specifically in time \( O(1 + |\tilde{E}|) \).

Finally, the case \( i = j \) is special. With the algorithm described above, for any \( u, v \in V_i \) we sample whether they form an edge twice, once for \( x_u \in A, x_v \in B \) (for some \((A, B) \in \mathcal{P}_v(i, j)\)) and once for \( x_v \in A', x_u \in B' \) (for some \((A', B') \in \mathcal{P}_v(i, j)\)). To fix this issue, in the case \( i = j \) we only accept a sampled edge \((u, v) \in V_i^A \times V_j^B\) if \( u < v \); then only one way of sampling edge \( \{u, v\} \) remains. This changes the expected running time only by a constant factor.

### 4.5.4 Analysis

Correctness of our algorithm follows immediately from the above explanations. In the following we show that Algorithm 4.1 runs in expected linear time. This is clear for lines 1-2. For line 3, since building the data structure from Lemma 4.23 takes time \( O(|P| + 1/v) \), it takes total time \( \sum_{i=1}^{L} O(|V_i| + W/(w_i w_0)) \). Clearly, the first summand \( |V_i| \) sums up to \( n \). Using \( w_0 = w_{\min} = \Omega(1), W = O(n) \), and that \( w_i \) grows exponentially with \( i \), implying \( \sum_i 1/w_i = O(1) \), also the second summand sums up to \( O(n) \). For line 5, all invocations in total take time \( O(\sum_{i,j} W/(w_i w_j)) \), which is \( O(n) \), since again \( W = O(n) \) and \( \sum_i 1/w_i = O(1) \). We claim that for any weight layers \( V_i, V_j \) the expected running time we spend on any \((A, B) \in \mathcal{P}_v(i, j)\) is \( O(1 + E[|E_{i,j}^{A,B}|]) \), where \( E_{i,j}^{A,B} \) is the set of edges in \( V_i^A \times V_j^B \). Summing up the first summand \( O(1) \) over all \((A, B) \in \mathcal{P}_v(i, j)\) sums up to \( 1/v(i, j) = W/(w_i w_j) \). As we have seen above, this sums up to \( O(n) \) over all \( i, j \). Summing up the second summand \( O(E[|E_{i,j}^{A,B}|]) \) over all \((A, B) \in \mathcal{P}_v(i, j)\) and weight layers \( V_i, V_j \) yields the total expected number of edges \( O(E[|E|]) \), which is \( O(n) \), since the average weight \( W/n = O(1) \) and thus the expected average degree is constant.

It is left to prove the claim that for any weight layers \( V_i, V_j \) the expected time spent on \((A, B) \in \mathcal{P}_v(i, j)\) is \( O(1 + E[|E_{i,j}^{A,B}|]) \). If \((A, B)\) is of type I, then any pair of vertices \((u, v) \in V_i^A \times V_j^B\) has probability \( \Theta(1) \) to form an edge: Since the volume of \( A \) and \( B \) is \( w_i w_j/W \), their diameter is \((w_i w_j/W)^{1/d} \) and we obtain
\|x_u - x_v\| \leq (w_i w_j / W)^{1/d} = O((w_u w_v / W)^{1/d}), which yields

\[ p_{uv} = \Theta\left( \min\left\{ \left( \frac{w_u w_v}{\|x_u - x_v\|^{d W}} \right)^{\alpha}, 1 \right\} \right) = \Theta(1). \]

As we spend time \( O(1) \) for any \((u, v) \in V_i^A \times V_j^B\), we stay in the desired running time bound \( O(\mathbb{E}[|E_{i,j}^{A,B}|]) \).

If \((A, B)\) is of type II, we first sample edges \(\tilde{E}\) with respect to the larger edge probability \(\tilde{p}\), and then for each edge \(e \in \tilde{E}\) sample whether it belongs to \(E\). This takes total time \( O(1 + |\tilde{E}|) \). Note that any edge \(e \in \tilde{E}\) has constant probability \(p_{uv} = \Theta(1)\) to survive: it follows from \(\|x_u - x_v\| = \Theta(d(A, B)), w_u = \Theta(w_i), \) and \(w_v = \Theta(w_j)\) that \(p_{uv} = \Theta(\tilde{p})\). Hence, we obtain \(\mathbb{E}[|\tilde{E}|] = O(\mathbb{E}[|E_{i,j}^{A,B}|])\), and the running time \(O(1 + |\tilde{E}|)\) is therefore in expectation bounded by \(O(1 + \mathbb{E}[|E_{i,j}^{A,B}|])\). This finishes the proof of the claim.

### 4.5.5 Sampling Threshold GIRGs

For the threshold case \(\alpha = \infty\), by (3.6) an edge between two vertices \(u\) and \(v\) can only exist if the distance satisfies \(\|x_u - x_v\| < c(w_u w_v / W)^{1/d}\). We change Algorithm 4.1 by setting \(v(i, j) = \max\{1, c\}^d \cdot w_i w_j / W\). Then for any \(u \in V_i, v \in V_j\) and \((A, B) \in P_{v(i, j)}\) of type II we have \(d(A, B) \geq \text{vol}(A)^{1/d} \geq \nu^{1/d} \geq c(w_u w_v / W)^{1/d}\), so there are no edges between \(V_i^A\) and \(V_j^B\) for type II. This allows to simplify the algorithm by completely ignoring type II pairs; the rest of the algorithm stays unchanged.

Additionally, we have to slightly change the running time analysis, since it no longer holds that all pairs of vertices \((u, v) \in V_i^A \times V_j^B\) satisfy \(p_{uv} = \Theta(1)\). However, a variant of this property still holds: If we only uncovered that \(x_u \in A\) and \(x_v \in B\), but not yet where exactly in \(A, B\) they lie, then the marginal probability of \((u, v)\) forming an edge is \(\Theta(1)\), since for any \(\epsilon > 0\) a constant fraction of all pairs of points in \(A \times B\) are within distance \(\epsilon(w_{i-1} w_{j-1} / W)^{1/d}\), guaranteeing edge probability \(\Theta(1)\) for sufficiently small \(\epsilon\). This again allows to check all pairs of vertices in \(V_i^A \times V_j^B\) whether they form an edge, which yields expected linear running time.
4.6 Separators, Entropy, and Compression Algorithm

In this section we prove Theorems 4.6 and 4.7. More precisely, we show that w.h.p. the graph (and its giant) has separators of sublinear size, and we make use of these small separators to devise a compression algorithm that can store the graph using a linear number of bits in expectation. Note that the compression maintains only the graph up to isomorphism, not the underlying geometry. The main idea is to enumerate the vertices in an ordering that reflects the geometry, and then storing for each vertex $i$ the differences $i - j$ for all neighbors $j$ of $i$. We start with a technical lemma that gives the number of edges intersecting an axis-parallel, regular grid: for $\gamma > 0$ with $1/\gamma \in \mathbb{N}$, the axis-parallel, regular grid with side length $\gamma$ is the union of all $d - 1$-dimensional hyperplanes that are orthogonal to an axis and that are in distance $k\gamma$ from the origin for a $k \in \mathbb{Z}$. Both the existence of small separators and the efficiency of the compression algorithm follow easily from that formula.

**Lemma 4.25.** Let $\eta > 0$. Let $1 \leq \mu \leq n^{1/d}$ be an integer, and consider an axis-parallel, regular grid with side length $1/\mu$ on $[0,1]^d$. Then the expected number of edges intersected by the grid is at most

$$O(n \cdot (n/\mu^d)^{2-\beta+\eta} + (n^{2-a} \mu^{d(a-1)} + n^{1-1/d} \mu)(1 + \log(n/\mu^d))).$$

We defer the proof of Lemma 4.25 to the end of this chapter. For the regime $2 < \beta < 3$, we immediately obtain that there is a sublinear set of vertices that disconnects the giant component.

**Proof of Theorem 4.6.** By Lemma 4.25 for $\mu = 2$, the number of edges intersecting a grid of side length $1/2$ is in expectation $m = O(n^{\max(3-\beta,2-a,1-1/d)} + \eta)$, and two hyperplanes of this grid suffice to split $[0,1]^d$ into two halves. W.h.p. there are $\Omega(n)$ vertices in each grid cell, and w.h.p. the weights of the vertices in each half satisfy a power law. If $2 < \beta < 3$ then w.h.p. each halfspace gives rise to a giant component of linear size, this follows analogously to Lemma 3.12 (as we only need the lower bound of (PL2) for the giant component). Hence, w.h.p. the two hyperplanes split the giant of $G$ into two parts of linear size, although almost surely they only intersect $n^{1-\Omega(1)}$ edges. Finally, since the bound
$m = O(n^{\max\{3-\beta, 2-a, 1-1/d\}+\eta})$ holds for all $\eta > 0$ we may conclude that it also holds with $\eta$ replaced by $o(1)$. ■

**Compression algorithm:** With Lemma 4.25 at hand, we are ready to give a compression algorithm that stores the graph with $O(n)$ bits, i.e., with $O(1)$ bits per edge, proving Theorem 4.7. We remark that our result does not directly follow from the general compression scheme on graphs with small separators in [BBK03], since our graphs only have small separators in expectation, in particular, small subgraphs of size $O(\sqrt{\log n})$ can form expanders and thus not have small separators. However, our algorithm loosely follows their algorithm as well as the practical compression scheme of [BV04], see also [Chi+09b].

We first enumerate the vertices as follows. Recall the definition of cells from Section 4.5.1, and consider all cells of volume $2^{-\ell_0 d}$, where $\ell_0 := \lfloor \log n/d \rfloor$. Note that the boundaries of these cells induce a grid as in Lemma 4.25. Since each such cell has volume $\Theta(1/n)$, the expected number of vertices in each cell is constant. We fix a geometric ordering of these cells as in Lemma 4.22, and we enumerate the vertices in the order of the cells, breaking ties (between vertices in the same cell) arbitrarily. For the rest of the section we assume that the vertices are enumerated in this way, i.e., we identify $V = [n]$, where $i \in [n]$ refers to the vertex with index $i$.

Having enumerated the vertices, for each vertex $i \in [n]$ we store a block of $1 + \deg(i)$ sub-blocks. The first sub-block consists of a single dummy bit (to avoid empty sequences arising from isolated vertices). In the other $\deg(i)$ sub-blocks we store the differences $i - j$ using $\log_2 |i - j| + O(1)$ bits, where $j$ runs through all neighbors of $i$. We assume that the information for all vertices is stored in a big successive block $B$ in the memory. Moreover, we create two more blocks $B_V$ and $B_E$ of the same length. Both $B_V$ and $B_E$ have a one-bit whenever the corresponding bit in $B$ is the first bit of the block of a vertex, and $B_E$ has also a one-bit whenever the corresponding bit in $B$ is the first bit of an edge (i.e., the first bit encoding a difference $i - j$). All other bits in $B_V$ and $B_E$ are zero.

It is clear that with the data above the graph is determined. To handle queries efficiently, we replace $B_V$ and $B_E$ each with a rank/select data structure. This
data structure allows to handle in constant time queries of the form “\texttt{RANK} (b)”, which returns the number of one-bits up to position \( b \), and “\texttt{SELECT} (i)”, which returns the position of the \( i \)-th one-bit [Jac89; CM96; Păt08]. Given \( i, s \in \mathbb{N} \), we can find the index of the \( s \)-th neighbor of \( i \) in constant time by Algorithm 4.2, and the degree of \( i \) by Algorithm 4.3. In particular, it is also possible for Algorithm 4.2 to first check whether \( s \leq \text{deg} (i) \).

\begin{algorithm}
\caption{Finding the \( s \)-th neighbor of vertex \( i \)}
\begin{algorithmic}[1]
\State \( b := \text{SELECT}(i, B_V) \) \Comment{starting position of vertex \( i \)}
\State \( k := \text{RANK}(b, B_E) \) \Comment{number of edges and vertices before \( b \)}
\State \( b_1 := \text{SELECT}(k + s, B_E) \) \Comment{starting position of \( s \)-th edge of vertex \( i \)}
\State \( b_2 := \text{SELECT}(k + s + 1, B_E) \) \Comment{bit after ending position of \( s \)-th edge of vertex \( i \)}
\State \textbf{return} \( B[b_1 : b_2 - 1] \) \Comment{block that stores \( s \)-th edge of vertex \( i \)}
\end{algorithmic}
\end{algorithm}

\begin{algorithm}
\caption{Finding the degree of vertex \( i \)}
\begin{algorithmic}[1]
\State \( b := \text{SELECT}(i, B_V) \) \Comment{starting position of vertex \( i \)}
\State \( b' := \text{SELECT}(i + 1, B_V) \) \Comment{starting position of vertex \( i + 1 \)}
\State \( \Delta := \text{RANK}(b', B_E) - \text{RANK}(b, B_E) \) \Comment{block in \( B_E \) contains \text{deg}(i) + 1 \) one-bits}
\State \textbf{return} \( \Delta - 1 \)
\end{algorithmic}
\end{algorithm}

We need to show that the data structure needs \( O(n) \) bits in expectation. There are \( n \) dummy bits, so we must show that we require \( O(n) \) bits to store all differences \( i - j \), where the indices \( i \) and \( j \) run through all edges of the graph.

We need \( 2 \log_2 |i - j| + O(1) \) bits for each edge, as we store the edge \( ij \) both in the block of \( i \) and in the block of \( j \). The \( O(1) \) terms sum up to \( O(|E|) \), which is \( O(n) \) in expectation. Thus, it remains to prove the following.

\textbf{Lemma 4.26.} \textit{Let the vertices in} \( V \) \textit{be enumerated by the geometric ordering. Then,}

\begin{equation}
(4.11) \quad \mathbb{E} \left[ \sum_{ij \in E} \log(|i - j|) \right] = O(n).
\end{equation}

\textbf{Proof.} We abbreviate the expectation in (4.11) by \( R \). Note that the geometric ordering puts all the vertices that are in the same cell of a \( 2^{-\ell} \)-grid in a consecutive block, for all \( 1 \leq \ell \leq \ell_0 \). Therefore, if \( e = ij \) does not intersect the \( 2^{-\ell} \)-grid
then $|i - j| \leq |\{\text{vertices in the cell of } e\}|$. For $1 \leq \ell \leq \ell_0$, let $\mathcal{E}_\ell$ be the set of edges intersecting the $2^{-\ell}$-grid. For convenience, let $\mathcal{E}_0 := \emptyset$, and let $\mathcal{E}_{\ell_0 + 1} := E$ be the set of all edges. Then

$$R \leq \mathbb{E}\left[ \sum_{\ell=0}^{\ell_0} \sum_{e = i j \in \mathcal{E}_{\ell + 1} \setminus \mathcal{E}_\ell} \log(|\{\text{vertices in the cell of } e\}|) \right]$$

$$= \sum_{\ell=0}^{\ell_0} \sum_{u < v} \mathbb{P}[uv \in \mathcal{E}_{\ell + 1} \setminus \mathcal{E}_\ell] \cdot \mathbb{E}[\log(|\{\text{vertices in the cell of } u\}|) | uv \in \mathcal{E}_{\ell + 1} \setminus \mathcal{E}_\ell]$$

Now, we use concavity of $\log$ and apply Jensen’s inequality (Theorem 2.6) to infer that

$$R \leq \sum_{\ell=0}^{\ell_0} \sum_{u < v} \mathbb{P}[uv \in \mathcal{E}_{\ell + 1} \setminus \mathcal{E}_\ell] \cdot \mathbb{E}\left[ \log(|\{\text{vertices in the cell of } u\}|) \bigg| uv \in \mathcal{E}_{\ell + 1} \setminus \mathcal{E}_\ell \right].$$

The term $T_\ell$ is at most $T_\ell \leq 2 + (n - 2)2^{-\ell d} \leq 3n2^{-\ell d}$ for $\ell \leq \ell_0$ (where we count 2 for $u$ and $v$ and use independence of the other vertex positions). Thus it remains to show that $\mathbb{E}\left[ \sum_{\ell=0}^{\ell_0} |\mathcal{E}_{\ell + 1}| \log(3n2^{-\ell d}) \right] = O(n)$. Let $\eta > 0$ be sufficiently small. From Lemma 4.25 we know that $\mathbb{E}[|\mathcal{E}_\ell|] \leq E_\ell$, where

$$E_\ell = n \cdot (2^{d\ell}/n)^{\beta - 2 - \eta} + (n^{2-\alpha}2^{d\ell(\alpha-1)} + n^{1-1/d}2^{\ell})(1 + \log(n2^{-d\ell})).$$

Since $E_\ell$ increases exponentially in $\ell$, we obtain

$$\mathbb{E}\left[ \sum_{\ell=0}^{\ell_0} |\mathcal{E}_{\ell + 1}| \log(3n2^{-d\ell}) \right] \leq O\left( \sum_{\ell=1}^{\ell_0+1} E_\ell \log(3n2^{-d\ell+d}) \right)$$

$$= O(E_{\ell_0+1} \log(3n2^{-d\ell_0})) = O(n),$$

where the last equality follows since $1/n \leq 2^{-d\ell_0} \leq O(1/n)$ by our choice of $\ell_0$. This proves the lemma, and hence shows that we need $O(n)$ bits in expectation to store the graph.  

This concludes the proof of Theorem 4.7, and it only remains to verify Lemma 4.25. In the proof of this lemma, we will use the following technical statement, which is a consequence of Fubini’s theorem and allows us to replace certain sums by integrals.
Lemma 4.27. Let \( f : \mathbb{R} \to \mathbb{R} \) be a continuously differentiable function. Then for any weights \( 0 \leq w_0 \leq w_1 \),
\[
\sum_{\nu \in V, w_0 \leq w_\nu \leq w_1} f(w_\nu) = f(w_0) \cdot |V \geq w_0| - f(w_1) \cdot |V > w_1| + \int_{w_0}^{w_1} f'(w) \cdot |V \geq w| \, dw.
\]
In particular, if \( f(0) = 0 \), then
\[
\sum_{\nu \in V, w_0 \leq w_\nu \leq w_1} f(w_\nu) = \int_{0}^{w_{\text{max}}} |V \geq w| f'(w) \, dw = \int_{0}^{\infty} |V \geq w| f'(w) \, dw.
\]

Proof. Let \( \nu \) be the sum of all Dirac measures given by the vertex weights between \( w_0 \) and \( w_1 \), i.e., for every set \( A \subseteq \mathbb{R} \) we put
\[
\nu(A) := |\{ \nu \in V : w_\nu \in A, w_0 \leq w_\nu \leq w_1 \}|.
\]
Then
\[
\sum_{\nu \in V, w_0 \leq w_\nu \leq w_1} f(w_\nu) = \int_{0}^{w_{\text{max}}} f(w) \, dv(w)
\]
\[
= \int_{0}^{w_{\text{max}}} \int_{0}^{w} f'(x) \, dx \, dv(w) + \int_{0}^{w_{\text{max}}} f(0) \, dv(w)
\]
\[
= \int_{0}^{w_{\text{max}}} \int_{0}^{\infty} f'(x) \cdot \mathbb{1}_{[x \leq w]} \, dx \, dv(w) + f(0) \cdot |V \geq w_0 \setminus V > w_1|.
\]
Notice that \([0, w_{\text{max}}]\) is a compact set and \( f'(x) \) is continuous by assumption. Hence the function \(|f'(x) \cdot \mathbb{1}_{[x \leq w]}|\) is globally bounded on \([0, w_{\text{max}}]\) and always zero for \( x > w_{\text{max}} \). Thus, \( f'(x) \cdot \mathbb{1}_{[x \leq w]} \) is integrable and we can apply Fubini’s theorem (see, e.g., [Hal74]), which yields
\[
\sum_{\nu \in V, w_0 \leq w_\nu \leq w_1} f(w_\nu) = \int_{0}^{\infty} f'(x) \int_{0}^{w_{\text{max}}} \mathbb{1}_{[w \geq x]} \, dv(w) \, dx + f(0) \cdot |V \geq w_0 \setminus V > w_1|
\]
\[
= \int_{0}^{\infty} f'(x) \cdot |V \geq \text{max} \{ x, w_0 \} \setminus V > w_1| \, dx + f(0) \cdot |V \geq w_0 \setminus V > w_1|.
\]
The first summand (i.e. the integral) can be rewritten as
\[
\int_{0}^{w_0} f'(x) \cdot |V \geq w_0 \setminus V > w_1| \, dx + \int_{w_0}^{w_1} f'(x) \cdot |V \geq x \setminus V > w_1| \, dx,
\]
and then we obtain the first statement by using \(|V \geq w_0 \setminus V > w_1| = |V \geq w_0| - |V > w_1|\), calculating the integrals, and combining the resulting terms. The second statement follows directly by choosing \( w_0 = 0 \) and \( w_1 > w_{\text{max}} \).
Proof of Lemma 4.25. Let $\beta > 2$. We first observe that we can assume $\mu \geq 2$ as this implies the statement for smaller $\mu$ immediately. Thus let $2 \leq \mu \leq n^{1/d}$, and consider an axis-parallel, regular grid with side length $1/\mu$ on $\mathbb{T}^d$. For $u, v \in V$, let $\rho_{uv}$ be the probability that the edge $uv$ exists and cuts the grid. Let $r_{\max} = \Theta(1)$ be the geometric diameter of $\mathbb{T}^d$. Using $\mathcal{S}_{u,v,\mu}$ for the event that $x_u$ and $x_v$ are in different cells of the $\mu$-grid, we write

$$
\rho_{uv} = \int_0^{r_{\max}} \Pr[\|x_u - x_v\| = r] \cdot p_{uv}(r) \cdot \Pr[\mathcal{S}_{u,v,\mu} | \|x_u - x_v\| = r] dr.
$$

Starting with the first factor of the product, we observe that $u$ and $v$ have distance $r$ with probability density $\Pr[\|x_u - x_v\| = r] = O(r^{d-1})$. Furthermore, setting $\gamma_{uv} := \min\{(w_u w_v / W)^{1/d}, r_{\max}\}$ we have

$$
p_{uv}(r) = \begin{cases} 
\Theta(1), & \text{if } r \leq \gamma_{uv}, \text{ and} \\
\Theta((\gamma_{uv}/r)^{d\alpha}), & \text{otherwise.}
\end{cases}
$$

Additionally, in the case $\alpha = \infty$, by increasing $\gamma_{uv}$ by at most a constant factor we may assume $p_{uv}(r) = 0$ for all $r \geq \gamma_{uv}$. For the last term in (4.12), for a fixed axis of $\mathbb{T}^d$ consider the hyperplanes $\{h_i\}_{1 \leq i \leq \mu}$ of the grid perpendicular to that axis. If the edge $e = uv$ has length $\|x_u - x_v\| = r$, then after a random shift along the axis, the edge $e$ intersects one of the $h_i$ with probability at most $\min\{\mu r, 1\}$. By symmetry of the underlying space, a random shift does not change the probability to intersect one of the $h_i$, so any edge of length $r$ (where we only condition on the length but not on the coordinates of the nodes) has probability at most $\min\{\mu r, 1\}$ to intersect one of the $h_i$. By the union bound over all (constantly many) axes, the probability for $u, v$ to lie in different cells of the grid is $O(\min\{\mu r, 1\})$.

Now we distinguish several cases. For $\gamma_{uv} > 1/\mu$ and $\alpha < \infty$, we may estimate

$$
\rho_{uv} \leq O\left(\int_0^{1/\mu} r^{d-1} \cdot \mu rdr + \int_{1/\mu}^{\gamma_{uv}} r^{d-1} dr + \int_{\gamma_{uv}}^{r_{\max}} r^{d-1-d\alpha} \gamma_{uv}^{d\alpha} dr \right) = O(\gamma_{uv}^d), \text{ since } d-d\alpha < 0
$$

$$
(4.13)
$$

$$
\leq O(\mu^{-d} + \gamma_{uv}^d) \leq O(\gamma_{uv}^d).
$$
For $\gamma_{uv} > 1/\mu$ and $\alpha = \infty$, equation (4.13) remains true, except that the third integral is replaced by 0 by our choice of $\gamma_{uv}$. So we still get $\rho_{uv} \leq O(\gamma_{uv}^d)$ in this case.

The case $\gamma_{uv} \leq 1/\mu$ is a bit more complicated. Again we consider first $\alpha < \infty$. Then we may bound

\begin{equation}
\rho_{uv} \leq O\left(\int_0^{\gamma_{uv}} r^{d-1} \cdot \mu r dr + \int_{\gamma_{uv}}^{1/\mu} r^{d-1} \cdot \mu r \cdot r^{-d\alpha} \gamma_{uv}^d dr + \int_{1/\mu}^{r_{max}} r^{d-1-d\alpha} \gamma_{uv}^d dr\right).
\end{equation}

Similarly as before, $I_1 \leq O(\gamma_{uv}^{d+1} \mu)$ and $I_3 \leq O(\mu^{d\alpha-d} \gamma_{uv}^d)$. Note that both terms are bounded from above by $O((\gamma_{uv} \mu)^{d\tilde{\alpha} \mu^{-d}})$, where $\tilde{\alpha} := \min\{\alpha, 1 + 1/d\}$, since $\gamma_{uv} \mu \leq 1$. For $I_2$, the inverse derivative of $r^{d-d\alpha}$ is either $\Theta(r^{1+d-d\alpha})$, or $\log r$, or $-\Theta(r^{1+d-d\alpha})$, depending on whether $1 + d - d\alpha$ is positive, zero, or negative, respectively. Therefore, we obtain

\begin{align*}
I_2 \leq \begin{cases} 
O(\gamma_{uv}^{d\alpha} \mu^{d\tilde{\alpha} \mu^{-d}}) = O((\gamma_{uv} \mu)^{d\alpha} \mu^{-d}) & \text{if } d - d\alpha > -1, \\
O(\gamma_{uv}^{d+1} \mu) = O((\gamma_{uv} \mu)^{d+1} \mu^{-d}) & \text{if } d - d\alpha < -1,
\end{cases}
\end{align*}

and, in the critical case $d - d\alpha = -1$,

\begin{equation*}
I_2 \leq O(\gamma_{uv}^{d\alpha} \mu (\log(1/\mu) - \log(\gamma_{uv}))) = O((\gamma_{uv} \mu)^{d+1} \mu^{-d} |\log(\gamma_{uv} \mu)|).
\end{equation*}

In particular, we can upper-bound all terms (including $I_1$ and $I_3$) in a unified way by $O((\gamma_{uv} \mu)^{d\tilde{\alpha} \mu^{-d}})(1 + |\log(\gamma_{uv} \mu)|)$. Moreover, since

\begin{equation*}
\gamma_{uv} \geq (w_{\min}^2/W)^{1/d} = \Omega(n^{-1/d}),
\end{equation*}

the second factor is bounded by $O(1 + \log(n^{1/d} / \mu)) = O(1 + \log(n / \mu^d))$. Also, in the case $\alpha = \infty$ the same calculation applies, except that $I_2$ and $I_3$ are replaced by 0. Note that naturally $\tilde{\alpha} = 1 + 1/d$ for $\alpha = \infty$. So altogether we have shown that

\begin{equation*}
\rho_{uv} \leq \begin{cases} 
O(\gamma_{uv}^d) & \text{, if } \gamma_{uv} \geq 1/\mu, \\
O((\gamma_{uv} \mu)^{d\tilde{\alpha} \mu^{-d}}(1 + \log(n^d / \mu))) & \text{, if } \gamma_{uv} \leq 1/\mu.
\end{cases}
\end{equation*}
Therefore, the expected number of edges intersecting the grid is in $O(S_1 + S_2)$, where
\[
S_1 := \sum_{u,v \in V, \gamma_{uv} > 1/\mu} \gamma_{uv}^d \quad \text{and} \quad S_2 := \sum_{u,v \in V, \gamma_{uv} \leq 1/\mu} (\gamma_{uv} \mu)^d \tilde{\alpha} \mu^{-d} (1 + \log(n^d / \mu)).
\]

Let $0 < \eta' < \eta < \beta - 2$ be sufficiently small constants. Then we may use the power-law assumption (PL2), Lemma 3.1, and Lemma 4.27 to bound $S_1$:
\[
S_1 \leq \sum_{u,v \in V, w_u w_v > W/\mu^d} \frac{w_u w_v}{W} = \sum_{u \in V} \frac{w_u}{W} \sum_{w \geq W / (\mu^d w_u)} 3.1 \leq O \left( \sum_{u \in V} \frac{w_u}{W} \cdot n \left( \frac{W}{\mu^d w_u} \right)^{2-\beta+\eta} \right) \leq O \left( \left( \frac{\mu^d}{n} \right)^{\beta-2-\eta} \sum_{u \in V} w_u^{\beta-1-\eta} \right) 4.27 \leq O \left( \left( \frac{\mu^d}{n} \right)^{\beta-2-\eta} \int_{w_{\min}}^{\infty} n w^{1-\beta+\eta} w^{\beta-2-\eta} d w \right) = O(n \cdot (\mu^d / n)^{\beta-2-\eta}).
\]

To tackle $S_2$, we again use Lemma 4.27, define $\lambda_u := W/(w_u \mu^d)$ and obtain
\[
S_2' := \sum_{u,v \in V, \gamma_{uv} \leq 1/\mu} (\gamma_{uv})^d \tilde{\alpha} = \sum_{u \in V} \sum_{v \in V \leq \lambda_u} \left( \frac{w_u w_v}{W} \right)^{\tilde{\alpha}} \leq 4.27 O \left( \sum_{u \in V} \left( \frac{w_u}{n} \right)^{\tilde{\alpha}} \int_{w_{\min}}^{\lambda_u} n w^{1-\beta+\eta} w^{\tilde{\alpha}-1} d w \right).
\]

Now we distinguish two cases, because the integral behaves differently for exponents larger or smaller than $-1$. If $\tilde{\alpha} \geq \beta - 1$, then for $0 < \eta' < \eta$ equation (4.15) evaluates to
\[
S_2' \leq O \left( \sum_{u \in V} \left( \frac{w_u}{n} \right)^{\tilde{\alpha}} n \lambda_u^{1-\tilde{\alpha}+\beta-\eta} \right) = O \left( \frac{n^{2-\beta+\eta}}{\mu^{d(1-\tilde{\alpha}+\beta+\eta)}} \sum_{u \in V} w_u^{\beta-1-\eta} \right) 4.27 \leq O \left( \frac{n^{2-\beta+\eta}}{\mu^{d(1-\tilde{\alpha}-\beta+\eta)}} \int_{w_{\min}}^{\infty} n w^{1-\beta+\eta} w^{\beta-2-\eta} d w \right) = O \left( n \cdot \frac{n^{2-\beta+\eta}}{\mu^{d(1-\tilde{\alpha}-\beta+\eta)}} \right).
\]

Therefore, $S_2 = \mu^{d\tilde{\alpha}-d} (1 + \log(n^d / \mu)) S_2' \leq O(n \cdot (n / \mu^d)^{2-\beta+\eta})$, which is one of the terms in (4.10). On the other hand, if $\tilde{\alpha} < \beta - 1$ then for $0 < \eta < \beta - \tilde{\alpha} - 1$ we obtain from (4.15)
\[
S_2' \leq O \left( n^{1-\tilde{\alpha}} \sum_{u \in V} w_u^{\tilde{\alpha}} \right) 4.27 \leq O \left( n^{1-\tilde{\alpha}} \int_{w_{\min}}^{\infty} n w^{1-\beta+\eta} w^{\tilde{\alpha}-1} d w \right) \leq O(n^{2-\tilde{\alpha}}),
\]
and again $S_2 = \mu^{d\tilde{\alpha}-d} (1 + \log(n^d / \mu)) S_2'$ corresponds to terms of (4.10) after plugging in $\tilde{\alpha}$. This concludes the proof. \qed
The algorithmic small-world phenomenon, empirically established by Milgram's letter forwarding experiments from the 60s [Mil67; TM69], was theoretically explained by Kleinberg in 2000 [Kle00a]. However, from today's perspective his model has several severe shortcomings that limit the applicability to real-world networks. In order to give a more convincing explanation of the algorithmic small-world phenomenon, we study decentralized greedy routing in geometric inhomogeneous random graphs which overcome all previous shortcomings.

We analyze a purely distributed routing protocol where each vertex only needs to know information about its direct neighbors. We prove that greedy routing succeeds with constant probability, and in case of success almost surely finds an almost shortest path of length at most $\frac{2+o(1)}{\log |\beta-2|} \log \log n$, where our bound is tight including the leading constant as it matches the average distance of the random graph. Moreover, we study natural local patching methods which augment greedy routing by backtracking and which do not require any global knowledge. We show that such methods can ensure success probability 1 in an asymptoti-
cally tight number of steps.

These results also explain several experimental findings [BK09; BKC09; CC09; Kri+09; Pap+10] and address the question of Krioukov et al. [Kri+07] whether there are efficient local routing protocols for the internet graph. There were promising experimental studies, but the question remained unsolved theoretically. Our results give for the first time a rigorous and analytic affirmative answer.

The entire chapter is joint work with Karl Bringmann, Johannes Lengler, Yannic Maus, and Anisur Molla [Bri+17; Bri+16].

5.1 Introduction

The idea that everyone is connected to everybody else through six degrees of separation, also known as the small world phenomenon, was empirically established by the classic experiments of Milgram [Mil67]. In the last 50 years, this phenomenon inspired researchers in various disciplines such as sociology, physics, and experimental and theoretical computer science to study networks with small-world properties. The first theoretical explanation of the algorithmic small-world phenomenon was the seminal work of Kleinberg [Kle00b] who showed that greedy routing succeeds in $O(\log^2 n)$ steps on certain random graphs. However, as we will discuss below, his model has several severe shortcomings, that have not been resolved simultaneously in subsequent work: routing succeeds only if a perfect lattice structure is contained in the network, the result is fragile with respect to small changes in the model, and the graph model is unrealistically homogeneous.

In order to give a more cogent theoretical explanation for Milgram’s result, we study greedy routing in GIRGs as we strongly believe that this model is more flexible and more convincing. On the one hand, our results of Chapter 4 establish GIRGs as a realistic model for real-world networks from a theoretical point of view. On the other hand, hyperbolic random graphs are special case of GIRGs (Theorem 3.15) and this particular model has also been extensively experimentally validated, cf. [ST08; Zha+11; PPK15; GP+16]. E.g., Boguñá et al. [BPK10]
computed and studied a (heuristic) maximum likelihood fit of the internet graph into the hyperbolic random graph model, based on ideas from [Kle07], and demonstrated its quality by showing that greedy routing in the underlying geometry of the fit finds near-optimal shortest paths. This successful embedding of a real-world network validates the model experimentally, and was repeated and extended in [PPK15; Bl16].

Furthermore, studying greedy routing in GIRGs is beneficial for the following reasons.

(1) GIRGs are a natural extension of Kleinberg’s model that overcome its three main shortcomings. We discuss this below on page 113 in more details.

(2) In general it is unclear how greedy routing generalizes to inhomogeneous geometric graphs. An ideal neighbor should optimize two objectives: being close to the target, and having large weight. Many possibilities have been suggested to resolve this conflict (e.g., in the physics community [TAK16; HCW14]). Since GIRGs have a specific connection probability for each pair of vertices, there is a natural interpretation of Milgram’s instruction to route to an “acquaintance who is more likely than you to know the target person” [TM69]: pick the neighbor most likely being adjacent to the target. Thus, Milgram’s experiment has a natural analogon on GIRGs.

(3) In the digital age, the area of routing protocols plays a more and more important role. Krioukov et al. [Kri+07] identified as a major open problem in this area “whether we can devise routing protocols for the internet that, having no full view of the network topology, can still efficiently route messages”. This question stimulated further research on the embeddings from [BPK10; Kle07], and by now routing in such graphs is experimentally well-studied [PPK15; Bl16; BK09; BKC09; CC09; Pap+10; Kri+09; Kri+10]. It turned out that greedy routing works experimentally surprisingly well on hyperbolic random graphs and GIRGs [Kri+09; Kri+10; Pap+10]. For example, the resulting greedy paths are only marginally longer than shortest paths in the network. Explaining these experimental findings theoretically was a driving question in the theory community, but it remained open.
We thus believe that studying greedy routing in GIRGs, in each step maximizing the connection probability to the target, gives a convincing theoretical explanation for Milgram’s experiments, and addresses at the same time the question of Krioukov et al. We obtain the following results for any fixed source \( s \) and target \( t \) (for exact statements see Section 5.3).

- Greedy routing succeeds with probability \( \Omega(1) \).

This is already surprising, as incorporating noisy positions in Kleinberg’s model results in a tiny success probability, as we discuss below. The result is also best possible, as in GIRGs the target node is isolated with constant probability.

- A.a.s. the stretch of greedy routing in case of success is \( 1 + o(1) \). (The stretch is the ratio of the lengths of the greedy path and the shortest path between source and target.) Hence, greedy routing is asymptotically optimal. In particular, since the average distance in the giant component of the GIRG model is ultra-small, i.e., \( \Theta(\log \log n) \) (Theorem 4.3), so are the lengths of the greedy paths.

Arguably, \( \Theta(\log \log n) \) is closer to six degrees of separation than Kleinberg’s \( O(\log^2 n) \), and average distances in modern networks are even smaller [Bac+12]. All previous analyses of greedy routing in small-world models yielded at least \( \Omega(\log n) \) steps, e.g. [Kle00b; FG10].

- Our results are robust in the model parameters, i.e., they hold for all parameter choices of the model, in particular for any power-law exponents in the natural range \((2, 3)\). Moreover, we show that nodes do not need to know exact weights and positions of their neighbors, but rather rough approximations suffice for all results.

Since constant success probability is too low for technical application, it is natural to ask whether there are patching methods that use backtracking to enforce success (if source and target are in the same component). Indeed, this is possible in a very strong sense.
Every patching method that satisfies three natural criteria has success probability 1 (conditional on \(s\) and \(t\) being in the same component), and still has a.a.s. stretch \(1 + o(1)\).

This result gives (in our model) a strongly affirmative answer to Krioukov et al.’s question. Crucially, each node only needs to know the positions and weights of its direct neighbors and the geometric position of \(t\), which we assume to be part of the message. Note that the greedy routing protocol is distributed and highly energy efficient since only one node needs to be awake at a time. Perhaps surprisingly, the same can be achieved for the patching protocols, so we can at the same time guarantee success and maintain the benefits of local protocols, with negligible additional costs.

Finally, our results explain many findings of the extensive experimental work which has been done on greedy routing in GIRGs. We postpone the discussion to Section 5.4.

**Kleinberg’s model:** In Section 3.6 we defined Kleinberg’s model formally. Informally speaking, we start with an \(n \times n\) lattice graph, and add for each node \(u\) a constant number of additional long-range edges, where the other endpoint \(v\) is chosen proportional to \(\|x_u - x_v\|^{-\alpha - 2}\) with decay parameter \(\alpha\). This model has several shortcomings that limit the applicability to real-world networks and that are also present in subsequent work [FM06; MN04; FGP06; MNW04; CFL08; Kle02].

As shown by Kleinberg, his model is very fragile with respect to the decay parameter in the exponent of the long-range probability distribution, since changing this distribution to \(\|x_u - x_v\|^{-\alpha - 2}\) for any \(\alpha \neq 1\) increases the expected number of steps to \(n^\Omega(1)\). This is reasonable for \(\alpha < 1\), as in this regime most neighbors of most vertices lie very far and thus we cannot expect greedy routing to utilize locality. However, for \(\alpha > 1\) there is no intuitive argument ruling out greedy routing in general. In particular, investigations of trajectories of bank notes in main cites of the USA suggest a value of \(\alpha = 1.59 \pm 0.02\) [BHG06].

Furthermore, Kleinberg’s model assumes an underlying perfect lattice substrate. In this way, every vertex knows a priori a path to the target, which is an
unrealistically strong assumption. In a more realistic model each vertex might choose a random position \( x_v \) in some geometric space, representing real-world coordinates, but also interests and occupation, followed by the same edge sampling procedure as in Kleinberg’s model. It can be shown that in this adapted model with high probability greedy routing does not reach the target, essentially since in each step the current vertex has constant probability not to have any neighbor with smaller distance to the target. Thus, the unnatural assumption of a perfect lattice (or a similar globally known structure) is crucial for Kleinberg’s result.

Finally, Kleinberg’s graphs are homogeneous, i.e., all vertices have the same degree. This is in strong contrast to real-world networks that have been found to be scale-free almost ubiquitously (see Section 1.1).

Despite some effort to remove these three shortcomings (see related work below), so far there was still no model which avoids all of them. Recall that on GIRGs, for fixed weights the edge probability \( p_{uv} \) is proportional to Kleinberg’s distribution \( \|x_u - x_v\|^{-\alpha} \cdot d \) for fixed weights, where we want to choose \( \alpha > 1 \). As our results do not depend on the concrete choice of \( \alpha \), we overcome the problem of fragile exponents. GIRGs also remove the implausible assumption of a perfect grid and are a natural scale-free variant of Kleinberg’s model. Moreover, the corresponding extensions of Kleinberg’s model were developed solely in the context of routing protocols. Thus there is a much larger body of evidence (theoretical and experimental) that GIRGs resemble real-world networks than there is for any other graph model in this area.

**Further related work:** For psychological work related to Milgram’s experiments we refer to the survey by Schnettler [Sch09]. The enormous amount of experimental and non-rigorous work on routing in real-world networks and random graphs is surveyed by Huang et al. [HCW14]. Regarding theoretical work, a monograph by Watts summarizes the early work [Wat04b]. Let us highlight a few more recent theory papers on greedy routing and related decentralized routing schemes: Watts et al. [WDN02] and Kleinberg [Kle02] studied hierarchical network models that are well motivated by sociological principles. Alternatively,
Fraigniaud et al. [FG14] augmented Kleinberg’s model with long-range edges according to a power law. These approaches make the model somewhat more realistic, however, they still suffer from similar shortcomings as Kleinberg’s original model, e.g., they either assume a perfect \( b \)-ary tree structure or a perfect lattice. On the other hand, after a long line of research [LN+05; Fra05; KLNT06; Duc+06; AG06; Fra07] Fraigniaud et al. [FG10] replaced the lattice structure in Kleinberg’s model with arbitrary connected base networks. They showed how these base networks can be augmented with long-range edges to make them navigable, which in the most general case means that variants of greedy routing succeed with a subpolynomial number of steps, provided that all nodes know the base network. Note that this available information is much less local than in our model, where each node only needs information about its direct neighbors. In particular, since every node knows in advance a path to the target, greedy routing succeeds with probability one.

Arguably, all these networks were designed specifically for studying greedy routing, while our approach is rather to study greedy routing in a model that is designed in the first instance to capture the structure of real-world networks. Also, none of the previous theoretical work showed ultra-small routing distance or considered patching strategies.

**Organization of the chapter:** After formally introducing the model in Section 5.2, we list and explain our main results in Section 5.3. In Section 5.4 we discuss related experimental results. In Section 5.5 we discuss the patching criteria (P1)-(P3) and related literature on patching algorithms, and give two simple examples for patching algorithms which satisfy (P1)-(P3). Section 5.6 contains a high level discussion of the typical trajectory of greedy routing, and it also contains a sketch of the main ideas of our proofs. Section 5.7 contains preparations for the proofs. In Section 5.8.1 we prove the main Lemma 5.16, which lies at the heart of all subsequent proofs, and finally we give the proofs of our main results in Sections 5.8 (basic algorithm), 5.9 (patching), 5.10 (relaxations), and 5.11 (hyperbolic random graphs).
5.2 Random Graph Model & Greedy Routing

For technical reasons, in this chapter we use a slightly modified version of GIRGs by applying a Poisson point process. The same model was used in [KL16] for studying bootstrap percolation. In this section, we describe the small modifications on the random graph model and define the routing model.

Poissonized GIRGs: As geometric ground space, we again use the torus $\mathbb{T}^d$, equipped with any norm. Note again that when embedding a real world network into the GIRG model, e.g., as in [BPK10], the real world coordinates do not necessarily resemble the coordinates in the geometric space. It rather resembles an ambient space which additionally includes more parameters such as occupation and interests. Then, let the set of vertices $V$ be given by a Poisson point process on $\mathbb{T}^d$ with intensity $n$, cf. [Str10] for a formal definition of Poisson point processes. In particular, the expected number of vertices is $n$, the positions of all vertices are uniformly at random in $\mathbb{T}^d$, and for each measurable $A \subseteq \mathbb{T}^d$, the number of vertices in $A$ is Poisson distributed with mean $n \cdot \text{vol}(A)$. Moreover, if $A, B \subseteq \mathbb{T}^d$ are disjoint measurable sets, then the number of vertices in $A$ and in $B$ are independent random variables.

Let $2 < \beta < 3$ and $w_{\text{min}} > 0$ be fixed constants. In this chapter we assume that the weights are chosen at random (cf. Lemma 3.5). More precisely each vertex $v \in V$ draws independently a weight according to the probability density function

$$f(w) := \begin{cases} \Theta(\frac{w^{\beta - 1}}{w_{\text{min}}^{\beta - 1}}) w^{-\beta} & \text{if } w \geq w_{\text{min}}, \\ 0 & \text{if } w < w_{\text{min}}. \end{cases}$$

For the edge probability function $p_{uv}$, we require for some constant $\alpha > 1$ the condition

$$p_{uv} = \Theta\left(\min\left\{ \frac{1}{\|x_u - x_v\|^\alpha d} \cdot \left(\frac{w_u w_v}{w_{\text{min}} n}\right)^{\alpha d}, 1 \right\} \right). \tag{5.1}$$

In particular, there exists a constant $c_1 > 0$ such that $p_{uv} = \Theta(1)$ if $u$ and $v$ are very close to each other, i.e., $\|x_u - x_v\|^d \leq c_1 \frac{w_u w_v}{w_{\text{min}} n}$. In the threshold case we
require instead of (5.1) that there are constants $0 < c_1 \leq c_2$ such that

\[
(5.2) \quad p_{uv} = \begin{cases} 
\Theta(1), & \text{if } \|x_u - x_v\| \leq c_1 \frac{w_u w_v}{w_{\min} n} \\
0, & \text{if } \|x_u - x_v\| > c_2 \frac{w_u w_v}{w_{\min} n}.
\end{cases}
\]

Note that there can be an interval $[(c_1 \frac{w_u w_v}{w_{\min} n})^{1/d}, (c_2 \frac{w_u w_v}{w_{\min} n})^{1/d}]$ for $\|x_u - x_v\|$ where the behavior of $p_{uv}$ is arbitrary. In many settings the connection probability is 1 if $u$ and $v$ are sufficiently close to each other.

In contrast to the previous definition of GIRGs as stated of Chapter 3, we make the dependency on the parameter $w_{\min}$ explicit. Note that in (5.1) and (5.2), the $W$ in the denominator is replaced by $n$. The formulas (5.1) and (5.2) depend on the parameter $w_{\min}$ in such a way that the expected degree of a vertex of weight $w$ is $\Theta(w)$. In all subsequent theorems and proofs, the $O$-notation may hide any dependence on the constant parameters $d, \alpha, \beta$ and on the hidden constants in (5.1) and (5.2), but not on $w_{\min}$ as Theorem 5.2 will link the failure probability of the routing protocol to $w_{\min}$.

**Routing protocol:** We evaluate the performance of greedy routing in a GIRG, which is the following process. A message should be sent from a starting vertex $s$ (source) to a target vertex $t$. The *address* of every vertex $v$ is the pair $(x_v, w_v)$. Every vertex has local information, i.e., it knows the address of itself and of its neighbors. In addition, the address of the target is written on the packet. Then the routing proceeds in rounds and in every hop, the packet is sent from the current vertex $v$ to a neighbor $u$ which maximizes a given objective function $\phi$, cf. below. When a dead end is reached, the process stops and the packet is dropped. If $u = t$, then we say that the routing was *successful*. A pseudocode description is given with Algorithm 5.1 below. Note that we will analyze the process for arbitrary vertices $s$ and $t$, i.e., we may explicitly choose weights and positions of $s$ and $t$, while everything else of the graph is still drawn randomly.

**Objective function:** It is natural and in spirit of Milgram’s experiment [TM69] that in every round, the packet should be sent to a neighbor $v$ which maximizes $p_{vt}$. Note that if $\alpha < \infty$, maximizing $p_{vt}$ is equivalent to maximizing
Algorithm 5.1 Greedy Routing Algorithm

1: function GREEDY($s$, $m$) \Comment{$s$: source, $m$: message (contains target information $t$)}
2: \hspace{1em} if $s == t$ then deliver message
3: \hspace{1em} else
4: \hspace{2em} $v \leftarrow \text{argmax}\{\phi(u) \mid u \in \Gamma(s)\}$
5: \hspace{2em} if $\phi(v) > \phi(s)$ then GREEDY($v$, $m$)
6: \hspace{1em} else return failure

$w_v/(w_{\min} n \|x_v - x_t\|^d)$. Hence, for our analysis we use the objective function

$$\phi(v) := \frac{w_v}{w_{\min} n \|x_v - x_t\|^d}$$

and observe that the target vertex $t$ globally maximizes $\phi$, which is a necessary condition for any reasonable objective function. In particular, if $\{s, t\} \in E$, then the algorithm sends the packet directly to the target. We keep the normalization factor $n$ since it will turn out to be the most natural normalization. In this way, $\phi(v)$ will always increase by the same exponent in the second phase of the algorithm, cf. Section 5.6. We use the additional $w_{\min}$-factor only for technical reasons, as it simplifies several calculations.

We emphasize that each vertex only needs to know the positions and weights of its direct neighbors, and the geometric position of $t$ (which we assume to be part of the message). This goes in line with Milgram’s experiment, where participants knew geometric positions and professions of their acquaintances and made their choices accordingly [Mil67; TM69].

5.3 Results

In this section we list and explain all our results about routing in GIRGs. Unless stated otherwise, we assume throughout the chapter that $s$ and $t$ are fixed, while the rest of the graph is drawn randomly. That is, an adversary may pick weights and positions of $s$ and $t$, while the remaining vertices and all edges are drawn randomly as described in Section 5.2. (Equivalently, we can first choose the random graph $G$, then draw $s$ and $t$ at random, and condition the resulting
probability space on $x_s, x_t, w_s,$ and $w_t$.) We then consider greedy routing from $s$ to $t$. Proof sketches and intuitive reasons for the results can be found in Section 5.6. The formal proofs can be found in Section 5.8 (basic algorithm), Section 5.9 (patching), Section 5.10 (relaxations), and Section 5.11 (hyperbolic random graphs).

Success probability: As our first result we show that greedy routing from $s$ to $t$ succeeds with constant probability. The proof of the following theorem is in Section 5.8.2.

**Theorem 5.1.** Greedy routing succeeds with probability $\Omega(1)$.

In general, we cannot expect anything better than constant success probability, since we only assumed a connection probability of $\Theta(1)$ in (5.1) and (5.2) even if two vertices are very close to each other. So we may end in a vertex which is extremely close to $t$, but which fails to actually connect to $t$. If we assume connection probability 1 in such cases, then we get a better connection probability. More precisely, assume that the constant $c_1$ in (5.1) or (5.2) is chosen such that the following condition holds.

\begin{equation}
\tag{5.3}
p_{uv} = 1 \quad \text{if} \quad \|x_u - x_v\|^d \leq c_1 \frac{w_u w_v}{w_{\min}}.
\end{equation}

Then we can strengthen Theorem 5.1 by showing that the failure probability decays exponentially with $w_{\min}$.

**Theorem 5.2.** Suppose that (5.3) holds in addition to the model assumptions. Then:

(i) Greedy routing succeeds with probability $1 - O(e^{-w_{\min}^\Omega(1)})$, where the hidden constants do not depend on $w_{\min}$.

(ii) If $\min\{w_s, w_t\} = \omega(1)$ then with probability $1 - \min\{w_s, w_t\}^{-\Omega(1)}$ greedy routing succeeds. In particular, in this case greedy routing succeeds a.a.s.

We will prove this result in Section 5.8.3. The first part of the theorem is also almost optimal. Indeed, the degree of a vertex $v \in V$ is Poisson distributed with
parameter $\Theta(w_v)$, as we will see in Section 5.7.1. Therefore a vertex $v$ of weight $w_v$ is isolated with probability $e^{-\Theta(w_v)}$. If for example $w_t = \Theta(w_{\text{min}})$, then the success probability cannot be higher than $1 - e^{-\Theta(w_{\text{min}})}$.

For both statements we remark that we only need (5.3) for the very last step. Even without (5.3), with sufficiently high probability greedy routing finds a vertex $u$ such that $u$ and $t$ satisfy the precondition of (5.3). However, without (5.3) the vertex $u$ only has probability $\Theta(1)$ to connect to $t$ directly, and there is a considerable chance that there is no better neighbor of $u$, i.e., that $u$ is a local optimum.

**Stretch:** One of the most crucial efficiency measures for routing protocols is the stretch, i.e., the ratio of the routing paths length compared to the length of a shortest path. Clearly, routing can only succeed if the source $s$ and the target $t$ are in the same component of the graph. By Theorem 4.2 and Theorem 4.3, a.a.s. a GIRG possesses a giant connected component of linear size in which the average distance is $\frac{2+o(1)}{|\log(\beta-2)|}$ log log $n$. The next theorem shows that the number of steps of the greedy routing algorithm (either the target is found or the routing stops) is a.a.s. equal to shortest paths, up to a stretch of $1 + o(1)$. The proof is given in Section 5.8.4.

**Theorem 5.3.** A.a.s., greedy routing stops after at most $\frac{2+o(1)}{|\log(\beta-2)|}$ log log $n$ steps and either finds $t$ or ends in a dead end. More precisely, for every $f_0(n) = \omega(1)$ a.a.s. it holds the stronger bound

$$1 + o(1) \left( \frac{1}{|\log(\beta-2)|} (\log \log w_s, \phi(s)^{-1} + \log \log w_t, \phi(s)^{-1}) + O(f_0(n)) \right).$$

Note that Theorems 5.1 and 5.3 together imply that even if we condition on greedy routing being successful, then still a.a.s. greedy routing needs at most $\frac{2+o(1)}{|\log(\beta-2)|}$ log log $n$ steps. Since this agrees with the average distance in the giant, it is optimal. In particular, for two vertices $s, t$ with random weights and positions, if greedy routing between $s$ and $t$ is successful then a.a.s. the stretch is $1 + o(1)$. Here, we use statement (v) of Theorem 4.20 that a.a.s. a $(1 - o(1))$-fraction of all pairs of vertices in the giant have distance at most $\frac{2+o(1)}{|\log(\beta-2)|}$ log log $n$. 
In the most typical situation (e.g., if \( s \) and \( t \) are chosen randomly), (5.4) evaluates to the simpler expression \( \frac{2+o(1)}{\log(\beta-2)} \log n \) in Theorem 5.3. For example, this is the case if \( s \) and \( t \) have constant weights and are far apart from each other: \( w_s, w_t = O(1) \), and \( \|x_s - x_t\| = \Omega(1) \). However, if \( s \) and \( t \) have large weights or are close to each other, then (5.4) gives a much better bound (but never a worse bound, since \( \phi(s) = \Omega(1/n) \) is always true).

Patching: As outlined in Section 5.1, greedy routing may have practical applications. For these applications, it is not acceptable that the algorithm simply fails if it enters a local maximum. Therefore, several patching methods have been proposed in the literature. We show that any patching algorithm \( A \) is efficient if it satisfies the following three basic conditions. For further discussions, in particular of the conditions, see Section 5.5.

(P1) (Greedy choices) If \( A \) is in some vertex \( v \) and decides to visit an unexplored neighbor then it must choose the unexplored neighbor of largest objective. If \( A \) visits a vertex \( v \) for the first time, and \( v \) has a neighbor of larger objective, then \( A \) proceeds to the neighbor of \( v \) with largest objective.

(P2) (Poly-time exploration) If \( A \) has explored \( k \) vertices, and at least one of these vertices has an unexplored neighbor, then \( A \) visits an unexplored vertex in \( k^{O(1)} \) steps.

(P3) (Poly-time exhaustive search) Assume that \( A \) visits a vertex \( v \) that has larger objective than all previously visited vertices, and let \( S \) be the connected component of \( v \) in \( G[V_{\geq \phi(v)}] \), i.e., \( S \) is the set of all vertices that can be reached from \( v \) without touching vertices of worse objective than \( v \). Then \( A \) visits the vertices in \( S \) in the next \( |S|^{O(1)} \) steps (or finds \( t \) in one of these steps). In the exceptional case \( \phi(s) = \Omega(1) \), we require in addition that there exists a sufficiently slowly falling \( \phi_0 = o(1) \) such that \( A \) explores the connected component \( S \) of \( s \) in \( G[V_{\geq \phi_0}] \) in \( |S|^{O(1)} \) steps.

Condition (P2) ensures that greedy routing is always successful if source \( s \) and target \( t \) are in the same component. The surprising result is that (P1)-(P3)
are already sufficient to ensure efficient greedy routing, i.e., a.a.s. the stretch is $1 + o(1)$. The proof can be found in Section 5.9.

**Theorem 5.4.** Assume that $A$ is a routing algorithm that satisfies (P1)-(P3). If $s$ and $t$ are in the same component then $A$ always successfully routes from $s$ to $t$, and a.a.s. this takes at most $2 + o(1) \cdot \log \log n$ steps.

**Relaxations:** So far we assumed that the objective $\phi$ can be computed exactly, which may be unrealistic in practice. For example, in the Milgram experiment the participants can only estimate the objective values of their neighbors. Hence, we also study an approximate version of greedy routing, where the greedy algorithm uses an objective function $\tilde{\phi}$ which is only an approximation of $\phi$. This adds quite some flexibility: for any of the best $\min\{w_v, \phi(v)^{-1}\} \cdot o(1)$ neighbors of a vertex $v$, there is a “good enough” approximation $\tilde{\phi}$ which puts this vertex to the top. In particular, it is not necessary to compute the optimal neighbor w.r.t. $\phi$, which may be costly. Moreover, the routing process is also very robust: for example, it is no problem if some of the edges fail during execution of the routing, since the current vertex can send the message to any other good neighbor instead.

**Theorem 5.5.** Let $\tilde{\phi} : V \to \mathbb{R}$ be a function which is maximized at $t$ such that $\tilde{\phi}(v) = \Theta(\phi(v))$ holds for all $v \in V$. Then Theorems 5.1, 5.2, 5.3, and 5.4 also hold for greedy routing with respect to $\tilde{\phi}$ instead of $\phi$.

We may also replace $\tilde{\phi}(v) = \Theta(\phi(v))$ by the weaker assumption

$$\tilde{\phi} = \Theta(\phi(v) \cdot \min\{w_v, \phi(v)^{-1}\} \cdot o(1)).$$

We prove Theorem 5.5 in Section 5.10. A consequence of Theorem 5.5 is that the results of this chapter also hold for variations of the model. For example, most of the experimental work is performed on hyperbolic random graphs (see Section 5.4 for details). By Theorem 3.15, one-dimensional GIRGs ($d = 1$) contain the model of hyperbolic random graphs as a special case, but still geometric routing in hyperbolic space is not exactly the same as greedy routing in our sense. Theorem 5.5 shows that the differences are negligible. Formally, geometric
routing in hyperbolic random graphs induces an objective function $\phi_H$ on the corresponding GIRG which falls into the class of objective functions considered in Theorem 5.5. The proof is contained in Section 5.11.

Corollary 5.6. All theorems about success probability and path length apply as well for geometric routing in hyperbolic random graphs. Moreover, if $t$ is a random vertex, then Theorem 5.4 for patching algorithms also holds for hyperbolic random graphs.

5.4 Comparison with Experimental Results

There are several recent papers which experimentally study “greedy processes” in graphs with hidden metric spaces [BK09; BKC09; BPK10; CC09; Kri+09; Pap+10]. In these papers, often hyperbolic random graphs are sampled for various parameter combinations, sometimes under different names; or given real networks (e.g., the internet graph) are embedded into the hyperbolic space. Afterwards, in both settings, the authors experimentally study how greedy routing algorithms perform on these graphs. Thus we give theoretical explanations for the effects that were observed in these papers, which we discuss in detail below.

In contrast, the authors of [BK09] and [BKC09] study geometric greedy processes, i.e., the routing is degree-agnostic and only uses the distance in a homogeneous and isotropic space. It turns out that this geometric type of routing still works in some settings, but is far less efficient and robust (e.g., it completely fails for some values $2 < \beta < 3$). This suggests that greedy routing as considered in this chapter is superior to geometric routing. It remains an intriguing open task to understand theoretically how purely geometric routing behaves for realistic network models.

Success probability: As main contribution the aforementioned papers observe a large success probability of greedy routing for most studied combinations of parameters. This is explained by Theorem 5.1, which gives a general lower bound $\Omega(1)$ on the success probability. Often, the experimentally observed
success probability is surprisingly large (i.e., larger than 97% in [BPK10]). We believe that we can give the reason for this observation with Theorem 5.2 (i), where we show that the failure probability drops exponentially with the parameter $w_{\text{min}}$ controlling the minimum expected degree. Thus, even with very moderate values of $w_{\text{min}}$ it is not surprising to obtain large success probabilities.

**Trajectory of a greedy path:** In their experiments the authors in [BPK10; Kri+09; Kri+10; Pap+10] observed that a successful greedy path starting at a small degree node first visits nodes of larger degrees and continues along this path until it reaches the core of the network; afterwards, the greedy path leaves the core towards nodes with smaller degrees but but much shorter geometric distance to the target until it reaches its destination. We prove this observation formally. In our proofs, we partition our network into layers and show that the greedy path follows closely this layer structure. In the first half of the process, the layers are characterized by increasing weights, in the second by decreasing geometric distances to the target. We prove that a.a.s. the greedy algorithm visits each layer at most once, and that it visits a $(1 - o(1))$-fraction of all layers. Thus we characterize accurately the trajectory that greedy routing takes.

**Stretch:** Besides the success probability, the stretch of a greedy path, i.e., the quotient of a greedy paths length vs. the length of a shortest path, is one of the most meaningful measures to evaluate the usefulness of greedy routing in applications. Experiments revealed a surprisingly small stretch close to 1, both for embedded real graphs [BPK10] and for sampled random graphs [Kri+09; Kri+10; Pap+10]. In Theorem 5.3 we give the theoretical explanation and formally prove that successful greedy routing has a stretch of $1 + o(1)$ in the case of success.

**Patching:** Patching has been experimentally found to be highly efficient, and to maintain an average stretch close to 1 [CC09; HCW14; Kri+09; Kri+10; Pap+10]. However, most experiments in sampled models have been performed with a small number of vertices (e.g., $n = 1000$). Also, not all methods guarantee successful routing if $s$ and $t$ are in the same component. Interestingly, the
gravity-pressure method considered in [CC09] shows good results, although
it does not fall into our class of patching algorithms, and it suffers from the
problems described in Section 5.5. So it remains open whether this algorithm is
also theoretically efficient, or whether the good results were simulation artefacts
(see also [Sah+13]).

As Theorem 5.4 shows, a broad class of natural patching algorithms are
highly efficient, even if they boost the success probability to 100% for vertices
in the same component. Most notably, we prove in Theorem 5.4 that a stretch
of $1 + o(1)$ is not only achievable, but that it is guaranteed under the weak and
natural assumptions (P1)-(P3).

5.5 Patching Algorithms: Discussion and Examples

The naive greedy protocol gets stuck if it enters a local optimum, which is
not acceptable for technical applications. Several different patching strategies
have been suggested to overcome this issue, and experimental results showed
that many of them work quite well in practice [CC09; HCW14; Kri+09; Kri+10;
Pap+10; Sah+13]. We formally prove that in fact any patching strategy works well
if it satisfies some basic conditions, namely (P1) greedy choices, (P2) poly-time
exploration, and (P3) poly-time exhaustive search (cf. page 121).

Note that all three conditions are natural: the first one is that the algorithm
makes greedy choices whenever there is an obvious decision to be made. Note
that still there are many choices left to the algorithm. For example, if it enters
a local optimum, then it is free to decide if and how much it backtracks be-
fore continuing exploration. The second condition ensures that $A$ never gets
stuck (or only for polynomial time). Finally, the third condition (P3) forces the
algorithm to search exhaustively a connected set of “good” candidates in rea-
sonable time before turning to less promising candidates. The second part of
(P3) is concerned with the exceptional case $\phi(s) = \Omega(1)$ (for random $s$ and $t$ this
only happens with probability $O(1/n)$) and is necessary to avoid trivialities: if
$\phi(s) = \Omega(1)$ then there may be no (or only few) vertices of better objective than
$s$, and the precondition of the first part of (P3) may never be fulfillled. We remark
that the requirement that \( \phi_0 \) is “sufficiently slowly falling” can be specified to \( \phi_0 = (\log \log n)^{-o(1)} \).

The last condition (P3) does exclude some algorithms like the gravity-pressure algorithm introduced in [Pap+10]. This algorithm always visits the best neighbor \( v' \) of the current search point \( v \), even if \( \phi(v') < \phi(v) \), and thus does not satisfy (P3). E.g., assume that the algorithm visits two consecutive vertices \( v_1, v_2 \) of increasing objective, and assume further that \( v_2 \) has exactly one more neighbor \( u \) of very bad objective (\( \phi(u) = o(\phi(v_1)) \)), while \( v_1 \) has other neighbors of better objective. Then continuing the search from \( u \) before exploring the other neighbors of \( v_1 \) may be a bad strategy, and this is what (P3) rules out. For example, the gravity-pressure algorithm will prefer to visit any unexplored vertex over returning to \( v_1 \). So if \( v_1 \) lies on the only path to \( t \) (which happens with probability \( \Omega(1) \)), then the algorithm may potentially explore large parts of the giant before returning to \( v_1 \), and thus before finding \( t \). This may explain that the algorithm is especially vulnerable in sparse networks [Sah+13].

Generally, there are two relevant types of algorithms if the global structure is unknown to the vertices: either information about the routing history is stored in the message (e.g., the protocol SMTP for emails [Pos82]); or for every message a small amount of information is stored in each vertex, yielding rather exploration than classical routing (e.g., flooding algorithms like [HABR05; KV00], the gravity pressure algorithm [Pap+10], or tree-based approaches [Sah+13]). In both cases, it is very easy to design a routing algorithm that satisfies conditions (P1)-(P3), as we outline in the following. For simplicity, we assume that \( \phi(s) = o(1) \).

For the first case, we may simply store the list of visited vertices in the message, and for each vertex \( v \) we additionally store the objective of the best unexplored incident edge in the message (i.e., the objective of the best neighbor \( u \) of \( v \) for which the algorithm did not traverse the edge \( uv \)). Compared to SMTP this only increases the required memory by a single value per visited node. With this information, a trivial way to satisfy conditions (P1)-(P3) is to use the greedy algorithm if possible (i.e., if we are not in a local optimum), and otherwise explore the best unexplored edge that goes out from any visited vertex.

For the second case, a variant of depth-first search satisfies condition (P1)-
5.5. PATCHING ALGORITHMS: DISCUSSION AND EXAMPLES

(P3), in which the message and each visited vertex only need to store a constant number of pointers and objective values. More precisely, for a vertex $v \in V$ and a value $\Phi \in \mathbb{R}^+$ consider a greedy depth first search on the subgraph $G[V_{\geq \Phi}]$ of vertices which have objective at least $\Phi$, starting in $v$. “Greedy” here means that if there are several unexplored edges going out from a vertex, then the DFS algorithm picks the edge that leads to the vertex of highest objective. We call this algorithm (greedy) $\Phi$-DFS for short. Note that since $G[V_{\geq \Phi}]$ does not need to be connected, the $\Phi$-DFS does not necessarily visit all vertices in $G[V_{\geq \Phi}]$.

The idea is now the following. Whenever we encounter a vertex $v$ which has strictly larger objective than all previously visited vertices, then we start a $\phi(v)$-DFS at $v$. We do this recursively, so when in this $\phi(v)$-DFS we encounter another vertex $v'$ which has larger objective than all previously visited vertices (including the vertices visited during the $\phi(v)$-DFS), then we pause the $\phi(v)$-DFS, and start a $\phi(v')$-DFS in $v'$. It may happen that the $\phi(v')$-DFS is completed without success, i.e., that we return to $v'$ after recursively exploring all its better neighbors without finding $t$. In this case, we simply discard the $\phi(v')$-DFS and resume the paused $\phi(v)$-DFS. Note that we treat all vertices visited during the $\phi(v')$-DFS as unvisited for the resumed $\phi(v)$-DFS. However, we do store the best objective value that we have ever seen (regardless in which DFS we have seen them) in the message. A pseudocode description is provided in [Bri+16]. The algorithm is distributed, i.e., there is no shared global memory, and no global stack of function calls is required to execute the algorithm. Moreover, at each time only one vertex is active, and each vertex only needs to know positions and weights of its direct neighbors.

Let us argue that this algorithm only needs to store a constant number of pointers and objective values in the message and the visited vertices. During a $\Phi$-DFS, we store at each visited vertex the identity of the parent vertex (to allow backtracking) together with the value $\Phi$, and we call this pair the $\Phi$-information. In the message we keep track of the current value of $\Phi$, the identity of the last visited vertex, and the best seen objective. Note that this information already allows us to perform the $\Phi$-DFS in a greedy depth-first manner, and also to decide whether we want to start a new DFS. If we start a new DFS in a vertex $v$,
we need to store in $v$ all information that is necessary to resume the previous DFS: the parent vertex, the previous value of $\Phi$, and a flag indicating that we started a new DFS in $v$. So for each value of $\Phi$, the $\Phi$-DFS requires a constant memory in each vertex. It only remains to show that no vertex ever needs to store the $\Phi$-information for two different values of $\Phi$ simultaneously. So assume we enter a vertex during a $\Phi$-DFS, and it already contains the $\Phi'$-information for some $\Phi' > \Phi$. Then we started the $\Phi'$-DFS after the $\Phi$-DFS, and the only way to resume the $\Phi$-exploration is that the $\Phi'$-DFS terminated without finding $t$. So in this case we can safely delete the $\Phi'$-information. On the other hand, if we run a $\Phi$-DFS then we claim that we can never encounter a vertex $v$ which carries a $\Phi'$-information for some $\Phi' < \Phi$. Indeed, this could only happen if $\phi(v) \geq \Phi$ (to visit $v$ during the $\Phi$-DFS), and if moreover we visited $v$ during the $\Phi'$-DFS, i.e., before the start of the $\Phi$-DFS. But if such a vertex exists then we wouldn’t have started the $\Phi$-DFS in the first place, since $\Phi$ wouldn’t have been the best encountered objective at this point. So indeed we can not visit a vertex $v$ which carries some $\Phi'$-information for $\Phi' < \Phi$. Summarizing, we never need to store the $\Phi$-information for more than one value of $\Phi$ in the same vertex.

Finally we argue that the algorithm satisfies (P1)-(P3). It is evident that the algorithm satisfies conditions (P1) and (P2). Moreover, it also satisfies (P3): if $x \in V$ and $S \subseteq V$ is the set of all vertices of objective at least $\phi(v)$, then a $\phi(v)$-DFS takes at most $O(|E(S, S)|) = O(S^2)$ steps. For every vertex $u \in S$, we may interrupt the $\phi(v)$-DFS for another $\phi(u)$-DFS, but each of these also takes at most $O(S^2)$ steps. So in total we will explore the set $S$ in at most $O(S^3)$ steps.

5.6 Proof Sketches and the Typical Trajectory

In this section we describe the evolution of the basic greedy process and give a very rough outline of the proof ideas and the main difficulties. At the same time, we describe the trajectory of a typical greedy path, which is also depicted in Figure 5.1 below. Full proofs can be found in the subsequent sections.
Mean field analysis and typical trajectories: For simplicity, assume that both the source $s$ and the target $t$ have constant weight and $\|x_s - x_t\| = \Omega(1)$, which is the typical case. For the sake of readability, we ignore $w_{\min}$-factors in this sketch. As $s$ has small weight, a.a.s. all its neighbors (if any) have relatively small weight as well and are located in a ball of small geometric distance around $x_s$, which we call “region of influence” for illustrative purposes. Since the region of influence of $s$ is small and far away from $t$, all points in this region have the same distance from $t$, up to factors $(1 + o(1))$. In particular, all neighbors of $s$ have the same distance from $t$, up to factors $(1 + o(1))$. On the other hand, the weights of the neighbors fluctuate by non-negligible constant factors. (So do the distances from $s$, but they do not influence the objective values.) Therefore, the fluctuations of the weights dominate the fluctuations of the distances from $t$, and consequently, routing proceeds to a neighbor of $s$ with higher weight. Let $\overline{w}$ be a given weight. By Lemma 3.2, the expected number of neighbors of $s$ with weight at least $\overline{w}$ is

$$
\Theta\left(n \int_0^\infty f(w) \min\left\{ \frac{w w_s}{n}, 1 \right\} \, dw \right) \approx \Theta\left( \int_0^\infty w_s w^{1-\beta} \, dw \right) = \Theta(w_s \overline{w}^{2-\beta}).
$$

Hence, setting $\overline{w} = w_s^{1/(\beta-2)}$ we can expect that $s$ has a neighbor of weight roughly $\overline{w}$. We repeat the argument and observe that the routing exhibits a first phase in which the weight of the current node increases by an exponent $\approx 1/(\beta - 2) > 1$ within every step. This phase stops when a node $v$ with $\phi(v) \geq w_v^{1/(\beta-2)}$ is reached. Since the weight increases by an exponent of $1/(\beta - 2)$ with every hop, this first phase needs only $\approx \log\log n/|\log(\beta-2)|$ steps.

Once the routing reaches a node $v$ such that $\phi(v) \geq w_v^{1/(\beta-2)}$, the relation between the current weight and the current objective changes, since now the region of influence of $v$ contains $t$. That does not mean that $v$ is adjacent to $t$ (in fact, this is very unlikely), but $v$ does have neighbors which are significantly closer to $t$. These neighbors typically have smaller weight than $v$, but the gain in distance is enough to possess higher objective $\phi$. More precisely, let $\delta := (\phi(v)^{1-\beta} n^{-1})^{1/d}$ and observe that, by the definition of $\phi$,

$$
\|x_v - x_t\|^d = w_v \phi(v)^{-1} n^{-1} \geq \phi(v)^{1-\beta} n^{-1} = \delta^d
$$
holds. Thus we have $\delta \leq \|x_v - x_t\|$. Next, we consider the set

$$A(v) := \{ u \in V : w_{u} \geq \phi(v)^{-1} \land \|x_u - x_t\| \leq \delta \}.$$  

We claim that with constant probability, $v$ has a neighbor in $A(v)$. By the Poisson point process, the expected number of vertices located in the ball of radius $\delta$ around $t$ is $\delta d$. It follows that the expected number of vertices in the set $A(v)$ is

$$\Omega\left(n \int_{\phi(v)^{-1}}^{\infty} f(w) \delta^d \, dw\right) = \Omega\left(\phi(v)^{1-\beta} \int_{\phi(v)^{-1}}^{\infty} w^{-\beta} \, dw\right) = \Omega(1),$$

and it remains to verify that vertices in $A(v)$ are connected to $v$ with constant probability. Let $u \in A(v)$. We have $\|x_v - x_u\| = O(\|x_v - x_t\|)$ by the triangle inequality. Next we use that the assumption $\phi(v) \gtrsim w^{-1/(\beta-2)}$ is equivalent to $w_v \gtrsim \phi(v)^{2-\beta}$, and deduce

$$\frac{w_u w_v}{n} \geq \frac{\phi(v)^{-1} w_v}{n} \gtrsim \delta^d \geq \|x_v - x_t\|^d = \Omega(\|x_v - x_u\|^d).$$

Then by (5.1) and (5.2), indeed $v$ is connected to $u$ with constant probability and it follows that with constant probability, $v$ has a neighbor in $A(v)$. Finally, every node $u \in A(v)$ has significantly larger objective than $v$, i.e., it satisfies $\phi(u) \geq (\phi(v) \delta^d n)^{-1} \geq \phi(v)^{\beta-2}$, and in this way we can hope for a neighbor $u$ of objective $\approx \phi(v)^{\beta-2}$. (Note that $\phi(v) < 1$, so indeed the objective increases.) Moreover, it can be shown that the best neighbor $u'$ again satisfies the relation $\phi(u') \gtrsim w_{u'}^{-1/(\beta-2)}$, so the argument can be repeated. Thus in this second phase, within every step the objective increases by the same exponent.

After $\approx \log \log n/|\log(\beta-2)|$ rounds, our second phase reaches a node $v$ of objective $\phi(v) = \Omega(1)$. At this point, in the case $\alpha < \infty$, we have

$$p_{vt} = \Theta\left(\min\{1, (\phi(v) w_t^0)\}\right) = \Omega(1),$$

so $v$ connects directly to $t$ with probability $\Omega(1)$. This sketch gives an idea why greedy routing succeeds with probability $\Omega(1)$ (Theorem 5.1), and why it only needs $\frac{2+o(1)}{|\log(\beta-2)|} \log \log n$ steps (Theorem 5.3). It also yields a typical trajectory of the greedy path, i.e., the routing consists of a first phase in which it merely increases weights and of a second phase in which the objective significantly increases and the distance to the target decreases.
Main difficulties: In the formal proofs of our statements we show that with sufficiently high probability, the routing does not deviate too much from this typical trajectory. For a single vertex $v$ it is not difficult to calculate the weight $w_u$ and the objective $\phi(u)$ of its best neighbor $u$. However, there is a major difficulty: the greedy algorithm only proceeds to $u$ when $u$ is the best neighbor of $v$. So if we want to compute the probability that the best neighbor of $u$ is again what we expect, then we need to compute a conditional probability, namely conditioned on $v$ having no better neighbor. Unfortunately, this introduces dependencies which are impossible to handle directly.

The main technical contribution of our proofs is in overcoming these dependencies. The basic idea is to uncover the vertices by increasing objective. However, note that the greedy path $P_{\phi}$ in the graph $G_{\leq \phi}$ induced by vertices of objective at most $\phi$ does not need to coincide with the greedy path $P$ in $G$, since the vertices in $P_{\phi}$ may have a neighbor of objective larger than $\phi$, thus shortcutting the rest of $P_{\phi}$. Unfortunately, the number of vertices which lie on a greedy path in some $G_{\leq \phi}$ is so large that many of them have undesired properties. Thus we can not prove iteratively that all nodes on the temporary greedy paths have good properties, because they do not. On the other hand, uncovering the next node on $P$ is impossible without introducing the aforementioned dependencies.

Layer technique: Instead of looking at individual vertices, we define carefully crafted layers $A_i$ for the first phase (defined via weights) and the second phase (defined via objectives) such that the typical trajectory contains at most one vertex per layer. Then we consider the greedy path $P_i$ induced by the first $i$ layers and prove that with sufficiently high probability either $P_i$ has no vertex in $A_i$ or the first vertex $v \in A_i \cap P_i$ has a neighbor $u$ outside the first $i$ layers with better objective than all neighbors inside the first $i$ layers (which we already have uncovered). Note that in this way we can mostly avoid dependencies: the node $v$ is defined only in terms of the first $i$ layers, so we can determine $v$ without uncovering any parts of the remaining graph. At the same time, this approach deals with the problem of shortcuts because we only need to consider one vertex in each layer, rather than all vertices on the temporary greedy paths. Moreover,
First phase

Second phase

**Figure 5.1.** A typical greedy path. With the main Lemma 5.16 we prove that if for some $\phi_0 = o(1)$ we uncover the graph $G[V \leq \phi_0]$ then the trajectory looks as depicted. Moreover, in the last layer we find a vertex which has in expectation (over the randomness in $V > \phi_0$) still many neighbors with objective at least $\phi_0$. Note that the greedy path in $G$ does not need to coincide with the greedy path on $G[V \leq \phi_0]$, but it can only be different when jumping to a vertex of objective larger than $\phi_0$. The special layer $A_{1,\infty}$ is tiny and will be needed only for technical reasons. The two phases are given by $V_1$ and $V_2$ and their layers $A_{1,i}$ resp. $A_{2,i}$.

this vertex $\nu$ is allowed to have shortcuts to higher layers, which helps us again to avoid dependencies.

Due to independence, the number of neighbors of $\nu$ in larger layers is Poisson distributed, and it suffices to show that its expectation is large and that indeed we find the best neighbor of $\nu$ in these larger layers. It is not hard to see that the above events together imply that greedy routing succeeds throughout all layers, and that it visits no layer twice. So we may apply a union bound over
all layers, and obtain that the routing algorithms succeeds with probability $\Omega(1)$ and visits at most one node per layer. Note that we do not necessarily visit every layer, but since a.a.s. the stretch is $1 + o(1)$, we visit a.a.s. a $(1 - o(1))$-fraction of all layers. This proves Theorem 5.1. A more careful analysis along the same lines gives the relaxation result, Theorem 5.5. All other theorems also rely heavily on the layer technique, but require more tricks to handle the start and the end phase.

**Start and end phases:** It turns out that the failure probability of greedy routing is dominated by the first few steps (when the weight is still constant) and the last few steps (when the objective is constant). In each of these steps the algorithm has probability $\Omega(1)$ to fail. For example, in the typical case that $s$ has a constant weight, there is a constant probability that $s$ has no neighbors at all. However, the number of neighbors of $s$ is Poisson distributed with mean $\Theta(w_s)$. In particular, the probability to have no neighbors decays exponentially with $w_s$. Similar considerations apply for the number of neighbors with better objective. This is the reason why the failure probability decays exponentially with $w_{\min}$ (Theorem 5.2). The actual proof is much more tricky and is carried out in Section 5.8.3.

**Patching:** Finally, to prove the patching result (Theorem 5.4) we show three intermediate results.

(i) If we have explored $k$ vertices, starting from $s$, then among the explored vertices, with probability $1 - \exp(-k^{\Omega(1)})$ at least one is adjacent to a vertex of weight at least $k^{\Omega(1)}$. By condition (P2) it takes only $k^{O(1)}$ steps to explore $k$ vertices, so after a short exploration phase of $o(\log \log n)$ steps, a.a.s. we find a vertex of weight $\omega(1)$.

(ii) As in the purely greedy case, starting from a vertex of weight $\omega(1)$ a.a.s. we follow a typical trajectory of the greedy algorithm as described above, until we find a vertex $v$ with almost constant objective (say, with objective $\phi(v) \approx (\log \log \log n)^{-1}$). This middle phase is purely greedy and does not
require patching. In particular, \( v \) has better objective than any previously visited vertices.

(iii) We study the graph \( G_{\geq \phi(v)} \) induced by vertices of objective larger than \( v \) and find that a.a.s. \( G_{\geq \phi(v)} \) contains at most \( O(\phi(v)^{-1}) \) vertices and a giant component, that contains both the vertices \( t \) and \( v \). Thus by condition (P3) the algorithm explores the giant component of \( G_{\geq \phi(v)} \) and finds \( t \) in additional \( o(\log \log n) \) steps.

\section{Preparations for the Proofs}

Before proving our main results, we make in this section several preparations. In Section 5.7.1 we collect several auxiliary properties of the GIRG model which hold in general and have no direct connection to greedy routing. Then Section 5.7.2 contains lemmas which are more directly related to greedy routing. In particular, there we calculate for a given vertex \( v \in V \) where we expect to find its neighbors with highest objective. As the proofs of these lemmas are rather technical, for the sake of readability we defer them to Section 5.12.

But first, let us introduce some important notations. For \( \phi \in \mathbb{R} \) we denote by \( V_{\leq \phi} \) and \( V_{> \phi} \) the set of vertices with objective at most \( \phi \) and larger than \( \phi \), respectively. We remind the reader that the Landau notation \( O, \Omega, \) etc. may hide the power-law exponent \( \beta \), the decay parameter \( \alpha \), the dimension \( d \), and the constants in (5.1) and (5.2), but not the minimal weight \( w_{\text{min}} \). Since the formal analysis of the routing process is quite technical, we collect the most important notations used throughout the proofs in Table 5.1 on page 135.

\subsection{Auxiliary Properties of GIRGs}

In this section we collect some auxiliary properties of GIRGs. Recall that compared to the original definition of GIRGs, the vertices are given by a Poisson point process instead of having \( n \) vertices that are placed uniformly at random in \( \mathbb{T}^d \). However, for any two given vertices \( u \) and \( v \), the distribution of weights and positions and the connection probabilities are the same. Moreover, the
### Notation and Definition Table

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>expected number of vertices, given by the intensity of the Poisson point process</td>
</tr>
<tr>
<td>$\alpha, \beta, d$</td>
<td>the standard parameters of the GIRG model</td>
</tr>
<tr>
<td>$x_v, w_v$</td>
<td>position and weight of vertex $v$</td>
</tr>
<tr>
<td>$w_{\min}$</td>
<td>the minimal weight of the vertices</td>
</tr>
<tr>
<td>$f(\cdot)$</td>
<td>the probability density function of the weights</td>
</tr>
<tr>
<td>$p_{uv}$</td>
<td>connection probability between two fixed vertices $u$ and $v$</td>
</tr>
<tr>
<td>$c_1, c_2$</td>
<td>constants in the connection probability $p_{uv}$</td>
</tr>
<tr>
<td>$s, t$</td>
<td>source and target nodes of the routing process</td>
</tr>
<tr>
<td>$\phi$</td>
<td>the objective function</td>
</tr>
<tr>
<td>$V_{\leq \phi}, V_{&gt; \phi}$</td>
<td>the set of vertices with objective $\leq \phi$, resp. $&gt; \phi$</td>
</tr>
<tr>
<td>$\tilde{\phi}$</td>
<td>an approximate objective function</td>
</tr>
<tr>
<td>$\gamma(\varepsilon)$</td>
<td>$\frac{1-\varepsilon}{\beta-2}$, a constant $&gt; 1$</td>
</tr>
<tr>
<td>$\varepsilon_1$</td>
<td>$\varepsilon_1(\alpha, \beta) &gt; 0$, a fixed global constant</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>$\max{\frac{3}{2}, \frac{2\alpha-1}{2\alpha+4-2\beta}}$ for $\alpha &lt; \infty$, $\frac{3}{2}$ for $\alpha = \infty$</td>
</tr>
<tr>
<td>$V_1$</td>
<td>vertex set corresponding to the first phase of routing: ${v \in V : \phi(v) \leq w_v^{-\gamma(\varepsilon)}}$</td>
</tr>
<tr>
<td>$V_2$</td>
<td>vertex set corresponding to the first phase of routing: ${v \in V : \phi(v) \geq w_v^{-\gamma(\varepsilon)}}$</td>
</tr>
<tr>
<td>$V^+(v, \varepsilon)$</td>
<td>the set of good vertices: large weight and larger objective</td>
</tr>
<tr>
<td>$V^-(v, \varepsilon)$</td>
<td>the set of bad vertices: small weight and large objective</td>
</tr>
<tr>
<td>$w_1(\varepsilon)$</td>
<td>$O(\varepsilon^{d/\varepsilon})$, a specific weight, chosen in terms of $\varepsilon$</td>
</tr>
<tr>
<td>$\phi_1(\varepsilon)$</td>
<td>$\Omega(\varepsilon^{-d/\varepsilon})$, a specific objective, chosen in terms of $\varepsilon$</td>
</tr>
<tr>
<td>$\overline{V}(w, \phi)$</td>
<td>${v \in V_1 \mid w_v \geq w \land \phi(v) \leq \phi} \cup {v \in V_2 \mid \phi(v) \leq \phi}$</td>
</tr>
</tbody>
</table>

**Table 5.1.** A general notation table for the formal analysis of greedy routing in geometric inhomogeneous random graphs.

models coincide if we condition our model on the number of vertices. In particular, if a statement holds a.a.s. in the usual GIRG then it also holds a.a.s. in the Poissonized variant.

First, from Lemma 3.2 we know that the marginal probability for an edge
between two vertices $u$ and $v$ with given weights is

$$\Pr_{x_u,x_v} \left[ \{u, v\} \in E \mid w_u, w_v \right] = \Theta \left( \min \left\{ \frac{w_u w_v}{w_{\min} n}, 1 \right\} \right).$$

Now we can directly deduce that for a fixed vertex $v \in V$ with weight $w_v$, the random variable $\deg(v)$ has Poisson distribution with parameter $\Theta(w_v)$, and in particular $\mathbb{E}[\deg(v)] = \Theta(w_v)$. Indeed, the fact that $\deg(v)$ is a Poisson distributed random variable follows immediately because the vertices are given by a Poisson point process. The formula $\mathbb{E}[\deg(v)] = \Theta(w_v)$ was shown in Lemma 3.3, and the expectation of this model agrees with the expectation in our model. In particular, the $w_{\min}$-terms in the expressions for the edge probabilities and the weight distribution are set up such that the constant hidden in the $\Theta$ does not depend on $w_{\min}$.

For a fixed vertex $v$, the following lemma shows that it is unlikely for $v$ to have a neighbor of much larger weight. We will need this auxiliary result only for the proof of Theorem 5.5, i.e., for studying the relaxations.

**Lemma 5.7.** Let $\epsilon > 0$ and let $v \in V$ be a fixed vertex of weight $w_v$. Then the expected number of neighbors of $v$ with weight at least $w^+ := w_v^{(1+\epsilon)/(\beta-2)}$ is $O(w_{\min}^{\beta-2} w_v^{1-\epsilon})$.

**Proof.** We calculate the expected value with a straightforward integral over all vertex weights larger than $w^+$. The probability measure is given by the density $f(w)$, and for a vertex $u$ of fixed weight $w_u$ and random coordinate $x_u$, the probability to have an edge with $v$ is at most $O\left(\frac{w_u w_v}{w_{\min} n}\right)$ by (5.5). Then, using the density function $f(w)$ for the weights, the expected number of such neighbors of $v$ is at most

$$n \cdot \int_{w^+}^{\infty} f(w) \cdot O\left(\frac{w_v w}{w_{\min} n}\right) \, dw = O\left(\frac{w_v^{\beta-2} w_v}{w_{\min} \int_{w^+}^{\infty} w^{1-\beta} \, dw}\right) = O\left(\frac{w_v^{\beta-2} w_v}{w_{\min}^{\beta-2} w_v^{1-\epsilon}}\right).$$

Next, we calculate the total number of vertices whose objective is larger than a given $\phi_0$. 

Lemma 5.8. Let \( \phi_0 = O(1) \) be a given objective, and denote by \( V_{\geq \phi_0} \) the set of vertices which have objective at least \( \phi_0 \). Then with probability \( 1 - e^{-\Omega(\phi_0^{-1})} \), it holds \( |V_{\geq \phi_0}| = \Theta(\phi_0^{-1}) \).

Proof. By definition of the objective function \( \phi \), a vertex \( u \) satisfies \( \phi(u) \geq \phi_0 \) if and only if \( \|x_u - x_t\|^d \leq \frac{w_u}{\phi_0 w_{\min} n} \). We integrate over all weights \( w \) and see that the expected number of vertices with objective at least \( \phi_0 \) is

\[
nw_{\min}^{\beta-1} \int_{w_{\min}}^{\infty} w^{-\beta} \frac{w}{\phi_0 w_{\min} n} dw = \Theta(\phi_0^{-1}).
\]

Then by a Chernoff bound, with probability \( 1 - e^{-\Omega(\phi_0^{-1})} \), this Poisson random variable is concentrated around its expectation and thus \( \Theta(\phi_0^{-1}) \). 

The following lemma is a consequence of the Poisson point process. It will help us below to deal with events which are not independent as its statement shows how certain random variables are correlated.

Lemma 5.9. Let \( A_1, A_2 \subset \mathbb{T}^d \) and let \( u_1, u_2 \notin A_1 \cup A_2 \) be two vertices. Then

\[
\Pr[\Gamma(u_2) \cap A_2 = \emptyset | \Gamma(u_1) \cap A_1 = \emptyset] \geq \Pr[\Gamma(u_2) \cap A_2 = \emptyset].
\]

Proof. Let \( A_1, A_2 \subset \mathbb{T}^d \) be two subsets and let \( u_1, u_2 \) be two vertices not contained in the sets \( A_1, A_2 \). Suppose we know that \( u_1 \) has no neighbors in \( A_1 \). Clearly, conditioning on this event decreases the expected number of vertices in \( A_1 \) and in every subset of \( A_1 \), hence

\[
\Pr[\Gamma(u_2) \cap (A_1 \cap A_2) = \emptyset | \Gamma(u_1) \cap A_1 = \emptyset] \geq \Pr[\Gamma(u_2) \cap (A_1 \cap A_2) = \emptyset].
\]

Furthermore, by the Poisson point process, for every set \( A_3 \) which is disjoint from \( A_1 \), the number of vertices in \( A_1 \) and \( A_3 \) is independent. We put \( A_3 := A_2 \setminus A_1 \) and deduce

\[
\Pr[\Gamma(u_2) \cap (A_2 \setminus A_1) = \emptyset | \Gamma(u_1) \cap A_1 = \emptyset] = \Pr[\Gamma(u_2) \cap (A_2 \setminus A_1) = \emptyset].
\]

Combining the two inequalities gives the desired property. 

■
Some of the proofs require that we split the geometric space into small subspaces. We do so by introducing grids, dividing the torus into smaller $d$-dimensional cells, similarly as in Section 4.5.1.

**Definition 5.10.** For a given weight $w = w(n)$, a $w$-grid is a grid which splits the geometric ground space $\mathbb{T}^d$ into $\frac{n}{w}$ equal and disjoint cubes of side lengths $(\frac{w}{n})^d$.

Here, for ease of notation, we assume that $w$ divides $n$. One application of $w$-grids is the following lemma. For a given $w$-grid and a given vertex set $S$, it states that if $S$ contains vertices of sufficiently many different cells of the grid, then at least one vertex $v \in S$ will have a neighbor of weight at least $w$.

**Lemma 5.11 (Bulk lemma).** Let $w_0 = w_0(n) = \omega(1)$ be a function growing in $n$ and let $S$ be a set of vertices of weight at most $w_0$ such that $S$ contains vertices in at least $w_0$ different cells of a fixed $w_0$-grid. Then a.a.s., there exists a vertex $v \in S$ and a vertex $u$ of weight at least $w_0$ such that $u$ and $v$ are neighbors and $u$ and $v$ are contained in the same cell of the $w_0$-grid.

**Proof.** We uncover the graph in two steps: First we only consider the vertices of weight less than $w_0$ and afterwards we uncover the remaining vertices. Let $S$ be a fixed set of vertices with weight less than $w_0$, and let $C_1, \ldots, C_{w_0}$ be cells of a $w_0$-grid which contain at least one vertex of $v_i \in S \cap C_i$.

Now, we insert the remaining vertices of high weight at random, according to the definition of the model. Together, the $w_0$ cells have volume $w_0 \cdot \frac{w_0}{n}$. Therefore, a single vertex $u$ of weight $w_u \geq w_0$ falls with probability $\frac{w_0^2}{n}$ into one of the $w_0$ cells. Furthermore, if $u$ falls into such a cell $C_i$, then $\|x_u - x_{v_i}\|^d = O(\frac{w_0}{n})$. By definition, $v_i$ has weight at least $w_{\text{min}}$, and (5.1) or (5.2) imply

$$\Pr[\{u, v_i\} \in E \mid u, v_i \in C_i] = \Omega(1).$$

Thus every vertex of weight at least $w_0$ is connected to the set $\{v_1, \ldots, v_{w_0}\}$ with probability at least $\Omega(\frac{w_0^2}{n})$. Note that the expected number of vertices of weight at least $w_0$ is $\Theta(nw_0^{\beta-1}w_0^{1-\beta})$, hence the expected number of vertices of weight at least $w_0$ which are connected to a vertex of $\{v_1, \ldots, v_{w_0}\}$ is at least $\Omega(w_0^{\beta-1}w_0^{3-\beta}) = \omega(1)$. By a Chernoff bound, it follows that a.a.s., there exists a vertex $v_i$ in a cell $C_i$ such that $v_i$ has at least one neighbor $u$ with $w_u \geq w_0$. ■
5.7.2 Where to expect Neighbors & Unlikely Jumps

We start by introducing technical notation. For all $\varepsilon > 0$, we put $\gamma(\varepsilon) := \frac{1-\varepsilon}{\beta-2}$. Next, let $\varepsilon_1 = \varepsilon_1(\alpha, \beta) > 0$ be a fixed constant which we choose sufficiently small during proofs, e.g., such that $\gamma(\varepsilon_1) > 1$. We define the following two classes of vertices:

\[ V_1 := \{ v \in V \mid \phi(v) \leq w_v^{-\gamma(\varepsilon_1)} \} \quad \text{and} \quad V_2 := \{ v \in V \mid \phi(v) \geq w_v^{-\gamma(\varepsilon_1)} \}. \]

In every hop, the process chooses a neighbor $u$ of the current vertex $v$ which maximizes $\phi(u)$. As discussed above in Section 5.6, it turns out that if $v \in V_1$, then typically $w_u$ is by an exponent $\gamma(\varepsilon)$ larger than $w_v$. On the other hand, if $v \in V_2$, then we expect that $\phi(u) \sim \phi(v)^{1/\gamma(\varepsilon)}$. Therefore, vertices in $V_1$ will correspond to the first phase of the routing, where the current weight increases by an exponent with every hop. The vertices in $V_2$ correspond to the second phase where the current objective is increased by an exponent with every hop (due to $\phi(v) < 1$, indeed the objective increases).

In order to find such a trajectory, we want to classify the vertices in $V$ which accelerate the routing as desired. Let $\zeta := \max\{\frac{3}{2}, \frac{2\alpha-1}{2\alpha+4-2\beta}\}$ for $\alpha < \infty$ and $\zeta := \frac{3}{2}$ in the threshold case $\alpha = \infty$. Then for $v \in V_1$ and $\varepsilon = \varepsilon(n) > 0$ we define

\[ V^+(v, \varepsilon) := \{ u \in V \mid w_u \geq w_v^{\gamma(\varepsilon)} \wedge \phi(u) \geq \phi(v)w_v^{\gamma(\varepsilon)-1} \}, \quad \text{and} \quad V^-(v, \varepsilon) := \{ u \in V \mid w_u \leq w_v^{\gamma(\varepsilon\cdot\varepsilon)} \wedge \phi(u) \geq \phi(v)w_v^{\gamma(\varepsilon)-1} \}. \]

In the first phase, the goal will be to increase the weight within every hop of the routing. Therefore, $\Gamma(v) \cap V^+(v, \varepsilon)$ is the set of good neighbors as every $u \in \Gamma(v) \cap V^+(v, \varepsilon)$ has significantly larger weight than $v$. On the other side, the set $\Gamma(v) \cap V^-(v, \varepsilon)$ contains the bad neighbors as its vertices have small weight but nevertheless large objective and therefore could force the routing to proceed with a low-weight-vertex.

In the second phase, the goal is to increase the objective by an exponent in every step of the routing. Furthermore, once we reach vertices of $V_2$, we want to stay in $V_2$. Let $v \in V_2$. Similarly as above, we introduce a set $V^+(v, \varepsilon)$ of good
vertices and a set $V^{-}(v, \varepsilon)$ of bad vertices. More precisely, for $\varepsilon = \varepsilon(n)$ we define

$$V^{+}(v, \varepsilon) := \{ u \in V_{2} \mid \phi(u) \geq \phi(v)^{1/\gamma(\varepsilon)} \}, \quad \text{and}$$

$$V^{-}(v, \varepsilon) := \{ u \in V_{1} \mid \phi(u) \geq \phi(v)^{1/\gamma(\varepsilon)} \}. \tag{5.7}$$

In the following we collect several technical statements about the expected number of good and bad neighbors. All proofs can be found in Section 5.12. For $0 < \varepsilon \leq \varepsilon_{1}$, let $w_{1}(\varepsilon) := O(e^{d/\varepsilon})$ and $\phi_{1}(\varepsilon) := \Omega(e^{-d/\varepsilon})$ with sufficiently large respectively small hidden constants. We start with the first lemma which considers vertices in $V_{1}$ and shows that the expected number of good neighbors is large whereas the expected number of bad neighbors is small. Notice that for $\varepsilon = \varepsilon_{1}$, every vertex $v \in V_{1}$ satisfies $\phi(v)w_{v}^{\gamma(\varepsilon_{1})} \leq 1$ by the definition of $V_{1}$, and in this case, the statement can be applied for all vertices in $V_{1}$.

**Lemma 5.12 (Neighborhoods in first phase).** Let $0 < \varepsilon \leq \varepsilon_{1}$ and let $v \in V_{1}$ be a vertex such that $\phi(v)w_{v}^{\gamma(\varepsilon_{1})} \leq 1$. Then

(i) $E[|\Gamma(v) \cap V^{+}(v, \varepsilon)|] = \Omega(w_{\min}^{\beta-2}w_{v}^{\varepsilon})$, and

(ii) if in addition $w_{v} \geq w_{1}(\varepsilon)$, then $E[|\Gamma(v) \cap V^{-}(v, \varepsilon)|] = O(w_{\min}^{\beta-2}w_{v}^{-\Omega(\varepsilon)})$.

Next, similarly to Lemma 5.12 for the first phase we give a lemma for the second phase in the vertex set $V_{2}$. We calculate for a vertex $v \in V_{2}$ the expected number of good and bad neighbors.

**Lemma 5.13 (Neighborhoods in second phase).** Let $0 < \varepsilon \leq \varepsilon_{1}$ and let $v \in V_{2}$ be a vertex such that $\phi(v) \leq 1$. Then

(i) $E[|\Gamma(v) \cap V^{+}(v, \varepsilon)|] = \Omega(w_{\min}^{\beta-2}\phi(v)^{-\Omega(\varepsilon)})$, and

(ii) if in addition $\phi(v) \leq \phi_{1}(\varepsilon)$, then $E[|\Gamma(v) \cap V^{-}(v, \varepsilon)|] = O(w_{\min}^{\beta-2}\phi(v)^{\Omega(\varepsilon)})$.

In order to prove our main results, we want to show that with sufficiently high probability, greedy routing follows our anticipated trajectory. If $v \in V_{1}$ is a vertex on the greedy path, we want that the best neighbor of $v$ is in $V^{+}(v, \varepsilon)$, i.e., that its weight is significantly larger than $w_{v}$. By Lemma 5.12 (ii) the expected number of vertices in $\Gamma(v)$ with small weight and relatively large objective is polynomially
small in \( w_v \). By Markov's inequality (Theorem 2.1), this will give us an error probability which is polynomially small in \( w_v \). If \( v \) has constant weight, this is not precise enough as for the proof of Theorem 5.2, we need an error probability which is \textit{exponentially} small in \( w_{\min} \). With our next lemma, we improve this as follows: We consider vertices \( v \in V_1 \) whose objective is relatively small compared to the weight \( w_v \), and calculate the expected size of \( V^{-}(v, \varepsilon) \) depending on \( \phi(v) \). Later on, this statement can be applied in particular to vertices of constant weight with large geometric distance to the target.

**Lemma 5.14.** There exists a constant \( \tilde{c} > 0 \) such that for all \( 0 < \varepsilon \leq \varepsilon_1 \) and all vertices \( v \in V_1 \) with \( w_v \geq w_1(\varepsilon) \) and \( \phi(v)w_v^{\gamma(\varepsilon)} \leq \tilde{c} \), it holds

\[
\mathbb{E}[|\Gamma(v) \cap V^{-}(v, \varepsilon)|] = O(\phi(v)^{\Omega(1)}w_v^{O(1)}).
\]

We turn to the last lemma of this section. In order to prove Theorem 5.3 and Theorem 5.4, we need to control the trajectory in particular at the start and at the end of the routing. For technical reasons, we require that the routing doesn't traverse edges on which the objective increases too fast. Statement (i) of this lemma states that a.a.s. there are no vertices of both high weight and high objective at all. Statements (ii) and (iii) upper-bound the probability that a vertex of small weight has an incident edge of unexpected length, i.e., a neighbor which is located geometrically far away. Finally (iv) upper-bounds this probability for a vertex of high weight. In the proofs, we will apply (ii) with \( K \leq 7 \) and (iii) with \( K = 3 \).

**Lemma 5.15** (Unlikely jumps). Let \( w_0 = w_0(n) = \omega(1) \) be a growing function in \( n \) and let \( v \in V \). Then the following statements hold.

(i) With prob. \( 1 - w_0^{-\Omega(1)} \), there exists no vertex \( u \) (except potentially \( t \)) s.t. \( w_u \geq w_0 \) and \( \phi(u) \geq w_0^{-\Omega(1)} \).

Let \( w_v \leq w_0 \).

(ii) For any fixed \( K > 0 \), with probability \( 1 - O(w_0^{-K}) \) there exists no neighbor \( u \) of \( v \) with \( w_u \leq w_0 \) and \( \|x_u - x_v\|^d \geq \frac{w_0^{O(1)}}{n} \).
(iii) Let \( \phi(v) \leq w_0^{-\Omega(1)} \). Then for any fixed \( K > 0 \), with probability \( 1 - O(w_0^{-K}) \) there exists no neighbor \( u \) of \( v \) with \( w_u \geq w_0 \) and \( \phi(u) \leq \phi_0^{O(1)} \).

Let \( w_v \geq w_0 \).

(iv) Let \( \phi(v) \leq w_0^{-1-\Omega(1)} \), and let \( M := \min\{w_v, \phi(v)^{-1}\} \). Then with probability \( 1 - O(M^{-\Omega(1)}) \), \( v \) has no neighbor \( u \) with \( w_u \leq w_0 \) and \( \phi(u) \geq w_0^{-\Omega(1)} \).

### 5.8 Success Probability and Length of Greedy Routing

In this section we analyze the basic greedy process and prove Theorem 5.1 and Theorem 5.2 about the success probability and Theorem 5.3 about the running time. We advise the reader to first read Section 5.6 where we describe the high-level ideas that we use to prove our main Lemma 5.16. In a nutshell, we distinguish two phases of the process (governed by weights and by objectives, respectively), and for each of the phases we divide the space \( V \) into layers. Then we prove that with sufficiently large probability a greedy path never visits more than one vertex per layer, and a.a.s. it does not fail in any layer except for the very first or the very last ones. This observation is formalized in Lemma 5.16 and lies at the heart of all our main results, including the results for patching methods and for the relaxations. All algorithms follow very similar trajectories, only the starting phase and the end phase differ for the various algorithms. The layer structure is depicted in Figure 5.1 on page 132.

#### 5.8.1 Main Lemma

Let us start by repeating the notations which we introduced in Section 5.7.2. For all \( \varepsilon > 0 \), we denote \( \gamma(\varepsilon) := \frac{1-\varepsilon}{p-2} \). Furthermore \( \varepsilon_1 = \varepsilon_1(\alpha, \beta) > 0 \) is a fixed constant, chosen sufficiently small, \( \zeta = \zeta(\alpha, \beta) \) is also a fixed constant, \( w_1(\varepsilon) = O(e^{d/\varepsilon}) \), and \( \phi_1(\varepsilon) = \Omega(e^{-d/\varepsilon}) \). Then we have the following two classes of vertices:

\[
V_1 := \{ v \in V \mid \phi(v) \leq w_v^{-\gamma(\varepsilon_1)} \} \quad \text{and} \quad V_2 := \{ v \in V \mid \phi(v) \geq w_v^{-\gamma(\varepsilon_1)} \}.
\]

Finally, for a given vertex \( v \in V \) and a given \( \varepsilon > 0 \) we defined good sets \( V^+(v, \varepsilon) \) and bad sets \( V^-(v, \varepsilon) \) above in (5.6) and (5.7) on 139-140.
In Section 5.6 we claimed that the expected trajectory of the routing process is such that soon after the start the current weight is $\omega(1)$. We also expect that when reaching rather large objectives at the end of the routing, very few additional hops are needed to reach $t$. Therefore the main part of the process plays in-between. In this section we state and prove a lemma which describes the routing process in this main part. For any weight-objective-pair $(w, \phi)$ we define

$$V(w, \phi) := \{v \in V_1 | (1): w_v \geq w; (2): \phi(v) \leq \phi\} \cup \{v \in V_2 | \phi(v) \leq \phi\}.$$ 

Suppose that both $w$ and $\phi$ are constants. Then the main part of the routing process will happen in this set $V(w, \phi)$. With the following lemma, we study the routing process inside such a set and show that with sufficiently high probability, the routing is ultra-fast in traversing this set and never gets stuck. Thereby, the probabilities for the failure events depend on $w$ and $\phi$, which can be both growing functions as well. Later, we will apply the lemma for both the basic greedy algorithm and the patching algorithms.

**Lemma 5.16 (Main lemma).** Let $w_0 \geq w_1(\epsilon_1)$ and $\phi_0 \leq \phi_1(\epsilon_1)$. Furthermore let $f_0(n) = \omega(1)$ be any growing function such that $f_0(n) = o(\log \log n)$, and let $A$ be a routing protocol which satisfies (P1), i.e., it makes greedy choices. Suppose $\phi(s) \leq \phi_0$, denote by $P$ the greedy path induced by all vertices of objective at most $\phi_0$ starting from $s$, and suppose that there exists a first vertex $u_1 \in P \cap V(w_0, \phi_0)$. Then, for $M := \min\{w_0, \phi_0^{-1}\}$, with probability

$$1 - O\left(\frac{1}{w_{\min}^{1/2} M^{\Omega(1)}}\right),$$

there exists a subpath $P' = (u_1, \ldots, u_{\ell})$ of $P$ starting in $u_1$ with the following properties:

(i) $P'$ is contained in $V(w_0, \phi_0)$.

(ii) Either $P' \subset V_1$, or $P' \subset V_2$, or there exists a vertex $u_{\ell'}$ such that

$$\{u_1, \ldots, u_{\ell'}\} \in V_1 \quad \text{and} \quad \{u_{\ell'+1}, \ldots, u_{\ell}\} \in V_2.$$
(iii) Let \( \{u_i, u_{i+1}, u_{i+2}\} \) be three subsequent vertices on \( P' \cap V_1 \). Then
\[
\omega_{u_{i+2}} \geq \omega_{u_i}^{\gamma(\varepsilon_1)}.
\]

(iv) Let \( \{u_i, u_{i+1}, u_{i+2}\} \) be three subsequent vertices on \( P' \cap V_2 \). Then
\[
\phi(u_{i+2}) \geq \phi(u_i)^{1/\gamma(\varepsilon_1)}.
\]

(v) The length \( \ell \) of the path \( P' \) is at most \( 2 + o(1) \log \log n \). Moreover and more precisely, the length of \( P' \) is upper-bounded by
\[
\frac{1 + o(1)}{\log(\beta - 2)} \left( \log \log_{\omega_0} (\phi(u_1)^{-1}) + \log \log_{\phi_0^{-1}} (\phi(u_1)^{-1}) \right) + O(f_0(n)).
\]

(vi) It holds
\[
E_{>\phi_0} \left[ |\Gamma(u_\ell) \cap V^+(u_\ell, \varepsilon_1) \cap V_{>\phi_0}| \right] = \Omega(\omega_0^{\beta-2} M^{\Omega(1)}).
\]

Here \( E_{>\phi_0} \) means expectation w.r.t. uncovering the vertices of objectives larger than \( \phi_0 \), conditioned on position and weight of \( u_\ell \).

Note that Lemma 5.16 naturally applies to the situation that we have uncovered all vertices of objective at most \( \phi_0 \), but not the rest of the graph. More precisely, conditions (i)-(v) are independent of \( V_{>\phi_0} \), and (vi) makes a statement about the marginal expectation after uncovering \( V_{\leq \phi_0} \) and before uncovering \( V_{>\phi_0} \).

**Proof.** Let \( P \) be the greedy path induced by all vertices of objective at most \( \phi_0 \). We analyze the structure of \( P \) by partitioning the set \( \overline{V}(\omega_0, \phi_0) \) into several small layers. Then the idea will be to show that with sufficiently high probability, there exists a vertex \( u_\ell \) on the path \( P \) such that until reaching \( u_\ell \), \( P \) visits every layer at most once, and such that we can expect that \( u_\ell \) has neighbors of objective at least \( \phi_0 \). There will be two classes of layers: The layers \( A_{1,j} \) divide the area \( V_1 \cap \overline{V}(\omega_0, \phi_0) \) and are defined via weights, and the layers \( A_{2,j} \), defined via objectives, divide the area \( V_2 \cap \overline{V}(\omega_0, \phi_0) \).

Let \( f_0(n) = o(\log \log n) \). Recall that we already put \( \varepsilon_1 = \Theta(1) \). In addition, we put \( \varepsilon_2 := (\log \log f_0(n))^{-1} = o(1) \), then in particular \( \varepsilon_2 < \varepsilon_1 \). Furthermore, we
define the landmarks

\[ w'_0 := w_0^{(\gamma(\epsilon_1)f_0(n))} \quad \text{and} \quad \phi'_0 := \phi_0^{(\gamma(\epsilon_1)f_0(n))}. \]

These landmarks will be thresholds for slightly different definitions of the layers. For later reference we note that, since \( \gamma(\epsilon_1) > 1 \),

\[ w_0^{\Omega(\epsilon_2)} = w_0^{\Omega(\gamma(\epsilon_1)f_0(n)/\log \log f_0(n))} = \omega(w_0^{\Omega(1)}) \quad \text{and} \quad \phi_0^{\Omega(\epsilon_2)} = o(\phi_0^{\Omega(1)}), \]

where the second equation follows analogously to the first.

**Layers by weight:** We define a sequence \( y_0 := w_0 < y_1 < \ldots < y_j < \ldots \) for the weights which is growing doubly exponentially. More precisely, we require

\[
y_{j+1} = \begin{cases} 
y_j^{\gamma(\epsilon_1)}, & \text{if } y_j < w'_0, \\
y_j^{\gamma(\epsilon_2)}, & \text{if } y_j \geq w'_0.
\end{cases}
\]

Then we partition the set \( V'_1 := \{ v \in \overline{V}(w_0, \phi_0) \mid \phi(v)w_v^{(\epsilon_2)} \leq 1 \} \) into layers defined via weight,

\[ A_{1,j} := \{ v \in V'_1 \mid y_{j-1} \leq w_v < y_j \} \quad \forall j \geq 1. \]

Note that by definition, \( V'_1 \subset V_1 \) and no vertex of \( V_1 \) has larger weight than \((w_{\min}n)^{(1+\gamma(\epsilon_1))}^{-1} \). Thus only the sets \( A_{1,j} \) need to be considered for which it holds \( y_{j-1} \leq (w_{\min}n)^{(1+\gamma(\epsilon_1))}^{-1} \). Furthermore, and only for technical reasons we introduce the additional layer

\[ A_{1,\infty} := \{ v \in \overline{V}(w_0, \phi_0) \mid \phi(v)w_v^{(\epsilon_1)} \leq 1 \leq \phi(v)w_v^{(\epsilon_2)} \}, \]

which covers the remaining vertices of \( V_1 \cap \overline{V}(w_0, \phi_0) \).

**Layers by objective:** Similarly, define a sequence \( \psi_0 = \phi_0 > \psi_1 > \ldots > \psi_j > \ldots \) for the objectives such that this sequence falls doubly exponentially, i.e.,

\[
\psi_{j+1} = \begin{cases} 
\psi_j^{(\epsilon_1)}, & \text{if } \psi_j > \phi'_0, \\
\psi_j^{(\epsilon_2)}, & \text{if } \psi_j \leq \phi'_0.
\end{cases}
\]
Next, we define similar layers for the vertices in \( V_2 \cap \overline{V}(w_0, \phi_0) \).

\[
A_{2,j} := \{ v \in V_2 \cap \overline{V}(w_0, \phi_0) \mid \psi_{j-1} \geq \phi(v) > \psi_j \} \quad \forall j \geq 1.
\]

For this second set of layers, note that every vertex in the graph has objective at least \( n^{-1} \) and we do not need to consider layers \( A_{2,j} \) which would contain vertices of even smaller objective.

**Definition of the events \( \mathcal{E} \) and \( \mathcal{E}_{i,j} \):** The layers \( A_{i,j} \) as defined above classify all vertices of objective at most \( \phi_0 \). We want to show that with sufficiently high probability, \( P \) contains at most one vertex per layer \( A_{i,j} \) until reaching a vertex \( u_\ell \) for which the expected number of neighbors in \( V_{>\phi} \) is large. We will do this by considering the layers in the following consecutive order:

\[
A_{1,1} < ... < A_{1,j} < A_{1,j+1} < ... < A_{1,\infty} < ... < A_{2,j} < A_{2,j-1} < ... < A_{2,1}.
\]

Given this ordering of the layers, we denote by \( B_{i,j} \) the union of the layer \( A_{i,j} \), all previous layers, and the set \( V_{\leq \phi_0} \setminus \overline{V}(w_0, \phi_0) \). Finally, by \( P_{i,j} \) we denote the greedy path induced by the set \( B_{i,j} \subset V \).

We now consider, for a given pair \((i,j)\), the first vertex \( v \in P_{i,j} \cap A_{i,j} \), if such a vertex exists. We will first show that with sufficiently high probability the neighbor of \( v \) with highest objective is located outside \( B_{i,j} \), which we consider a “good” event.

More precisely, let \( i \in \{1,2\} \), \( j \geq 1 \), and let \( \epsilon \in \{\epsilon_1, \epsilon_2\} \) be the \( \epsilon \) which was used for the definition of \( A_{i,j} \). Then we denote by \( \mathcal{E}_{i,j} \) the event that either \( P_{i,j} \cap A_{i,j} = \emptyset \), or the first vertex \( v \in P_{i,j} \cap A_{i,j} \)

- satisfies condition (vi), or
- has at least one good neighbor, i.e, a neighbor \( v' \in \overline{V}(w_0, \phi_0) \setminus B_{i,j} \) with \( \phi(v') \geq \phi(v) \) such that \( \phi(v') > \phi(u) \) holds for all \( u \in \Gamma(v) \cap B_{i,j} \).

Finally, we denote by \( \mathcal{E} := \cap_{i,j} \mathcal{E}_{i,j} \) the intersection of all good events.
**Lower bound for Pr[\(\mathcal{E}_{i,j}\):** In this part, our goal is to lower-bound \(\Pr[\mathcal{E}_{i,j}]\) for every pair \((i, j)\) in order to lower-bound \(\Pr[\mathcal{E}]\). Thereby, our construction of layers gives rise to several different subcases, corresponding to different subphases of the routing.

Let \((i, j)\) be a given pair for which we want to bound \(\Pr[\mathcal{E}_{i,j}]\) from below. Clearly we have \(\Pr[\mathcal{E}_{i,j} \mid P_{i,j} \cap A_{i,j} = \emptyset] = 1\), and we only need to consider \(\Pr[\mathcal{E}_{i,j} \mid P_{i,j} \cap A_{i,j} \neq \emptyset]\). Let \(v\) be the first vertex on \(P_{i,j} \cap A_{i,j}\). Again, if \(v\) satisfies condition (vi) then \(\mathcal{E}_{i,j}\) holds and there is nothing to show, so we assume otherwise,

\[
(5.9) \quad \mathbb{E}_{\geq \phi_0} \left[ |\Gamma(v) \cap V^+(v, \epsilon_1) \cap V_{> \phi_0}| \right] < E,
\]

where \(E := \eta w_{\min}^{\beta-2} M^\eta\) where we choose \(\eta\) small enough such that for a given \(c > 0\), we can achieve for all \(v \in A_{i,j}\) that \(cw_{\min}^{\beta-2} w_v^\epsilon > 2E\) and \(cw_{\min}^{\beta-2} \phi(v)^{-c\epsilon} > 2E\). In all cases let \(\epsilon \in \{\epsilon_1, \epsilon_2\}\) be the same \(\epsilon\) as used for the definition of the considered layer.

**Case \(i = 1\) and \(j < \infty\):** We are in the first phase of the routing, and the layer \(A_{1,j}\) is contained in \(V'_1 \subset V_1\). Then by Lemma 5.12 (i) the expected number of neighbors of \(v\) in \(V^+(v, \epsilon)\) is large, i.e., \(\mathbb{E}[|\Gamma(v) \cap V^+(v, \epsilon)|] = \Omega(w_{\min}^{\beta-2} w_v^\epsilon)\). Since \(\epsilon \leq \epsilon_1\) we have \(V^+(v, \epsilon) \subseteq V^+(v, \epsilon_1)\), so in particular we do not expect many of these neighbors in \(V_{> \phi_0}\), i.e.,

\[
(5.10) \quad \mathbb{E}_{\geq \phi_0} \left[ |\Gamma(v) \cap V^+(v, \epsilon) \cap V_{> \phi_0}| \right] \leq \frac{1}{2} \mathbb{E}[|\Gamma(v) \cap V^+(v, \epsilon)|],
\]

where the second inequality comes from our choice of the constant \(\eta\) in \(E\). Thus we expect most of these good neighbors in \(V_{\leq \phi_0}\), i.e.,

\[
\mathbb{E}[|\Gamma(v) \cap V^+(v, \epsilon) \cap V_{\leq \phi_0}|] = \Omega(w_{\min}^{\beta-2} w_v^\epsilon).
\]

By a Chernoff bound (Theorem 2.2 (ii)), using \(w_v \geq y_{j-1}\), with probability at least \(1 - \exp(-\Omega(w_{\min}^{\beta-2} y_{j-1}^\epsilon))\) there exists a neighbor \(v'\) of \(v\) in \(V^+(v, \epsilon) \cap V_{\leq \phi_0}\). By the definition of \(V^+(v, \epsilon)\), such a neighbor fulfills \(v' \notin B_{1,j}\) and \(\phi(v') \geq \phi(v)w_{\min}^{\gamma(\epsilon)-1}w_v^\epsilon\).

Now, we want to show that \(v'\) actually is a good neighbor of \(v\), i.e., every \(u \in \Gamma(v) \cap B_{i,j}\) has objective less than \(\phi(v)w_{\min}^{\gamma(\epsilon)-1}\). We observe that every vertex with objective at least \(\phi(v)w_{\min}^{\gamma(\epsilon)-1}\) and weight at most \(y_j\) is contained in \(V^-(v, \epsilon)\),
thus we want to upper-bound $\mathbb{E}[|\Gamma(v) \cap V^-(v, \varepsilon)|]$. Note that here we need to condition on that no earlier vertex on the path $P_{1,j}$ had a neighbor $u \in V^-(v, \varepsilon)$ (because otherwise, $u$ would be the first vertex in $A_{i,j} \cap P_{i,j}$ and not $v$). By Lemma 5.9, this condition decreases the expected number of vertices in $V^-(v, \varepsilon)$, and therefore also the expected number of neighbors of $v$ in $V^-(v, \varepsilon)$. Hence, in order to give an upper bound on this value, we are allowed to neglect the dependencies. Since $y_{j-1} \geq w_1(\varepsilon) = O(e^{d_j/\varepsilon})$, by Lemma 5.12 (ii) we deduce

$$\mathbb{E}[|\Gamma(v) \cap V^-(v, \varepsilon)|| v \in A_{1,j} \cap P_{1,j}] \leq \mathbb{E}[|\Gamma(v) \cap V^-(v, \varepsilon)|] = O(w_{\min j}^{\beta-2}y_{j-1}^{-\Omega(\varepsilon)}).$$

Then we apply Markov’s inequality (Theorem 2.1) to see that with probability at least $1 - O(w_{\min j}^{\beta-2}y_{j-1}^{-\Omega(\varepsilon)})$, $v$ has no neighbor in $V^-(v, \varepsilon)$ and therefore every $u \in \Gamma(v, \varepsilon) \cap B_{1,j}$ has objective at most $\phi(u) < \phi(v)w_v^{\gamma(\varepsilon)} \leq \phi(v')$ as desired. It follows

$$\Pr[\mathcal{E}_{1,j}] \geq 1 - O(w_{\min j}^{\beta-2}y_{j-1}^{-\Omega(\varepsilon)}).$$

Note that depending on how the weight $y_{i-1}$ compares to $w'_0$, we are either using $\varepsilon_1$ or $\varepsilon_2$, and the last exponent of the above expression is either $-\Omega(1)$ or $-\Omega((\log\log f_0(n))^{-1})$.

Case $i = 1$ and $j = \infty$: Next we give a lower bound for $\Pr[\mathcal{E}_{1,\infty}]$ in exactly the same way. Let $v$ be a vertex in this extra layer $A_{1,\infty}$, then by definition it has the property $w_v^{\gamma(\varepsilon_2)} \leq \phi(v) \leq w_v^{\gamma(\varepsilon_1)}$. Therefore, we can apply Lemma 5.12 (i) only with $\varepsilon_1$ but not with $\varepsilon_2$. We observe that every vertex $v'$ in the considered set $V^+(v, \varepsilon_1)$ satisfies

$$\phi(v') \geq \phi(v)w_v^{\gamma(\varepsilon_1)-1} \geq w_v^{\gamma(\varepsilon_1)-1-\gamma(\varepsilon_2)} \geq w_{v'}^{-(\gamma(\varepsilon_1)-1-\gamma(\varepsilon_2))/\gamma(\varepsilon_1)} \geq w_{v'}^{-\gamma(\varepsilon_1)}$$

for $\varepsilon_1$ small enough. Then indeed $v' \not\in B_{1,\infty}$, and thus $V^+(v, \varepsilon_1) \subset V_2$. This allows us to argue exactly in the same arguments as above in the case $j < \infty$. Then we apply Lemma 5.12 (ii) and deduce that

$$\Pr[\mathcal{E}_{1,\infty}] \geq 1 - O(w_{\min \infty}^{\beta-2}w_0^{-\Omega(\varepsilon)}).$$

Case $i = 2$: We continue with the events $\mathcal{E}_{2,j}$ for the layers $A_{2,j}$. Here, the argument works similar: Again we can assume that the set $A_{2,j} \cap P_{2,j}$ is non-empty,
and that the first vertex \( v \in A_{2,j} \cap P_{2,j} \) satisfies (5.9). By Lemma 5.13 (i) we have \( \mathbb{E}[|\Gamma(v) \cap V^+(v, \varepsilon)|] = \Omega(w_{\min}^{\beta-2} \phi(v)^{-\Omega(\varepsilon)}) \), and as in (5.10) our choice of \( E \) ensures that also \( \mathbb{E}[|\Gamma(v) \cap V^+(v, \varepsilon) \cap V_{\leq \phi_0}|] = \Omega(w_{\min}^{\beta-2} \phi(v)^{-\Omega(\varepsilon)}) \). By the Chernoff bound, and since \( \psi_{j-1} \geq \phi(v) > \psi_j \), with probability at least \( 1 - \exp(-\Omega(w_{\min}^{\beta-2} \psi_{j-1}^{-\Omega(\varepsilon)})) \) the set \( \Gamma(v) \cap V^+(v, \varepsilon) \cap V_{\leq \phi_0} \) is non-empty. In this case, there exists a neighbor \( v' \) in \( \bar{\Gamma}(w_0, \phi_0) \cap B_{2,j} \) with objective larger than \( \psi_{j-1} \). It remains to show that there exists no neighbor in \( v_1 \) with higher objective. By Lemma 5.9 and the same arguments as above we can use Lemma 5.13 (ii) to see that the expected number of neighbors of \( v \) inside \( B_{2,j} \) with objective larger than \( \psi_{j-1} \) is \( O(w_{\min}^{\beta-2} \psi_{j-1}^{-\Omega(\varepsilon)}) \). Then Markov’s inequality shows that with probability \( 1 - O(w_{\min}^{\beta-2} \psi_{j-1}^{-\Omega(\varepsilon)}) \), there exists no such neighbor. Thus

\[
(5.13) \quad \Pr[E_{2,j}] \geq 1 - O(w_{\min}^{\beta-2} \psi_j^{-\Omega(\varepsilon)}).
\]

Here, the last exponent is either \( \Omega(1) \) or \( \Omega((\log \log f_0(n))^{-1}) \), depending on which \( \varepsilon \) we consider.

**Lower bound for \( \Pr[E] \):** We computed a lower bound for \( \Pr[E_{i,j}] \) for every layer \( A_{i,j} \). Combining equations (5.11), (5.12), and (5.13) yields

\[
\Pr[E] \geq 1 - \sum_{j \geq 1} \Pr[\neg E_{1,j}] - \Pr[E_{1,\infty}] - \sum_{j \geq 1} \Pr[\neg E_{1,j}] - \Pr[\neg E_{1,j}]
\]

\[
= 1 - O(w_{\min}^{\beta-2} (w_0^{-\Omega(1)} + w_0'^{-\Omega(2)} + \phi_0^{-\Omega(1)} + \phi_0'^{-\Omega(2)}))
\]

\[(5.8) \quad = 1 - O(w_{\min}^{\beta-2} \min\{w_0, \phi_0^{-1}\}^{-\Omega(1)}).
\]

**\( E \) implies (i)-(iv):** Having proven that \( E \) occurs with sufficiently high probability, it remains to show that this event implies all desired properties (i)-(vi). Hence suppose that \( E \) occurs and that \( P \) visits at least one vertex of \( \bar{\Gamma}(w_0, \phi_0) \). Then the first vertex \( u_1 \in P \cap \bar{\Gamma}(w_0, \phi_0) \) is contained in some layer \( A_{i,j} \). We show that there exists a subpath \( P' = \{u_1, \ldots, u_{\ell}\} \subseteq P \) which visits every layer at most once and then implies (i)-(vi).

**First Phase (i = 1):** Suppose that \( u_1 \in V_1 \). Then there exists a maximal subpath

\[
P_1 = \{u_1, \ldots, u_{\ell'}\} \subseteq P
\]
starting at $u_1$ such that $P_1 \subset V_1$ and no vertex before $u_{\ell'}$ satisfies condition (vi). We start with an analysis of this subpath $P_1$ and show by induction that every vertex of $P_1$ is located in a higher layer than its predecessor. Clearly, the induction hypothesis holds for the base case $\{u_1\}$. Assume that the hypothesis holds for $\{u_1, \ldots, u_i\}$. Then $u_i$ is contained in layer $A_{1,j}$, and by the induction hypothesis, $u_i$ is the first vertex that $P$ visits in this layer. Moreover, by the induction hypothesis it follows that $\{u_1, \ldots, u_i\}$ is also a subpath of the induced greedy path $P_{1,j}$. Then by the event $\mathcal{E}_{1,j}$ we have one of the following two cases:

- $u_i$ satisfies condition (vi): Then set $P' = (u_1, \ldots, u_{\ell'}$ with $u_{\ell'} := u_i$.

- $u_i$ has a good neighbor $v'$: Let $u_{i+1}$ be the best neighbor of $u_i$. Then $u_{i+1} \in V_2$, or $u_{i+1}$ is located in a layer $A_{1,j'}$ where $j' > j$ because $u_{i+1}$ has higher weight than every vertex in $A_{1,j}$. Since the protocol makes greedy choices, $u_{i+1}$ follows $u_i$ on the path $P$.

This proves the induction hypothesis. We see that $P_1$ traverses the layers $A_{1,j}$ according to our ordering and visits every layer at most once. Furthermore, applying the same argument for the last vertex $u_{\ell'} \in P_1$ shows that either $u_{\ell'}$ satisfies (vi), or it has a good neighbor in a layer $A_{2,j}$. Note that in the former case, we would choose $u_{\ell} := u_{\ell'}$ and $P' = P_1$, and (i), (ii), (vi) follow directly. Otherwise, we claim that we will find a path $P'$ such that $P' \cap V_1 = P_1$.

**Proof of (iii):** If $\{u_i, u_{i+1}, u_{i+2}\}$ are three subsequent vertices on the subpath $P_1$, then the weight increases at least by an exponent $\gamma(\zeta \varepsilon_1)$ between $u_i$ and $u_{i+2}$ since there exists at least one layer in-between containing $u_{i+1}$. Then (iii) follows as we are assuming that $P'$ will not visit any vertex in $V_1 \setminus P_1$.

**Second Phase ($i = 2$):** Suppose that either the last vertex $u_{\ell'} \in P_1$ does not satisfy (vi) or $s \in V_2$. Then there exists a first vertex $u_{\ell'+1} \in P \cap V_2'$, and a maximal subpath

$$P_2 = P_1 \cup \{u_{\ell'+1}, \ldots, u_{\ell}\} \subseteq P$$

starting at $u_1$ such that $(P_2 \setminus P_1) \subset V_2$ and no vertex before $u_{\ell}$ satisfies condition (vi). We show by induction that every vertex of $P_2 \setminus P_1$ is located in a higher layer (w.r.t. the objective $\phi$) than its predecessor. Again, the hypothesis holds for
$P_1 \cup \{u_{\ell'+1}\}$. Assume that the induction hypothesis holds for $P_1 \cup \{u_{\ell'+1}, \ldots, u_i\}$. Then $u_i$ is contained in a layer $A_{2,j}$, and by the induction hypothesis, $u_i$ is the first vertex that $P$ visits in this layer. Furthermore, $P_1 \cup \{u_{\ell'+1}, \ldots, u_i\}$ is a subpath of the induced greedy path $P_{2,j}$. Then by the event $E_{2,j}$, we have one of the following two cases:

- $u_i$ satisfies condition (vi): then set $P' = (u_1, \ldots, u_\ell)$ with $u_\ell := u_i$.

- $u_i$ has a good neighbor $v'$: let $u_{i+1}$ be the best neighbor of $u_i$. Then $u_i \in A_{2,j'}$ holds for $j' < j$. Since the protocol makes greedy choices, $u_{i+1}$ follows $u_i$ on the path $P$.

This proves the induction hypothesis, and we observe that $P_2$ traverses the layers $A_{2,j}$ according to our ordering and visits every layer at most once. However, by repeating the argument for the last vertex $u_\ell$ it follows that the best neighbor of $u_\ell$ can not be in $V_1$. Then the only possibility is that $u_\ell$ satisfies condition (vi). Hence we put $P' = P_2$, and properties (i), (ii), and (vi) follow. In particular, we see that $P' \cap V_1 = P_1$ as claimed above.

Proof of (iv): For three subsequent vertices $\{u_i, u_{i+1}, u_{i+2}\}$ on the subpath $\{u_{\ell'+1}, \ldots, u_\ell\}$, the objective increases always at least by an exponent $\gamma(\varepsilon_1)$ between $u_i$ and $u_{i+2}$ as at least one layer, containing $u_{i+1}$, lies in-between. This proves (iv), and the only remaining property is (v).

Proof of (v): We already know that the path $P'$ visits every layer at most once. Therefore we can upper-bound the length $|P'|$ by counting the total number of potentially visited layers. By construction we have $f_0(n)$ first layers $A_{1,j}$ which are defined via $\varepsilon_1$, and similarly there are $f_0(n)$ final layers $A_{2,j}$ defined via $\varepsilon_1$. Let $L_2$ denote the number of layers $A_{2,j}$ defined via $\varepsilon_2 = o(1)$ which the routing potentially visited. In order to upper-bound $L_2$, we recall that between two neighboring vertices of $P'$, the objective is always increasing. Therefore $P'$ only contains vertices of objective at least $\phi(u_1)$. Clearly, this implies that we only need to consider layers $A_{2,j}$ where $\psi_{j-1} \geq \phi(u_1)$. Therefore $L_2$ is upper-bounded.
by the solution of $\phi_0^{-\gamma(\epsilon_2) L_2} = \phi(u_1)^{-1}$, and we obtain

$$L_2 = \frac{\log \log \phi_0^{-1}(\phi(u_1)^{-1})}{\log(\gamma(\epsilon_2))} \leq \frac{\log \log \phi_0^{-1}(\phi(u_1)^{-1})}{\log((\beta - 2)^{-1/(1+o(1)))}}$$

$$= \frac{1 + o(1)}{|\log(\beta - 2)|} \log \log \phi_0^{-1}(\phi(u_1)^{-1}).$$

It remains to upper-bound the number $L_1$ of layers $A_{1,j}$, defined via $\epsilon_2$, which $P'$ potentially visits. We observe that every vertex $v \in P'$ satisfies $\phi(v) \geq \phi(u_1)$. Thus, every vertex $v \in P'$ with weight at least $\phi(u_1)^{-1/\gamma(\epsilon_1)}$ belongs to $V_2$, and we only need to count layers $A_{1,j}$ which contain vertices of weight at most $\phi(u_1)^{-1/\gamma(\epsilon_1)} < \phi(u_1)^{-1}$. Then $L_1$ is bounded from above by the solution of the equation $w_0^{\gamma(\epsilon_2) L_1} = \phi(u_1)^{-1}$. It follows

$$L_1 = \frac{\log \log w_0(\phi(u_1)^{-1})}{\log \gamma(\epsilon_2)} \leq \frac{\log \log w_0(\phi(u_1)^{-1})}{\log((\beta - 2)^{-1/(1+o(1)))}}$$

$$= \frac{1 + o(1)}{|\log(\beta - 2)|} \log \log w_0(\phi(u_1)^{-1}).$$

Then the length of $P'$ is at most $L_1 + L_2 + O(f(n))$, which proves (v). 

### 5.8.2 Proof of Theorem 5.1

In this subsection we prove that for all initial choices of $s$ and $t$, greedy routing succeeds with probability $\Omega(1)$. We do this in three steps: In a first step we show that with probability $\Omega(1)$, the best neighbor $u_1$ of $s$ has weight $w_{u_1} \geq w_0$, where $w_0$ is a given constant. In a second step we use Lemma 5.16 to see that then with probability $\Omega(1)$, from $u_1$ the greedy algorithm on $V_{\leq \phi_0}$ finds a path to a vertex $u_\ell$ which has in expectation $\Omega(1)$ neighbors of objective at least $\phi_0$. Since the number of such neighbors is Poisson distributed, with probability $\Omega(1)$ there is at least one such neighbor, and we call the best of them $v_1$. Finally, in a last step we will show that with constant probability, the routing will reach $t$ from $v_1$ via at most one intermediate vertex. For all steps, the success probability is a constant larger than zero, but unfortunately not large enough such that we can do a union bound over the error events. Therefore, we always compute the probabilities conditioned on the previous steps being successful.
Proof of Theorem 5.1. We assume that the parameter $w_{\text{min}}$ is lower-bounded by a small constant, say $\epsilon_1$. Let us start by choosing a weight $w_0 = \Theta(1)$ and an objective $\phi_0 = \Theta(1)$ such that they satisfy the preconditions of Lemma 5.16 and such that the events (i)-(vi) of Lemma 5.16 occur with constant probability $\Omega(1) > 0$.

Start of routing process: Suppose for now that $\phi(s) \leq \phi_0$. We define the event $\mathcal{E}_s: s$ itself or its best neighbor has weight at least $w_0$.

If $w_s \geq w_0$, the event $\mathcal{E}_s$ happens with probability 1. Thus we assume $w_s \leq w_0$. Let $r := \left( \frac{c_1 w_s w_0}{w_{\text{min}} n} \right)^{1/d}$, where $c_1$ is the constant given by (5.1) and (5.2). Since $\phi(s) \leq \phi_0$, we have

$$r^d = c_1 w_0 \phi(s) \|x_s - x_t\|^d \leq c_1 w_0 \phi_0 \|x_s - x_t\|^d = O(\|x_s - x_t\|^d).$$

Next we define a subset of $V$ in which we want to find a neighbor of $s$. Let $A_s := \{u \in V \mid (1) : w_u \geq w_0; (2) : \|x_u - x_t\| \leq \|x_s - x_t\|; (3) : \|x_u - x_s\| \leq r\}$.

By (5.1) and (5.2), every vertex $u \in A_s$ is connected to $s$ with probability $\Omega(1)$. Furthermore, by equation (5.14) a random vertex which satisfies (3) has property (2) with probability $\Omega(1)$ as well. Therefore a random vertex satisfies (2) and (3) with probability $\Omega(r^d)$, and by integrating over all weights larger than $w_0$ it follows that the expected number of vertices in $A_s$ is at least

$$\Omega \left( \int_{w_0}^{\infty} n w_{\text{min}}^{\beta - 1} w^{-\beta} r^d d w \right) = \Omega \left( n w_{\text{min}}^{\beta - 1} w_0^{1 - \beta} r^d \right) = \Omega \left( w_{\text{min}}^{\beta - 2} w_0^{2 - \beta} w_s \right) = \Omega(1).$$

Hence with probability $\Omega(1)$, there exists at least one vertex $u \in \Gamma(s) \cap A_s$. For every $u \in \Gamma(s) \cap A_s$ it holds $\phi(u) \geq \phi(s)$, as the distance to $t$ decreases and the weight increases. Then we take $u_1$ as the neighbor of $s$ in $A_s$ with maximal objective. By definition of $A_s$ we have $w_{u_1} \geq w_0$.

Since $w_s = O(1)$, the degree distribution of $s$ is a Poisson random variable with rate $\Theta(1)$, and thus with probability $\Omega(1)$, $s$ has no neighbor of weight less than $w_0$ with higher objective than $\phi(u_1)$. In this case, the routing indeed proceeds with the vertex $u_1$ and we obtain $\Pr[\mathcal{E}_s] = \Omega(1)$. 


Main part of routing process: Our aim is to show that, conditioned on $E$, greedy routing proceeds to a vertex of objective larger than $\phi_0$. If $\phi(u_1) > \phi_0$ then there is nothing to show, so assume otherwise. We want to show that on the graph induced by the vertex set $V_{\leq \phi_0}$, greedy routing finds a vertex $u_\ell$ which we expect to have neighbors in $V_{> \phi_0}$. More formally, we consider the event

$$E_m : \text{on } G[V_{\leq \phi_0}], \text{the routing finds a vertex } u_\ell \text{ with } \mathbb{E}_{\phi_0}[|\Gamma(u_\ell) \cap V_{> \phi_0}|] = \Omega(1).$$

Then by our choice of $w_0$ and $\phi_0$, Lemma 5.16 (vi) gives the lower bound

$$\Pr[E_m \mid E_s] = \Omega(1).$$

End of routing process: Suppose that either both events $E_s$ and $E_m$ hold, or that already $\phi(s) \geq \phi_0$. We want to prove that with constant probability, the routing finds the target $t$ while visiting at most two other vertices of $V_{> \phi}$. Let $\bar{w}$ be a constant weight chosen large enough and let

$$A_t := \{u \in V \mid (1) : w_u \geq \bar{w}; (2) : \|x_u - x_t\|^d \leq (n\phi_0)^{-1}\}.$$ 

We claim that the expected number of vertices in the set $A_t$ is $\Omega(1)$. Note that here, we need to condition on the events $E_s$ and $E_m$. However, $A_t$ is independent of $E_m$ since $A_t \subseteq V_{> \phi_0}$, and conditioning on $E_s$ only increases the expected size of $A_t$ (in the situation where the best neighbor of $s$ has already very large objective). Thus

$$\mathbb{E}[|A_t| \mid E_s \land E_m] \geq \mathbb{E}[|A_t|] = \Omega(nw_{\max}^{\beta-1}\bar{w}^{1-\beta}(n\phi_0)^{-1}) = \Omega(1)$$

as we have both $\phi_0^{-1} = \Omega(1)$ and $\bar{w}^{1-\beta} = \Omega(1)$. Now, we uncover all remaining vertices, and all edges that involve at least one vertex from $V_{\leq \phi_0}$, but not the edges within $V_{> \phi_0}$. Due to the Poisson point process, both $|A_t|$ and $|\Gamma(u_\ell) \cap V_{> \phi_0}|$ are Poisson random variables with constant mean, therefore with constant probability both sets are non-empty. In this case, the set $\Gamma(u_\ell) \cap V_{> \phi_0}$ is non-empty and there exists a vertex $v_1 \in \Gamma(u_\ell) \cap V_{> \phi_0}$ with maximal objective. Note that if $v_1 = t$, then we are already done. Furthermore, in the case $\alpha < \infty$ by (5.1) the edge $\{v_1, t\}$ is present with probability $p_{v_1 t} = \Omega(\min\{1, (\phi(v_1)w_t)^\alpha\}) = \Omega(1)$, which is already sufficient for finding $t$ with constant probability.
It remains to investigate the case $\alpha = \infty$ for which we constructed the set $A_t$. Recall that we condition on that there exists at least one vertex $v_2 \in A_t$. We claim that for every $u' \in V_{> \phi_0}$ and $v' \in A_t$ it holds either $u' = v'$ or $p_{u'v'} = \Omega(1)$. Indeed, by the triangle inequality we have

$$\|x_{u'} - x_{v'}\| \leq 2 \max\{\|x_{u'} - x_t\|, \|x_{v'} - x_t\|\},$$

and for $\overline{w}$ large enough we deduce

$$\frac{w_{u'}w_{v'}}{w_{\min}n\|x_{u'} - x_{v'}\|^d} \geq \frac{1}{2} \max\{\phi(u')w_{v'}, \phi(v')w_{u'}\} = \Omega(\overline{w}\phi_0) \geq \frac{1}{c_1},$$

where $c_1$ is the constant provided by (5.2). Then by (5.2) it follows $p_{u'v'} = \Omega(1)$. In particular, it also holds $p_{tv'} = \Omega(1)$.

We can finish the proof as follows: we toss the coins for all edges in $V_{> \phi_0}$ except the edges incident to $t$. By the above observation, with constant probability either $v_1 = v_2$ or the edge $\{v_1, v_2\}$ is present. We now consider an additional vertex $v_3$ that is defined as follows: if $\phi(v_1) < \phi(v_2)$, let $v_3$ be the best neighbor of $v_1$, if $\phi(v_1) \geq \phi(v_2)$ put $v_3 := v_1$. Note that in the first case, we have $\phi(v_3) \geq \phi(v_2) \geq \phi(v_1)$. We see that

$$\frac{w_{v_3}w_t}{w_{\min}n\|x_{v_3} - x_t\|^d} = \phi(v_3)w_t \geq \phi(v_2)w_t \geq \frac{1}{c_1}.$$ 

Finally, we uncover the edge $\{v_3, t\}$ and independently of the previous events, this edge is also present with constant probability and with probability $\Omega(1) > 0$, the routing finds $t$ via $v_3$. ■

5.8.3 Proof of Theorem 5.2

In this section, in addition to (5.1) and (5.2) we also assume condition (5.3), i.e., whenever two vertices $u$ and $v$ satisfy $\|x_u - x_v\|^d \leq \frac{c_1w_uw_v}{w_{\min}n}$, then the edge $\{u, v\}$ is present deterministically. This assumption is very natural and in particular satisfied by hyperbolic random graphs. We first prove the second statement of the theorem which is a direct consequence of Lemma 5.16.

Proof of Theorem 5.2 (ii). Let $s$ and $t$ be two vertices such that both $w_s = \omega(1)$ and $w_t = \omega(1)$. Then we apply Lemma 5.16 for $w_0 := w_s$ and $\phi_0 := (c_1w_t)^{-1}$. 

By statement (vi) of Lemma 5.16 and Markov’s inequality it follows that with probability
\[ 1 - O(w_0^{\beta - 2} \min \{ w_s, w_t \}^{-\Omega(1)}) = 1 - \min \{ w_s, w_t \}^{-\Omega(1)} \]
the greedy routing reaches a first vertex \( u' \) of objective at least \( \phi_0 = (c_1 w_t)^{-1} \). Then either \( u' = t \), or
\[ c_1 \frac{w'_u w_t}{w_0 n^d \| x'_u - x_t \|^d} \geq c_1 \phi_0 w_t = 1. \]
In this case, \( \{ u', t \} \in E \) and thus the routing finds \( t \) via \( u' \).

For part (i) of Theorem 5.2, note that the statement becomes trivial if \( w_{\text{min}} \) is small. So we can assume that \( w_{\text{min}} \) is larger than a given constant \( c \) which depends on \( \alpha, \beta, d \) and \( \epsilon \). In particular, this allows us to assume \( w_{\text{min}} > w_1(\epsilon_1) \).

For the proof of Theorem 5.2 (i), we cannot directly apply Lemma 5.16 because this yields only a failure probability which is polynomially small in \( w_0 \geq w_{\text{min}} \). Instead, we will apply Lemma 5.14 for the very first steps, which will allow us to increase the lower bound on the success probability. Afterwards, we will apply Lemma 5.16 for values \( w_0, \phi_0 \) which depend exponentially in \( w_{\text{min}} \) such that the lemma indeed yields a sufficiently small failure probability. Finally, we need to ensure that after seeing vertices of objective at least \( \phi_0 \), the routing does not die out and continues until it reaches vertices which connect deterministically to \( t \).

For the first step, we will apply the following lemma.

**Lemma 5.17.** Let \( w_0 := e^{w_{\text{min}}^{\Omega(1)}} \) and assume that \( w_{\text{min}} \) is sufficiently large. Then there exists a constant \( c' > 0 \) such that with probability
\[ 1 - O(e^{-w_{\text{min}}^{\Omega(1)}}) \]
the greedy routing visits at least one vertex \( u \) which satisfies either (i) \( u \geq w_0 \), (ii) \( \phi(u) \geq w_0^{-c'} \), or (iii) \( u \in V_2 \).

**Proof.** We use similar techniques as in the proof of Lemma 5.16. Notice that if the starting vertex \( s \) satisfies at least one of the three properties, there is nothing
to show. Let $\phi_0 := w_0^{-c'}$ and assume that $w_s \leq w_0$, $\phi(s) \leq \phi_0$, and $s \in V_1$. We start by defining a weight sequence $y(i)$ as follows: We put $y_0 := w_s$, and for $i \geq 1$ we put $y_i := y_{i-1}^{\gamma(\epsilon_1)}$. Notice that there exists a first $y_j$ such that $y_j \geq w_0$. Moreover, for all $1 \leq i \leq j$ we define

$$A_i := \{v \in V_1 : (1): y_{i-1} \leq w_v < y_i; (2): \phi(v) \leq \phi_0\}$$

and

$$B_i := \{v \in V_1 : (1): w_v \leq y_i; (2): \phi(v) \leq \phi_0\}.$$ 

Furthermore, denote by $P_i$ the greedy path induced by the set $B_i$ and by $E_i$ the event that either $P_i \cap A_i$ is empty or the first vertex $v \in P_i \cap A_i$ has at least one neighbor $v' \notin B_i$ such that $\phi(v') \geq \phi(v)$ and $\phi(v') \geq \phi(u)$ holds for all $u \in \Gamma(v) \cap B_i$. Let $1 \leq i \leq j$. Clearly, we have $\Pr[E_i | P_i \cap A_i = \emptyset] = 1$, and in order to lower-bound $\Pr[E_i]$ we can assume $P_i \cap A_i \neq \emptyset$. Let $v$ be the first vertex of $P_i \cap A_i$. Then by Lemma 5.12 (i), it holds $\mathbb{E}[|\Gamma(v) \cap V^+(v, \epsilon_1)|] = \Omega(w_{\min}^{\beta-2}y_{i-1}^{\epsilon_1})$. By a Chernoff bound, with probability at least $1 - \exp(-\Omega(w_{\min}^{\beta-2}y_{i-1}^{\epsilon_1}))$ there exists a neighbor $v'$ of $v$ in the set $V^+(v, \epsilon_1)$. In particular $v' \notin B_i$ and $\phi(v') \geq \phi(v)\gamma^{\epsilon_1-1}(v)\gamma(\epsilon_1)$. 

Next we observe that every vertex $u \in B_i$ with objective at least $\phi(v)\gamma^{\epsilon_1-1}(v)$ is contained in $V^-(v, \epsilon_1)$. Hence in order to verify that $v'$ is better than every neighbor of $v$ in $B_i$, it is sufficient that $v$ has no neighbor in $V^-(v, \epsilon_1)$. We observe that by Lemma 5.9, our conditioning on $v$ being the first vertex in $P_i \cap A_i$ decreases the expected size of $V^-(v, \epsilon_1)$. Furthermore, for $c'$ large enough we have $\phi(v)\gamma^{\epsilon_1}(v) \leq \phi_0 \gamma^{\epsilon_1}(v) \leq c$, where $c$ is the constant given by Lemma 5.14. Hence by Lemma 5.14, we obtain

$$\mathbb{E}[|\Gamma(v) \cap V^- (v, \epsilon_1)|] = O(\phi(v)^{\Omega(1)} \gamma^{O(1)}(v)) = O(\phi_0^{\Omega(1)} y_{i-1}^{O(1)}).$$

By Markov’s inequality, with probability $1 - O(\phi_0^{\Omega(1)} y_{i-1}^{O(1)})$ this set is empty and in this case, $\phi(v')$ is higher than the objective of every neighbor of $v$ in $B_i$. It follows that

$$\Pr[E_i] \geq 1 - O(\phi_0^{\Omega(1)} y_{i-1}^{O(1)}) - \exp(-\Omega(w_{\min}^{\beta-2}y_{i-1}^{\epsilon_1})).$$
Recall that the sequence \( y_{(i)} \) grows exponentially, and we obtain

\[
\Pr \left[ \land_{i=1}^j \mathcal{E}_i \right] = 1 - O(\phi_0^{\Omega(1)} y_0^{O(1)}) - \exp(-\Omega(w_{\min}^{\beta-2} y_0^{\epsilon_1}))
= 1 - O(w_0^{-c'\cdot \Omega(1)} + \exp(-\Omega(w_{\min}^{\beta-2} w_0^{\Omega(1)})))
= 1 - O(w_0^{-\Omega(1)})
\]

if \( c' \) is chosen sufficiently large. This allows us to assume that all good events \( \mathcal{E}_i \) occur. Let us consider the final greedy path \( P = \{s = v_0, v_1, v_2, \ldots\} \) and suppose that all good events \( \mathcal{E}_1 \) occur. By construction \( s \in A_1 \), and \( \mathcal{E}_1 \) implies that the routing does not die out at \( s \) (i.e., \( s \) is not isolated) and \( v_1 \notin B_1 \). Then we distinguish two cases: If \( v_1 \notin B_j \), then we are done since \( v_1 \) would satisfy (i), (ii), or (iii). Else, \( v_1 \) is contained in a layer \( A_i \), and in particular it is also the first vertex of \( P_i \cap A_i \). Then we can repeat the argument, apply the event \( \mathcal{E}_i \) and see again that there exists a subsequent vertex \( v_2 \) on the greedy path which is either located outside \( B_j \) or contained in a layer \( A_{i'} \) for \( i' > i \). We can repeat this argument inductively. However, at the latest when we reach a vertex \( v_k \in P \cap A_j \), the event \( \mathcal{E}_j \) implies that the best neighbor \( u \) of \( v_k \) is located outside \( B_j \). Then \( u \) must fulfill at least one of the properties (i)-(iii), which proves the lemma. \( \blacksquare \)

Next, we study the end-phase of the routing and show that once the routing arrives at a vertex of high objective, with sufficiently high probability it continues to a vertex \( u \) such that \( p_{uv} = \Omega(1) \).

**Lemma 5.18.** Let \( \phi_0 = e^{-w_{\min}^{O(1)}} \) and assume that \( w_{\min} \) is sufficiently large. Then if the greedy routing reaches a vertex \( v \) with \( \phi(v) \geq \phi_0 \), with probability

\[
1 - O(e^{-\phi_{\min}^{O(1)}})
\]

it also visits a vertex \( u \) with \( \phi(u) \geq (c_1 w_i)^{-1} \).

**Proof.** If \( \phi_0 \geq \phi_1 := \frac{1}{c_1 w_i} \), there is nothing to show, so we can assume \( \phi_0 < \phi_1 \). In the following we partition the vertex set with objective between \( \phi_0 \) and \( \phi_1 \) in small layers. Let \( \psi_0 := \phi_0 \), and for \( i \geq 1 \) put \( \psi_i := \psi_{i-1} \cdot w_{\min}^{(1-\gamma(\epsilon_i))/2} \). We assume w.l.o.g. that there exists an index \( j \) such that \( \psi_j = \phi_1 \). Then for all \( 1 \leq i \leq j \) we define \( A_i := \{v \in V \mid \psi_{i-1} \leq \phi(v) < \psi_i\} \). Our goal is to show that with sufficiently
high probability, the greedy routing visits at most one vertex per layer \( A_i \) and never gets stuck thereby. Therefore, we denote by \( P_i \) the greedy path induced by all vertices of objective less than \( \psi_i \) and by \( \mathcal{E}_i \) the event that either \( P_i \cap A_i \) is empty or the first vertex \( v \in P_i \cap A_i \) has a neighbor \( v' \) such that \( \phi(v') \geq \psi_i \).

Let \( 1 \leq i \leq j \). If \( P_i \cap A_i = \emptyset \), then \( \mathcal{E}_i \) holds with probability 1. Hence suppose that \( P_i \cap A_i \) is non-empty and let \( v \) be the first vertex on \( P_i \cap A_i \). We distinguish two cases: First suppose \( v \in V_1 \). Note that for \( \epsilon_1 \) small enough it holds

\[
\gamma(\epsilon_1) - 1 = \frac{3 - \beta - \epsilon_1}{\beta - 2} \geq \frac{3 - \beta - \epsilon_1}{2 - 2\epsilon} = \frac{1 - 1/\gamma(\epsilon_1)}{2}.
\]

Then by Lemma 5.12 (i) and a Chernoff bound (Theorem 2.2 (ii)), with probability at least

\[
1 - \exp(-\Omega(w_{\min}^{\Omega(1)}))
\]

there exists a neighbor \( v' \) of \( v \) with objective

\[
\phi(v') \geq \phi(v)w_{\min}^{\gamma(\epsilon_1)-1} \geq \phi(v)w_{\min}^{\gamma(\epsilon_1)-1} \geq \phi(v)w_{\min}^{(1-1/\gamma(\epsilon_1))/2} \geq \psi_i.
\]

It remains the case \( v \in V_2 \). Since \( w_t \geq w_{\min} \) and since we are assuming that \( w_{\min} \) is sufficiently large, it holds \( \phi(v) \leq (c_1w_t)^{-1} \leq (c_1w_{\min})^{-1} \leq 1 \). This allows us to apply Lemma 5.13 (i), and we see that the expected number of neighbors of \( v \) with objective at least \( \phi(v)^{1/\gamma(\epsilon_1)} \) is at least \( \Omega(w_{\min}^{3-\beta-\epsilon}/(2-2\epsilon)) \). By a Chernoff bound, the probability that such a neighbor \( v' \) exists is at least

\[
1 - \exp(-\Omega(w_{\min}^{\Omega(1)})).
\]

By definition of \( V^+(v, \epsilon_1) \) we have at least \( \phi(v') \geq \phi(v)^{1/\gamma(\epsilon_1)} \).

In addition, we use that \( \phi(v) \geq \phi_1 \) and \( w_{\min} \) is a sufficiently large constant, and deduce

\[
\phi(v') \geq \phi(v)^{1/\gamma(\epsilon_1)} = \phi(v)^{1-(3-\beta-\epsilon)/(1-\epsilon)} \geq \phi(v)(c_1w_{\min})^{(3-\beta-\epsilon)/(1-\epsilon)}
\]

\[
\geq \phi(v)w_{\min}^{(3-\beta-\epsilon)/(2-2\epsilon)} = \phi(v)w_{\min}^{(1-1/\gamma(\epsilon_1))/2} \geq \psi_i.
\]

It follows that \( \Pr[\mathcal{E}_i] \geq 1 - O(e^{-w_{\min}^{\Omega(1)}}) \) holds for all \( 1 \leq i \leq j \). Note that by construction we have \( \phi_0 w_{\min}^{j(1-1/\gamma(\epsilon_1))/2} = \phi_1 \), which implies

\[
j = O\left(\frac{\log \phi_0}{\log w_{\min}}\right) = O\left(\frac{w_{\min}^{\Omega(1)}}{\log w_{\min}}\right).
\]

By a union bound it follows

\[
\Pr[\land_{i=1}^{j} \mathcal{E}_i] = 1 - O(w_{\min}^{\Omega(1)} e^{-w_{\min}^{\Omega(1)}}) = 1 - O(e^{-w_{\min}^{\Omega(1)}}).
\]
Therefore we are allowed to assume that all good events occur and finish the proof as follows: Let $P$ be the final greedy path and let $v \in P$ be the first vertex contained in a layer $A_i$. Then $v \in P_i \cap A_i$, and the event $\mathcal{E}_i$ implies that $v$ has a neighbor $v'$ with higher objective than $\psi_i$. Then either $\phi(v') \geq \phi_1$ or $v' \in A_{i'}$ for some $i' > i$. We can repeat the argument until we reach the last layer $A_j$, and it follows that indeed $P$ must contain a vertex of objective at least $\phi_1$.

**Proof of Theorem 5.2 (i).** We prove the statement by combining the two previous lemmas with Lemma 5.16. Let $c' > 0$ be a sufficiently large constant and let $P$ be the resulting greedy path. We put $w_0 := \exp(w_{\min})$ and $\phi_0 := w_0^{-c'}$. First suppose that $\phi(s) \leq \phi_0$ and $s \notin V(w_0, \phi_0)$. In this case, we apply Lemma 5.17 and see that if $c'$ is large enough, then with probability $1 - O(e^{-w_{\min}^{\Omega(1)}})$ there exists a vertex $u \in P$ which is either contained in the large set $V(w_0, \phi_0)$ or satisfies $\phi(u) > \phi_0$.

Next suppose that there exists a first vertex $u_1 \in P \cap V(w_0, \phi_0)$ (in the case $s \in V(w_0, \phi_0)$, we would put $u_1 = s$). If there exists such a vertex, by Lemma 5.16 with probability

$$1 - O\left(\frac{w_0^{\beta-2} \min\{w_0, \phi_0^{-1}\}^{-\Omega(1)}}{w_0^{\Omega(1)}}\right) = 1 - O(e^{-w_{\min}^{\Omega(1)}})$$

we have a vertex $u_\ell$ on the greedy path induced by $V_{\leq \phi_0}$ such that $u_\ell \in V(w_0, \phi_0)$ and

$$\mathbb{E}_{\phi_0} \left[ |\Gamma(u_\ell) \cap V_{> \phi_0}| \right] = \Omega\left(\frac{w_0^{\beta-2} \min\{w_0, \phi_0^{-1}\}^{\Omega(1)}}{w_0^{\Omega(1)}}\right) = \Omega(e^{w_{\min}^{\Omega(1)}}).$$

So far, we uncovered the graph $G[V_{\leq \phi_0}]$. We continue by tossing the coins for all remaining vertices and edges. By a Chernoff bound, with probability $1 - \exp(-\Omega(\exp(w_{\min}^{\Omega(1)})))$ the set $\Gamma(u_\ell) \cap V_{> \phi_0}$ is non-empty, and $u_\ell$ has a neighbor $u'$ of objective at least $\phi_0$. Then there are two possible situations: If $P$ visits $u_\ell$, then the vertex following after $u_\ell$ in $P$ has objective at least $\phi_0$. If $P$ does not visit $u_\ell$, then there must exist an earlier vertex $u_i \in P$ from which $P$ jumps to $G[V_{\geq \phi_0}]$. In both cases, the routing arrives at vertex $u'$ satisfying $\phi(u') \geq \phi_0$. 

Finally suppose that the routing visits a vertex $u'$ of objective at least $\phi_0$ or even starts at such a vertex. Then by Lemma 5.18, with probability $1 - O(e^{-w_{\min}^{\Omega(1)}})$ it also visits a vertex $u''$ such that $\phi(u'') \geq (c_1w_t)^{-1}$.

By a union bound over all error probabilities it follows that with probability $1 - O(e^{-w_{\min}^{\Omega(1)}})$ the greedy path $P$ arrives at a vertex $u''$ fulfilling $\phi(u'') \geq (c_1w_t)^{-1}$. If $u'' = t$, we are done, otherwise we have

$$\|x_{u''} - x_t\|^d = \frac{w_{u''}}{w_{\min}n\phi(u'')} \leq \frac{c_1w_{u''}w_t}{w_{\min}n}$$

and by the additional assumption (5.3) it follows that $\{u'', t\} \in E$ and therefore $t \in P$.

### 5.8.4 Proof of Theorem 5.3

Before bounding the length of the resulting greedy path, we look at the very first steps of the routing process. With the following lemma, we show for both the basic routing algorithm and its patching variants that they can not visit too many vertices of low weight before reaching a vertex of growing weight. We will use the lemma both for proving Theorem 5.3 and later in Section 5.9 for analyzing patching protocols.

**Lemma 5.19.** Let $w_0 = w_0(n) = \omega(1)$ be a function growing in $n$, and suppose that the starting vertex $s$ satisfies $w_s \leq w_0$ and $\phi(s) \leq e^{-w_0}$. Let $A$ be a routing protocol which satisfies (P1), i.e., it makes greedy choices. Then a.a.s., either $A$ visits at most $O(w_0^6)$ different vertices in total, or after visiting at most $O(w_0^6)$ different vertices, $A$ reaches a vertex of weight at least $w_0$.

**Proof.** We denote by $G'$ the subgraph induced by all vertices of weight less than $w_0$ and by $G''$ the subgraph induced by all vertices of weight less than $w_0^3$. Let $r = (w_0^c/n)^{1/d}$, where we take $c > 0$ sufficiently large.

We first prove that a.a.s., during the first few hops the algorithm stays close to the starting position $x_s$, unless it finds a large-weight vertex very soon. Suppose that we first uncover all vertices of weight less than $w_0$ and only see $G'$. Let $s' = v_0, v_1, v_2, \ldots$ be the sequence of different vertices visited by algorithm $A$
when we run it on \( G' \) (so if a vertex is visited several times, we only list it once). By Lemma 5.15 (ii), with probability \( 1 - O(w^{-7}_0) \) there exists no neighbor of \( s \) in \( G' \) with geometric distance at least \( r \) to \( x_s \), i.e., there is no jump outgoing from \( s \). For the following vertices \( v_1, v_2, \ldots \), we argue similarly: If we condition on that there was no outgoing jump on the previous vertices \( s, v_1, \ldots, v_i \), then by Lemma 5.9, the probability that there is an outgoing jump from \( v_{i+1} \) is even smaller than without conditioning, and by Lemma 5.15 (ii) for \( v_{i+1} \) this probability is at most \( 1 - O(w^{-7}_0) \), regardless whether we condition on the same event for the previous vertices or not. We apply a union bound over the first \( 2w^6_0 \) vertices that the routing algorithm finds in \( G' \) and see that with probability \( 1 - O(w^{-7}_0) \), we do not see such an outgoing jump during the first \( 2w^6_0 \) steps. Note that if the algorithm visits less vertices in \( G' \), we can apply the union bound just over all visited vertices and obtain the same result.

As a next step, let us consider a \( w^3_0 \)-grid as given by Definition 5.10 and the ball \( B(2w^6_0 r, s) \) of radius \( 2w^6_0 r \) around \( x_s \). This ball has volume at most \( 4^d w^{6d} r^d = 4^d w^{6d+c} n^{-1} \) and intersects \( \Theta(w^{6d+c-3}_0) = w^{O(1)}_0 \) cells of the \( w^3_0 \)-grid. Let \( \mathcal{C} \) be the collection of all these cells contained in our ball \( B(2w^6_0 r, s) \). The expected number of vertices of weight at most \( w_0 \) contained in a single cell \( C_i \in \mathcal{C} \) is at most \( w^3_0 \), and by a Chernoff bound, with probability \( 1 - \exp(-\Omega(w^3_0)) \) there are not more than \( 2w^3_0 \) vertices in \( C_i \). The same is true for all \( C_i \in \mathcal{C} \), and by a union bound, with probability \( 1 - O(w^{O(1)}_0 \exp(-\Omega(w^3_0))) = 1 - o(1) \) every cell \( C_i \in \mathcal{C} \) contains at most \( 2w^3_0 \) vertices.

It follows that after visiting \( 2w^6_0 \) different vertices, the algorithm either has escaped from the ball \( B(2w^6_0 r, s) \) or it has visited at least \( w^3_0 \) different cells. However, from the above calculation we know that a.a.s. none of the first \( 2w^6_0 \) visited vertices has a neighbor inside \( G' \) with geometric distance at least \( r \). Thus for all \( i \leq 2w^6_0 \) it holds

\[
\|x_s - x_{v_i}\| \leq 2w^6_0 r.
\]

Therefore a.a.s. the routing does not escape the ball \( B(2w^6_0 r, s) \) during the first \( 2w^6_0 \) hops, and in this case, the algorithm running on \( G' \) either stops before visiting \( 2w^6_0 \) vertices or it visited vertices in at least \( w^3_0 \) different cells.

In particular, we can lower-bound the geometric distance to \( t \) for all these
first vertices. Observe that \( \|x_s - x_t\|_d = \frac{w_s}{\phi(s) w_{\min}} \geq \frac{e^{-w_0}}{n} \) as we are assuming \( \phi(s) \leq e^{-w_0} \). On the other hand it holds
\[
(2w_0^6 r)^d = 2^d w_0^{O(1)} n^{-1} = o(e^{w_0 n^{-1}}),
\]
and by the triangle inequality it follows that for all \( i \leq 2w_0^6 \) we have
\[
\|x_{v_i} - x_t\|_d \geq e^{w_0} \frac{2n}{n}.
\]

Now we add the vertices of weights between \( w_0 \) and \( w_3^0 \) to the graph in order to obtain \( G'' \). Notice that due to the Poisson point process, the distribution of the new vertices is completely independent from what we have observed in \( G' \). The vertices with even higher weight will be uncovered later. Then we rerun the algorithm \( A \) on \( G'' \), yielding a sequence \( s'' \) of different visited vertices. There are three cases. First, if \( |s''| < 2w_0^6 \) then we are done, since then on \( G \) the algorithm either visits less than \( 2w_0^6 \) different vertices, or it visits a vertex that is not in \( G'' \). Similarly, if \( |s''| \geq 2w_0^6 \), and among the first \( 2w_0^6 \) vertices in \( s'' \) there is at least one with weight at least \( w_0 \), then we are also done: on \( G \) the algorithm must either visit a vertex that is not in \( G'' \), or it visits the same first \( 2w_0^6 \) different vertices as in \( s'' \), thus visiting a vertex of weight at least \( w_0 \).

It remains the case where \( |s''| \geq 2w_0^6 \), and the first \( 2w_0^6 \) vertices in \( s'' \) all have weight at most \( w_0 \). We have seen that in this case, it visits vertices in at least \( w_0^3 \) different cells of the \( w_0^3 \)-grid. By the Bulk Lemma 5.11, we deduce that when adding the heavy vertices of weight at least \( w_0^3 \) to the graph, a.a.s. there exists a smallest index \( i < 2w_0^6 \) such that \( v_i \) has a neighbor \( u \) with \( w_u \geq w_0^3 \) which is located in the same cell. We know that a.a.s. \( \|x_{v_i} - x_t\|_d \geq e^{w_0} \frac{2n}{2n} \). Since \( u \) is in the same cell as \( v_i \), it holds \( \|x_u - x_t\|_d \leq (1 + o(1))\|x_{v_i} - x_t\|_d \) and therefore
\[
\phi(u) \geq (1 - o(1)) w_0^2 \phi(v_i).
\]

On the other hand, we know that a.a.s. \( v_i \) has no neighbor \( u' \) satisfying both \( w'_u < w_0 \) and \( \|x_{v_i} - x'_u\| \geq r \), therefore for every neighbor \( u'' \) of \( v \) with weight less than \( w_0 \) we have
\[
\|x_{v_i} - x''_u\| < o(\|x_{v_i} - x_t\|).
\]
Again by the triangle inequality it follows that such a neighbor \( u'' \) satisfies

\[
\|x_{u''} - x_t\| = (1 - o(1))\|x_{v_i} - x_t\|.
\]

We conclude that in this case we have

\[
\phi(u'') \leq (1 + o(1)) w_0 \phi(v_i)
\]

and indeed the objective of the vertex \( u \) is large enough such that by (P1) the algorithm will prefer \( u \) to all neighbors of \( v_i \) of weight at most \( w_0 \). Hence, the neighbor of \( v_i \) with maximal objective has weight at least \( w_0 \) and at the latest at this point, \( A \) visits a vertex of weight at least \( w_0 \).

\[\blacksquare\]

**Proof of Theorem 5.3.** Let \( f_0 = f_0(n) = \omega(1) \) be a growing function. Furthermore let \( w_0 = w_0(n) := \max\{\log f_0(n), w_s\} \) and \( \phi_0 := \min\{(c_1 w_t)^{-1}, f_0^{-1}\} \). We first assume \( \phi(s) \leq \phi_0 \). Suppose \( w_s \leq w_0 \). Then by Lemma 5.19, a.a.s. either greedy routing ends in a dead end after visiting \( O(w_0^\delta) = o(f_0) \) vertices, or one of the first \( O(w_0^\delta) \) vertices has weight at least \( w_0 \). If \( w_s \geq w_0 \), we already start at such a vertex. Let \( u_1 \) be the first visited vertex of weight at least \( w_0 \). Clearly \( \phi(u_1) \geq \phi(s) \). Suppose \( \phi(u_1) \leq \phi_0 \). Then we apply Lemma 5.16 with \( w_0 \) and \( \phi_0 \). By statements (v) and (vi), a.a.s. after at most

\[
L = \frac{1 + o(1)}{\log \beta - 2} \left( \log \log w_0 (\phi(s)^{-1}) + \log \log \phi_0^{-1} (\phi(s)^{-1}) \right) + O(f_0(n))
\]

steps, the routing process arrives at a vertex of objective at least \( \phi_0 \). Notice that by construction, \( w_0 \geq w_s \) and \( \phi_0^{-1} \geq w_t \), thus \( L \) is not larger than the claimed length of the greedy path.

Hence we can assume that the routing reaches a vertex \( u' \) with \( \phi(u') \geq \phi_0 \). Note that in the case \( \phi(s) \geq \phi_0 \), we would use \( s = u' \). If \( \phi_0 = f_0^{-1} \), then we apply Lemma 5.8 to see that a.a.s., there exist at most \( O(f_0) \) vertices with objective at least \( \phi_0 \). Even if the algorithm visits every single vertex with this property, this only needs \( O(f_0(n)) \) steps. It remains to treat the case \( \phi_0^{-1} = c_1 w_t \). Then

\[
c_1 \frac{w_{u'w_t}}{w_{\min n\|x_{u'} - x_t\|^d}} = c_1 \phi(u') w_t \geq c_1 \phi_0 w_t = 1,
\]
and we observe that every vertex $u$ that the algorithm visits after $u'$ has probability $p_{ut} = \Omega(1)$ to directly connect to the target $t$. Suppose for contradiction that from now on, the routing process visits more than $f_0(n) = o(1)$ additional vertices. Since each of these vertices independently connects to $t$ with constant probability, it follows that with probability $1 - o(1)$ at least one of the $f_0$ additional vertices connects to $t$. However, in this case the algorithm indeed goes to $t$ and stops. We conclude that after reaching a vertex of objective at least $\phi_0$, a.a.s. the routing stops before visiting $f_0(n)$ additional vertices. Adding up all steps then proves the second statement of the Theorem. For the first, general upper bound of $2 + o(1) \log \beta - 2 \log \log n$, we pick $f_0(n) = o(\log \log n)$. Then the statement follows immediately as $\phi(s) = \Omega(n^{-1})$ and $w_s, w_t = \Omega(1)$.

5.9  Patching Algorithms: Proof of Theorem 5.4

Similarly as in the analysis of the basic routing process, we prove Theorem 5.4 in three steps. We first use Lemma 5.19 in order to show that the algorithm soon finds vertices of high weight. Next we can apply Lemma 5.16 for analyzing the process until an objective $\phi_0$ is reached. For the end of the routing process, we need a preparatory result. We do this with the following lemma which considers the very last phase before the algorithm hits the target vertex $t$.

**Lemma 5.20.** There is a constant $c > 0$ such that the following statement is true. Let $w_0 = w_0(n) = \omega(1)$ be a function growing in $n$ such that $w_0(n) = O(\log \log n)$, and suppose that the target vertex $t$ satisfies $w_t \leq w_0$ and the start vertex satisfies $\phi(s) \geq w_0^{-c}$. Then a.a.s., either $s$ and $t$ are not in the same component, or there exists a path $P$ of length $O(w_0^2)$ from $t$ to a vertex $u$ with $w_u \geq w_0$ such that every vertex $v \in P$ (including $u$) fulfills $\phi(v) \geq w_0^{-O(1)}$.

**Proof.** Let $G'$ be the subgraph induced by vertices of weight at most $w_0$, and suppose that we first uncover this subgraph. Furthermore let $C$ be the connected component of $G'$ which contains $t$, and let $r := (w_0^c / n)^{1/d}$, where $c > 0$ is a sufficiently large constant. By Lemma 5.15 (ii), with probability $1 - O(w_0^{-3})$ the
target vertex $t$ has no neighbor in $G'$ with geometric distance at least $r$ to $x_t$. Next we explore $C$ by depth first search (selecting vertices in arbitrary order) until either $2w_0^2$ vertices have been visited or $C$ has been completely traversed. Let $v$ be any vertex which has been explored in this process. By Lemma 5.9 and Lemma 5.15 (ii), with probability $1 - O(w_0^{-3})$ the vertex $v$ has no neighbor in $G'$ with geometric distance at least $r$ to $x_v$. As we explored at most $2w_0^2$ vertices, by a union bound it follows that with probability $1 - O(w_0^{-1}) = 1 - o(1)$, we didn’t explore any edge of geometric length at least $r$. Therefore, every explored vertex $v$ satisfies $\|x_v - x_t\|^d \leq 2w_0^2r^d = w_0^{O(1)}/n$ and thus $\phi(v) \geq w_0^{-O(1)}$.

We proceed by considering a $w_0$-grid as given by Definition 5.10, and uncovering the vertices of weight at least $w_0$. Similarly as in the proof of Lemma 5.19, we deduce that there are $w_0^{O(1)}$ cells of the grid which are in distance at most $2w_0^2$ to $x_t$, and a.a.s. all of these cells contain at most $2w_0$ vertices. It follows that if we explored $2w_0^2$ vertices, then they cover at least $w_0$ different cells. In this case, by Lemma 5.11 a.a.s. there exists a vertex $u$ and a vertex $v$ such that (i) $w_u \geq w_0$, (ii) $u$ and $v$ are in the same cell, (iii) $v$ is among the explored vertices, and (iv) $u$ and $v$ are connected with an edge. From (ii) it follows in particular that $\phi(u) \geq w_0^{-O(1)}$, and clearly there exists a $u$-$t$-path among the explored vertices. Hence we find a $u$-$v$-$t$-path with all desired properties.

Suppose that the DFS stops after exploring less than $2w_0^2$ vertices, and let $v$ be any explored vertex. Recall that $\|x_v - x_t\|^d \leq 2w_0^2r^d = w_0^{O(1)}/n$. Then by Lemma 5.15 (iii), with probability $1 - O(w_0^{-3})$ there exists no neighbor $u$ of $v$ such that $\phi(u) \leq w_0^{-c}$ and $w_u \geq w_0$, for a suitable choice of $c$. By a union bound we deduce that a.a.s. this holds for all explored vertices. However, this is only possible if either $C$ is itself the whole connected component of $t$ (and in this case $s$ is not in the same component as $t$) or there exists a vertex $u$ with $w_u \geq w_0$ and $\phi(u) \geq w_0^{-O(1)}$ which has an explored neighbor $v$. This yields again an $u$-$v$-$t$-path with all desired properties.

**Proof of Theorem 5.4.** We first study the main case $\phi(s) = o(1)$, and only come to the exceptional case $\phi(s) = \Omega(1)$ at the very end.
The case $\phi(s) = o(1)$ - finding a vertex $u_1$ of large weight: Let us start by fixing a weight $w_0 = O(\log \log \log n)$. We assume for now that the starting vertex $s$ satisfies $\phi(s) = o(1)$, which happens a.a.s. if $s$ is a random vertex. Note that we may choose $w_0$ such that $\phi(s) \leq e^{-w_0}$. Furthermore let $A$ be an algorithm satisfying $(P1)$-$(P3)$. We first show that either $s$ is not in the same connected component as $t$, or $A$ finds $t$ quickly, or $A$ visits soon a vertex of weight at least $w_0$. If $w_s \geq w_0$, this is trivial. If $w_s \leq w_0$, either (i) $A$ finds $t$ among the first $O(w_0^6)$ (different) visited vertices, or (ii) the connected component of $s$ has only $O(w_0^6)$ vertices, or (iii) the algorithm visits $O(w_0^6)$ without finding $t$. Note that by condition $(P2)$, the algorithm $A$ spends at most $w_0O(1) = o(\log \log n)$ steps in this phase, which settles the first two cases. In the last case, $A$ visits $O(w_0^6)$ vertices without finding $t$, and by Lemma 5.19 a.a.s. at least one of these vertices has weight at least $w_0$. Let us call the first such vertex $u_1$.

The length of middle part - application of the main lemma: As a next step, we want to apply Lemma 5.16 to bound the number of steps in the middle phase. Let $f_1(n) = \omega(1)$ be a function which grows sufficiently slowly in $n$, and let $w_1 := \log w_0$, $w_2 := \log w_1 = \log \log w_0$, and $\phi_0 := (w_2)^{-1/f_1} = o(1)$. We consider the induced graph $G[V_{\leq \phi_0}]$ of vertices of objective at most $\phi_0$, and denote by $P = \{s, \ldots, u_1, \ldots\}$ the path obtained from running $A$ on $G[V_{\leq \phi_0}]$. We apply Lemma 5.16 for $w_0$ and $\phi_0$. It follows that a.a.s., after $u_1$ the patching algorithm always finds an objective-improving neighbor until it reaches a vertex $u_\ell$ which has a.a.s. $\omega(1)$ neighbors in $V^+(u_\ell, \varepsilon_1)$ of objective at least $\phi_0$. Let $P'$ be the subpath $\{u_1, \ldots, u_\ell\}$. By statement (v) of Lemma 5.16, and by our choice of $w_0$ and $\phi_0$, a.a.s. it holds

$$|P'| \leq \frac{1 + o(1)}{|\log(\beta - 2)|} (\log \log w_0 \phi(u_1)^{-1} + \log \log \phi_0^{-1} \phi(u_1)^{-1}) + o(\log \log n)$$

$$\leq \frac{2 + o(1)}{|\log(\beta - 2)|} \log \log n.$$ 

Existence of $u_j, u_i$ (vertices of objective $\phi(u_j) \approx w_1^{-1}, \phi(u_i) \approx w_2^{-1}$): Next, we claim that a.a.s. there exists a vertex $u_j \in P'$ such that

(5.15) $w_1^{-\Omega(1)} \geq \phi(u_j) \geq w_1^{-1-\Omega(1)}$,
and such that \( A \) also visits \( u_j \) if we let it run on \( G \setminus \{ t \} \) instead of \( G[V_{\leq \phi_0}] \).

By Lemma 5.15 (i), there exists a constant \( \varepsilon' > 0 \) such that a.a.s. the graph has no vertices of weight at least \( w_0 \) and objective at least \( w_0^{-\varepsilon'} \) (except potentially \( t \), if we fix \( t \) and set \( w_t \) to be large). Moreover, we already know that the vertices which \( A \) visited before \( u_1 \) have no neighbor within geometric distance at least \( r \) and weight less than \( w_0 \), thus for \( \varepsilon' \) small enough, they have no neighbor of objective at least \( w_0^{-\varepsilon'} < w_1^{-1-\Omega(1)} \) (except potentially \( t \)). In particular it follows that \( \phi(u_1) \leq w_1^{-1-\Omega(1)} \).

Now let \( u \in P' \) be a vertex such that \( \phi(u) \leq w_1^{-1-\Omega(1)} \). By Lemma 5.15 (iv), with probability \( 1 - O(\min\{w_u, \phi(u)^{-1}\}^{-\Omega(1)}) \) the vertex \( u \) has no neighbor of objective at least \( w_1^{-\Omega(1)} \) and weight at most \( w_1 \). Moreover, by Lemma 5.16 (iii) the weights of the vertices in \( P' \) increase doubly exponentially during the first phase, and by Lemma 5.16 (iv) the objectives increase doubly exponentially during the second phase. This allows us to apply a union bound over all vertices on \( P' \) with objective at most \( w_1^{-1-\Omega(1)} \). We deduce that a.a.s., no vertex \( u \in P' \) with \( \phi(u) \leq w_1^{-1-\Omega(1)} \) has a neighbor of objective at least \( w_1^{-\Omega(1)} \) and weight at most \( w_1 \). By Lemma 5.15 (i), we know that a.a.s. there are no vertices except \( t \) with objective at least \( w_1^{-\Omega(1)} \) and weight at least \( w_1 \). Hence a.a.s., the early vertices of \( P' \) don’t have any neighbors in \( V_{\geq w_1^{-\Omega(1)}} \) and the only possible situation that remains is that indeed, a.a.s. the path \( P' \) contains at least one vertex \( u_j \) satisfying (5.15). Moreover, for sufficiently large \( n \) we have \( w_1^{-\Omega(1)} \leq \phi_0 \). Then, since none of the preceding vertices has a neighbor in \( V_{> \phi_0} \), this vertex \( u_j \) is also visited if \( A \) runs on \( G \). This proves that \( u_j \) exists.

Let \( w_2 = \log(w_1) \). We claim that analogously to (5.15), a.a.s. there exist a vertex \( u_i \in P' \) such that

\[
(5.16) \quad w_2^{-\Omega(1)} \geq \phi(u_i) \geq w_2^{-1-\Omega(1)},
\]

and such that \( A \) also visits \( u_i \) if we let it run on \( G \setminus \{ t \} \) instead of \( G[V_{\leq \phi_0}] \). Indeed, the proof is completely analogous to (5.15).

**End phase - finding a path from \( u_i \) to \( t \) in \( G[V_{> \phi(u_i)}] \):** To finish the argument, we consider the graph \( G[V_{> \phi(u_i)}] \). Clearly, if \( t \) is already connected to a vertex
on the path $P$ which is located before $u_i$ on $P$, then the algorithm succeeds. Otherwise, we claim that it is sufficient to show that either $t$ is not in same connected component as $s$, or that among all vertices of objective at least $\phi(u_j)$ there exists a path from $u_j$ to $t$. Suppose we find such a path. Then $u_j$ and $t$ are in the same connected component $S$ in $G[V_{>\phi(u_j)}]$. In particular, by (P3) the algorithm $A$ visits all vertices of $S$ (including $t$) during the next $|S|^{O(1)}$ steps. By Lemma 5.8, a.a.s. there are at most $O(\phi(u_j)^{-1}) = w_0^{O(1)}$ vertices of objective at least $\phi(u_j)$, therefore $|S|^{O(1)} = w_0^{O(1)} = o(\log \log n)$ as desired. Hence it remains to find a $u_j$-$t$-path in $G[V_{>\phi(u_j)}]$. Since we already know that $u_i$ and $u_j$ are connected by $P'$, it also suffices to find a path from $u_i$ to $t$ in $G[V_{>\phi(u_j)}]$. (Mind the two different indices $i, j$ here.)

We start by finding a path from $t$ to a vertex $u$ of weight at least $w_u \geq w_2^\varepsilon$, where $\varepsilon > 0$ is a sufficiently small constant. If $w_t \geq w_2^\varepsilon$, this is trivial by setting $u = t$, otherwise we apply Lemma 5.20 with weight $w_2^\varepsilon$. Note that the condition $\phi(s) < w_2^{-c}$ is satisfied for sufficiently large $n$. Then either $s$ and $t$ are in different components, or there is a path from $t$ to a vertex $u$ with $w_u \geq w_2^\varepsilon$ such that every vertex in the path has objective at least $w_2^{-O(1)} \geq \phi(u_j)$, where the latter inequality holds for sufficiently large $n$. So it only remains to show that $u$ and $u_i$ are in the same connected component of $G[V_{>\phi(u_j)}]$. For this last step, fix a sufficiently large constant $C > 0$ and consider the set

$$B := \{ v \in V \mid \|x_v - x_t\|^d \leq w_2^C/n \}.$$ 

Clearly, $B$ contains $t$. If $u \neq t$, then a.a.s. we have $\phi(u) \leq w_u^{-\Omega(1)}$ by Lemma 5.15 (i), and thus $w_2^{-O(1)} \leq \phi(u) \leq w_u^{-\Omega(1)} \leq w_2^{-\Omega(1)}$. This upper-bounds both $w_u$ and $\phi(u)^{-1}$, so we get $\|x_u - x_t\|^d = \phi(u)^{-1}w_u/n \leq w_2^{O(1)}/n$. Hence, if $C$ is large enough then $u \in B$. Similarly, using (5.16) we find that $u_i \in B$.

The set $B$ contains in expectation $n' = w_2^C$ vertices. We apply Lemma 3.12 and see that $G[B]$ is itself a (Poissonized) GIRG with probability 1. (The a.a.s.-statement of Lemma 3.12 is only present because we required that the weight sequence of the subgraph follows a power law. Here, in the setting with an underlying Poisson point process, this is not necessary.) By Lemma 4.13 (ii) and (iii) a.a.s. the induced graph $G[B]$ has a giant component which contains
all vertices of weight at least \((\log n')^{O(1)}\). In particular, \(u\) and \(u_i\) are in the giant component of \(G[B]\), so there is a path from \(u\) to \(u_i\) in \(G[B]\). Moreover, every vertex \(v\) in \(B\) has objective at least \(\phi(v) \geq w_u/(n\|x_v - x_t\|^d) \geq w_{\min}w_2^{-C} > \phi(u_j)\) (if \(n\) is sufficiently large), so \(B \subseteq V_{>\phi(u_j)}\). Altogether, we have found a path from \(u_i\) to \(u\) in \(V_{>\phi(u_j)}\), and thus a path from \(u_j\) to \(t\) in \(V_{>\phi(u_j)}\). This concludes the proof under the initial assumption \(\phi(s) = o(1)\).

**The Case \(\phi(s) = \Omega(1)\):** In the remaining case \(\phi(s) = \Omega(1)\), condition (P3) ensures that there exists an objective \(\phi_1 = o(1), \phi_1 \geq (\log\log n)^{o(1)}\), such that we explore the connected component \(S\) of \(s\) in \(G[V_{>\phi_1}]\) in \(|S|^{O(1)}\) steps. We note that by Lemma 5.8, a.a.s. \(|V_{>\phi_1}| = O(\phi_1^{-1})\), so in particular we need at most \(|S|^{O(1)} = o(\log\log n)\) steps to explore \(S\). It only remains to show that a.a.s. either \(s\) and \(t\) are in different connected components in \(G\), or they are in the same component in \(G[V_{>\phi_1}]\). This argument resembles closely the argument in the previous paragraph, and we only repeat it briefly. By Lemma 5.20 either \(s\) and \(t\) are in different components in \(G\), or there is path from \(s\) to \(t\) among vertices of objective at least \(\phi_1\) and of weight at most \(\phi_1^{-\epsilon}\), or in \(G[V_{>\phi_1}]\) there are paths from \(s\) and from \(t\) to vertices \(u_s\) and \(u_t\) of weights at least \(\phi_1^{-\epsilon}\), where \(u_s = s\) or \(u_t = t\) are possible if \(w_s\) or \(w_t\) are large. Then by the same argument as we used for \(u\) and \(u_j\) above, a.a.s. \(u_s\) and \(u_t\) are in the same connected component of \(G[V_{>\phi_1}]\). Summarizing, either there is no path from \(s\) to \(t\) in \(G\), or there is a path from \(s\) to \(t\) in \(G[V_{>\phi_1}]\), and in the latter case the path is explored in \(o(\log\log n)\) steps, as required.

**5.10 Relaxations: Proof of Theorem 5.5**

In our analysis of the basic greedy algorithm and the patching algorithm we always assumed that a node is aware of the exact objectives of its neighbors and packets are always forwarded to the neighbor with largest objective. In practice the exact objective of a neighbor might be unavailable. In the following we show that all our theorems also hold if objective values are only approximately known.
In particular, we generalize from our original objective function $\phi$ to a class of objective functions, and prove Theorem 5.5.

We first remark without formal proof that Theorem 5.5 is essentially best possible.

**Remark 5.21.** The relaxations in Theorem 5.5 are best possible. In Section 5.6 we have described typical trajectories of greedy routing. In particular, it turned out that there are two phases in which the weight and the objective improve by an exponent of $1/(\beta - 2)$, respectively. If we replace the exponent $o(1)$ in Theorem 5.5 by a constant, then in each step the algorithm could pick vertices which would increase the weight (in the first phase) and the objective (in the second phase) only by an exponent $c < 1/(\beta - 2)$. This would increase the number of steps to at least $2^{2+o(1)} \log \log n$ instead of $2^{2+o(1)/(\log(\beta - 2))}$, thus making the algorithm considerably slower.

In fact we prove a slightly stronger statement than Theorem 5.5 which we will be useful in Section 5.11 for transferring our results from GIRGs to hyperbolic random graphs.

**Proposition 5.22.** Let $\tilde{\phi} : V \to \mathbb{R}$ be a function which is maximized with $t$ such that

$$
\tilde{\phi}(v) = \Theta(\phi(v) \cdot \min\{w_{v}, \phi(v)^{-1}\}^{o(1)}), \quad (5.17)
$$

holds for all $v \in V$. Then Theorems 5.1, 5.2, 5.3, and 5.4 also hold for greedy routing with respect to $\tilde{\phi}$ instead of $\phi$.

Moreover, let $\delta > 0$ be a small constant and let $\lambda_1 := \frac{2}{3}$ if $w_t \geq 1$ and $\lambda_1 := \frac{4}{3}$ if $w_t < 1$. Then for Theorems 5.1, 5.2, and 5.3 the previous statement is still true if for vertices $v \in V$ with $\phi(v) \geq O(w_t^{-1+\delta})$ we relax Condition (5.17) to the weaker condition

$$
\tilde{\phi}(v) = \Omega(w_t^{-1+\lambda_1\delta}). \quad (5.18)
$$

Finally, for transferring Theorem 5.4 this weaker condition is also sufficient if $t$ is a random vertex.
Note that \( \lambda_1 \) is chosen such that in both cases, (5.18) is indeed weaker than (5.17) and the original objective \( \phi(v) \) fulfils (5.18). Below we show that our analysis of the greedy routing w.r.t. \( \phi \) applies as well for such a relaxed objective function \( \tilde{\phi} \) which fulfills the preconditions of Theorem 5.22. Note that all technical definitions of sets such as \( V_1, V_2, V^+(v, \varepsilon), V^-(v, \varepsilon), \ldots \) are still defined via \( \phi \) and not via \( \tilde{\phi} \).

**Proof of Proposition 5.22.** We first study how Theorems 5.1, 5.2, and 5.3 for the basic routing process can be transferred to relaxed objective functions, and defer the corresponding analysis for patching algorithms to the end of this section. Let \( f_0(n) = \omega(1) \) be a growing function, let \( \delta > 0 \), and recall from our basic analysis that \( \varepsilon_1 \) is a constant chosen sufficiently small. Moreover we put \( \varepsilon_2 = (\log \log f_0(n))^{-1} \). We want to prove for every \( g_0(n) = o(1) \) that the results about greedy routing w.r.t. \( \phi \) transfer to every relaxed objective function \( \tilde{\phi} \) which is maximized at the target \( t \) and satisfies

\[
\tilde{\phi}(v) = \Theta(\phi(v) \cdot \min\{w_v^{\gamma(\varepsilon_1)}, \phi(v)^{-1}\} \pm \varepsilon_2 / c)
\]

for all \( v \) with \( \phi(v) \leq c w_t^{-1+\delta} \), where \( c > 0 \) is a constant chosen sufficiently large later. Furthermore we assume that all other vertices satisfy \( \tilde{\phi}(v) = \Omega(w_t^{-1+\lambda_1 \delta}) \).

We observe that by choosing \( f_0(n) \) sufficiently small compared to \( g_0(n) \), we can assume as well that every vertex \( v \) with \( \phi(v) \leq c w_t^{-1+\delta} \) satisfies

\[
\tilde{\phi}(v) = \Theta(\phi(v) \cdot \min\{w_v^{\gamma(\varepsilon_1)}, \phi(v)^{-1}\} \pm \varepsilon_2 / c)
\]

for \( c = 4\gamma(\varepsilon_1) \), which is a weaker restriction and thus gives a stronger statement. Notice that for vertices \( v \in V_1 \), the minimum will be attained by \( w_v^{\gamma(\varepsilon_1)} \), and vice-versa for vertices \( v \in V_2 \), the maximum will be attained by \( \phi(v)^{-1} \). In other words, unless \( \phi(v) > \bar{c} w_t^{-1+\delta} \), for \( v \in V_1 \) the relaxation is \( \tilde{\phi}(v) = \Theta(\phi(v)w_v^{\pm \gamma(\varepsilon_1) \varepsilon_2 / c}) \) and for \( v \in V_2 \) the relaxation is \( \tilde{\phi}(v) = \Theta(\phi(v)^{1 \pm \varepsilon_2 / c}) \).

Let \( \tilde{A} \) be a greedy routing algorithm which performs routing w.r.t. \( \tilde{\phi} \). In the following, we show that we can modify our analysis of the routing process such that whenever \( \tilde{A} \) proceeds to a new vertex \( v \), this vertex still has all desired properties and the routing trajectory is essentially the same. We will describe
how the proofs of Theorem 5.1, 5.2, and 5.3 can be adjusted for the relaxed routing algorithm $\tilde{A}$, but we will not prove this in full detail. Recall that all these proofs basically contain three parts: first we investigated the start of the routing, which contains the very first steps until a certain weight $w_0$ is reached. Then in the second step we applied Lemma 5.16 in order to reach a vertex of objective at least $\phi_0$, and in the final part we studied how the process finds $t$. When describing how our results can be transferred to the relaxed algorithm $\tilde{A}$, we therefore handle these three parts separately.

Let $\lambda_2 := \frac{1}{3}$ if $w_t \geq 1$ and $\lambda_2 := \frac{5}{3}$ if $w_t < 1$. In the following, we always assume that $\phi_0 \leq c w_t^{-1+\lambda_2 \delta}$, where $c > 0$ is a sufficiently small constant. As it turns out, it is possible to transfer the main theorems to $\tilde{A}$ without being obliged to take $\phi_0$ larger than $c w_t^{-1+\lambda_2 \delta}$. We start by studying the main part of the routing process.

**Main part of routing process:** In all proofs we applied Lemma 5.16 for analyzing the main part of the routing process. Here, we indicate why this main lemma still applies for the relaxed algorithm $\tilde{A}$. (Note that we still apply it to vertices with real objective at most $\phi_0$.) In fact, we will slightly strengthen part (vi) of the lemma such that for the end phase we can still guarantee that $\tilde{A}$ reaches a vertex of real objective at least $\phi_0$. Recall that we intend to apply Lemma 5.16 only for values $\phi_0 \leq c w_t^{-1+\lambda_2 \delta}$. We claim that as long as $c$ is sufficiently small, then for every vertex $v \in V_{\leq \phi_0}$ and every vertex $u$ with $\phi(u) \geq c w_t^{-1+\lambda_2 \delta}$, it holds

$$\tilde{\phi}(v) < \tilde{\phi}(u). \quad (5.20)$$

Indeed, first suppose $w_t \geq 1$. Then, by the weak restriction (5.18),

$$\tilde{\phi}(v) \leq O(\phi_0^{1+\varepsilon_2}) \leq O\left((c w_t^{-1+\delta/3})^{1-\varepsilon_2}\right) \leq \Omega(w_t^{-1+2\delta/3}) \leq \tilde{\phi}(u).$$

On the other hand, if $w_t < 1$ we have

$$\tilde{\phi}(v) \leq O(\phi_0^{1+\varepsilon_2}) \leq O\left((c w_t^{-1+5\delta/3})^{1+\varepsilon_2}\right) \leq \Omega(w_t^{-1+4\delta/3}) \leq \tilde{\phi}(u).$$

Lemma 5.16 deals with two phases of the routing. In the first phase of the routing process, our goal was to prove for a given vertex $v \in V_1$ that its best neighbor w.r.t. $\phi$ has weight at least $w_v^\phi(\varepsilon \varepsilon)$, where $\varepsilon \in \{\varepsilon_1, \varepsilon_2\}$ and this choice was
depending on the weight \( w_v \). Recall that our analysis assumed that all good events occur. Then in particular \( v \) has no neighbor in the set \( V^-(v, \varepsilon) \), and every neighbor \( u \) with weight at most \( w_v^{\gamma(\varepsilon)} \) can have objective at most \( \phi(v)w_v^{\gamma(\varepsilon)-1} \).

Let \( u \) be such a neighbor. We claim that \( \tilde{\phi}(u) \) is small enough such that \( u \) is still not the best neighbor of \( v \) w.r.t. \( \tilde{\phi} \). Due to the relaxation (5.19), the vertex \( u \) satisfies

\[
\tilde{\phi}(u) \leq \phi(u)w_u^{\gamma(\varepsilon)\epsilon_2/c} \leq O(\phi(v)w_v^{\gamma(\varepsilon)-1+\gamma(\varepsilon)\gamma(\varepsilon_1)\epsilon_2/c}).
\]

In order to guarantee that the routing does not proceed with such a vertex, we need to find a neighbor of weight at least \( w_v^{\gamma(\varepsilon)} \) and higher relaxed objective \( \tilde{\phi} \). Before, we found such good neighbors in the set \( V^+(v, \varepsilon) \). Instead, we could as well consider the set \( V^+(v, \varepsilon/3) \). We apply Lemma 5.12 to see that

\[
\mathbb{E}[|\Gamma(v) \cap V^+(v, \varepsilon/3)|] = \Omega(w_{\min}^{\beta-2}w_v^{\Omega(\varepsilon)})
\]

which is fine for all error bounds. Let \( \gamma' := \frac{1+\varepsilon/3}{\beta-2} \). By Lemma 5.7, the expected number of neighbors with weight larger than \( w_v^{\gamma'} \) is \( O(w_{\min}^{\beta-2}w_v^{-\Omega(\varepsilon)}) \). Hence we expect \( \Omega(w_{\min}^{\beta-2}w_v^{\Omega(\varepsilon)}) \) neighbors in \( V^+(v, \varepsilon/3) \) with weight at most \( w_v^{(1+\varepsilon/3)/\beta-2} \).

By our relaxations, a vertex \( u' \) in this set satisfies

\[
\tilde{\phi}(u') \geq \phi(u')w_u^{-(\gamma(\varepsilon)\epsilon_2/c)} \geq \Omega\left(\phi(v)w_v^{\gamma(\varepsilon/3)-1-\gamma'(\gamma(\varepsilon_1)\epsilon_2/c)}\right).
\]

Since we assume that \( \varepsilon_1 \) is small enough and \( \varepsilon \leq \varepsilon_1 \), we observe that

\[
\gamma(\varepsilon/3) - \gamma(\varepsilon) = \frac{2\varepsilon}{3(\beta-2)} > \frac{(2+\varepsilon-\zeta\varepsilon)\epsilon_2}{4(\beta-2)} = (\gamma(\varepsilon) + \gamma') \frac{\gamma(\varepsilon_1)\epsilon_2}{c}.
\]

Combining the last inequality with (5.21) and (5.22), we see that \( \tilde{\phi}(u') \) is larger than the relaxed objective of \( v \) itself and of every neighbor of \( v \) with weight less than \( w_v^{\gamma(\varepsilon)} \). It follows that our analysis of the first phase is valid for routing w.r.t. \( \tilde{\phi} \) as well and thus applies to \( \tilde{A} \).

For the second phase in \( V_2 \), we can apply the same arguments in order to generalize our analysis. For \( v \in V_2 \), we consider again \( V^+(v, \varepsilon/3) \) instead of \( V^+(v, \varepsilon) \). As long as all good events occur, the set \( V^-(v, \varepsilon) \) is again empty and every vertex \( u \in \Gamma(v) \cap V_1 \) has relaxed objective at most \( \tilde{\phi}(u) = O(\phi(v)^{(1-\epsilon_2/\gamma(\varepsilon))}). \)
However, by (5.19) every vertex \( u' \in \Gamma(v) \cap V^+(v, \varepsilon/3) \) with \( \phi(u') \leq \overline{c}w_t^{1-\delta} \) satisfies \( \tilde{\phi}(u') = \Omega(\phi(v)^{(1+\varepsilon_2/c)/\gamma(\varepsilon/3)}) \). Note that since \( c > 4 \) and \( \varepsilon_2 \leq \varepsilon \), we have \( \varepsilon_2/c \leq \varepsilon/3 \).

With a short calculation we deduce

\[
(1 + \varepsilon_2/c)(1 - \varepsilon) < (1 - \varepsilon_2/c)(1 - \varepsilon/3),
\]

which is equivalent to

\[
\frac{1 + \varepsilon_2/c}{\gamma(\varepsilon/3)} < \frac{1 - \varepsilon_2/c}{\gamma(\varepsilon)}.
\]

Indeed, \( \tilde{\phi}(u') > \tilde{\phi}(u) \) and the routing algorithm \( \tilde{A} \) favors vertices in \( V^+(v, \varepsilon) \) as desired.

Moreover, regarding statement (vi) of the main lemma, the previous arguments for \( V_1 \) and \( V_2 \) imply that every vertex \( v \in \Gamma(u_\ell) \cap V^+(u_\ell, \varepsilon/3) \) with \( \phi(v) \leq \overline{c}w_t^{1+\delta} \) satisfies \( \tilde{\phi}(v) > \tilde{\phi}(u_\ell) \). For every vertex \( v \in \Gamma(u_\ell) \cap V^+(u_\ell, \varepsilon/3) \) with \( \phi(v) > \overline{c}w_t^{1+\delta} \) the same is true by (5.20). It follows

\[
\{ v \in \Gamma(u_\ell) \cap V^+(u_\ell, \varepsilon/3) \cap V_{>\phi_0} | \tilde{\phi}(v) > \tilde{\phi}(u_\ell) \} = \Gamma(u_\ell) \cap V^+(u_\ell, \varepsilon/3) \cap V_{>\phi_0},
\]

and using \( M := \min\{w_0, \phi_0^{-1}\} \) we obtain

\[
(5.23) \quad \mathbb{E}_{>\phi_0} [||\{v \in \Gamma(u_\ell) \cap V^+(u_\ell, \varepsilon/3) \cap V_{>\phi_0} | \tilde{\phi}(v) > \tilde{\phi}(u_\ell)\}||] = \Omega(M^{\Omega(1)})
\]

for a suitable choice of \( \varepsilon \in \{\varepsilon_1, \varepsilon_2\} \). Note that the difference to part (vi) of Lemma 5.16 is the additional condition \( \tilde{\phi}(v) > \tilde{\phi}(u_\ell) \). In particular, if \( \tilde{A} \) visits \( u_\ell \) and the set described in (5.23) is non-empty, then \( \tilde{A} \) proceeds to a vertex in \( V_{>\phi_0} \) which has higher relaxed objective than \( u_\ell \) and every neighbor of \( u_\ell \) in \( V_{<\phi} \) by (5.20). We remark that it is not guaranteed that this next vertex lies in \( V^+(u_\ell, \varepsilon/3) \), because vertices with real objective at least \( \overline{c}w_t^{1+\delta} \) can have arbitrarily large relaxed objective, thus vertices in \( V^+(u_\ell, \varepsilon) \) don’t need to be the best neighbors of \( u_\ell \) w.r.t. \( \tilde{\phi} \).

**Start of routing process:** We can assume that all considered vertices are contained in the set \( V_{\leq\phi_0} \), as the analysis of the end phase will handle all other vertices. Let us start by considering the proof of Theorem 5.1. There, we used that if \( w_s < w_1(\varepsilon_1) \), then with constant probability \( s \) has a neighbor in a superior
set $A_s$. More precisely, $A_s$ contains vertices of weight at least $w_1(\varepsilon_1)$, and the proof required that with constant probability, $s$ has at least one neighbor $u_1 \in A_s$ and no neighbor with smaller weight than $w_1$ but higher objective than $u_1$. We observe that this argument can be transferred to the analysis $\tilde{A}$ as long as $\tilde{\phi}(u_1)$ is large enough and in particular larger than $\tilde{\phi}(s)$. However, due to (5.19) it holds $\tilde{\phi}(u_1) = \Omega(\phi(u_1)w_{u_1}^{-o(1)}) = \Omega(\phi(s))$. If we make the set $A_s$ slightly smaller and require that the weight of its vertices is by a constant factor larger than $w_1(\varepsilon_1)$, then still with constant probability there exists a vertex $u_1 \in \Gamma(s) \cap A_s$, but now we ensured the property $\tilde{\phi}(u_1) > \tilde{\phi}(s)$, so the starting argument applies as well for $\tilde{\phi}$.

The proof of Theorem 5.2 used Lemma 5.17 for the very first steps. There we constructed a set of layers $A_i$, defined via weights, and proved that with sufficiently high probability, the routing process traverses these layers and thereby the weight of the current vertex increases as desired. By the same arguments as seen above for the main lemma and the $V_1$-part, we can apply Lemma 5.17 for the relaxations.

When proving Theorem 5.3, we used Lemma 5.19 for the starting phase. The crucial argument behind Lemma 5.19 was that after visiting many vertices of weight at most $w_0(n) = \omega(1)$, the routing will arrive at a vertex $v$ which has a neighbor $u$ in the same cell of a $w_0$-grid such that $w_u \geq w_0^3$. We observe that due to (5.19), $u$ still has the property $\tilde{\phi}(u) \geq (1 - o(1))w_0^{2-o(1)}\phi(v)$. On the other hand, every neighbor $v'$ of $v$ which has weight less than $w_0$ has relaxed objective $\tilde{\phi}(v') \leq (1 + o(1))w_0^{1+o(1)}\phi(v)$. We see that since $w_0 = \omega(1)$, the algorithm $\tilde{A}$ still prefers $u$ to $v'$. Therefore Lemma 5.19 can be applied as well for studying the start of the routing algorithm $\tilde{A}$.

**End of routing process:** We have seen that as long as the visited vertices have real objective at most $\phi_0 \leq \omega_{t}^{-1+\lambda_2\delta}$, the routing w.r.t. $\tilde{\phi}$ has the same properties as the original routing. It remains to study the end phase where the vertices have higher objective. We start by considering the proof of Theorem 5.2 (ii). There, we used $\phi_0 = \frac{1}{c_1w_t}$. Since $w_t = \omega(1)$ by assumption, we have $w_t^{-1} = o(w_t^{-1+\lambda_2\delta})$. Hence, there is nothing to change. For adapting the proof of Theorem 5.2 (i),
we recall that we are assuming that $w_{\text{min}}$ is sufficiently large and thus $w_t$ as well. In particular, we can assume that $(c_1w_t)^{-1} \leq c w_t^{-1+\lambda_2 \delta} = c w_t^{-1+\delta/3}$. Then, we slightly adjust this proof by setting

$$
\phi_0 := \min\{(c_1w_t)^{-1}, \exp(-w_{\text{min}}^{c'})\},
$$

where $c' > 0$ is the same large constant used previously in the proof of Theorem 5.2 (i). We distinguish two cases: If $\phi_0 = (c_1w_t)^{-1}$, then our choice of $\phi_0$ is valid for applying the main lemma as well with relaxations. We also know by (5.23) that in this case the algorithm $\tilde{A}$ reaches a vertex $u$ satisfying $\phi(u) \geq (c_1w_t)^{-1}$. In this case, $p_{ut} = 1$, and the routing finds $u$. On the other hand, if $\phi_0 = \exp(-w_{\text{min}}^{c'})$, then it follows that $w_t = O(1)$. Here, we need the sequence of additional layers defined via $\phi$, as given in the proof of Theorem 5.2 (i), until we reach real objective $(c_1w_t)^{-1}$. Notice that for all these vertices in-between, in the additional layers, the relaxation (5.19) holds, but not (5.18). That is, every vertex $v$ with $\phi_0 \leq \phi(v) \leq \frac{1}{c_1w_t}$ satisfies $\tilde{\phi}(v) = \Theta(\phi(v))$. So the two objective functions $\phi$ and $\tilde{\phi}$ are very similar in this additional set of layers, all arguments can be transferred to $\tilde{\phi}$, and then Theorem 5.2 holds as well for the algorithm $\tilde{A}$.

For adapting the proof of Theorem 5.1, we assume again that $\zeta$ is chosen sufficiently small and $\overline{c}$ is chosen sufficiently large. We distinguish the following two cases. If $c_1\zeta w_t^{\lambda_2 \delta} \geq 1$, we choose $\phi_0 := (c_1w_t)^{-1}$. First, by assumption we have $\zeta w_t^{-1+\lambda_2 \delta} \geq (c_1w_t)^{-1}$, thus our choice is valid in terms of applying the main lemma with relaxations. Second, for $\zeta$ small enough it also follows that the error probability given by the main lemma is a sufficiently small constant. So our choice of $\phi_0$ has all desired properties, and by the same arguments as above (where we adjusted the proof of Theorem 5.2 (ii)) we see that once the routing reaches a vertex $u$ of real objective at least $\phi_0$, it directly connects to $t$ with constant probability, which is already sufficient for the theorem. On the other hand, if $c_1\zeta w_t^{\lambda_2 \delta} < 1$, then $w_t = O(1)$. In this case, we choose $\phi_0 = \Omega(1)$ small enough such that (i) $\phi_0$ is still a valid choice, i.e., $\phi_0 \leq \zeta w_t^{-1+\lambda_2 \delta}$, and such that (ii) the error probability when applying Lemma 5.16 is again small enough. Then we know that $\tilde{A}$ finds at least a vertex $u$ such that $\phi(u) \geq \phi_0 = \Omega(1)$. However, since $c_1\zeta w_t^{\lambda_2 \delta} < 1$ we have $(c_1\zeta)^{(1-\delta)/(\lambda_2 \delta)} \leq w_t^{-1+\delta}$ and it follows that all vertices
with real objective at most $\overline{c}(c_1 \xi)^{(1-\delta)/(\lambda_2 \delta)} \leq \overline{c}w_t^{-1+\delta}$ satisfy the relaxation (5.19) but not (5.18), which means that in this regime, $\phi$ and $\tilde{\phi}$ deviate by at most a bounded constant factor. Recall that in the proof of Theorem 5.1, we constructed a superior set $A_t$. By picking $\overline{c}$ large enough, with constant probability $A_t$ contains vertices that fall into the regime where (5.19) is satisfied. Then again, all arguments can be transferred to $\tilde{\phi}$ since $\phi$ and $\tilde{\phi}$ are almost equal in this regime.

For adapting the proof of Theorem 5.3, we slightly modify our choice of $\phi_0$ compared to the original proof and use $\phi_0 := \min\{(c_1 w_t)^{-1+\epsilon_2/c}, f_0^{-1+\epsilon_2/c}\} = o(1)$. Note that since $\epsilon_2 = o(1)$, we satisfy $\phi_0 \leq c w_t^{-1+\lambda_2 \delta}$ and thus the choice is valid. If the minimum is attained with the first term, then as soon as $\tilde{A}$ reaches a vertex of real objective at least $\phi_0$, it will never visit a vertex of real objective less than $(c_1 w_t)^{-1}$ later on, and then the argument from the proof can be adapted. On the other hand, if the minimum is attained with the second term, then again $\tilde{A}$ will stay at vertices of real objective at least $f_0^{-1}$, once it reached a vertex of real objective at least $\phi_0$. And then again the proof of Theorem 5.3 can be adapted in a straight-forward manner.

**Patching algorithms:** We first want to explain why the situation for patching algorithms is slightly different than for the basic routing algorithm, and why we can not transfer the proof of Theorem 5.4 in the case where $\tilde{\phi}$ only satisfies the weaker relaxation and $t$ can be an arbitrary vertex. Above, we have seen that if we only assume the weak relaxation (5.18), then we need to take $\phi_0 \leq \xi w_t^{-1+\lambda_2 \delta}$ when applying the main lemma. However, our proof of Theorem 5.4 requires that $\phi_0 = o(1)$ falls arbitrarily slow in $n$. We see that in a model where $w_t$ can be chosen arbitrarily large, our choice of $\phi_0$ is no longer valid when considering relaxations. Vice-versa, if $t$ is a random vertex, then a.a.s. the weight $w_t$ is small enough such that this conflict does not appear. Similarly, if we assume the stronger relaxation (5.17) for all vertices, then it is not difficult to observe that we get rid of the nasty restriction $\phi_0 \leq \xi w_t^{-1+\lambda_2 \delta}$ and our original choice of $\phi_0$ is also valid when considering the relaxations.

For the start, we analyzed patching algorithms by applying Lemma 5.19.
Above, we have already seen that this lemma can be transferred to relaxed objectives \( \tilde{\phi} \). For the main part of the routing, we argued before that under the additional assumptions, we can still apply the main lemma. Thus it remains to study the end of the routing process. There, we proved the existence of two specific vertices \( u_j \) and \( u_i \) on the greedy path. We required that \( \phi(u_j) \) and \( \phi(u_i) \) lie in a specific range. We observe that since \( u_j, u_i \in V_{\leq \phi_0} \), the stronger relaxation (5.17) holds in the considered regimes. Therefore, \( \tilde{\phi} \) deviates only by an \( o(1) \)-exponent from \( \phi \), and we don’t need to adjust the arguments. Finally, we completed the proof by finding a path from \( u_i \) to \( t \) inside a certain subset \( B \). Clearly, this path is still present in the graph, and again the relaxations ensure that \( B \subset V_{> \phi(u_j)} \). This allows us to transfer the analysis of patching algorithms to relaxed objective functions \( \tilde{\phi} \).

5.11 Greedy Routing in Hyperbolic Random Graphs

In this chapter we show how all result about greedy routing in GIRGs can be transferred to hyperbolic random graphs, by using the relaxations provided by Theorem 5.5 and Proposition 5.22. Recall the definition of hyperbolic random graphs (Definition 3.13) and of our mapping \( g \) between one-dimensional GIRGs and hyperbolic random graphs. In particular, a vertex with coordinates \((r_v, \nu_v)\) on the hyperbolic disk is mapped to a vertex with weight \( w_v = ne^{-r_v/2} \) and position \( x_v = \frac{\nu_v}{2\pi} \). Greedy routing in hyperbolic random graphs is completely geometric: a vertex \( v \) that receives the packet sends it to the neighbor \( u \in \Gamma(v) \) which minimizes the (hyperbolic) distance \( d_H(u, t) \) (see (3.7) on page 56). Notice that minimizing \( d_H(u, t) \) is equivalent to maximizing \( (\cosh(d_H(u, t)))^{-1/2} \), as the function \( \cosh \) is monotone increasing on positive values, and we are free to multiply this expression by any terms which are independent of \( u \).

Let us assume that \( d = 1 \) and that a target vertex \( t \) is given. Then on such a GIRG, we can define the objective function

\[
\phi_H(v) := \frac{n}{w_tw_{\min}\sqrt{\cosh(d_H(g(v), g(t)))}},
\]

where \( g(v) \) and \( g(t) \) are the vertices on the hyperbolic disk that correspond to
Suppose that the one-dimensional GIRG was obtained by embedding a hyperbolic random graph, as described above. Then, maximizing $\phi_H$ on the GIRG is the same as minimizing the hyperbolic distance to the target on the original hyperbolic graph. Thus routing w.r.t. $\phi_H$ on the GIRG is equivalent to geometric greedy routing in hyperbolic random graphs. The following lemma states that for $d = 1$, with high probability, for most vertices the objective function $\phi_H$ differs by at most a constant factor from the original objective function $\phi$.

**Lemma 5.23.** Let $d = 1$ and let $\delta > 0$ be a sufficiently small constant. Then with probability $1 - n^{-\Omega(1)}$, for every vertex $v$ with $\phi(v) \leq O(w_t^{-1+\delta})$ it holds

$$\phi_H(v) = \Theta(\phi(v)),$$

and for every vertex $v$ with $\phi(v) = \Omega(w_t^{-1+\delta})$ it holds

$$\phi_H(v) = \Omega(w_t^{-1+\delta}).$$

The statement of this lemma directly implies Corollary 5.6.

**Proof of Corollary 5.6.** By Lemma 5.23, with probability $1 - n^{-\Omega(1)}$ the objective function $\phi_H$ falls into the general class of objective functions considered in Proposition 5.22. Note that the small error probability $n^{-\Omega(1)}$ coming from Lemma 5.23 is negligible compared to the error probabilities of all previous theorems. Then the statement of the corollary follows directly from Proposition 5.22, that is, our results transfer from GIRGs to hyperbolic random graphs. □

**Proof of Lemma 5.23.** Let $\delta > 0$ and put $\phi_0 := O(w_t^{-1+\delta})$. For the proof of this lemma, we assume that there exists no vertex with weight larger than $n^{1-2\delta}$, which happens with probability $1 - n^{-\Omega(1)}$ if $\delta$ is chosen small enough. Let $v \in V$ such that $\phi(v) \leq \phi_0$, and denote by $(r_v, v_v) := g(w_v, x_v)$ and $(r_t, v_t) := g(w_t, x_t)$ the mappings of $v$ and $t$ onto the hyperbolic disk, and assume without loss of generality that $v_t = 0$ and $v_v \leq \pi$. We claim that on the hyperbolic disk, the assumption $\phi(v) \leq \phi_0$ implies the geometric property

(5.24) $$v_v \geq e^{-\min\{r_v, r_t\}}.$$
We first show $\nu_v \geq e^{-r_v}$. By assumption $\phi(v) = O(1)$, and then for $n$ large enough our mapping implies

$$
\nu_v = 2\pi \|x_v - x_t\| = \frac{2\pi w_v}{w_{\min} n \phi(v)} \geq w_v n^{-1-2\delta}.
$$

Furthermore $w_v \leq n^{1-2\delta}$ as there are no vertices of higher weight, and thus

$$
\nu_v \geq w_v n^{-1-2\delta} \geq w_v^2 n^{-2} = e^{-r_v}.
$$

Next, we observe that because there are no vertices of weight higher than $n^{1-2\delta}$, we have $w_v^{1+\delta} \leq n^{(1+\delta)(1-2\delta)} = o(n)$ and thus $w_v^{1-\delta} n^{-1} = o(w_v^2 n^{-2})$. On the other hand we are assuming $\phi(v) = O(w_v^{-1+\delta})$. Then for $n$ large enough, we obtain

$$
\nu_v = \frac{2\pi w_v}{w_{\min} n \phi(v)} \geq \frac{2\pi}{n \phi(v)} \geq \frac{1}{n \phi(v)} \geq w_v^2 n^{-2} = e^{-r_t}
$$

which proves (5.24).

In the following we estimate the hyperbolic cosine in terms of $w_v$, $w_t$ and $\|x_v - x_t\|$. We apply the definition of $d_H$ (see equation (3.7)) and the identity

$$
cosh(x - y) = \cosh(x) \cosh(y) - \sinh(x) \sinh(y),
$$

which yields

$$
cosh(d_H(g(v), g(t))) = \cosh(r_v) \cosh(r_t) - \sinh(r_v) \sinh(r_t) \cos(\nu_v)
$$

$$
= \cosh(r_v - r_t) + (1 - \cos(\nu_v)) \sinh(r_v) \sinh(r_t).
$$

Our assumption $w_v \leq n^{1-2\delta}$ is equivalent to $r_v \geq 2\delta \log n$ on the hyperbolic geometry, therefore we get $\sinh(r_v) = (1 - o(1)) e^{r_v / 2}$. Clearly, the same estimate is true for $\sinh(r_t)$. Then

$$
cosh(d_H(g(v), g(t))) = \cosh(r_v - r_t) + (1 - \cos(\nu_v))(1 - o(1)) e^{r_v + r_t / 4}
$$

$$
= \cosh(r_v - r_t) + \Theta(v_v^2 e^{r_v + r_t}),
$$

where we applied the Taylor approximation $1 - \cos(\nu_v) = \nu_v^2 / 2 - \nu_v^4 / 24 + O(\nu_v^6)$. Finally, equation (5.24) implies $\cosh(r_v - r_t) = O(v_v^2 e^{r_v + r_t})$. It follows (5.25)

$$
\phi_H(v) = \Theta\left(\frac{e^{r_t / 2}}{w_{\min} v_v e^{(r_v + r_t) / 2}}\right) = \Theta\left(\frac{ne^{-r_v / 2}}{w_{\min} n \nu_v}\right) = \Theta\left(\frac{w_v}{w_{\min} n \|x_v - x_t\| d}\right) = \Theta(\phi(v)).
$$
This proves the first part of the lemma. For the second part, let \( v \) be a vertex such that \( \phi(v) \geq \phi_0 \). We try to add an artificial vertex \( u \) of weight \( w_u = w_v \) in the graph such that \( \phi(u) = \phi_0 \). If this is not possible because \( w_v \) is too large, then we place \( u \) such that \( \|x_u - x_t\| \) is maximized. In both cases \( u \) has larger geometric distance to \( t \) than \( v \). If \( \phi(u) = \phi_0 \), we know by the first part of the proof that \( \phi_H(u) = \Theta(\phi_0) \). In the second case, \( \nu_u = \Omega(1) \) and therefore it holds

\[
\cosh(d_H(g(u), g(t))) = \cosh(r_{u} - r_{t}) + \Theta(\nu_{u}^{2}e^{r_{u}+r_{t}}) = \Theta(\nu_{u}^{2}e^{r_{u}+r_{t}}),
\]

and by (5.25) we obtain \( \phi_H(u) = \Theta(\phi(u)) = \Omega(\phi_0) \). Considering \( g(v) \) and \( g(u) \), we see that \( \nu_u \geq \nu_v \). Since in the interval \( \nu \in [0, \pi] \), the hyperbolic distance in monotone decreasing in \( \nu \), it follows

\[
d_H(g(v), g(t)) \leq d_H(g(u), g(t))
\]

and thus in both cases it holds \( \phi_H(v) \geq \phi_H(u) = \Theta(\phi_0) \).

\[\blacksquare\]

### 5.12 Proofs of Auxiliary Lemmas of Section 5.7.2

In this section we proof the four auxiliary lemmas that have been stated above in Section 5.7.2.

**Proof of Lemma 5.12.** Let \( 0 < \epsilon \leq \epsilon_1 \) be such that \( \phi(v)w_v^{\gamma(\epsilon)} \leq 1 \). Let \( v \in V_1 \).

**Proof of (i):** In order to prove the first statement, we want to lower-bound the expected number of neighbors of \( v \) in \( V^+(v, \epsilon) \), i.e., the neighbors with larger weight and slightly larger objective. For this, define \( \delta := \left( \frac{c_1w_v^{1+\gamma(\epsilon)}}{w_{min}^{n}} \right)^{1/d} \), where \( c_1 \) is the constant given by (5.1) and (5.2). Then let \( A(v, \epsilon) \) be the following set of vertices:

\[
A(v, \epsilon) := \{u \in V \mid (1) : w_u \geq w_v^{\gamma(\epsilon)}; (2) : \|x_u - x_t\| \leq \|x_v - x_t\|; (3) : \|x_u - x_v\| \leq \delta \}.
\]

Note that \( A(v, \epsilon) \subseteq V^+(v, \epsilon) \), because for \( u \in A(v, \epsilon) \)

\[
\phi(u) = \frac{w_u}{w_{min}n\|x_u - x_t\|^d} \geq \frac{w_v^{\gamma(\epsilon)}}{w_{min}n\|x_u - x_t\|^d} \geq \phi(v)w_v^{\gamma(\epsilon)-1}
\]
holds. Hence for proving the first statement it is sufficient to lower-bound $\mathbb{E}[|\Gamma(v) \cap A(v, \varepsilon)|]$. By the probability distribution of the weights there are in expectation $n w_{\min}^{-1} W_v^\gamma(1-\beta)$ vertices satisfying Condition (1) and Condition (3) is satisfied independently with probability $\delta^d$. Now, observe that by assumption the vertex $v$ satisfies $\phi(v) W_v^\gamma \leq 1$, which implies

$$\delta^d = \frac{c_1 W_v^{1+\gamma} \|x_v - x_t\|^d}{w_{\min} n} = c_1 \phi(v) W_v^\gamma \|x_v - x_t\|^d \leq c_1 \|x_v - x_t\|^d.$$ 

Hence a random vertex $v$ fulfilling (3) satisfies (2) too with constant probability. Furthermore, Condition (3) ensures that by (5.1) and (5.2), every vertex $u \in A(v, \varepsilon)$ is connected to $v$ with constant probability. Combining this yields

$$\mathbb{E}[|\Gamma(v) \cap A(v, \varepsilon)|] = \Omega\left(n w_{\min}^{-1} W_v^\gamma(1-\beta) \delta^d\right) = \Omega\left(W_v^{\beta-2} w_{\min}^{-1} W_v^\gamma(\beta-2)\right) = \Omega(W_v^{\beta-2} \varepsilon).$$

**Proof of (ii):** We want to upper-bound the expected number of bad neighbors of $v$, i.e., the neighbors in $V^-(v, \varepsilon)$. First, we will show that the geometric distance between a bad node $u \in V^-(v, \varepsilon)$ and $v$ is not much less than the geometric distance between $v$ and $t$. This will imply a sufficiently small connection probability between $u$ and $v$. For this purpose, define $\delta(w) := \|x_v - x_t\|(w W_v^{-\gamma})^{1/d}$ for all weights $w \leq W_v^{\gamma(c \varepsilon)}$. Due to $w_v \geq w_1(\varepsilon) = 2^{O(d/\varepsilon)}$, for $w_1(\varepsilon)$ large enough every weight $w \leq W_v^{\gamma(c \varepsilon)}$ satisfies $\delta(w)^d \leq \|x_v - x_t\|^d W_v^{\gamma(c \varepsilon)-\gamma(\varepsilon)} \leq (0.5 \|x_v - x_t\|^d)$ (the second inequality follows since $c > 1$). Moreover, for a vertex $u \in V^-(v, \varepsilon)$ we observe that the condition $\phi(u) \geq \phi(v) W_v^{\gamma(e)-1}$ implies $\|x_u - x_t\| \leq \delta(w_u)$ and with the triangle inequality, we obtain that $u$ fulfills $\|x_u - x_v\| \geq 0.5 \|x_v - x_t\|$. In order to upper-bound $\mathbb{E}[|\Gamma(v) \cap V^-(v, \varepsilon)|]$, we proceed with a case distinction regarding the parameter $\alpha$.

**Case $\alpha = \infty$:** Let $c_2 > 0$ be the constant given by (5.2), and let $w_1(\varepsilon)$ be large enough. Then

$$0.5^d \geq c_2 w_1(\varepsilon)^{-\varepsilon(\xi-1)/(\beta-2)} \geq c_2 w_v^{-\varepsilon(\xi-1)/(\beta-2)} = c_2 W_v^{\gamma(c \varepsilon)-\gamma(\varepsilon)}.$$ 

Furthermore, $\phi(v) W_v^{\gamma(\varepsilon)} \leq 1$ implies

$$\frac{W_v^{1+\gamma(\varepsilon)}}{w_{\min} n} = \phi(v) W_v^{\gamma(\varepsilon)} \|x_v - x_t\|^d \leq \|x_v - x_t\|^d.$$
Together with the above observation we deduce
\[\|x_u - x_v\| \geq 0.5d \|x_v - x_t\| \geq c_2 \|x_v - x_t\| d \|w^{\gamma(\xi) - \gamma(\epsilon)}_v\| \geq \frac{c_2 w_v^{1 + \gamma(\xi) \epsilon}}{w_{\min} n} \geq \frac{c_2 w_v w_t}{w_{\min} n}.\]

But then, by (5.2) the two vertices \(u\) and \(v\) can not be connected, hence \(v\) has deterministically no neighbor in \(V^-(v, \epsilon)\) in the case \(\alpha = \infty\).

**Case \(\alpha < \infty\):** By the geometrical property \(\|x_u - x_v\| \geq 0.5 \|x_v - x_t\|\) and by (5.1) the probability that a vertex \(u \in V^-(v, \epsilon)\) is connected to \(v\) is at most \(O\left(\left(\frac{w_u w_v}{\|x_u - x_v\| d w_{\min} n}\right)^{\alpha}\right) = O\left(\left(\frac{w_u w_v}{\|x_v - x_t\| d w_{\min} n}\right)^{\alpha}\right) = O((w_u \phi(v))^{\alpha})\).

Furthermore, we have already seen that such a vertex of weight \(w\) needs to be in distance at most \(\delta(w)\) to \(t\), which a random vertex of weight \(w\) does with probability \(\delta(w)^d\). We integrate over all weights between \(w_{\min}\) and \(w^{\gamma(\xi) \epsilon}_v\), where we use the density function \(f(w)\) for the distribution of the weights. We deduce that
\[
E[|\Gamma(v) \cap V^-(v, \epsilon)|] = O\left(n \int_{w_{\min}} w^{\gamma(\xi) \epsilon}_v f(w) \delta(w)^d (w_u \phi(v))^{\alpha} dw\right)
= O\left(w_{\min}^{\beta - 2} \phi(v)^{\alpha - 1} w^{1 - \gamma(\epsilon)}_v \int_{w_{\min}} w^{\alpha + 1 - \beta} dw\right)
= O\left(w_{\min}^{\beta - 2} \phi(v)^{\alpha - 1} w^{1 - \gamma(\epsilon) + \gamma(\xi) \epsilon (\alpha + 2 - \beta)}\right).
\]

Next we use that \(\phi(v) w^{\gamma(\epsilon)}_v \leq 1\) holds. Then the above term is at most \(O(w_{\min}^{\beta - 2} w^{\tau}_v),\) where
\[
\tau = 1 - \alpha \gamma(\epsilon) + \gamma(\xi \cdot \epsilon)(\alpha + 2 - \beta) = \frac{\alpha \epsilon - \xi \epsilon (\alpha + 2 - \beta)}{\beta - 2} = -\Omega(\epsilon)
\]
by our choice of \(\zeta\).

**Proof of Lemma 5.13.** We prove the statement in a similar way as we proved Lemma 5.12. Let \(0 < \epsilon \leq \epsilon_1\).

**Proof of (i):** Let \(v \in V_2\) such that \(\phi(v) \leq 1\). In order to prove the first statement, we want to lower-bound the expected number of neighbors of \(v\) in the set

\[\]
$V^+(v, \varepsilon)$. In particular, such neighbors have objective at least $\phi(v)^{1/\gamma(\varepsilon)}$. Let $\delta := (w_{\min}^{-1} n^{-1} \phi(v)^{-1 - 1/\gamma(\varepsilon)})^{1/d}$ and let $c' = \min\{1, c_1/2^d\}$, where $c_1$ is the constant given by (5.1) and (5.2). We consider the set

$$A(v, \varepsilon) := \{ u \in V \mid \|x_u - x_\ell\| \leq \delta; (2) : w_u \geq (c' \phi(v))^{-1} \geq \phi(v)^{-1}\}.$$  

By the definition of $\delta$ and Condition (2), a vertex $u \in A(v, \varepsilon)$ satisfies

$$\phi(u) \geq \frac{w_u}{\delta d w_{\min}} \geq w_u \phi(v)^{1 + 1/\gamma(\varepsilon)} \geq \phi(v)^{1/\gamma(\varepsilon)}.$$  

On the other hand, by (2) and the assumption $\phi(v) \leq 1$ it holds $w_u \geq 1$, and then we have $w_u \phi(v)^{1 + 1/\gamma(\varepsilon)} \geq w_u^{1/\gamma(\varepsilon)} \geq w_u^{-\gamma(\varepsilon)}$. Thus, $A(v, \varepsilon) \subseteq V^+(v, \varepsilon)$.

Our next goal is to upper-bound the geometric distance between vertices of $A(v, \varepsilon)$ and the vertex $v$ itself. We observe that the assumptions $v \in V_2$ and $\varepsilon \leq 1$ imply $\phi(v)^{-1/\gamma(\varepsilon)} \leq w_v^{\gamma(\varepsilon)1/\gamma(\varepsilon)} \leq w_v$, and it follows

$$nw_{\min}\delta^d = \phi(v)^{-1 - 1/\gamma(\varepsilon)} \leq \phi(v)^{-1} w_v = nw_{\min}\|x_u - x_\ell\|^d.$$  

By the triangle inequality we obtain $\|x_u - x_v\| \leq 2\|x_v - x_\ell\|$ for all $u \in A(v, \varepsilon)$.

Next we want to verify that all vertices in $A(v, \varepsilon)$ are connected to $v$ with constant probability. We notice that for all $u \in A(v, \varepsilon)$, by Condition (2) we have

$$\frac{c_1 w_u w_v}{w_{\min} n} = c_1 w_u \phi(v) \|x_v - x_\ell\|^d \geq \frac{c_1}{c'} \|x_v - x_\ell\|^d \geq \frac{c_1}{2^d c'} \|x_u - x_v\|^d \geq \|x_u - x_v\|^d.$$  

by our choice of $c'$. Then, indeed by (5.1) and (5.2), all vertices of $A(v, \varepsilon)$ are connected to $v$ with constant probability. The expected number of vertices with weight at least $(c' \phi(v))^{-1}$ is $\Omega(n w_{\min}(\phi(v))^{\beta - 1})$, and thus

$$\mathbb{E}[|\Gamma(v) \cap A(v, \varepsilon)|] = \Omega\left(n w_{\min}(\phi(v))^{\beta - 1} \delta^d\right) = \Omega\left(w_{\min}^{\beta - 2}(\phi(v))^{\beta - 2 - 1/\gamma(\varepsilon)}\right) = \Omega\left(w_{\min}^{\beta - 2}(\phi(v))^{-\Omega(\varepsilon)}\right),$$  

as $\varepsilon$ is chosen small enough.

**Proof of (ii):** Let $v \in V_2$ such that $\phi(v) \leq \phi_1$ and let $u \in V^-(v, \varepsilon)$. We first prove that $u$ has the property $\|x_u - x_v\| \geq 0.5\|x_v - x_\ell\|$. Since $u \in V^-(v, \varepsilon)$, we have $\phi(u) \geq \phi(v)^{1/\gamma(\varepsilon)}$, and by the assumption $\phi(v) \leq \phi_1$ for $\varepsilon_1$ small enough this is at
least $\phi(v)^{1-\epsilon} \geq 2^d \phi(v)$, where we used $\phi(v) \leq \phi_1$ in the last step. Because $u \in V_1$, it follows

$$
x_u - x_f \parallel^d = \phi(u)^{-1} \frac{w_u}{w_{\min}} n \leq \frac{\phi(u)^{-1-\frac{1}{\gamma(\epsilon)}}}{w_{\min}} \leq \frac{(2^d \phi(v))^{-1-\frac{1}{\gamma(\epsilon)}}}{w_{\min}}.
$$

On the other hand, $v \in V_2$, and then

$$
x_u - x_f \parallel^d = \frac{(2^d \phi(v))^{-1-\frac{1}{\gamma(\epsilon)}}}{w_{\min}} \leq \frac{0.5^d \phi(v)w_v}{w_{\min}} = 0.5^d \|x_v - x_f\|^d.
$$

By the triangle inequality $\|x_u - x_v\| \geq 0.5\|x_v - x_f\|$ holds as desired. Moreover, $u \in V_1$ and thus $w_u \leq \phi(u)^{-1/\gamma(\epsilon)}$. Since $\phi(u) \geq \phi(v)$, for $\epsilon_1$ small enough it follows

$$
w_u \phi(v) \leq \phi(u)^{-1/\gamma(\epsilon)} \phi(v) \leq \phi(v)^{1-1/\gamma(\epsilon)} \leq \phi(v)^\epsilon.
$$

**Case $\alpha = \infty$:** We obtain

$$
\frac{w_u w_v}{\|x_u - x_v\|^d} \leq \frac{w_u w_v}{0.5^d \|x_v - x_f\|^d} = 2^d w_u \phi(v) w_{\min} n \leq 2^d \phi(v)^\Omega(\epsilon) w_{\min} n.
$$

Together with the assumption $\phi(v) \leq \phi_1$ this implies $\frac{c_2 w_u w_v}{\|x_u - x_v\|^d w_{\min}} \leq c_2 2^d \phi_1^{\Omega(\epsilon)}$, and by taking $\phi_1$ small enough, this expression is smaller than 1. Then (5.2) yields $u \notin \Gamma(v)$ deterministically.

**Case $\alpha < \infty$:** We have

$$
p_{uv} = O\left(\left(\frac{w_u w_v}{\|x_u - x_v\|^d w_{\min}}\right)^\alpha\right) = O\left(\left(\frac{w_u w_v}{\|x_v - x_f\|^d w_{\min}}\right)^\alpha\right) = O(\phi(u) w_{\min})^\alpha.
$$

Recall that we observed already above that $w_u \leq \phi(v)^{-1/\gamma(\epsilon)}$, and note that every vertex $u \in V^-(v, \epsilon)$ has the geometric property $\|x_u - x_f\|^d \leq \phi(v)^{-1/\gamma(\epsilon)} w_u w_{\min}$, which a random vertex of given weight $w$ satisfies with probability at most $\phi(v)^{-1/\gamma(\epsilon)} \frac{w}{w_{\min}}$. We conclude that

$$
\mathbb{E}[|\Gamma(v) \cap V^-(v)|] = O\left(\int_{\phi(v)^{-1/\gamma(\epsilon)}} \frac{nw_{\min}^{-\frac{1}{\gamma(\epsilon)}}}{\phi(v)^{-1/\gamma(\epsilon)}} w^{-\beta} \phi(v)^{-1/\gamma(\epsilon)} \frac{w}{w_{\min}} (w\phi(v))^\alpha d\phi(v)\right)
$$

$$
= O\left(\phi(v)^{-1/\gamma(\epsilon)} + \alpha \int_{\phi(v)^{-1/\gamma(\epsilon)}} w^{\alpha - 1/\gamma(\epsilon)} d\phi(v)\right)
$$

$$
= O\left(\phi(v)^{-1/\gamma(\epsilon)} \frac{\tau}{\alpha - 1/\gamma(\epsilon)}\right),
$$

where $\tau = (\alpha + 2 - \beta)(1 - 1/\gamma(\epsilon)) - \alpha \epsilon = \Omega(\epsilon)$ for $\epsilon \leq \epsilon_1$ and $\epsilon_1$ small enough. This proves (ii).
Proof of Lemma 5.14. Let $0 < \varepsilon < \varepsilon_1$ and let $v \in V$ be a vertex such that $\phi(v)w_v^{\gamma(\varepsilon-v)} \leq \bar{c} := 1/(2^d c_2)$, where $c_2$ is the constant given by (5.2). We want to show that with relatively high probability, every neighbor of $v$ with small weight is located geometrically close to $v$. For all weights $w \leq w_v^{\gamma(\varepsilon-v)}$ we define a critical radius

$$
(5.26) \quad r_0(w) := 0.5 \left( \frac{w_v^{1-\gamma(\varepsilon-v)} \omega}{w_{\min} n \phi(v)} \right)^{1/d}.
$$

Next we define

$$
B(v, \varepsilon) := \{ u \in V \mid (1) : w_u \leq w_v^{\gamma(\varepsilon-v)}; (2) : \|x_v - x_u\| \geq r_0(w_u) \}.
$$

We first show that the expected size of the set $\Gamma(v) \cap B(v, \varepsilon)$ is very small. Afterwards, we will prove that $V^-(v, \varepsilon) \subseteq B(v, \varepsilon)$ holds under the assumptions of the lemma.

As we have $\phi(v) \leq \bar{c}w_v^{\gamma(\varepsilon-v)}$ by assumption, $r_0(w)^d \geq \frac{w_v^{\omega w_v}}{2^d \bar{c} w_{\min} n} w_v^{w_v^{\omega w_v}}$ holds for all weights $w \leq w_v^{\gamma(\varepsilon-v)}$. Then in the case $\alpha = \infty$, by (5.2) the vertex $v$ can not have any neighbors in $B(v, \varepsilon)$. If $\alpha < \infty$, then by (5.1) a vertex $u \in B(v)$ with distance $r$ to $v$ is connected to $v$ with probability $O((\frac{w_v^{\omega w_v}}{r^d w_{\min} n})^\alpha)$. We calculate $\mathbb{E}[|\Gamma(v) \cap B(v, \varepsilon)|]$ by integrating over all weights smaller than $w_v^{\gamma(\varepsilon-v)}$ and all distances $r \geq r_0(w)$. Using the density function $f(w) = \Theta(w_v^{\beta-1} w^{-\beta})$ we deduce

$$
\mathbb{E}[|\Gamma(v) \cap B(v, \varepsilon)|] = O \left( \int_{w_{\min}} w_v^{\gamma(\varepsilon-v)} \, \frac{nw_{\min}^{-\beta} w^{-\beta}}{\int_{0}^{1} r_0(w)^{d(1-\alpha)} w^{-\beta} dr dw} \right)
= O \left( \int_{w_{\min}} w_v^{\gamma(\varepsilon-v)} \, \left( \frac{w_v w}{w_{\min} n} \right)^\alpha \int_{0}^{1} r_0(w)^d w^{-\beta} dr dw \right)
= O \left( \int_{w_{\min}} w_v^{\gamma(\varepsilon-v)} \, \frac{w_v w}{w_{\min} n} \phi(v)^{\alpha-1} w^{-\beta} dw \right)
= O \left( \int_{w_{\min}} w_v^{\gamma(\varepsilon-v)(\alpha-1)+1} \phi(v)^{\alpha-1} w^{-\beta} dw \right)
= O \left( \int_{w_{\min}} w_v^{\gamma(\varepsilon-v)(\alpha-1)+1} \phi(v)^{\alpha-1} \right).
$$

We finish the proof of the lemma by showing that $V^-(v, \varepsilon) \subseteq B(v, \varepsilon)$ holds under the assumption that $v$ satisfies $w_v \geq w_1(\varepsilon)$. By assumption, $v \in V_1$ and
thus by definition of the set $V^-(v, \varepsilon)$, all $u \in V^-(v, \varepsilon)$ fulfill $w_u \leq w_v^{\gamma(\varepsilon)}$. Next we observe that again by the definition of the set $V^-(v, \varepsilon)$, all $u \in V^-(v, \varepsilon)$ satisfy

$$\|x_u - x_t\|^d = \frac{w_u}{nw_{\min}\phi(u)} \leq \frac{w_v^{\gamma(\varepsilon)}}{nw_{\min}\phi(v)w_v^{\gamma(\varepsilon)-1}} = \|x_v - x_t\|^d w_v^{-\Omega(\varepsilon)}.$$ 

Hence $\|x_u - x_t\| \leq \|x_v - x_t\| w_v^{-\Omega(\varepsilon)/d} \leq 0.5\|x_v - x_t\|$ as we assume $w_v \geq w_1(\varepsilon)$. By the triangle inequality we deduce $\|x_v - x_u\| \geq 0.5\|x_v - x_t\|$. On the other side, since $w_u \leq w_v^{\gamma(\varepsilon)}$, the definition of $r_0(w)$ implies

$$r_0(w_u)^d = 0.5d \frac{w_v^{-\gamma(\varepsilon)}}{w_{\min}n\phi(v)} \leq 0.5d \|x_v - x_t\|^d.$$ 

It follows $\|x_v - x_u\| \geq 0.5\|x_v - x_t\| \geq r_0(w_u)$, which proves $V^-(v, \varepsilon) \subseteq B(v, \varepsilon)$.

**Proof of Lemma 5.15.** Let $c > 0$ be a constant chosen sufficiently large, $\varepsilon$ be a constant chosen sufficiently small, let $r_0 := (w_0^c/n)^{1/d}$ and for every weight $w \geq w_{\min}$ define $\delta(w) := (\frac{ww_0^c}{w_{\min}n})^{1/d}$.

**Proof of (i):** We first observe that a vertex $u$ of weight $w_u$ satisfies $\phi(u) \geq w_0^{-\varepsilon}$ if and only if $\|x_u - x_t\|^d \leq \delta(w_u)^d$. Then excluding $t$ which is potentially fixed, the expected number of vertices with weight at least $w_0$ and objective at least $w_0^{-\varepsilon}$ is at most

$$\int_{w_0}^{\infty} nw_{\min}^{\beta-1} w^{-\beta} \delta(w)^d d w = \int_{w_0}^{\infty} nw_{\min}^{\beta-2} w_0^\varepsilon w^{-1-\beta} d w = O(w_{\min}^{\beta-2} w_0^{2-\beta+\varepsilon}) = w_0^{-\Omega(1)},$$

where we used that for $\varepsilon$ sufficiently small it holds $\beta - 2 \geq 2\varepsilon$. By Markov’s inequality (Theorem 2.1) with probability $1 - w_0^{-\Omega(1)}$ there exist no such vertices.

**Proof of (ii):** Let $v$ be a vertex of weight at most $w_0$ and define

$$A_0(v) := \{u \in V \mid w_u \leq w_0 \text{ and } \|x_u - x_v\| \geq r_0\}.$$ 

For $c$ sufficiently large we have $\frac{w_u w_v}{w_{\min}n} \leq \frac{w_0^2}{w_{\min}n} = o(r_0^d)$ for any vertex $u \in A_0(v)$.

In the case $\alpha = \infty$, the claim follows directly as by (5.2) $u$ and $v$ can not share
an edge, i.e., $\Gamma(v) \cap A_0(v) = \emptyset$. So assume $\alpha < \infty$. By (5.1) and the assumption $w_v \leq w_0$, the probability $p_{uv}$ is at most $O\left(\left(\frac{w_0^2}{n}\right)^{\alpha}\right)$. Thus we have

$$
\mathbb{E}[|\Gamma(v) \cap A_0(v)|] = O\left(n \int_{r_0}^{1/2} r^{d-1} \left(\frac{w_0^2}{nr^d}\right)^{\alpha} dr\right) = O\left(w_0^{2\alpha n^{1-\alpha} r_0^{d(1-\alpha)}}\right)
$$

If we take $c \geq \frac{K+2\alpha}{\alpha-1}$, then we obtain $\mathbb{E}[|\Gamma(v) \cap A_0(v)|] = O(w_0^{-K})$, and by Markov’s inequality with probability at least $1 - O(w_0^{-K})$, the set $\Gamma(v) \cap A_0(v)$ is empty.

**Proof of (iii):** Let $v$ be a vertex of weight at most $w_0$ which satisfies $\phi(v) \leq w_0^{-\epsilon}$, and define

$$A_1(v) := \{u \in V \mid (1) : w_u \geq w_0; (2) : \phi(u) \leq \phi(v)^\epsilon\}.$$  

We want to show that a.a.s. the vertex $v$ has no neighbor in the set $A_1(v)$. Define $\delta'(w) := \left(\frac{w}{\phi(v)^\epsilon w_{\text{min}}/n}\right)^{1/d}$ and observe that a vertex $u \in V$ with $w_u \geq w_0$ is contained in $A_1(v)$ if and only if $\delta'(w_u) \leq \|x_u - x_v\|$. However, due to $w_v \leq w_0$, for $c > 1/\epsilon$ we have $\delta'(w_u)^\alpha = \omega(\frac{w_u w}{w_{\text{min}} n})$. Again, in the threshold case $\alpha = \infty$ the claim follows directly. So assume $\alpha < \infty$. We observe that for $w \geq w_0$ we have $\delta'(w) = \omega(\|x_v - x_v\|)$, and by the triangle inequality we deduce that for a vertex $u \in A_1(v)$ it holds $\|x_u - x_v\| \geq 0.5\|x_u - x_v\|$. After this preparation, we can calculate the expected number of vertices in $\Gamma(v) \cap A_1(v)$ by integrating over the complete set $A_1(v)$. We obtain

$$
\mathbb{E}[|\Gamma(v) \cap A_1(v)|] = O\left(\int_{w_0}^{\infty} n w_{\text{min}}^{\beta-1} w^{-\beta} \int_{\delta'(w)}^{1/2} r^{d-1} \left(\frac{w w_v}{(0.5r)^d w_{\text{min}} n}\right)^{\alpha} dr dw\right)
$$

$$
= O\left(\int_{w_0}^{\infty} n w_{\text{min}}^{\beta-1} w^{-\beta} \delta'(w)^{d(1-\alpha)} dw\right)
$$

$$
= O\left(w_{\text{min}}^{\beta-2} w^{\alpha} \phi(v)^{c(\alpha-1)} \int_{w_0}^{\infty} w^{1-\beta} dw\right)
$$

$$
= O\left(w_{\text{min}}^{\beta-2} w^{\alpha} \phi(v)^{c(\alpha-1)} w_0^{2-\beta}\right) = O\left(w_v^{\alpha} w_0^{c(\alpha-1)}\right).
$$

Since $w_v \leq w_0$ by assumption, we can choose $c$ large enough such that this value is at most $O(w_0^{-K})$. Then Markov’s inequality implies that indeed the set $\Gamma(v) \cap A_1(v)$ is empty with probability $1 - O(w_0^{-K})$. This proves statement (iii).
Proof of (iv): Let \( v \) be a vertex with \( w_v \geq w_0 \) and objective \( \phi(v) \leq w_0^{-1-\varepsilon} \), and define \( M := \min\{w_v, \phi(v)^{-1}\} \). We want to upper-bound the probability that \( v \) has a neighbor in

\[
A_2(v) := \{ u \in V \mid (1) : w_u \leq w_0; (2) : \phi(u) \geq w_0^{-\varepsilon}\}.
\]

We know that every vertex \( u \in A_2(v) \) satisfies

\[
\|x_u - x_t\|_d^d \leq \delta(w_u)^d = \frac{w_u w_0^\varepsilon}{w_{\min} n} = o\left(\frac{w_u w_v}{w_{\min} n}\right).
\]

On the other hand, the vertex \( v \) satisfies

\[
\|x_v - x_t\|_d^d = \frac{w_v}{\phi(v) w_{\min} n} \geq \frac{w_v w_0^{1+\varepsilon}}{w_{\min} n} = \omega\left(\frac{w_u w_v}{w_{\min} n}\right).
\]

By the triangle inequality it follows \( \|x_v - x_t\| = \omega\left(\frac{w_u w_v}{w_{\min} n}\right) \). Then for \( \alpha = \infty \) the claim follows by (5.2) because there exists deterministically no edge between \( u \) and \( v \), so we can assume \( \alpha < \infty \) where we have \( p_{uv} = O((\phi(v) w_u)^{\alpha}) \). By integrating over all weights \( w \in [w_{\min}, w_0] \) we obtain

\[
\mathbb{E}[|\Gamma(v) \cap A_2(v)|] = O\left(\int_{w_{\min}}^{w_0} n w_{\min}^{\beta-1} w^{-\beta} \delta(w)^d (\phi(v) w_u)^\alpha \, dw\right)
\]

\[
= O\left(\int_{w_{\min}}^{w_0} w_{\min}^{\beta-2} w_0^{\varepsilon} (\phi(v) w_0)^\alpha \, dw\right) = O\left(w_0^{\varepsilon} (\phi(v) w_0)^\alpha\right).
\]

We continue with a case distinction. First suppose \( \phi(v) \leq w_0^{-1-2\varepsilon} \). Since \( \alpha > 1 \), \( \phi(v) \leq w_0^{-1} \), and \( w_v \geq w_0 \), in this case the above expression is at most

\[
O\left(w_0^{1+\varepsilon} \phi(v)\right) = O\left(w_0^{1+\varepsilon} w_0^{-1-2\varepsilon}\right) = O\left(w_v^{-\varepsilon}\right) = O\left(M^{-\Omega(1)}\right).
\]

On the other hand, if \( \phi(v) \geq w_0^{-1-2\varepsilon} \) we use \( \phi(v) \leq w_0^{-1-\varepsilon} \) \( \Leftrightarrow w_0 \leq \phi(v)^{-1/(1+\varepsilon)} \) and get

\[
\mathbb{E}[|\Gamma(v) \cap A_2(v)|] = O\left(w_0^{\varepsilon+\alpha} \phi(v)^{\alpha}\right) = O\left(\phi(v)^{\alpha-\frac{\alpha+\varepsilon}{1+\varepsilon}}\right) = O\left(\phi(v)^{\frac{\alpha-\varepsilon}{1+\varepsilon}}\right) = O\left(M^{-\Omega(1)}\right).
\]

Then again by Markov's inequality, the set \( \Gamma(v) \cap A_2(v) \) is empty with probability

\[
1 - O(M^{-\Omega(1)}),
\]

which finishes the proof. \( \blacksquare \)
PART II

GRAPH COLORING GAMES
Chapter 6

The Game Chromatic Number of Random Graphs

We study the game chromatic number of the random graph model $G(n, p)$. This topic was initiated in 2008 by Bohman, Frieze, and Sudakov. In their paper they provided almost matching lower and upper bounds on $\chi_g(G(n, p))$ [BFS08] and conjectured that for all $p = \omega(n^{-1})$, the game chromatic number $\chi_g(G(n, p))$ is asymptotically twice as large as the ordinary chromatic number $\chi(G(n, p))$. In this chapter, we verify their conjecture for sufficiently dense random graphs. Concretely, we prove for all $p \geq e^{-o(\log n)}$ that

$$\chi_g(G_{n,p}) = (1 + o(1)) \frac{n}{\log_b(np)} = (2 + o(1)) \chi(G_{n,p}),$$

where $b = \frac{1}{1-p}$.

This chapter is based on joint work with Angelika Steger [KS14].

6.1 Introduction

Let $G$ be a graph with initially uncolored vertices and suppose there is a set of $k$ different colors. Maker and Breaker play the following game: they alternately...
take turns, in each move assigning one of the \( k \) colors to an uncolored vertex such that the coloring remains proper, i.e., two neighbors never receive the same color. Maker’s goal is to ensure that all vertices get colored. Vice-versa, Breaker wins if there exists an uncolored vertex where no color is available and thus the partial coloring cannot be extended to a coloring of the full vertex set. We assume that Maker has the first move. The \textit{game chromatic number} \( \chi_g(G) \) is defined as the smallest \( k \) for which Maker has a winning strategy, no matter how Breaker plays.

Clearly, \( \chi_g(G) \) is at least as large as the chromatic number \( \chi(G) \). On the other hand, Maker always wins the game if the number of colors is larger than the maximum degree of \( G \), because then no vertex can run out of colors. Therefore the parameter \( \chi_g(G) \) is well-defined. Usually, \( \chi_g(G) \) is larger than \( \chi(G) \). In fact, the difference between the two parameters can be as large as \( \Omega(|V|) \). Consider for example the complete bipartite graph \( B_{n,n} \) minus a perfect matching \( M \). The chromatic number of this graph is two, whereas Breaker has a winning strategy whenever the number of colors is less than \( n \): if Maker colors some vertex \( v \) with color \( i \), then Breaker uses the same color \( i \) on the vertex \( w \) which is matched to \( v \) in the matching \( M \). Color \( i \) can henceforth not be used on any other vertex, and the claim follows by induction.

This graph coloring game was first proposed in Martin Gardner’s column in the Scientific American in 1981, after an idea of Brams [Gar81]. The original motivation behind the game was the hope of finding a proof for the four color theorem of planar graphs that is \textit{not} computer-based. Indeed, suppose that by using only four colors, Maker has a winning strategy for every planar graph. Then the bound \( \chi(G) \leq \chi_g(G) \) would immediately imply that all planar graphs admit a proper coloring with four colors. Unfortunately, it turns out that this is not possible: Figure 6.1 provides an example of a planar graph with game chromatic number at least six, cf. [KT94].

The game remained unnoticed until Bodlaender re-invented it [Bod91]. Since then, it has been studied for many different graph classes and several variants have been invented. Regarding the original problem on planar graphs, Kierstead and Trotter [KT94] proved (1) that there are planar graphs with game
FIGURE 6.1. A planar graph $G$ on 12 vertices. Note that each of the six pairs $\{a_j, b_j\}$ forms a dominating set. Whenever Maker colors a vertex of a pair $\{a_j, b_j\}$ with a color $i$, Breaker can answer by using $i$ on the other vertex of the pair. Then color $i$ is forbidden on all uncolored vertices. We see that Maker needs at least six colors for winning the game, thus $\chi_g(G) \geq 6$.

chromatic number at least 8 and (2) that for every planar graph, the game chromatic number is at most 33. Their paper started a race of decreasing this number: it was first improved to 30 [DZ99], then to 19 [Zhu99], to 18 [Kie00], and finally by Zhu to 17 [Zhu08b] which is currently the best-known upper bound.

Besides planar graphs, the game has been analyzed on forests too. Another natural graph class are forests. It is easy to observe that the game chromatic number is at least 3 as soon as the graph contains a path of length 3 (i.e., 3 edges and 4 vertices). On the other hand, Faigle, Kern, Kierstead, and Trotter
proven that $\chi_g(T) \leq 4$ holds for every forest $T$ [Fai+93]. For both values 3 and 4, there are classes of forest attaining it [Dun+15]. Further graph classes on which the game chromatic number has been analyzed include cartesian products of graphs [Zhu08a]. For a survey on early results on the game chromatic number we refer to the paper of Bartnicki et al. [Bar+07].

In this chapter, we study the game chromatic number of the Erdős-Rényi random graph model $G(n, p) = (V, E)$. We assume throughout the chapter that $p = p(n) \leq 1 - \eta$, where $\eta > 0$ is an arbitrarily small, fixed constant. Bohman, Frieze, and Sudakov [BFS08] initiated the analysis of $\chi_g(G(n, p))$ and provided upper and lower bounds for $\chi_g(G(n, p))$ for a wide range of edge probabilities $p$. Let $b = \frac{1}{1-p}$ and note that for $p = o(1)$ we have

$$\log_b(np) = \frac{\log(np)}{\log b} = (1 + o(1)) \frac{\log(np)}{p}.$$ 

Their main result is the following.

**Theorem 6.1** (Theorem 1.1 in [BFS08]).

(i) There exists a constant $K > 0$ such that for $\varepsilon > 0$ and $p \geq (\log n)^{K\varepsilon^{-3}} / n$, a.a.s. it holds

$$\chi_g(G_{n,p}) \geq (1 - \varepsilon) \frac{n}{\log_b np}.$$ 

(ii) If $\alpha > 2$ is a constant, $K = \max\{\frac{2\alpha}{\alpha-1}, \frac{\alpha}{\alpha-2}\}$ and $p \geq (\log n)^K / n$, then a.a.s.

$$\chi_g(G_{n,p}) \leq \alpha \frac{n}{\log_b np}.$$ 

Furthermore, they conjectured that their lower bound of Theorem 6.1 (i) is tight.

**Conjecture 6.2** (Conjecture 5.1 in [BFS08]). If $p = \omega(n^{-1})$, then a.a.s.

$$\chi_g(G(n, p)) = (1 + o(1)) \frac{n}{\log_b(np)}.$$ 

In 2013, Frieze et al. investigated the game chromatic number of sparse random graphs and sparse $d$-regular graphs [FHL13]. For $\chi_g(G(n, p))$ they obtained the following bounds.
Theorem 6.3 (Theorem 1.2 in [FHL13]). Let $p = \frac{d}{n}$, where $d$ is larger than some absolute constant and $d \leq n^{1/4}$.

(i) If $\alpha < \frac{4}{7}$ is a constant then a.a.s.

$$\chi_g(G_n, p) \geq \alpha d \frac{d}{\log d}.$$ 

(ii) If $\alpha > 12$ is a constant then a.a.s.

$$\chi_g(G_n, p) \leq \alpha d \frac{d}{\log d}.$$ 

Note that for the ordinary chromatic number it is well-known by the results of Bollobás and Luczak [Bol88; Luc91] that $\chi(G_n, p) = (1 + o(1)) \frac{n}{\log_b(np)}$ a.a.s. Hence on random graphs, the game chromatic number is by a constant factor larger than the ordinary chromatic number and the two parameters have the same order of magnitude.

Our contribution is an improved upper bound for dense random graphs that matches the lower bound provided by Theorem 6.1 (i).

Theorem 6.4. Let $p \geq e^{-o(\log n)}$ and $b := \frac{1}{1-p}$. Then a.a.s.

$$\chi_g(G_n, p) \leq (1 + o(1)) \frac{n}{\log_b(np)}.$$ 

Together with the lower bound of Theorem 6.1 it implies that the game chromatic number of dense random graphs is asymptotically twice as large as the ordinary chromatic number.

Corollary 6.5. Let $p \geq e^{-o(\log n)}$ and $b := \frac{1}{1-p}$. Then a.a.s.

$$\chi_g(G_n, p) = (1 + o(1)) \frac{n}{\log_b(np)} = (2 + o(1)) \chi(G_n, p).$$ 

In Theorem 6.4 we assumed the lower bound $p \geq e^{-o(\log n)}$. Note that for any constant $\xi > 0$, this implies in particular $(np)^\xi p = \omega(\log(np) \cdot \log^2 n)$ and thus

$$(6.1) \quad (np)^\xi = \omega(\log_b(np) \cdot \log^2 n).$$
In the proof of Theorem 6.4 we will need (6.1) several times.

One of the main ingredients of our proof is an advanced version of Box games. In Section 6.2 we study Box games in detail as we use them not only in this chapter but also below in Chapter 7. Afterwards in Section 6.3 we present an overview of our proof strategy and describe the most important concepts. Finally Section 6.4 contains the main part of the technical work and finishes the proof of Theorem 6.4.

6.2 Box Games

The concept of Box games has been introduced in the seminal paper of Chvátal and Erdős [CE78]. Two players, which we call here Alice and Bob, take turns in claiming previously unclaimed elements of some given, pairwise disjoint, non-empty sets $A_1, \ldots, A_s$. Alice wins if she can claim all elements of at least one set $A_i$ whereas Bob wants to claim at least one element from each set $A_i$. We assume that Bob has the first move. The original version of Chvátal and Erdős is called a canonical game because it was assumed that the sets are almost equally large, i.e., $|A_i|$ and $|A_j|$ differ by at most one for all choices of $i$ and $j$.

A natural strategy for Bob is to play greedily, i.e., to always claim an element from the currently smallest set. If $|A_i| \geq 2$ holds for all $1 \leq i \leq s$, Then Bob easily wins the game by using the greedy strategy. To increase Alice’s chances to win the game, we add a bias $q$: Alice is now allowed to claim up to $q$ elements per turn. We then denote by $B(A_1, \ldots, A_s; q)$ the Box game on the sets $A_1, \ldots, A_s$ with bias $q$. Note that if Bob claims an element from a set $A_i$, we can remove this set from the game as it is not dangerous anymore from Bob’s perspective and Alice has no possibility to claim all elements of $A_i$. On the other hand, we can also remove all elements claimed by Alice. (But only the elements, not the sets where she claimed elements from!)

We now state some winning conditions for the two players. As a convention we use $\sum_{i=1}^{s} 1_i = 0$.

**Lemma 6.6.** Let $q \in \mathbb{N}$.
(i) If \( \sum_{i \in I} |A_i| > q|I| \sum_{i=1}^{|I|-1} \frac{1}{i} \) holds for all non-empty subsets \( I \subseteq \{1, \ldots, s\} \), then Bob wins the Box game \( B(A_1, \ldots, A_s; q) \) by playing greedily.

(ii) If \( B(A_1, \ldots, A_s; q) \) is a canonical Box game and \( \sum_{i=1}^s |A_i| \leq (q-1)s \sum_{i=1}^{s-1} \frac{1}{i} \), then Alice has a winning strategy.

The original statement (Theorem 2.1 in [CE78]) gives more precise formulas, but unfortunately the proof turns out to be incomplete. The proof has been completed in [HLV87], together with some generalizations. For the sake of readability, we restrict ourselves to slightly weaker statements that are absolutely sufficient for our purposes.

**Proof.** We start with (i) and prove the statement by induction over \( s \). For the base case \( s = 1 \) we observe that the assumption yields \( |A_1| > 0 \), thus Bob can claim an element of the only set \( A_1 \) and wins the game. Now suppose the statement is true for \( s-1 \) and consider a Box game \( B(A_1, \ldots, A_s; q) \) where \( \sum_{i \in I} |A_i| > q|I| \sum_{i=1}^{|I|-1} \frac{1}{i} \) is satisfied for all non-empty subsets \( I \subseteq \{1, \ldots, s\} \). Bob chooses greedily an element of the smallest set \( A_j \) and Alice answers by claiming \( q \) arbitrary elements. W.l.o.g. let \( j = s \). We remove the set \( A_j \), the elements claimed by Alice, and obtain a reduced Box game \( B(A_1', \ldots, A_{s-1}'; q) \) on \( s-1 \) sets.

Let \( I \subseteq \{1, \ldots, s-1\} \) be non-empty index set and put \( I' := I \cup \{s\} \), so \( |I'| = |I| + 1 \). By assumption it holds \( \sum_{i \in I'} |A_i| > q|I| \sum_{i=1}^{|I|-1} \frac{1}{i} \). For the index set \( I \) we then deduce

\[
\begin{align*}
\sum_{i \in I} |A_i'| &\geq \sum_{i \in I'} |A_i| - |A_j| - q \geq \sum_{i \in I'} |A_i| - \left[ \sum_{i \in I'} |A_i| \right] - q \\
&\geq \frac{|I'|-1}{|I'|} \sum_{i \in I'} |A_i| - q > \frac{|I'|-1}{|I'|} \left( q|I'| \sum_{i=1}^{|I'|-1} \frac{1}{i} \right) - q = q|I| \sum_{i=1}^{|I|-1} \frac{1}{i},
\end{align*}
\]

where we used in the second step that Bob is playing greedily. This bound holds for every choice of \( I \). By induction it follows that Bob wins the reduced Box game \( B(A_1', \ldots, A_{s-1}'; q) \) and thus as well the original Box game \( B(A_1, \ldots, A_s; q) \).

We now turn to (ii) where we assume that the Box game \( B(A_1, \ldots, A_s; q) \) is canonical. As we are assuming that all sets \( A_i \) are non-empty, the precondition implies \( s \geq 2 \). Bob claims an arbitrary element from any set \( A_j \) which we remove immediately from the board. Since the game is canonical we have
$|A_j| \geq \left\lfloor \frac{1}{s} \sum_{i=1}^{s} |A_i| \right\rfloor$. Furthermore, we can assume again $j = s$ w.l.o.g. Afterwards Alice claims $q$ elements which she clearly can do in a way such that the remaining Box game $B(A'_1, \ldots, A'_{s-1}; q)$ on $s-1$ sets stays canonical. We now distinguish two cases: if there exists $1 \leq i \leq s-1$ such that $A'_i$ is empty, then Alice has claimed each element of $A_i$ and thus wins the game. On the other hand, if $\min\{|A'_1|, \ldots, |A'_{s-1}|\} \geq 1$, we have

$$\sum_{i=1}^{s-1} |A'_i| \leq \sum_{i=1}^{s} |A_i| - |A_j| - q \leq \sum_{i=1}^{s} |A_i| - \left[ \frac{\sum_{i=1}^{s} |A_i|}{s} \right] - q \leq \frac{s-1}{s} \sum_{i=1}^{s} |A_i| + 1 - q$$

$$< \frac{s-1}{s} \left((q-1)s \sum_{i=1}^{s-1} \frac{1}{i}\right) + 1 - q = (q-1)(s-1) \sum_{i=1}^{s-2} \frac{1}{i}.$$  

We see that in this case, the reduced Box game is not only canonical, but also satisfies the other precondition of the statement, so the argument can be repeated. At the latest when $s = 2$, the board of the game is such that Alice can claim all elements of a set $A_i$ (because $q$ is sufficiently large) and thus wins the game. ■

Next, we introduce a more general version of the Box game: we allow Alice to steal moves of Bob. More precisely, for some parameter $z \in \mathbb{N}$, we allow Alice after every real move of Bob to steal the next $z - 1$ moves of Bob. Instead of Bob choosing an element, we allow Alice each time to claim an element from an arbitrary set, label it as Bob’s element, remove the corresponding set from the game (as Bob would do it if his move hadn’t been stolen by Alice!), and then proceed to her own move and eliminate at most $q$ elements from the remaining sets. This essentially means that after one real move of Bob, Alice may remove $z - 1$ sets of her choice plus $qz$ arbitrary elements from the remaining sets. We denote this modified Box game by $B(A_1, \ldots, A_k; q, z)$. Note that we still assume that Bob starts the game, with a real move, but afterwards Alice immediately starts stealing moves.

With the next lemma we formulate a criterion whether Bob can win the game $B(A_1, \ldots, A_k; q, z)$ even though Alice steals many of his turns. In contrast to Lemma 6.6, we even assume that Bob is not playing greedily but only “$d$-greedily”, meaning that he always chooses a set that contains at most $d$ elements more than the currently smallest set, where $d$ is some additional parameter.
Lemma 6.7. Let \( q, d, z \in \mathbb{N} \), \( f(1, q, d) := zq + d \) and
\[
f(s, q, d) := (zq + d)s\left(1 + \sum_{i=1}^{s-1} \frac{1}{i}\right) \text{ for } s > 1.
\]
If \( \sum_{i \in I} |A_i| > f(|I|, q, d) \) holds for all non-empty subsets \( I \subseteq \{1, \ldots, s\} \), then Bob wins the game \( B(A_1, \ldots, A_s; q, z) \) by playing \( d \)-greedily.

Proof. We prove this statement by induction on \( s \), always considering periods of \( z \) moves of Bob, the first of these moves being a real Bob’s move and the remaining \( z - 1 \) turns stolen by cruel Alice. Playing \( d \)-greedily, in his real moves Bob always claims an element from a set that contains at most \( d \) elements more than the currently smallest set.

For the base case of the induction, we show that Bob cannot lose in the first period. The first period consists of \( z \) rounds where the last \( z - 1 \) of Bob’s moves get stolen by Alice. In this period Alice can claim at most \( zq \) elements from the \( s - z \) remaining sets, after the \( z \) moves of Bob (real or stolen). By definition of the function \( f \) we know that \( |A_i| > zq \) holds for all \( 1 \leq i \leq s \), therefore no set \( A_i \) runs out of elements during the first moves of Alice, and Bob doesn’t lose in the first period of the game.

If \( s \leq z \), this proves the statement. (Recall that when Alice steals a move of Bob, nevertheless her claimed element is marked as Bob’s element.) So assume \( k > z \) and let \( \tilde{I} \) be the remaining index set at the moment where Bob plays his next real move, so \( |\tilde{I}| = s - z \). For \( i \in \tilde{I} \) we denote by \( A_i' \) the remaining set after the first \( z \) rounds of the game. Let \( I' \subseteq \tilde{I} \) be any non-empty subset of size \( \ell = |I'| > 0 \), and let \( I = I' \cup \{j\} \), where \( A_j \) denotes the set from which Bob claimed an element in his very first real move. Recall that we are assuming that Bob plays \( d \)-greedily. We thus deduce for \( t := \sum_{i \in I} |A_i| \) that \( t^* := \sum_{i \in I'} |A_i'| \) satisfies
\[
t^* \geq t - |A_j| - zq \geq t - \left\lfloor \frac{t}{\ell + 1} \right\rfloor - d - zq \geq \frac{\ell}{\ell + 1} t - d - zq
\[
> \frac{\ell}{\ell + 1} f(\ell + 1, q, d) - d - zq = (zq + d)\ell\left(1 + \sum_{i=1}^{\ell-1} \frac{1}{i}\right) - d - zq
\]
\[
= (zq + d)\ell\left(1 + \sum_{i=1}^{\ell-1} \frac{1}{i}\right) = f(|I'|, q, d).
\]
Hence the reduced Box game satisfies again the preconditions, and the argument can be repeated until \( s \leq z \) and Bob wins the final game. \( \blacksquare \)

### 6.3 Outline of Proof Strategy

For the remainder of the chapter we assume \( p \geq e^{-o(\log n)} \) and put \( b = \frac{1}{1-p} \). We need to show that for any constant \( \alpha > 1 \), arbitrarily close to one (but not equal), and a number of colors \( k = \alpha \frac{n \log_b (np)}{\log n} \), Maker has a strategy so that he wins the game with probability \( 1 - o(1) \). For the rest of the chapter we assume that \( p \) and \( \alpha \) are fixed and \( k \) is chosen as above.

We denote by \( \mathcal{C} = (\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_k) \) a collection of pairwise disjoint sets, where \( \mathcal{C}_i \) is the set of all vertices to which color \( i \) has been assigned. Note that we do not require that the sets \( \mathcal{C}_i \) partition the full vertex set, i.e., \( \mathcal{C} \) can be interpreted as the partial coloring obtained until some point of the game process.

For a vertex \( v \) we denote by \( A(v, \mathcal{C}) \) the set of all colors which are still available at \( v \) with respect to the partial coloring \( \mathcal{C} \). That is,

\[
A(v, \mathcal{C}) = \{ i \in \{1, \ldots, k\} | \Gamma(v) \cap \mathcal{C}_i = \emptyset \}.
\]

Furthermore, we put

\[
a(v, \mathcal{C}) = |A(v, \mathcal{C})|.
\]

During the evolution of the game, more and more vertices get colored and the sets \( A(v, \mathcal{C}) \) shrink. Maker needs to avoid that a set \( A(v, \mathcal{C}) \) of an uncolored vertex \( v \) gets empty. This indicates that the coloring game bears some relation to the Box games as discussed in Section 6.2. If Maker colors a vertex \( v \) with color \( i \) in the coloring game, then we remove color \( i \) from the sets \( A(w, \mathcal{C}) \) for all neighbors \( w \in \Gamma(v) \). Hence, the coloring of \( v \) corresponds to a move of Bob in the Box game, while the removals correspond to a move of Alice. Assume we know that each color appears in at most \( q + 1 \) sets \( A(v, \mathcal{C}) \). Then we could enhance the power of Breaker by allowing him to remove an arbitrary color from at most \( q \) sets \( A(w, \mathcal{C}) \). With that Breaker has more power, but if we can show that Maker wins this generalized variant, then he wins the original coloring game too.
In the coloring game, Breaker colors vertices as well in his turns, which should be translated to a move of Bob in the Box game. We can model this in the Box game by allowing Alice to steal every other move from Bob. Instead of Bob choosing an element (ideally from the smallest set), we allow Alice to decide which set gets removed from the game board. In such a round, Alice again plays her own move and eliminates at most \( q \) elements from the remaining sets of the Box game. Actually, when applying Lemma 6.7, we even allow Alice to steal more of Bob's moves.

We are now ready to define Maker's strategy for the coloring game. One of Maker's goals is to ensure that the color classes grow almost uniformly. Clearly, Maker cannot achieve this completely, as Breaker can play arbitrarily. But at least he can make sure that no color class is too small. Let \( N \) be a constant chosen appropriately later. Then Maker's strategy is the following.

- In every \( N \)-th move, Maker chooses an uncolored vertex \( v \) such that \( a(v, \mathcal{C}) \) is minimal (\( \mathcal{C} \) being given by the current color classes) and assigns any color \( i \in A(v, \mathcal{C}) \) to \( v \). We call this a move of first type.

- In all his other turns, Maker chooses a color \( i \) such that \( |\mathcal{C}_i| \) is minimal among all colors that can still be used somewhere, and assigns \( i \) to an uncolored vertex \( v \) with \( i \in A(v, \mathcal{C}) \). This is a move of second type.

Note that we may assume that Maker's strategy is deterministic: we fix some arbitrary ordering on the vertices, so that we can break ties uniquely. With these preliminaries at hand we are now ready to outline the main idea of our proof strategy.

Assume Breaker wins at time \( t \), i.e., assume that when \( t - 1 \) vertices have been colored, there exists a vertex \( v_0 \) that is still uncolored but for which all \( k \) colors appear in the neighborhood \( \Gamma(v_0) \). To reach a contradiction we will then define a time \( t' < t \) and argue that we may view the coloring game between times \( t' \) and \( t \) as a Box game, where the sets \( A_i \) of the Box game are given by the sets \( A(v, \mathcal{C}) \) of the vertices that have been colored between times \( t' \) and \( t \), plus the vertex \( v_0 \) that ran out of colors. To see that Bob wins this Box game (and that
therefore the coloring game could not have stopped at time $t$), we will argue that the conditions of Lemma 6.7 are satisfied. For this we need that the sets $A(v, \mathcal{C})$ are large compared to the power of Alice. Recall that the power of Alice (the parameter $q$ in Lemma 6.7) corresponds to how often a color appears in the sets $A(v, \mathcal{C})$, maximized over all colors. Clearly, the larger $t$ (and thus $t'$), the smaller $q$ and the sets $A(v, \mathcal{C})$. To carefully balance these effects we partition the game into phases, parametrized by a parameter $h$. In each phase we will use different bounds for the size of the sets $A(v, \mathcal{C})$ and the power of Alice. Let $\xi > 0$ be a constant which we will define later. We define three functions as follows.

$$
\beta(h) = \frac{\alpha \xi n(np)^{-h}}{10 \log_b(np)} = \Theta\left(\frac{(np)^{1-h}}{\log np}\right),
$$

(6.2)

$$
\gamma(h) = \frac{10 n \log n}{\beta(h)} = \Theta\left(\frac{(np)^h}{p \log n \log(np)}\right), \text{ and}
$$

$$
q(h) = \frac{\beta(h)}{(\log n)^2} = \Theta\left(\frac{(np)^{1-h}}{(\log n)^2 \log(np)}\right).
$$

With these definitions at hand we can now verify that the conditions of Lemma 6.7 are satisfied under various assumptions.

**Lemma 6.8.** Let $\alpha, p, k, N, \beta(\cdot)$ and $q(\cdot)$ be as defined above and assume that Maker plays according to our proposed strategy. Let $t' \leq n$ be a point in time, let $U$ be a set of uncolored vertices at time $t'$ and denote by $\mathcal{C}'$ the coloring after the first $t' - 1$ vertices have been colored. Furthermore, assume that there exists a constant $h < 1$ such that the following conditions are satisfied.

(i) At time $t'$ Maker colors some vertex by using a move of his first type.

(ii) $a(v, \mathcal{C}') \geq \frac{\beta(h)}{2}$ for all $v \in U$.

(iii) For all $v \in U$ there exists $S(v) \subseteq A(v, \mathcal{C}')$ such that $|S(v)| \leq q(h)$ and such that

$$
|\{v \in U \mid i \in A(v, \mathcal{C}') \setminus S(v)\}| \leq q(h) \text{ for all colors } i = 1, \ldots, k.
$$

Then Maker does not lose the coloring game in the interval $[t', t' + |U|]$, if within this time interval both players color only vertices of $U$. 
Proof. We have already seen the connection between the coloring game and the Box game. We will show that Maker wins the Box game induced by the vertex set $U$ and the color sets $A(v, \mathcal{C}') \setminus S(v)$.

Every color appears at most $q(h)$ times in the sets $A(v, \mathcal{C}') \setminus S(v)$, which implies an Alice-power of at most $q(h)$ in our Box game translation. Recall that in his moves of first type, Maker chooses a vertex $v$ where $a(v, \mathcal{C})$ is minimal. Since $|S(v)| \leq q(h)$ for all $v \in U$, we know that on the sets $A(v, \mathcal{C}') \setminus S(v)$, Maker plays $q(h)$-greedily with his moves of first type. Note that if one player colors a vertex $v$ using a color $i$, we anyway remove the whole box $A(v, \mathcal{C}')$ from the Box game and don’t care whether $i \in S(v)$. This allows us to only consider the Box game of the restricted sets $A(v, \mathcal{C}') \setminus S(v)$.

Maker uses his move of first type in every $N$-th of his turns. In the Box game this corresponds to the setting where Alice steals all but every $2N$-th of Bob’s moves. We conclude that between $t'$ and $t$, Breaker and Maker have played the Box game

$$B(A(v_1, \mathcal{C}') \setminus S(v_1), A(v_2, \mathcal{C}') \setminus S(v_2), \ldots, A(v_{|U|}, \mathcal{C}') \setminus S(v_{|U|}); q(h), 2N).$$

Next, observe that by definition of $q(h)$ and $\log |U| \leq \log n$ it holds

$$(2Nq(h) + q(h)) \left(1 + \sum_{i=1}^{\frac{|U|-1}{2}} \frac{1}{i}\right) = \Theta(q(h)\log |U|) = o(\beta(h))$$

For $n$ large enough we thus have for all $v \in U$ that

$$|A(v, \mathcal{C}') \setminus S(v)| \geq \frac{\beta(h)}{2} - q(h) \geq (2Nq(h) + q(h)) \left(1 + \sum_{i=1}^{\frac{|U|-1}{2}} \frac{1}{i}\right).$$

By Lemma 6.7 Bob wins this Box game, hence no vertex of the set $U$ could run out of available colors until time $t' + |U|$. We see that Maker does not lose the coloring game in this period when applying the proposed strategy.  

Lemma 6.8 shows that it is not essential to bound how often colors are available at uncolored vertices. It suffices if we can gather the colors that appear too often (and therefore enlarge Breaker’s power) with the sets $S(v)$. The critical and most technical part of our proof will be to show that it is actually possible to
find such sets $S(v)$ such that the preconditions ofLemma 6.8 are fulfilled. In the remainder of this section we give an outline of the key steps.

First we study how the sizes of our color classes behave during the coloring process. For doing so we introduce some notation. We call a color $i$ active if there exists at least one uncolored vertex that has $i$ in its color set $A(v, \mathcal{C})$. The level of the game, given a partial coloring $\mathcal{C}$, is then defined as the size of the smallest active color class,

$$\ell(\mathcal{C}) := \min \{ |\mathcal{C}_i| \mid i \text{ active} \}.$$  

Similarly, we define $\ell(t) = \ell(\mathcal{C})$, where $\mathcal{C}$ is the partial coloring obtained at time $t$. Due to Maker’s moves of second type, $\ell(t)$ is increasing during the game. In Section 6.4.2 we show that if Maker uses our proposed strategy, then his moves of the second type imply that as long as we have enough uncolored vertices, there are always many active colors classes whose size is close to the current level of the game.

**Lemma 6.9.** Let $\alpha$, $p$, and $k$ be defined as above and define $\xi = \min\{\frac{1}{2\alpha}, \frac{1}{10}(1 - \frac{1}{\alpha})\}$. Then a.a.s., there exists a constant $N = N(\alpha)$ such that the following statement is true for all $t \leq n - (np)^{1-4\xi}$.

If Maker plays according to our proposed strategy with parameter $N$, then the total number of active colors $i$ with size $|\mathcal{C}_i| \leq \ell(t) + \xi \log_b(np)$ is at least $\frac{\xi}{8}k$.

Recall that Maker’s strategy is deterministic. The event in the above lemma thus depends solely on properties of the random graph $G(n, p)$. Moreover, note that since $\alpha > 1$, $\xi$ is a small but positive constant.

In Section 6.4.3 we prove the following lemma.

**Lemma 6.10.** Let $\alpha$, $p$, $k$, $\xi$ and $\gamma(\cdot)$ be defined as above. Let $t' \leq n$ be a point in time, $U$ be a set of uncolored vertices and $4\xi < h \leq \frac{1}{\alpha} + 3\xi$ be a constant such that $|\mathcal{C}_i| \geq (h - 4\xi) \log_b(np)$ holds for every active color $i$ at time $t'$ and such that $|U| \leq 2N\gamma(h) + 1$. If conditions (i) and (ii) of Lemma 6.8 are satisfied and Maker uses the proposed strategy, with parameter $N$ as given by Lemma 6.9, then a.a.s. condition (iii) is satisfied as well.
6.4. PROOFS

Basically the statement of this lemma is that if we can control the quantities of the game process with a single constant $h$, then we can apply Lemma 6.8 in order to show that Maker wins the coloring game. It is important to note that the a.a.s.-statements only hold for a fixed constant $h$. In order to complete the proof of Theorem 6.4 we do, however, need to apply these lemmas for different values of $h$. We achieve this by dividing the game process into a constant number of periods, which are defined via the level $\ell(t)$ of the game. Using Lemma 6.9 we will then show that for each period it suffices to consider a single constant $h$. Thus we need to apply Lemma 6.10 only a constant number of times, which is fine. Section 6.4.4 contains the details of these arguments and how we can use them to complete the proof of Theorem 6.4.

6.4 Proofs

6.4.1 Properties of Random Graphs

We start this proof section by collecting some properties of partial colorings $\mathcal{C}$ of a random graph $G(n, p)$. Recall that a color $i$ is called active whenever it occurs in at least one set $A(v, \mathcal{C})$ of an uncolored vertex $v$. We say that a color is eliminated if it is not active anymore. Clearly, a color $i$ may be eliminated if $|\mathcal{C}_i|$ is very large. For instance Breaker could use the same color over and over again until it is nowhere possible. It turns out that a.a.s., as long as we have enough uncolored vertices, the total number of colors which are eliminated before they are heavily used is relatively small. The following lemma formalizes this. (Think of $A$ as a set of uncolored vertices where none of the colors $1, \ldots, d$ can be used.)

**Lemma 6.11.** Let $e^{-\alpha(\log n)} \leq p \leq 1 - \eta$, let $b = \frac{1}{1-p}$, let $\alpha > 1$ be a constant and let $\xi = \min\{\frac{1}{2\alpha}, \frac{1}{10(1-\frac{1}{\alpha})}\}$. Put $d := \lceil (np)^{1-4\xi} \rceil$. Denote by $\mathcal{A}$ the event that there are pairwise disjoint vertex sets $A, \mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_d$ in $G(n, p)$ such that

(i) $|A| \geq d,$

(ii) $|\mathcal{C}_i| \leq (\frac{1}{\alpha} + \xi) \log_b(np)$ for all $1 \leq i \leq d$, and

(iii) for all $1 \leq i \leq d$ and for all $v \in A$ we have $\mathcal{C}_i \cap \Gamma(v) \neq \emptyset$. 
Then
\[ \Pr[\mathcal{A}] = o(1). \]

**Proof.** Choose a set \( A \) with \(|A| = d\) and sets \( \mathcal{C}_i \) with \(|\mathcal{C}_i| = (\frac{1}{\alpha} + \xi) \log_b(np)\). (Note that we ignore floors and ceilings for ease of notation.) For a fixed vertex \( v \in A \) we have
\[ \Pr \left[ \bigwedge_{i=1}^{d} \{\Gamma(v) \cap \mathcal{C}_i \neq \emptyset\} \right] = (1 - (1 - p)^{|\mathcal{C}_i|})^d = (1 - (np)^{-1/\alpha - \xi})^d \leq e^{-(np)^{-1/\alpha - \xi}d}, \]
and a union bound implies
\[ \Pr[\mathcal{A}] \leq \left( \frac{n}{|A|} \right) \left( \frac{n}{\frac{1}{\alpha} + \xi} \log_b(np) \right)^d e^{-|A|(np)^{-1/\alpha - \xi}d} \]
\[ \leq n^{\frac{|A|}{\frac{1}{\alpha} + \xi} + d \log_b(np) - |A|(np)^{-1/\alpha - \xi}d} e^{-|A|(np)^{-1/\alpha - \xi}d} \]
\[ = n^{(np)^{1-4\xi} + (\frac{1}{\alpha} + \xi)(np)^{1-4\xi} \log_b(np) e^{-(np)^{2-9\xi-1/\alpha}}} \]
\[ = o(1), \]
as \( 2 - 9\xi - 1/\alpha \geq 1 + \xi \) and \( p \) satisfies (6.1).

A direct consequence of Lemma 6.11 is that we can bound the level \( \ell(t) \) of the game (cf. definition (6.3)) from above, assuming that there are enough uncolored vertices.

**Corollary 6.12.** Let \( \alpha > 1, \xi = \min\left\{ \frac{1}{2\alpha}, \frac{1}{10} (1 - \frac{1}{\alpha}) \right\}, e^{-o(\log n)} \leq p \leq 1 - \eta, \) and \( b = \frac{1}{1-p} \). Let \( t \leq n - (np)^{1-4\xi} \). Then a.a.s., \( \ell(t) < \left( \frac{1}{\alpha} + \xi \right) \log_b(np) \).

**Proof.** Suppose by contradiction that \( \ell(t) \geq \left( \frac{1}{\alpha} + \xi \right) \log_b(np) \). By our assumption on \( t \) we know that there are at least \((np)^{1-4\xi}\) uncolored vertices. From Lemma 6.11 it follows that a.a.s., at most \( d = \lceil (np)^{1-4\xi} \rceil \) colors have been eliminated before they reached size at least \( \left( \frac{1}{\alpha} + \xi \right) \log_b(np) \). All other color classes have size at least \( \left( \frac{1}{\alpha} + \xi \right) \log_b(np) \) at time \( t \). But this would immediately imply that the total number of colored vertices is at least
\[ \left( \frac{\alpha n}{\log_b(np)} - \lceil (np)^{1-4\xi} \rceil \right) \left( \frac{1}{\alpha} + \xi \right) \log_b(np) = (1 - o(1))(1 + \alpha \xi)n > n, \]
which is not possible.
One of the crucial steps in the proof of Theorem 6.4 is to find sets $S(v)$ that satisfy condition (iii) of Lemma 6.8. Intuitively, this is easier if for a given partial coloring $\mathcal{C}$, the color lists $A(v, \mathcal{C})$ look almost randomly, i.e., if they contain different colors for different vertices $v$. Our next lemma establishes some bounds on how similar the sets $A(v, \mathcal{C})$ can be.

**Lemma 6.13.** Let $p, \xi, q(\cdot), k$ be defined as above and let $4\xi < h \leq \frac{1}{a} + 3\xi$ and $0 < c < 1$ be constants. Denote by $\mathcal{B}$ the event that there exist disjoint vertex sets $S, \mathcal{C}_1, \ldots, \mathcal{C}_d$ in $G(n, p)$, where $d = d(n) \leq k$, with $|\mathcal{C}_i| \geq (h - 4\xi)\log_b(np)$ for all $i \in \{1, \ldots, d\}$ and $|S| \leq (np)^{h+\xi}$, such that one of the following two conditions is satisfied:

(i) $d \leq cq(h)|S|^{-6\xi}$ and for all $v \in S$, at least $cq(h)$ sets $\mathcal{C}_i$ of the family $\mathcal{C}_1, \ldots, \mathcal{C}_d$ are such that $\Gamma(v) \cap \mathcal{C}_i = \emptyset$, or

(ii) $d \geq |S|q(h)^{-1}(np)^{6\xi}$ and for all $i \in \{1, \ldots, d\}$ there exist at least $q(h)$ vertices $v \in S$ such that $\Gamma(v) \cap \mathcal{C}_i = \emptyset$.

Then

$$\Pr[\mathcal{B}] = o(1).$$

**Proof.** Note that without loss of generality we may assume that the sets $\mathcal{C}_i$ all have size equal to $C := (h - 4\xi)\log_b(np)$. We can then apply a union bound over the choices of $d, S$ and $\mathcal{C}_1, \ldots, \mathcal{C}_d$ to observe that

$$\Pr[\mathcal{B}] \leq \sum_{d,s} \binom{n}{s} \binom{n}{C}^d \cdot \Pr[\mathcal{B}(d, S, \mathcal{C}_1, \ldots, \mathcal{C}_d)],$$

where $\Pr[\mathcal{B}(d, S, \mathcal{C}_1, \ldots, \mathcal{C}_d)]$ denotes the probability that $\mathcal{B}$ holds for some fixed sets $S$ and $\mathcal{C}_1, \ldots, \mathcal{C}_d$, and $s$ denotes the size of the set $S$. Observe that for a vertex $v \in S$ and a given set $\mathcal{C}_i$ we have

$$\Pr[\Gamma(v) \cap \mathcal{C}_i = \emptyset] = (1 - p)^{|\mathcal{C}_i|} = (np)^{-h+4\xi},$$

as we assumed that $|\mathcal{C}_i| = C$. 

We first consider property (i). For every \( v \in S \) we define a random variable \( X(v) \) that counts the number of sets \( C_i \in \{C_1, \ldots, C_d\} \) such that \( v \) has no neighbors in \( C_i \). By (6.4) and the upper bounds on \( d \) and \( |S| \) we deduce

\[
\mathbb{E}[X(v)] \leq d \cdot (np)^{-h+4\xi} \leq |S| \cdot cq(h)(np)^{-6\xi} \leq cq(h)(np)^{-h+4\xi} = o(cq(h)).
\]

\( X(v) \) is a sum of independent Bernoulli random variables. By a Chernoff bound (Theorem 2.2 (iii)) it follows that

\[
\Pr[X(v) \geq cq(h)] \leq 2^{-cq(h)}.\]

Note that the random variables \( X(v) \) are independent for all \( v \in S \). Hence for \( n \) sufficiently large we have

\[
\binom{n}{C} \cdot \Pr[(i) \text{ holds}] \leq n^{dC} \cdot 2^{-cq(h)|S|} = n^{(h-4\xi)\log_b(n)p\cdot cq(h)|S|(np)^{-6\xi} \cdot 2^{-cq(h)|S|} \leq 2^{-\frac{1}{2} cq(h)|S|} \tag{6.1}
\]

Now consider property (ii). For every \( i \in \{1, \ldots, d\} \) define a random variable \( X(i) \) that counts the number of vertices in \( S \) that have no neighbor in \( C_i \). Then by (6.4), the upper bounds on \( |S| \) and \( h \) and the property \( \xi \leq \frac{1}{10} (1 - \frac{1}{a}) \),

\[
\mathbb{E}[X(i)] \leq |S| \cdot (np)^{-h+4\xi} \leq (np)^{5\xi} \leq (np)^{1-h-2\xi} = o(q(h)).
\]

Similarly as above we apply Chernoff bounds to obtain \( \Pr[X(i) \geq q(h)] \leq 2^{-q(h)} \). Then using the independence of the random variables \( X(i) \) and (6.1) to deduce that for \( n \) sufficiently large,

\[
\binom{n}{C} \cdot \Pr[(ii) \text{ holds}] \leq n^{dC} \cdot 2^{-q(h)d} \leq 2^{-\frac{1}{2} q(h)d} \leq 2^{-\frac{1}{2} |S|(np)^{6\xi}}.
\]

Combining both cases we conclude

\[
\Pr[\mathcal{B}] \leq \sum_{d,s} n^s \left( 2^{-\frac{1}{2} cq(h)s} + 2^{-\frac{1}{2} s(np)^{6\xi}} \right) = o(1).
\]

\[
\square
\]

Given a partial coloring \( \mathcal{C} \) of \( G_{n,p} \), from Maker’s perspective a vertex \( v \) is dangerous if the set \( A(v, \mathcal{C}) \) is small. These sets \( A(v, \mathcal{C}) \) shrink during the game.
process, while the level of the game increases. Our last lemma of this section puts the total number of small sets \( A(v, \mathcal{C}) \) in relation with the level \( \ell(\mathcal{C}) \) and shows that with high probability, every partial coloring has the property that there are not many dangerous vertices with respect to \( \ell(\mathcal{C}) \).

**Lemma 6.14.** Let \( p, \xi, k, \beta(\cdot), \gamma(\cdot) \) be defined as above and let \( h < 1 \) be a constant. For all partial colorings \( \mathcal{C} \), define

\[
B(h, \mathcal{C}) := \{ v \in V \mid a(v, \mathcal{C}) < \beta(h)/2 \}.
\]

Denote by \( \mathcal{E} \) the event that there exists a partial coloring \( \mathcal{C} \) of the random graph \( G(n, p) \) such that

1. \( \ell(\mathcal{C}) \leq (h - \xi) \log_b(np) \),
2. \( |B(h, \mathcal{C})| \geq \gamma(h) \), and
3. \( |C_i| \leq \ell(\mathcal{C}) + \xi \log_b(np) \) holds for at least \( \frac{\xi}{10} k \) color classes \( \mathcal{C}_i \).

Then

\[
\Pr[\mathcal{E}] = o(1).
\]

**Proof.** Let \( h < 1 \) and let \( \mathcal{C} \) be a partial coloring of \( G(n, p) \) which satisfies

\[
\ell(\mathcal{C}) \leq (h - \xi) \log_b(np).
\]

Denote by \( I \) the set of all color classes \( \mathcal{C}_i \) which fulfill

\[
|\mathcal{C}_i| \leq \ell(\mathcal{C}) + \xi \log_b(np) \leq h \log_b(np).
\]

Assume that \( |I| \geq \frac{\xi}{10} k \). Then for any vertex \( v \in V \) it holds

\[
\mathbb{E}[a(v, \mathcal{C})] = \sum_{i=1}^{k} (1 - p)^{|\mathcal{C}_i|} \geq \sum_{i \in I} (1 - p)^{|\mathcal{C}_i|} \geq \frac{\xi k}{10} (1 - p)^{h \log_b(np)} = \frac{\alpha \xi n(np)^{-h}}{10 \log_b(np)} = \beta(h).
\]
Note that the number of colors available at a fixed vertex \( v \) is the sum of independent indicator variables \( X_i \), where \( X_i = 1 \) if and only if \( v \) has no neighbors in \( C_i \). Chernoff bounds thus imply

\[
\Pr[a(v, C) < \beta(h)/2] \leq e^{-\beta(h)/8},
\]

and therefore

\[
\Pr[|B(h, C)| \geq \gamma(h)] \leq \left( \frac{n}{\gamma(h)} \right) e^{-\beta(h)\gamma(h)/8} = \left( \frac{n}{\gamma(h)} \right) n^{-\frac{5}{4}n}.
\]

There are \((k + 1)^n\) different partial colorings of the graph. We finish the argument by applying a union bound over all partial colorings that satisfy the assumption of the lemma. This yields

\[
\Pr[\mathcal{E}] \leq (k + 1)^n \left( \frac{n}{\gamma(h)} \right) n^{-\frac{5}{4}n} \leq n^{n + o(n) - \frac{5}{4}n} = o(1).
\]

\[\blacksquare\]

### 6.4.2 Proof of Lemma 6.9

Before we consider the proof of Lemma 6.9 we study a balls-and-bins game. Suppose we have \( k \) bins and two players \( M \) and \( B \) who alternately put a ball into one of the bins. We don't know how \( B \) plays, but \( M \) chooses always the bin with minimum load. We have the following two exceptions: \( B \) steals every \( N \)-th ball of \( M \) and plays this ball himself, and \( B \) can remove bins at any point in time.

In this model we will use \( t \) and \( \ell(t) \) in a similar way as defined in the setting of the coloring game. That is, \( t \) denotes the time (number of balls played) and \( \ell(t) \) denotes the number of balls in the bin with minimal load at time \( t \). In addition, denote by \( t(\ell) \) the minimal time \( t \) such that \( \ell(t) = \ell \).

**Lemma 6.15.** Consider the ball-game described above with \( k \) bins and parameter \( N \). Let \( a \in \mathbb{N} \) and denote by \( C(\ell) \) the total number of balls which have been thrown at loads \( \ell' \), \( \ell < \ell' \leq a \), until time \( t(\ell) \). Then for all \( \ell < a \) it holds

\[
C(\ell) \leq \frac{k\ell(N + 1)(a - \ell)}{(N - 1)(a - 1)}.
\]
**Proof.** We prove this Lemma by induction over $\ell$. For $\ell = 0$, clearly $C(0) = 0$ which agrees with the formula. Let $\ell < a - 1$ and suppose the statement is true for $\ell$. By definition, all non-removed bins have load at least $\ell$ at time $t(\ell)$. Denote by $x$ the number of bins that have load exactly $\ell$ at time $t(\ell)$. Using the definition of $C(\ell)$ we observe that

$$x \leq k - \frac{C(\ell)}{a - \ell}.$$ 

$M$ chooses always minimum-loaded bins, therefore $M$ can throw at most $x$ balls between $t(\ell)$ and $t(\ell + 1)$. Note that if $B$ removes some bins in this time period, then the upper-bound on $M$-balls is even smaller. Since $B$ steals every $N$-th ball of $M$, $B$ can play at most $\frac{N+1}{N-1} x$ balls between $t(\ell)$ and $t(\ell + 1)$.

Recall that $C(\ell + 1)$ counts the number of balls which have been thrown at loads between $\ell + 1$ and $a$ until time $t(\ell + 1)$. We may count all $\frac{N+1}{N-1} x$ $B$-balls, but the balls at load $\ell + 1$ at time $t(\ell)$ don’t count for $C(\ell + 1)$. It follows

$$C(\ell + 1) \leq C(\ell) + \frac{N+1}{N-1} x - (k - x) = C(\ell) - k + \frac{2N}{N-1} x$$

$$\leq C(\ell) - k + \left( k - \frac{C(\ell)}{a - \ell} \right) \frac{2N}{N-1}$$

$$\leq k \frac{N+1}{N-1} + C(\ell) \left( 1 - \frac{2N}{(a - \ell)(N-1)} \right)$$

$$\leq k \frac{N+1}{N-1} + \frac{k\ell(N+1)(a - \ell)}{(N-1)(a - 1)} \cdot \frac{a - \ell - 2}{a - \ell}$$

$$= k \frac{N+1}{N-1} \left( 1 + \frac{\ell(a - \ell - 2)}{a - 1} \right) = k \frac{N+1}{N-1} \frac{(a - \ell - 1)(\ell + 1)}{a - 1}.$$

$\blacksquare$

A direct consequence of Lemma 6.15 is that there always exists a constant fraction of bins whose load is close to the actual level of the process.

**Corollary 6.16.** Let $0 < \xi < 1$ and let $a = a(\xi)$ be an integer large enough such that $\frac{(1-\xi)a+1}{a-1} \leq 1 - \frac{\xi}{2}$. Consider the ball-game described above with $k$ bins and parameter $N \geq \frac{8}{\xi}$, where we suppose that $B$ removes at most $\frac{\xi}{8} k$ bins with load less than $a$. Let $t$ be a point in time such that $\ell(t) \leq a(1 - \xi)$. Then there exist at least $\frac{\xi}{8} k$ non-removed bins which have load at most $a$ at time $t$. 
Proof. Let \( \ell' := \ell(t) + 1 \) and suppose that after time \( t \), the two players continue with the process until \( t' := t(\ell') \), that is, until the game reaches a new level. Then
\[
C(\ell') \leq \frac{k\ell'(N + 1)(a - \ell')}{(N - 1)(a - 1)}
\]
holds by Lemma 6.15. There are at most \( C(\ell') \) bins which have load at least \( a \) at time \( t' \). However, since the bin loads are increasing, this property holds also at time \( t \). Taking into account also the removed bins, we obtain that the total number of non-removed bins which have load at most \( a \) at time \( t \) is at least
\[
k - \frac{C(\ell')}{a - \ell'} - \frac{\xi}{8} k \geq k \left( 1 - \frac{\ell'(N + 1)}{(N - 1)(a - 1)} - \frac{\xi}{8} \right)
\]
\[
\geq k \left( 1 - \frac{N + 1}{N - 1} \cdot \frac{(1 - \xi)a + 1}{a - 1} - \frac{\xi}{8} \right) \geq k \left( 1 - \frac{\xi}{8} - \frac{N + 1}{N - 1} \cdot \left( 1 - \frac{\xi}{2} \right) \right),
\]
where we used \( \ell' \leq a(1 - \xi) + 1 \) and our choice of \( a \). It remains to show that the inequality \( \frac{N + 1}{N - 1} \leq 1 - \frac{\xi}{4} \) is satisfied, but this is guaranteed by \( N \geq \frac{8}{\xi} \).

The connection between this model and the coloring game is straightforward: the bins correspond to the colors, and if a ball falls into a bin this means that some vertex has been assigned with this color. Maker’s moves of second type are equal to \( M \)’s strategy of playing the balls. In order to perform a worst-case analysis, we donate Maker’s moves of first type to player \( B \) in the ball game. At last, if a color is eliminated, we model this by a removed bin. This allows us to prove Lemma 6.9 using the ball model.

Proof of Lemma 6.9. Assume that the random graph does not satisfy the event \( \mathcal{A} \) from Lemma 6.11, which happens with probability \( 1 - o(1) \). Then Corollary 6.12 and the definition of \( \mathcal{A} \) imply for all \( t \leq n - (np)^{1 - 4\xi} \) that we have \( \ell(t) \leq (\frac{1}{a} + \xi) \log_b (np) \) and that at most \( \lceil (np)^{1 - 4\xi} \rceil = o(\xi k/8) \) colors have been eliminated before they reached size \( \ell(t) \). Using \( a := \ell(t) + \xi \log_b (np) \) we obtain
\[
\frac{\ell(t)}{a} = \frac{\ell(t)}{\ell(t) + \xi \log_b (np)} \leq \frac{1/\alpha + \xi}{1/\alpha + 2\xi} \leq 1 - \xi,
\]
which ensures \( \ell(t) \leq a(1 - \xi) \). In addition we have for \( n \) large enough that
\[
\frac{(1 - \xi)a + 1}{a - 1} = (1 - \xi)(1 + o(1)) \leq 1 - \frac{\xi}{2}.
\]
If we thus set \( N = N(\alpha) \) to an arbitrary integer of size at least \( 8/\xi \), then the assumptions of Corollary 6.16 are all satisfied and the lemma follows.

6.4.3 Proof of Lemma 6.10

Define

\[
L := L(h) = \frac{h + 2\xi}{2\xi} \quad \text{and} \quad c := c(h) = \frac{1}{L + 1}.
\]

Note that both \( L \) and \( c \) are constants that depend only on \( \alpha \) and \( h \). In the following we assume that the random graph does not satisfy the event \( B \) of Lemma 6.13 for our choice of \( h \) and \( c = c(h) \). We show that then the conclusions of Lemma 6.10 hold deterministically.

Proof of Lemma 6.10. Let \( t' \) be a point in time, let \( \mathcal{C}' \) be the partial coloring obtained until time \( t' \), let \( U \) be a set of uncolored vertices \( U \) and let \( \xi \) be a constant such that \( 4\xi < h \leq \frac{1}{\alpha} + 3\xi \). Suppose that the conditions (i) and (ii) of Lemma 6.8 are fulfilled. Furthermore, assume that \( |U| \leq 2N\gamma(h) + 1 \) and assume that at time \( t' \) we have \( |\mathcal{C}_i'| \geq (h - 4\xi) \log_b(np) \) for every active color \( i \).

We are interested in finding an arrangement of subsets \( S(v) \subset A(v, \mathcal{C}') \) such that \( |S(v)| \leq q(h) \) holds for all vertices \( v \in U \) and such that every color \( i \) appears in at most \( q(h) \) sets \( A(v, \mathcal{C}') \setminus S(v) \).

We use the Algorithm 6.1 to find the sets \( S(v) \). Note that this algorithm is essentially a greedy algorithm. We start with sets \( S(v) \) that are empty for every vertex. Then we consider the colors one by one in an arbitrary order. Let \( U(i) \) denote the set of all vertices in \( U \) where color \( i \) is available at time \( t' \). We choose a subset \( U'(i) \subseteq U(i) \) that consists of \( q(h) \) vertices \( v \in U(i) \) which (currently) have the largest set \( S(v) \). (We break ties arbitrarily.) For all \( v \in U(i) \setminus U'(i) \) we then add \( i \) to the set \( S(v) \).

When Algorithm 6.1 terminates, by construction every color \( i \) appears in at most \( q(h) \) sets \( A(v, \mathcal{C}') \setminus S(v) \). We need to show that the constructed arrangement also has the property \( s(v) := |S(v)| \leq q(h) \) for all \( v \in U \). In the reminder of this section we show that the assumption that the event \( B \) of Lemma 6.13 does not hold for \( h \) and \( c \) suffices to make this conclusion. Assume for contradiction
**Algorithm 6.1** Algorithm for constructing the sets $S(v)$

1: **for all** $v \in U$ **do**
2: \( S(v) := \emptyset \)
3: **for** $i = 1$ to $k$ **do**
4: \( U(i) := \{ v \in U \mid i \in A(v, \mathcal{C}') \} \)
5: **if** $|U(i)| > q(h)$ **then**
6: \( U'(i) := \{ \text{the } q(h) \text{ vertices } v \in U(i) \text{ with largest } |S(v)| \} \)
7: **for all** $v \in U(i) \setminus U'(i)$ **do**
8: \( S(v) := S(v) \cup \{ i \} \)

that after termination of the algorithm there exists a vertex $w \in U$ such that $s(w) > q(h)$. For $K \in \{0, \ldots, L\}$ let

\[ W_K := \{ v \in U \mid s(v) > (1 - Kc)q(h) \}. \]

By assumption we have $|W_0| \geq 1$. Note also that all sets $W_K$ are subsets of $U$ and we thus have

\[ \text{(6.6)} \quad |W_K| \leq |U| = O(\gamma(h)) = o((np)^{h+\xi}). \]

where we use (6.1). We will show by induction over $K$ that

\[ \text{(6.7)} \quad |W_K| \geq c^K q(h)^{2K}(np)^{-12L\xi} \quad \text{for all } K \in \{0, \ldots, L\}. \]

Observe that this completes the proof of Lemma 6.10, as

\[ c^L q(h)^{2L}(np)^{-12L\xi} = c^L (q(h)(np)^{-6\xi})^{2L} = \omega\left(\frac{c^L (np)^{(1-h-6\xi)2L}}{\log^{8L} n}\right) = \omega((np)^{h+\xi}), \]

where the last step follows from $1 - h - 6\xi \geq \xi$ (cf. upper bound on $h$ and the definition of $\xi \leq \frac{1}{10}(1 - \frac{1}{\alpha})$) and our choice of $L$. That is, the validity of (6.7) for $K = L$ contradicts (6.6), yielding the desired contradiction.

It remains to prove that (6.7) holds. Let $K \leq L$ and suppose (6.7) is true for $K - 1$. By definition we have

\[ s(v) \geq (1 - (K - 1)c)q(h) \]

for all $v \in W_{K-1}$. Thus, for every $v \in W_{K-1}$ there have to exist $c q(h)$ colors that were added to $S(v)$ at a time when $S(v)$ contained already at least $(1 - Kc)q(h)$...
colors. We denote by $I(v)$ the set of exactly these colors, and put

$$I_{K-1} := \bigcup_{v \in W_{K-1}} I(v).$$

As we assumed that the event $\mathcal{B}$ of Lemma 6.13 does not hold for $h$ and $c$, we deduce from property (i) that for $S = W_{K-1}$ and $d = |I_{K-1}|$ we have

$$(6.8) \quad |I_{K-1}| \geq c q(h) \cdot |W_{K-1}| \cdot (np)^{-6\xi}.$$ 

All colors of the set $I_{K-1}$ have been added to some set $S(v)$ which contained already at least $(1 - Kc) q(h)$ colors. By construction of the algorithm this means that for every $i \in I_{K-1}$ there were (at the time when color $i$ was processed) at least $q(h)$ other vertices $v'$ with $i \in A(v', \mathcal{C}')$ that also satisfied $s(v') > (1 - Kc) q(h)$, hence all these vertices $v'$ lie in the set $W_K$. Applying Lemma 6.13 with $S = W_K$ and $d = I_{K-1}$, we see that property (ii) implies

$$|W_K| \overset{(ii)}{\geq} q(h) (np)^{-6\xi} |I_{K-1}| \overset{(6.8)}{\geq} c q(h)^2 (np)^{-12\xi} |W_{K-1}| \overset{i.a.}{\geq} c^K q(h)^2 K (np)^{-12K\xi},$$

and we conclude that (6.7) holds for $W_K$ as well.

$\blacksquare$

### 6.4.4 Proof of Theorem 6.4

We define a set of constants $H := \{h_1, \ldots, h_J\}$ as follows. First let $h_1 := \frac{1}{2} - \xi$, where $\xi = \min\{\frac{1}{2a}, \frac{1}{10} (1 - \frac{1}{a})\}$ as before. Then for all $j > 1$ we put $h_j := \frac{1}{2} + (j-2)\xi$. Finally we denote by $J$ the smallest integer such that $h_J \geq \frac{1}{a} + 2\xi$.

**Proof of Theorem 6.4.** We assume that for all pairs $(h, c(h))$, none of the low-probability events of Section 6.4.1 occurs, where $h \in H$ and $c(h)$ as defined in (6.5). We show that in this case Maker wins the game deterministically. Since $H$ is a finite set, by a union bound this will conclude the proof of Theorem 6.4.

Assume that Maker applies the proposed strategy, but at time $t$ he loses the coloring game. I.e., we assume that after $t - 1$ rounds of the game, the two players obtain a partial coloring $\mathcal{C}$ such that there is at least one uncolored vertex $v_0$ where no color is available. Our goal is to apply Lemma 6.8 in order to realize that Maker could not lose the game at time $t$. 


Depending on the value of the time \( t \) at which Maker supposedly loses the game we define a constant \( h = h(\mathcal{C}) \in H \) as follows. If \( \mathcal{C} \) contains at least \( \lceil (np)^{1-4\xi} \rceil \) uncolored vertices, put
\[
h(\mathcal{C}) := \min\{h \in H : h \log_b(np) > \ell(\mathcal{C}) + \xi \log_b(np)\}.
\]
Note that \( \ell(\mathcal{C}) < \left(\frac{1}{a} + \xi\right) \log_b(np) \) holds by Corollary 6.12 and therefore \( h(\mathcal{C}) \) is well-defined.

If \( \mathcal{C} \) contains less than \( \lceil (np)^{1-4\xi} \rceil \) uncolored vertices, we subsequently remove the color assignment from the vertices which have been colored last in order to obtain a coloring \( \mathcal{C} \) with exactly \( \lceil (np)^{1-4\xi} \rceil \) uncolored vertices. We then let \( h(\mathcal{C}) := h(\mathcal{C}) \).

Let \( j \in \{1, \ldots, J\} \) and suppose that \( h(\mathcal{C}) = h_j \). If \( \mathcal{C} \) contains at least \( \lceil (np)^{1-4\xi} \rceil \) uncolored vertices, then it follows by Lemma 6.9 that \( |\mathcal{C}_i| \leq \ell(\mathcal{C}) + \xi \log_b(np) \) holds for at least \( \frac{\xi}{8} k \) color classes \( \mathcal{C}_i \). In the special case where \( \mathcal{C} \) contains less than \( \lceil (np)^{1-4\xi} \rceil \) uncolored vertices, the same holds for the partial coloring \( \overline{\mathcal{C}} \) which can be extended to \( \mathcal{C} \) by coloring at most \( \lceil (np)^{1-4\xi} \rceil \) additional vertices. Since \( (np)^{1-4\xi} = o(k) \), we deduce that in this special case, at least \( \frac{\xi}{10} k \) colors \( i \) satisfy \( |\mathcal{C}_i| \leq \ell(\mathcal{C}) + \xi \log_b(np) \).

We now define \( t' \) as the last time before \( t \) where Maker colored a vertex \( v \) with at least \( \beta(h(\mathcal{C}))/2 \) available colors in a move of his first type. By definition of such a move, we know that all uncolored vertices had at least \( \beta(h(\mathcal{C}))/2 \) available colors at this time. That is, \( t' \) and \( h(\mathcal{C}) \) satisfy conditions (i) and (ii) of Lemma 6.8. Moreover, by the definition of \( t' \), we know that between \( t' \) and \( t \) Maker always colored a vertex with less than \( \beta(h(\mathcal{C}))/2 \) available colors in his moves of the first type. Lemma 6.14 implies that even at time \( t \) the number of vertices for which \( a(v, \mathcal{C}) \) is less than \( \beta(h(\mathcal{C}))/2 \) is bounded by \( \gamma(h(\mathcal{C})) \). We thus deduce that Maker can have colored at most \( \gamma(h(\mathcal{C})) \) vertices with a move of his first type between time \( t' \) and time \( t \). With that we have
\[
t - t' + 1 \leq 2N \cdot \gamma(h(\mathcal{C})) + 1 = o(n),
\]
where \( N = N(a) \) is the parameter from Maker’s strategy. If we thus use \( U \) to denote the set of all vertices which have been colored in the period \([t', t - 1]\),
together with $v_0$, we arrive at

$$|U| \leq 2N\gamma(h(\mathcal{C})) + 1$$

and $U$ thus satisfies the prerequisite of Lemma 6.10. It remains to check that the partial coloring $\mathcal{C}'$ obtained after $t' - 1$ rounds satisfies

(6.10) \quad |\mathcal{C}'_i| \geq (h(\mathcal{C}) - 4\xi) \log_b(np)

for every active color $i \in \{1, \ldots k\}$. If this is the case, it follows from Lemma 6.10 that condition (iii) of Lemma 6.8 is fulfilled as well and Lemma 6.8 thus implies that Maker cannot lose the game at time $t$ as Bob wins the induced Box game.

Before verifying (6.10) for the case $j > 1$, let us first consider the case $j = 1$ where we directly show that Maker cannot lose the coloring game in this phase of the game. Indeed, suppose $j = 1$ and suppose Maker has colored at time $t'$ a vertex $w$ such that $a(w, \mathcal{C}') \geq \beta(h_1)/2$. Because Maker has used there a move of first type, it follows $a(v, \mathcal{C}') \geq \beta(h_1)/2$ for every uncolored vertex at time $t'$. Note that

$$\beta(h_1) = \Theta((np)^{1/2+\xi} \log^{-1}(np)) \quad \text{and} \quad \gamma(h_1) = \Theta((np)^{1/2-\xi} p^{-1} \log np \log n).$$

By assumption $p$ satisfies (6.1) and we deduce that $\beta(h_1) = \omega(\gamma(h_1))$. Hence, for $n$ large enough, (6.9) implies $\beta(h_1)/2 > t - t' + 1$. As with every move of the game a vertex can lose at most one of its available colors this thus means that until time $t$ no vertex can have lost all available colors, contradicting our assumption that Maker lost the game at time $t$.

Therefore we can assume from now on that $j > 1$. For this we define two additional points in time:

$$\hat{t} = \min\{t, n - [(np)^{1-4\xi}]\} \quad \text{and} \quad t^* := \min\{t \in \mathbb{N} : \ell(t) = (h(\mathcal{C}) - 4\xi) \log_b(np)\}.$$

Observe that $\hat{t} \leq t$. We will show below that $\hat{t} - t^* = \Theta(n)$. Thus, we have $t - t^* = \Theta(n)$ as well and (6.9) therefore implies that $t^* < t'$. By definition, $\ell(t^*) \geq (h(\mathcal{C}) - 4\xi) \log_b(np)$. Clearly, the level of the game is increasing in time and indeed (6.10) follows.
It remains to show that \( \hat{t} - t^* = \Theta(n) \). However, Lemma 6.9 implies that at time \( t^* \), the number of active colors classes of size at most \( (h(C) - 3\xi) \log_b(np) \) is at least \( \frac{\xi}{8} k \). On the other hand, note that \( j > 1 \) implies that

\[
\ell(\hat{t}) \geq (h(C) - 2\xi) \log_b(np).
\]

Hence all of these \( \frac{\xi}{8} k \) colors either reached size at least \( (h(C) - 2\xi) \log_b(np) \) at time \( \hat{t} \) or have been eliminated. From Lemma 6.11 we know that by time \( \hat{t} \) at most \( [(np)^{1-4\xi}] = o(k) \) color classes with size less than \( (h(C) - 2\xi) \log_b(np) \) have been eliminated. Thus, at least \( \frac{\xi k}{8} - o(k) \) color classes increased by at least \( \xi \log_b(np) \) in the period between \( t^* \) and \( \hat{t} \) and we deduce that

\[
\hat{t} - t^* \geq \left( \frac{\xi k}{8} - o(k) \right) \cdot \xi \log_b(np) = \Theta(n),
\]

which concludes the proof of Theorem 6.4.

\section*{6.5 Concluding Remarks}

In this chapter we obtained the asymptotic value of the game chromatic number of \( G(n, p) \) for sufficiently dense graphs. We proved that a.a.s.

\[
\chi_g(G(n, p)) = (2 + o(1))\chi(G(n, p))
\]

holds whenever \( p \geq e^{-o(\log n)} \). However, there is no a priori reason why a similar statement should not be true for smaller values of \( p \) as well. A key ingredient of our proof is Lemma 6.10 in which we show that Algorithm 6.1 yields sets \( S(v) \) such that the induced Box game is a win for Bob. For its correctness proof we need that \( p \geq e^{-o(\log n)} \). Expanding Theorem 6.4 to smaller edge probabilities \( p \) thus seems to require different arguments. We leave this to future work.
After having studied the game chromatic number in Chapter 6, we now turn to the analogous game where Maker and Breaker color the edges instead of the vertices. Here, the game chromatic index \( \chi'_g(G) \) denotes the smallest number of colors for which Maker has a winning strategy. For the game being played on an arbitrary graph \( G \), we always have the trivial bounds \( \Delta(G) \leq \chi'_g(G) \leq 2\Delta(G) - 1 \) on the game chromatic index. Beveridge, Bohman, Frieze, and Pikhurko (2008) proved that for every \( \delta > 0 \) there exists a constant \( c > 0 \) such that \( \chi'_g(G) \leq (2 - c)\Delta(G) \) holds for any graph with \( \Delta(G) \geq (\frac{1}{2} + \delta)n \), and conjectured that the same holds for every graph \([Bev+08]\). In this chapter, we improve their result and show that \( \chi'_g(G) \leq (2 - c)\Delta(G) \) is true for all graphs \( G \) with \( \Delta(G) \geq C \log n \). In addition, we consider a biased version of the game where Breaker is allowed to color \( b \) edges per turn and give bounds on the number of colors needed for Maker to win the biased variant of the game. This chapter is single-author work \([Keu17]\).
7.1 Introduction

Let $G = (V, E)$ be a graph and let $k$ be a positive integer. We study the game where two players, called Maker and Breaker, take turns in which they alternately assign a color $i \in \{1, \ldots, k\}$ to a previously uncolored edge $e \in E$ such that the partial coloring stays proper, i.e., no two adjacent edges get the same color. Maker’s goal is that at the end of the game, every edge is colored. Meanwhile, Breaker plays against Maker and aims to produce a partial coloring such that for at least one uncolored edge, all colors are forbidden and thus the partial coloring can not be extended to a proper edge-coloring of $G$. The game chromatic index $\chi'_g(G)$ is defined as the smallest integer $k$ for which Maker has a winning strategy.

This game is the most natural variation of the vertex coloring game that we discussed in Chapter 6. The game chromatic index of graphs was first studied by Lam, Shiu, and Xu [LSX99]. For any graph $G$ we have the two trivial bounds

$$\Delta(G) \leq \chi'_g(G) \leq 2\Delta(G) - 1,$$

where $\Delta(G)$ denotes the maximum degree of $G$. Clearly, the lower bound is tight for star graphs. On the other hand, the upper bound is tight for instance for $K_4$ (the complete graph on four vertices) or for cycles of odd length.

In their paper, Lam et al. proved that for every tree $T$ it holds $\chi'_g(G) \leq \Delta(G) + 2$. Erdős, Faigle, Hochstättler, and Kern [Erd+04] improved this result and showed that for forests $T$ of maximum degree $\Delta(T) \geq 6$ it holds $\chi'_g(T) \leq \Delta(T) + 1$, and that there are forests for which this bound is tight. Afterwards, Andres [And06] extended this result to the case $\Delta = 5$. Apart from forests, Cai and Zhu [CZ01] proved that $\chi'_g(G) \leq \Delta + 3k - 1$ holds for every $k$-degenerate graph $G$, which implies in particular an upper bound of $\Delta(G) + 14$ for every planar graph $G$. Later, Bartnicki and Grytczuk [BG08] obtained an upper bound of $\Delta(G) + 3k - 1$ for graphs $G$ of arboricity $k$, further results include wheels [LSX99; AHS11]. The question arised whether there exists a global constant $C$ such that for any graph $G$ it holds $\chi'_g(G) \leq \Delta(G) + C$. The answer was given by Beveridge, Bohman, Frieze, and Pikhurko who proved that for every sufficiently large $d$ there exists a graph $G$ with $\Delta(G) \leq d$ and $\chi'_g(G) \geq 1.008d$ [Bev+08].
Unfortunately, it is believed that in general the game is hard to analyze. In contrast to the vertex coloring game, it seems challenging to find powerful strategies even for only one of the two players. For example, a player’s move that looks clever at the start of the game can easily hurt the same player later on. That is why accurate bounds on the game chromatic index are only known for very few specific and sparse graph classes, and in general, knowledge on the game is rather scarce. For example, still no reasonable strategies are known for the complete graph $K_n$. Although it is desirable to determine $\chi'_g(G)$ precisely, Beveridge et al. proposed that one should first decide whether $\chi'_g(G)$ is bounded away by a constant factor from $\Delta(G)$, from $2\Delta(G)$, or from both. In this spirit, as their second result Beveridge, Bohman, Frieze, and Pikhurko proved the following statement.

**Theorem 7.1** (Theorem 4 in [Bev+08]). For every $\delta > 0$ there exists $c > 0$ such that every graph $G$ of maximum degree at least $(1/2 + \delta)n$ satisfies

$$\chi'_g(G) \leq (2 - c)\Delta(G).$$

Furthermore, they conjectured that the same is true for every graph $G$.

**Conjecture 7.2** (Conjecture 2 in [Bev+08]). There exists $c > 0$ such that for every graph $G$ it holds

$$\chi'_g(G) \leq (2 - c)\Delta(G).$$

The main result of this chapter provides a non-trivial upper bound on the game chromatic index for all graphs $G$ of maximum degree at least $C \log n$. We thus extend Theorem 7.1 and make a first step towards a proof of Conjecture 7.2.

**Theorem 7.3.** There exist $C, c > 0$ such that all graphs $G$ with $\Delta(G) \geq C \cdot \log n$ satisfy

$$\chi'_g(G) \leq (2 - c)\Delta(G).$$

Note that in particular for complete bipartite graphs or random graphs $G(n, p)$ with $\frac{C \log n}{n} \leq p \leq c'n$ for $c' < \frac{1}{2}$, Theorem 7.3 yields the first non-trivial bound on the game chromatic index. Our result also generalizes to the variant
of the game where Breaker is allowed to sit out during his turns. Consequently, the identity of the starting player does not matter.

As we have seen before (e.g., Section 6.2), in the context of Maker-Breaker games it is natural to also consider biased games where one player is allowed to claim not one but multiple elements per round. Let \( b \geq 1 \). We introduce the edge coloring game with bias \( b \) as follows. Maker still colors a single edge per round as before, but Breaker is now allowed to color any number of edges that is at most \( b \) in each of his turns. The winning conditions for the two players remain the same. For this biased variant of the game, we define \( \chi'_g(G, b) \) as the smallest number of colors such that Maker has a winning strategy. Clearly, the bounds of (7.1) are also valid for \( \chi'_g(G, b) \). We show that Theorem 7.3 can be generalized to the biased edge coloring game.

**Theorem 7.4.** There exists \( c > 0 \) such that for any bias \( b \geq 1 \), every graph \( G \) with \( \Delta(G) \geq C(b) \cdot \log n \) satisfies

\[
\chi'_g(G, b) \leq (2 - cb^{-4})\Delta(G),
\]

where \( C(b) > 0 \) only depends on \( b \).

Vice-versa, our last result verifies that there are graphs \( G \) where a bias \( b \geq 2 \) results in Breaker winning the game even with \( 2\Delta(G) - 2 \) colors, and Maker effectively needs \( 2\Delta(G) - 1 \) colors. Hence, an analogue of Conjecture 7.2 can not hold for the biased variant of the game. In particular, for regular graphs it follows that the precondition of Theorem 7.4 is almost optimal and that the value of \( \chi'_g(G, b) \) depends not only on the maximum degree but also on the number of vertices.

**Theorem 7.5.** Let \( b \geq 2 \) and \( \Delta \geq 2 \). Then for every \( \Delta \)-regular graph \( G \) with at least \( 2 \cdot \Delta^3 \cdot \exp\left(\frac{2\Delta - 2}{b-1}\right) \) vertices it holds

\[
\chi'_g(G, b) = 2\Delta(G) - 1.
\]

This chapter is structured as follows: after introducing some specific notations in Section 7.2.1, we present in Section 7.2.2 a randomized strategy for
7.2 Upper Bounds

7.2.1 Notations

We start with some notations. We consider the game as a process that evolves in rounds. In the first round, only Breaker is allowed to play. Afterwards, in every round $r$ it is first Maker’s and then Breaker’s turn. When proving upper bounds, we allow Breaker to sit out and not color any edge in his turns. This setting was first studied by Andres [And06] and makes the identity of the starting player irrelevant. Clearly, any upper bound on $\chi'_g(G)$ that holds for this modified variant also serves as an upper bound for the original game. Throughout the chapter we abbreviate $\Delta = \Delta(G)$. We denote by $v$-edge an edge that is incident to $v$ and by $\Gamma'_r(v)$ the set of all neighbors of $v$ in the subgraph of uncolored edges after round $r$. Furthermore, the load $\ell_r(v) := \text{deg}(v) - |\Gamma'_r(v)|$ counts the total number of colored $v$-edges after round $r$. Finally, let $A_r(e)$ be the set of available colors at an edge $e$ after round $r$, and let $U_r(v)$ be the set of colors that have been used at $v$-edges during the first $r$ rounds.

7.2.2 Maker’s Strategy

We prove Theorem 7.3 and Theorem 7.4 by using a random strategy for Maker. Note that as we study a complete information game without chance moves, there exists a winning strategy for exactly one of the two players. Hence, it is sufficient to prove that with the proposed random strategy, Maker wins with strictly positive probability against any fixed, deterministic strategy of Breaker. Then Breaker can not have a winning strategy, implying that there exists a deterministic winning strategy for Maker. This application of the probabilistic method was first used by Spencer [Spe91].
Let \( G = (V, E) \) be a given graph. We fix \( \lambda \) and \( c \) globally such that
\[
1 > \lambda > c > 0,
\]
i.e., \( \lambda \) is sufficiently small and \( c \) is sufficiently small compared to \( \lambda \). For the sake of readability, we always ignore roundings and assume that all considered quantities are integers. We do the proofs of Theorem 7.3 and Theorem 7.4 at once and therefore assume that there exists a fixed integer \( b \geq 1 \) such that in each of his turns, Breaker colors at least 0 and at most \( b \) edges. Note that \( \lambda \) and \( c \) do not depend on \( b \).

Given this set of constants, we assume that the game is played with a set of \( k := (2 - cb^{-4})\Delta \) colors.

Before defining the strategy, we make some further preparations. Suppose an uncolored edge \( e = \{u, v\} \) satisfies \( |U_r(u) \cap U_r(v)| > 2\Delta - k = cb^{-4}\Delta \) after some round \( r \). Then the edge \( e \) will never run out of available colors. For every vertex \( v \in V \) with \( \deg(v) \geq 2\lambda b^{-1}\Delta \), Maker uses this observation as follows. After the first round \( r \) where \( \ell_r(v) \geq 2\lambda b^{-1}\Delta \) holds, he looks at the set of uncolored \( v \)-edges, and defines a set \( D(v) \subseteq \Gamma(v) \) of dangerous neighbors, containing all vertices \( u \in \Gamma(v) \) that fulfill the following four conditions:

(i) the edge \( \{u, v\} \) is still uncolored, i.e., \( u \in \Gamma'_r(v) \),

(ii) \( \deg(u) + \deg(v) \geq k = (2 - cb^{-4})\Delta \),

(iii) \( |U_r(u) \cap U_r(v)| \leq 2\Delta - k = cb^{-4}\Delta \), and

(iv) \( \ell_r(u) \geq \lambda b^{-1}\Delta \) and \( u \) reached load \( \lambda b^{-1}\Delta \) not after its neighbor \( v \), i.e., for all \( r' \) such that \( \ell_{r'}(v) \geq \lambda b^{-1}\Delta \) it also holds \( \ell_{r'}(u) \geq \lambda b^{-1}\Delta \).

Clearly, an edge \( \{u, v\} \) can run out of available colors only if (i)-(iii) are fulfilled. Intuitively speaking, with condition (iv) we decide which vertex is responsible for such an edge. In case \( u \) and \( v \) reach load \( \lambda b^{-1}\Delta \) at the same round, the construction yields \( u \in D(v) \) and \( v \in D(u) \). Once we are at a round \( r \) such that \( v \) satisfies \( \ell_r(v) \geq 2\lambda b^{-1}\Delta \), the set \( D(v) \) is defined and Maker’s local goal for the remaining game process will be to color all edges between \( v \) and \( D(v) \) before \( v \) reaches load \( 3\lambda b^{-1}\Delta \). Note that as long as \( \ell_r(v) \leq 3\lambda b^{-1}\Delta \), there are still colors available for these edges.
We now start describing Maker’s strategy at an arbitrary round $r$ where there are still uncolored edges left. Let $f_0$ be the edge colored by Maker at round $r - 1$. (If it is Maker’s first move of the game, take an arbitrary edge for $f_0$.) Furthermore, let $F$ be the set of edges that Breaker colored in his turn at round $r - 1$. In the special case where Breaker didn’t color any edge in his last move, choose instead any uncolored edge $f_1$ and put $F := \{f_1\}$. Let $q := \frac{6c}{λ}$. We then propose Maker to play at random in the following way.

1. Choose $f \in \{f_0\} \cup F$ at random such that $\Pr[f = f_0] = \frac{1}{2}$ and $\Pr[f = f_i] = \frac{1}{2|F|}$ for all $f_i \in F$.

2. Let $v$ be one of the two vertices incident to $f$ chosen uniformly at random. If $Γ'_{r-1}(v)$ is empty, discard this first choice of $v$ and replace it by an arbitrary vertex $v$ such that $Γ'_{r-1}(v)$ is non-empty.

3. Choose a neighbor $u \in Γ'_{r-1}(v)$ uniformly at random. If $ℓ_{r-1}(v) \geq 2λb^{-1}Δ$ and $D(v) \cap Γ'_{r-1}(v)$ is non-empty, with probability $q$ discard the first choice of $u$ and replace it by $u \in D(v) \cap Γ'_{r-1}(v)$ chosen uniformly at random.

4. Let $e = \{u, v\}$ and color $e$ with a color $i \in A_{r-1}(e)$ chosen uniformly at random. We call $e$ a good $v$-edge.

Note that the strategy is well-defined, i.e., it always yields an uncolored edge $e$ for Maker to color. For every edge $f_i \in \{f_0\} \cup F$, the probability that it is chosen by Maker in the first step is at least $\frac{1}{2b}$, as we have $|F| \leq b$ by assumption.

Suppose Maker applies the proposed strategy. In case the strategy tells Maker to color an edge $e \in E$ at round $r$ but $A_{r-1}(e)$ is empty, Maker loses the game by definition. Then, we don’t yet abort the game. Instead, we let Maker play a color $i$ chosen uniformly at random among all colors and create a non-proper coloring, whereas Breaker is still forced to color edges properly. Consequently, if there is a left-over of uncolored edges where all colors are blocked, Breaker has no other option than sitting out for the remainder of the game (which is indeed possible for him). This yields a slightly different coloring process that always terminates with a full but not necessarily proper edge coloring. Observe that
as long as Maker doesn’t need to use forbidden colors in the modified process, the original and the modified process coincide. If Maker is never forced to use forbidden colors, then in both processes we obtain a proper edge coloring of $G$ and Maker wins the game.

### 7.2.3 Main Proof

Since Theorem 7.3 is a special case of Theorem 7.4, it suffices to prove the latter. By the precondition of Theorem 7.4, the maximum degree $\Delta$ is larger than $C(b)$, so we can always assume that $\Delta$ is sufficiently large.

We start the analysis by collecting several auxiliary results. Let $v \in V$ be any vertex. Recall the definitions of $v$-edges and good $v$-edges. Our first goal is to verify that while the load $\ell_r(v)$ increases during the game process, always a constant fraction of the colored $v$-edges are good $v$-edges. We specify this with the following lemma.

**Lemma 7.6.** Let $j \in \{1, 2, 3\}$ and let $v \in V$ be a vertex of degree at least $j\lambda b^{-1}\Delta$. Denote by $B_j(v)$ the bad event that among the $v$-edges that have been colored at rounds $r$ where $\ell_{r-1}(v) \geq (j - 1)\lambda b^{-1}\Delta$ and $\ell_r(v) < j\lambda b^{-1}\Delta$, less than $\frac{1}{5b^2} \lambda \Delta$ edges are good $v$-edges. Then

$$\Pr[B_j(v)] = \exp(-\Omega(\Delta)).$$

We defer the proof of Lemma 7.6 together with the proofs of the two subsequent lemmas to Section 7.2.4. Note that here and in the following, whenever we use the Landau-notation for estimating the probability of some events, we hide constant factors that may depend on $\lambda$, $c$, or $b$.

Next, we study how fast the load of a vertex of large degree grows compared to the average load of its neighbors. We show that it is unlikely that the average load among vertices in $\Gamma'_r(v)$ deviates by too much.

**Lemma 7.7.** Let $v \in V$ be a vertex of degree at least $k - \Delta = (1 - cb^{-4})\Delta$. Denote by $B_4(v)$ the bad event that there exists a round $r$ where $\ell_r(v) < 2\lambda b^{-1}\Delta$ but

$$\frac{1}{|\Gamma'_r(v)|} \sum_{u \in \Gamma'_r(v)} \ell_r(u) \geq 9\lambda \Delta.$$  

Then

$$\Pr[B_4(v)] = \exp(-\Omega(\Delta)).$$
As a next step, we also take colors into consideration. Let \( v \) be a vertex of degree at least \( \lambda b^{-1} \Delta \). Then we know that unless the bad event \( \mathcal{B}_1(v) \) occurs, among the first \( \lambda b^{-1} \Delta \) colors used at \( v \)-edges there are at least \( \frac{1}{5b^2} \lambda \Delta \) colors that were assigned by Maker to good \( v \)-edges. As long as \( \ell_r(v) < \lambda b^{-1} \Delta \), for all \( v \)-edges the set \( A_r(e) \) is non-empty, Maker is not forced to color \( v \)-edges non-properly, and indeed all colored \( v \)-edges use distinct colors.

For any vertex \( v \), let \( I'(v) \) be a subset of colors assigned to good \( v \)-edges and defined as follows. If Maker colors less than \( \frac{1}{5b^2} \lambda \Delta \) good \( v \)-edges before the load of \( v \) reaches \( \lambda b^{-1} \Delta \) (i.e., at rounds \( r \) such that \( \ell_r(v) < \lambda b^{-1} \Delta \)), then \( I'(v) \) is the set of colors that Maker used for these good \( v \)-edges. If there are at least \( \frac{1}{5b^2} \lambda \Delta \) such edges, \( I'(v) \) only contains the first \( \frac{1}{5b^2} \lambda \Delta \) colors that Maker used at such moves. For a vertex \( v \), we hope that colors are distributed rather randomly inside the sets \( \{ I'(u) \mid u \in \Gamma(v) \} \). We formalize such a distribution with the following lemma.

**Lemma 7.8.** Let \( v \in V \) be any vertex. Denote by \( \mathcal{B}_5(v) \) the event that there exists a subset of neighbors \( W \subseteq \Gamma(v) \) of size \( cb^{-2} \Delta \) and a set \( I^- \) of \( cb^{-2} \Delta \) colors such that for all \( i \in I' \), we have \( |\{ u \in W \mid i \in I'(u) \}| \geq \frac{1}{4b^4} c \lambda \Delta \). Then

\[
\Pr[\mathcal{B}_5(v)] = \exp(-\Omega(\Delta^2)).
\]

We now start proving the main theorem.

**Proof of Theorem 7.4.** By definition of the sets \( D(v) \), it is sufficient to verify for all vertices \( v \) of degree at least \( k - \Delta = (1 - cb^{-4}) \Delta \) that Maker is fast enough to color all edges between \( v \) and \( D(v) \) before \( v \) reaches load \( 3\lambda b^{-1} \Delta \). If this is possible for Maker, then the process yields a proper edge coloring of \( G \) and Maker wins the game.

Let \( v \in V \) be a fixed vertex of degree at least \( (1 - cb^{-4}) \Delta \). Further, denote by \( D'(v) \subseteq \Gamma(v) \) the set of all neighbors of \( v \) that reach load \( \lambda b^{-1} \Delta \) not after \( v \). More precisely, \( D'(v) \) contains all neighbors \( u \in \Gamma(v) \) such that \( \ell_r(v) \geq \lambda b^{-1} \Delta \) always implies \( \ell_r(u) \geq \lambda b^{-1} \Delta \). In the following we assume that for all \( u \in D'(v) \) the bad event \( \mathcal{B}_1(u) \) and for \( v \) itself the bad event \( \mathcal{B}_5(v) \) do not occur. By Lemma 7.6, Lemma 7.8, and a union bound, this happens with probability \( 1 - \exp(-\Omega(\Delta)) \).
Suppose $|D'(v)| \geq cb^{-2}\Delta$ and define $\mathcal{W} := \{W \subseteq D'(v) \mid |W| = cb^{-2}\Delta\}$. Let us first consider a fixed set $W \in \mathcal{W}$. As we are excluding the events $\{B_1(u) \mid u \in W\}$ and $B_5(v)$, for every vertex $u \in W$ there exists a set $I'(u)$ of exactly $\frac{1}{5b^2}\lambda\Delta$ colors with the property that at most $cb^{-2}\Delta$ colors $i$ satisfy

$$|\{u \in W \mid i \in I'(u)\}| \geq \frac{1}{4b^4}c\lambda\Delta.$$ 

Next, let us define $I_W$ as the set of all colors $i$ that fulfill

$$|\{u \in W \mid i \in I'(u)\}| \geq \frac{1}{24b^4}c\lambda\Delta.$$ 

We claim that $|I_W| \geq \frac{\Delta}{2}$. Indeed, if this is not the case, then there are at least $k - \frac{\Delta}{2} = (\frac{3}{2} - cb^{-4})\Delta$ colors that are contained in at most $\frac{1}{24b^4}c\lambda\Delta$ of the sets $\{I'(u) \mid u \in W\}$. On the other hand, except a small set of at most $cb^{-2}\Delta$ heavy colors, all colors $i \in I_W$ are contained in at most $\frac{1}{4b^4}c\lambda\Delta$ of the sets $\{I'(u) \mid u \in W\}$. If $c$ is sufficiently small compared to $\lambda$, this yields

$$\sum_{i=1}^{k} |\{u \in W \mid i \in I'(u)\}| \leq \left(\frac{3}{2} - cb^{-4}\right)\Delta \cdot \frac{1}{24b^4}c\lambda\Delta + \left(\frac{1}{2} - cb^{-2}\right)\Delta \cdot \frac{1}{4b^4}c\lambda\Delta + cb^{-2}\Delta \cdot cb^{-2}\Delta \leq \frac{3}{16b^4}c\lambda\Delta^2 + \frac{1}{b^4}c^2\Delta^2 < \frac{1}{5b^4}c\lambda\Delta^2 = \sum_{u \in W} |I'(u)|,$$

which is clearly a contradiction as the first and last term of the inequality chain are equal.

Now let us look at the period of the game process that contains all rounds $r$ where the load of $v$ fulfills $\ell_{r-1}(v) \geq \lambda b^{-1}\Delta$ and $\ell_r(v) < 2\lambda b^{-1}\Delta$. Denote by $I_v$ the set of colors assigned to good $v$-edges within this period and by $B_6(v)$ the bad event that there exists a set $W \in \mathcal{W}$ such that $|I_v \cap I_W| < \frac{1}{100b^2} \lambda\Delta$. In the following, we show that $\Pr[B_6(v)] = \exp(-\Omega(\Delta))$.

Consider a single round $r$ within this period and condition on that at round $r$, Maker colors a good $v$-edge. Then the color $i$ that Maker picks at random is added to the set $I_v$. Suppose that we have

$$\sum_{u \in \Gamma_{r-1}'(v)} \ell_{r-1}(u) \leq 9\lambda\Delta \cdot |\Gamma_{r-1}'(v)| \leq 9\lambda\Delta^2.$$
In this case, we infer from Markov’s inequality (Theorem 2.1) that there exist at most \( \frac{9}{16} \Delta \) vertices \( u \in \Gamma'_{r-1}(v) \) with the property \( \ell_{r-1}(u) \geq 14 \lambda \Delta \). Moreover,

\[
|\Gamma'_{r-1}(v)| \geq \deg(v) - 2\lambda b^{-1} \Delta \geq (1 - cb^{-4} - 2\lambda b^{-1}) \Delta \geq \frac{27}{28} \Delta,
\]

if \( c \) and \( \lambda \) are chosen sufficiently small. Hence at round \( r \), the vertex \( w \) that Maker chooses for his edge \( e = \{v, w\} \) uniformly at random in \( \Gamma'_{r-1}(v) \) satisfies \( \ell_{r-1}(w) \leq 14 \lambda \Delta \) with probability at least \( \frac{1}{3} \). If the random choice yields such a neighbor \( w \), it also follows

\[
|A_{r-1}(e)| \geq k - \ell_{r-1}(v) - \ell_{r-1}(w) \geq (2 - cb^{-4} - 16 \lambda) \Delta.
\]

Since Maker takes \( i \in A(e) \) uniformly at random and \( \lambda \) is chosen sufficiently small, we have

\[
\Pr[i \in I_W] \geq 1 - \frac{k - I_W}{|A_{r-1}(e)|} \geq 1 - \frac{3/2 - cb^{-4}}{2 - cb^{-4} - 16 \lambda} = \frac{1 - 32 \lambda}{4 - 2cb^{-4} - 32 \lambda} \geq \frac{1}{5}.
\]

We summarize that as long as we have

\[
\sum_{u \in \Gamma'_{r-1}(v)} \ell_{r-1}(u) \leq 9 \lambda \Delta \cdot |\Gamma'_{r-1}(v)|,
\]

a color \( i \) that is added to \( I_v \) at round \( r \) is also contained in \( I_W \) with probability at least \( \frac{1}{5} \cdot \frac{1}{5} = \frac{1}{15} \), independently of the success in previous rounds as we do not yet condition on any good or bad events concerning the actual time period.

Let \( m := \frac{1}{5b^2} \lambda \Delta \) and let \((X^W_1, X^W_2, \ldots)\) be an infinite 0/1-sequence where each entry is 1 independently with probability \( \frac{1}{15} \). We use the sequence \((X^W_i)_{i \geq 1}\) for a coupling as follows. Whenever Maker adds a color to \( I_v \) at a round \( r \) and \( \sum_{u \in \Gamma'_{r-1}(v)} \ell_{r-1}(u) \leq |\Gamma'_{r-1}(v)| \cdot 9 \lambda \Delta \), we read the next bit \( X^W_i \) of \((X^W_i)_{i \geq 1}\). Then the coupling is such that \( X^W_i = 1 \) implies \( i \in I_W \). Clearly

\[
\mu := \mathbb{E}\left[ \sum_{i=1}^{m} X^W_i \right] = \frac{m}{15} = \frac{1}{75b^2} \lambda \Delta,
\]

and by a Chernoff bound (Theorem 2.2 (ii)) we deduce

\[
\Pr\left[ \sum_{i=1}^{m} X^W_i \leq \frac{1}{100b^2} \lambda \Delta \right] \leq \Pr\left[ \sum_{i=1}^{m} X^W_i \leq \frac{3}{4} \mu \right] = \exp\left( - \frac{\mu}{\left( \frac{4^2}{2} \right)} \right) = \exp\left( - \frac{\lambda \Delta}{16 \cdot 2 \cdot 75b^2} \right).
\]
Next, we do a union bound over all sets \( W \in \mathcal{W} \). Using the inequality \( \binom{n}{k} \leq \left( \frac{ne}{k} \right)^k \) we obtain

\[
\Pr \left[ \bigwedge_{W \in \mathcal{W}} \left\{ \sum_{i=1}^{m} X_i^W \leq \frac{1}{100b^2} \lambda \Delta \right\} \right] \leq \left( \frac{\Delta}{cb^{-2}\Delta} \right) \cdot \exp \left( -\frac{\lambda \Delta}{16 \cdot 2 \cdot 75b^2} \right) \\
\leq \left( \frac{e}{cb^{-2}} \right)^{cb^{-2} \Delta} \cdot \exp \left( -\frac{\lambda \Delta}{2400b^2} \right) \\
= \exp(\Omega(-\Delta)),
\]

where the last step follows if \( c \) is chosen sufficiently small compared to \( \lambda \).

By Lemma 7.6 and Lemma 7.7 we have \( \Pr[\mathcal{B}_2(v) \cup \mathcal{B}_4(v)] = \exp(-\Omega(\Delta)) \). Hence with probability \( 1 - \exp(-\Omega(\Delta)) \) it holds \( |I_v| \geq \frac{1}{5b^2} \lambda \Delta \), and as long as \( \ell_r(v) < 2\lambda b^{-1} \Delta \), we also have

\[
\sum_{u \in \Gamma_r^r(v)} \ell_r(u) < 9\lambda \Delta \cdot |\Gamma_r^r(v)|.
\]

In this case, for all \( W \in \mathcal{W} \) the size of the set \( I_v \cap I_W \) is lower-bounded by \( \sum_{i=1}^{m} X_i^W \). By a union bound over all bad events, it follows that with probability \( 1 - \exp(-\Omega(\Delta)) \), all \( W \in \mathcal{W} \) satisfy

\[
|I_v \cap I_W| \geq \frac{1}{100b^2} \lambda \Delta.
\]

Therefore,

\[
\Pr[\mathcal{B}_6(v)] = \exp(-\Omega(\Delta)).
\]

Suppose now that the bad event \( \mathcal{B}_6(v) \) does not happen. Then for every set \( W \in \mathcal{W} \) we have

\[
\sum_{u \in W} |I'(u) \cap I_v| \geq \sum_{u \in W} |I'(u) \cap I_v \cap I_W| \geq \frac{1}{100b^2} \lambda \Delta \cdot \frac{1}{24b^4} c \lambda \Delta = \frac{1}{2400b^6} c \lambda^2 \Delta^2.
\]

Let \( s := \min\{r \mid \ell_r(v) \geq 2\lambda b^{-1} \Delta\} \). By an averaging argument, we see that for every \( W \in \mathcal{W} \) there exists a vertex \( u_W \in W \) such that

\[
|U_s(u_W) \cap U_s(v)| \geq |I'(u_W) \cap I_v| \geq \frac{1}{2400b^4} \lambda^2 \Delta > cb^{-4} \Delta,
\]

given that \( c \) is sufficiently small compared to \( \lambda \). Hence, for every set \( W \in \mathcal{W} \) there exists at least one vertex that does not belong to \( D(v) \), implying

\[
(7.2) \quad |D(v)| \leq cb^{-2} \Delta.
\]
Note that in the case $|D'(v)| < cb^{-2}\Delta$, (7.2) holds as well since $D(v) \subseteq D'(v)$.

Once having derived (7.2), we proceed by considering the rounds $r$ where $\ell_{r-1}(v) \geq 2\lambda b^{-1}\Delta$ and $\ell_r(v) < 3\lambda b^{-1}\Delta$. We want to show that within this period, Maker is fast enough to color all uncolored edges between $v$ and $D(v)$. Recall from Maker’s strategy that whenever he colors a good $v$-edge at a round $r$ and $D(v) \cap \Gamma_{r-1}'(v)$ is non-empty, with probability at least $q = \frac{6c}{\lambda}$ Maker chooses a vertex $w \in D(v)$ and colors the edge $\{v, w\}$. Again, we couple the process with an infinite 0/1-sequence $(X_1, X_2, \ldots)$ where each entry is 1 independently with probability $q$. Whenever $D(v) \cap \Gamma_{r-1}'(v)$ is non-empty and Maker is about to color a good $v$-edge at round $r$, we read the next bit $X_i$ of $(X_i)_{i \geq 1}$. If $X_i = 1$, we require that $w \in D(v)$. Recall that $m = \frac{1}{5b^{-1}}\lambda\Delta$. Clearly $\mu' := \mathbb{E}[\sum_{i=1}^{m} X_i] = m \cdot q \geq \frac{6}{5} cb^{-2}\Delta$. Denote by $\mathcal{B}_7(v)$ the bad event that $\sum_{i=1}^{m} X_i \leq cb^{-2}\Delta$. By a Chernoff bound it holds

$$\Pr[\mathcal{B}_7(v)] \leq \Pr \left[ \sum_{i=1}^{m} X_i \leq \frac{5}{6} \mu' \right] = \exp(-\Omega(\mu')) = \exp(-\Omega(\Delta)).$$

Assume $\mathcal{B}_3(v)$ and $\mathcal{B}_7(v)$ do not occur. Then in the considered period of the process where $\ell_r(v)$ increases from $2\lambda\Delta$ to $3\lambda\Delta$, Maker colors at least $m'$ good $v$-edges, implying that either $|D(v)| > cb^{-2}\Delta$ (which contradicts (7.2)) or Maker is fast enough and colors all edges between $v$ and $D(v)$ before the load of $v$ is above $3\lambda b^{-1}\Delta$. Hence, for all rounds $r$ such that $\ell_r(v) \geq 3\lambda b^{-1}\Delta$ we have $D(v) \cap \Gamma_{r-1}'(v) = \emptyset$, implying that indeed all $v$-edges can be colored properly.

We see that as long as for all $v \in V$ no bad event $\mathcal{B}_j(v)$ happens, Maker is never forced to use forbidden colors, meaning that the process yields a proper coloring of the complete edge set $E$. Finally, recall that we are assuming that the maximum degree $\Delta(G)$ is at least $C(b) \log n$. Then by a union bound we have

$$\Pr \left[ \bigvee_{v \in V} \bigvee_{j=1}^{7} \mathcal{B}_j(v) \right] \leq n \cdot \exp(-\Omega(\Delta)) = \exp(-\Omega(\Delta)) < 1$$

for $C$ sufficiently large. We conclude that (a) Maker wins with probability $1 - \exp(-\Omega(\Delta))$ when applying the proposed strategy and (b) Maker has a deterministic winning strategy. This finishes the main proof. ■
7.2.4 Missing Proofs

**Proof of Lemma 7.6.** Let $j \in \{1, 2, 3\}$ and let $v \in V$ be a vertex of degree at least $j\lambda\Delta$. We define $R = \{r_1, \ldots, r_{|R|}\}$ as the set of rounds satisfying

$$\ell_{r-1}(v) \geq (j-1)\lambda b^{-1}\Delta \quad \text{and} \quad \ell_r(v) < j\lambda b^{-1}\Delta - (b+1)$$

in which $v$-edges get colored by any of the two players. Maker's strategy is such that after every round $r_i \in R$, with non-zero probability Maker colors a good $v$-edge at round $r_i + 1$. Let $X(i)$ be the indicator random variable for this event. Furthermore, for $1 \leq i \leq |R|$ let $e(i) \in \{1, \ldots, b+1\}$ be the number of $v$-edges that have been colored by Maker and Breaker at round $r_i$. Note that the values $e(i)$ depend on Breaker's strategy, which may itself heavily depend on Maker's random answers in previous moves as Breaker might apply an adaptive strategy.

By definition of Maker's strategy, for all $1 \leq i \leq |R|$ we independently have

$$(7.3) \quad \Pr[X(i) = 1] \geq \frac{e(i)}{4b}. \quad \tag{7.3}$$

For all $1 \leq b' \leq b+1$ let $(Y_1^{b'}, Y_2^{b'}, \ldots)$ be an infinite 0/1-sequence where each entry is 1 independently with probability $\frac{b'}{4b}$. We use this set of 0/1-sequences for a coupling as follows. Whenever there is a new round $r_i \in R$, we read the next entry $Y_j^{e(i)}$ of the sequence $(Y_j^{e(i)})_{j \geq 1}$. We require that $Y_j^{e(i)} = 1$ implies $X(i) = 1$, i.e., Maker plays a good $v$-edge at round $r_i + 1$. By (7.3), this is a valid coupling. Let $1 \leq b' \leq b+1$ and $1 \leq m \leq \lambda b^{-1}\Delta$. We have $\mathbb{E}[\sum_{j=1}^{m} Y_j^{b'}] = \frac{b'm}{4b}$, and by a Chernoff bound,

$$\Pr \left[ \sum_{j=1}^{m} Y_j^{b'} < \frac{b'm}{4b} - \frac{\lambda\Delta}{25b^2(b+1)} \right] \leq \Pr \left[ \sum_{j=1}^{m} Y_j^{b'} < \left(1 - \frac{4\lambda\Delta}{25(b'm)b(b+1)}\right) \frac{b'm}{4b} \right]$$

$$\leq \exp \left(-\Omega \left( \frac{\Delta^2}{m} \right) \right) \leq \exp(-\Omega(\Delta)).$$

By a union bound, with probability $1 - \exp(-\Omega(\Delta))$, for all choices of $b'$ and $m$ it simultaneously holds

$$\sum_{j=1}^{m} Y_j^{b'} \geq \frac{b'm}{4b} - \frac{\lambda\Delta}{25b^2(b+1)}. \quad \tag{7.3}$$
Suppose this good event happens. For all $1 \leq b' \leq b + 1$, denote by $a(b')$ the total number of rounds $r_i \in R$ such that $e(i) = b'$. Clearly, the random variables $a(b')$ are upper-bounded by $\lambda b^{-1} \Delta$. Moreover, $\ell_{r_1}(v) \leq (j - 1)\lambda b^{-1} \Delta + (b + 1)$ and $\ell_{r|R_i}(v) \geq j\lambda b^{-1} \Delta - 2(b + 1)$, so

$$\sum_{b' = 1}^{b+1} b' \cdot a(b') \geq \lambda b^{-1} \Delta - 3(b + 1).$$

No matter how Breaker plays, it follows

$$\sum_{i=1}^{|R|} X(i) \geq \sum_{b' = 1}^{b+1} \sum_{j=1}^{a(b')} Y_j^{b'} \geq \sum_{b' = 1}^{b+1} \left( \frac{b' a(b')}{4b} - \frac{\lambda \Delta}{25b^2(b + 1)} \right) \geq \frac{\lambda \Delta}{4b^2} - 2 - \frac{\lambda \Delta}{25b^2} \geq \frac{\lambda \Delta}{5b^2}.$$

However, by construction $\ell_{r_{i+1}}(v) < j\lambda b^{-1} \Delta$ holds for all $r_i \in R$. Hence, the sum $\sum_{i=1}^{|R|} X(i)$ lower-bounds the total number of good $v$-edges in the considered period of the process, and with probability $1 - \exp(-\Omega(\Delta))$, Maker is sufficiently fast in coloring good $v$-edges.

**Proof of Lemma 7.7.** Let $v \in V$ be a vertex of degree at least $(1 - cb^{-4})\Delta$. We study how fast $\ell_r(v)$ grows compared to $L_r := \sum_{u \in \Gamma_r(v)} \ell_r(u)$ over time. First, note that whenever a player colors an edge $\{u, v\}$ at round $r$, we have $u \notin \Gamma_r'(v)$, and thus the edge $\{u, v\}$ does not contribute to $L_r$. Therefore, an edge $e = \{u, w\}$ that is played at round $r$ and contributes to $L_r$ is either (1) such that $u \in \Gamma_r(v)$ and $w \notin \Gamma_r(v) \cup \{v\}$, or (2) such that $u, w \in \Gamma_r(v)$. For a single edge $e$ of type (1), due to the proposed strategy, with probability at least $\frac{1}{4b}$ Maker answers by coloring a good $u$-edge in his next move at round $r + 1$, no matter whether $e$ was colored by Maker or Breaker. In this case, with probability at least $\frac{1-d}{\Delta}$ he colors the edge $\{u, v\}$. All together, for an edge of type (1) the probability that Maker’s next edge at round $r + 1$ increases the load of $v$ is at least $\frac{1-d}{4b\Delta}$. For an edge of type (2), the same argument yields that with probability at least $\frac{1-q}{2b\Delta}$, Maker answers by coloring a $v$-edge at round $r + 1$.

Let $R = \{r_1, \ldots, r_{|R|}\}$ be the set of rounds in which at least one edge of type (1) or (2) is played. Note that $|R| \leq \Delta^2$ is a random variable. For all $1 \leq i \leq |R|$ let $e_1(i)$ be the number of edges of type (1) played at round $r_i$, let $e_2(i)$ be the same for edges of type (2), and put $e(i) := e_1(i) + 2e_2(i) \in [1, \ldots, 2b + 2]$. Let $X(i)$
be the indicator random variable for the event that at round \( r_i + 1 \), Maker colors a \( v \)-edge. Then for all \( 1 \leq j \leq |R| \) the sum \( \sum_{i=1}^{j} X(i) \) lower-bounds \( \ell_{r_j+1} \), and by the previous observations we know that

\[
(7.4) \quad \Pr[X(i) = 1] = \mathbb{E}[X(i)] \geq \frac{(1 - q)\ell(i)}{4b\Delta}.
\]

For all \( 1 \leq b' \leq 2b + 2 \) let \((Y_1^{b'}, Y_2^{b'}, \ldots)\) be an infinite 0/1-sequence where each entry is 1 independently with probability \( \frac{(1 - q)b'}{4b\Delta} \). We build a coupling by using this set of 0/1-sequences. After every round \( r_i \in R \), we read the next bit \( Y_j^{e(i)} \) of the sequence \((Y_j^{e(i)})_{j \geq 1}\) and require that whenever \( Y_j^{e(i)} \) equals one, then at the next round \( r_i + 1 \), Maker plays a \( v \)-edge, implying \( X(i) = 1 \). By (7.4), indeed this coupling is possible. Let \( 1 \leq b' \leq 2b + 2 \) and \( 1 \leq m \leq \frac{17}{2} \lambda \Delta^2 \). We have \( \mathbb{E}[\sum_{j=1}^{m} Y_j^{b'}] = \frac{(1 - q)b'm}{4b\Delta} \), and by a Chernoff bound

\[
\Pr\left[ \sum_{j=1}^{m} Y_j^{b'} < \frac{(1 - q)b'm}{4b\Delta} - \frac{\lambda \Delta}{32b(b+1)} \right] \\
\leq \Pr\left[ \sum_{j=1}^{m} Y_j^{b'} < \left(1 - \frac{\lambda \Delta^2}{8(1 - q)(b'm)(b+1)}\right)(1 - q)b'm \frac{1}{4b\Delta} \right] \\
\leq \exp\left(-\Omega\left(\frac{\Delta^3}{m}\right)\right) \leq \exp(-\Omega(\Delta)).
\]

By a union bound, with probability \( 1 - \exp(-\Omega(\Delta)) \) the same holds for all choices of \( b' \) and \( m \) simultaneously.

We now always assume that this good event occurs. Let \( s \leq |R| \) be maximal such that

\[
(7.5) \quad \sum_{i=1}^{s} e(i) < \frac{17}{2} \lambda \Delta^2.
\]

We distinguish two cases. If \( s = |R| \), then for all rounds \( r \) of the game process we have

\[
(7.6) \quad L_r = \sum_{u \in \Gamma_{r}^+(v)} \ell_r(u) \leq \sum_{j=1}^{r} e(i) < \frac{17}{2} \lambda \Delta^2.
\]

In the case \( s < |R| \) we want to show that (7.6) holds at least for all rounds \( r \) where \( \ell_r(v) < 2\lambda b^{-1}\Delta \). For all \( 1 \leq b' \leq 2b + 2 \) let \( \alpha(b') \) count the number of rounds
\[ r_i \in R \text{ such that } i \leq s \text{ and } e(i) = b'. \text{ Since } e(s) \leq 2b + 2, \text{ inequality (7.5) implies} \]
\[ \sum_{b' = 1}^{2b+2} b' \cdot \alpha(b') \geq \frac{17}{2} \lambda \Delta^2 - (2b + 2). \]

Therefore,
\[ (7.7) \sum_{i=1}^{s} X(i) \geq \sum_{b' = 1}^{2b+2} \alpha(b') \sum_{j=1}^{b'} Y_j^{b'} \geq \sum_{b' = 1}^{2b+2} \left( (1 - q) \frac{b' \alpha(b)}{4b\Delta} - \frac{\lambda \Delta}{32b(b+1)} \right) \]
\[ \geq (1 - q) \frac{17\lambda \Delta}{8b} - \frac{1 - q}{\Delta} - \frac{\lambda \Delta}{16b}. \]

Recall that \( q = \frac{6c}{\lambda} \) and note that the \( v \)-edge that Maker eventually plays at round \( r_s + 1 \) does not contribute to \( \ell_{r_s}(v) \). Then for \( c \) sufficiently small and \( \Delta \) sufficiently large, (7.7) yields
\[ \ell_{r_s}(v) \geq \sum_{i=1}^{s} X(i) - 1 \geq 2\lambda b^{-1} \Delta. \]

On the other hand, by definition of \( s \) for all rounds \( r \leq r_s \) we have
\[ L_r = \sum_{u \in \Gamma'_r(v)} \ell_r(u) \leq \sum_{i=1}^{r} \alpha(i) < \frac{17}{2} \lambda \Delta^2. \]

We summarize that in both cases, for all \( r \) such that \( L_r \geq \frac{17}{2} \lambda \Delta^2 \) we also have \( \ell_r(v) \geq 2\lambda b^{-1} \Delta \). However, for all rounds \( r \) satisfying \( \ell_r(v) < 2\lambda b^{-1} \Delta \), the assumption \( \deg(v) \geq (1 - cb^{-4}) \Delta \) implies
\[ \frac{1}{|\Gamma'_r(v)|} \sum_{u \in \Gamma'_r(v)} \ell_r(u) \leq \frac{1}{(1 - 2\lambda b^{-1} - cb^{-4}) \Delta} L_r < \frac{17\lambda \Delta}{2(1 - 2\lambda b^{-1} - cb^{-4})} < 9\lambda \Delta, \]
given that \( \lambda \) and \( c \) are sufficiently small.

Hence, indeed with probability \( 1 - \exp(-\Omega(\Delta)) \) the load \( \ell_r(v) \) grows fast enough compared to the average load of the vertices in \( \Gamma'_r(v) \).

**Proof of Lemma 7.8.** Let \( v \) be any fixed vertex and consider the color sets
\[ \{I'(u) \mid u \in \Gamma(v)\}. \]

By definition, the sets \( I'(u) \) only contain colors that were chosen by Maker uniformly at random when coloring a good \( u \)-edge \( e \) at a round \( r \) such that
\( \ell_r(u) < \lambda b^{-1} \Delta \). Thus, when Maker is about to color such an edge \( e \), the set \( U_{r-1}(u) \) has size less than \( \lambda b^{-1} \Delta \). Then

\[
|A_{r-1}(e)| > k - \Delta - \lambda b^{-1} \Delta = (1 - cb^{-4} - \lambda b^{-1}) \Delta.
\]

(7.8)

Let \( W \subseteq \Gamma(u) \) be a fixed subset of size \( |W| = cb^{-2} \Delta \). For every color \( i \) we define \( \eta_i := |\{ u \in W \mid i \in I'(u) \}| \). Moreover, let \( I^- \) be any fixed set of \( cb^{-2} \Delta \) colors. Our plan is to show that it is unlikely that \( \eta_i \) is large for every color \( i \in I^- \). Let \( u \in W \). Whenever Maker is about to color a good \( u \)-edge \( e \) at a round \( r \) and the corresponding color will be added to \( I'(u) \), the probability that Maker chooses a color \( i \in I^- \) is maximal in a situation where \( I^- \subseteq A_{r-1}(e) \) but \( |A_{r-1}(e)| \) is as small as possible, i.e. the number of forbidden colors is as large as possible. If \( I^- \cap A_{r-1}(e) = \emptyset \), clearly the probability is zero. In general, by (7.8) the probability to hit a color of \( I^- \) is at most

\[
\frac{|I^- \cap A_{r-1}(e)|}{|A_{r-1}(e)|} \leq \frac{cb^{-2} \Delta}{(1 - cb^{-4} - \lambda b^{-1}) \Delta} = \frac{cb^{-2}}{(1 - cb^{-4} - \lambda b^{-1})}.
\]

We now want to upper-bound \( \eta_{I^-} := \sum_{i \in I^-} \eta_i \). We do a worst-case analysis and use a coupling where we assume that whenever Maker colors an edge \( e \) with some color \( i \) and the corresponding color will be contained in \( I'(u) \) for some \( u \in W \), the probability that \( i \in I^- \) is \( exactly \ cb^{-2}(1 - cb^{-4} - \lambda b^{-1})^{-1} \). We even allow that two neighboring edges get the same color \( i \in I^- \), which is fine regarding an upper bound of \( \eta_{I^-} \). The advantage of this coupling is that the probabilities for Maker choosing color of the set \( I^- \) in such rounds become \textit{independent}, which simplifies the analysis as we get rid of nasty dependencies and case distinctions. For a single vertex \( u \in W \) there are at most \( \frac{1}{5b \Delta} \lambda \Delta \) rounds in which Maker colors a good \( u \)-edge and adds a color to \( I'(u) \). Recall that \( |W| = cb^{-2} \Delta \). Hence in total, at most \( \frac{1}{5b} c \lambda \Delta^2 \) rounds have to be considered. It follows that \( \eta_{I^-} \) is upper-bounded by a random variable \( X \) with binomial distribution

\[
X \sim \text{Bin} \left( \frac{c \lambda \Delta^2}{5b^4}, \frac{cb^{-2}}{(1 - cb^{-4} - \lambda b^{-1})} \right).
\]

Clearly, the expected value of \( X \) is

\[
\mathbb{E}[X] = \frac{c^2 \lambda \Delta^2}{5b^6(1 - cb^{-4} - \lambda b^{-1})} \leq \frac{2}{9b^6} c^2 \lambda \Delta^2,
\]
and by a Chernoff bound we have
\[
\Pr\left[ X \geq \frac{1}{4b^6} c^2 \lambda \Delta^2 \right] \leq \Pr\left[ X \geq \left( 1 + \frac{1}{8} \right) \mathbb{E}[X] \right] = \exp(-\Omega(\Delta^2)).
\]
Thus, with probability \( 1 - \exp(-\Omega(\Delta^2)) \) it holds \( \eta_{I^-} \leq \frac{1}{4b^6} c^2 \lambda \Delta^2 \).

Regarding the values of the random variables \( \eta_i \), we observe that the event
\[
\bigwedge_{i \in I^-} \left\{ \eta_i \geq \frac{1}{4b^4} c \lambda \Delta \right\}
\]
can only occur when \( \eta_{I^-} \geq \frac{1}{4b^6} c^2 \lambda \Delta^2 \). Hence
\[
\Pr\left[ \bigwedge_{i \in I^-} \left\{ \eta_i \geq \frac{1}{4b^4} c \lambda \Delta \right\} \right] \leq \Pr\left[ \eta_{I^-} \geq \frac{1}{4b^6} c^2 \lambda \Delta^2 \right] \leq \exp(-\Omega(\Delta^2)).
\]

It remains to union bound over all choices of vertex sets \( W \) and color sets \( I^- \). We can assume that \( \deg(v) \geq cb^{-2} \Delta \), otherwise the statement is trivial. Using the inequality \( \binom{n}{k} \leq \left( \frac{Ne}{k} \right)^k \) we deduce
\[
\Pr[\mathcal{B}_5(v)] \leq \left( \frac{\deg(v)}{cb^{-2} \Delta} \right) \binom{k}{cb^{-2} \Delta} \exp(-\Omega(\Delta^2))
\]
\[
\leq \left( \frac{b^2 e}{c} \right)^{cb^{-2} \Delta} \left( \frac{2b^2 e}{c} \right)^{cb^{-2} \Delta} \exp(-\Omega(\Delta^2))
\]
\[
= \exp(-\Omega(\Delta^2)).
\]

### 7.3 Lower Bound for the Biased Game

The main idea for proving the lower bound of Theorem 7.5 is to use a reduction to Box games. For the formal definition of Box games we refer to Section 6.2. Let \( b \geq 2 \) be a bias for the game. We want to verify that there are graphs \( G = (V, E) \) that attain \( \chi'_g(G, b) = 2\Delta(G) - 1 \). In the following, we denote by \( d_G(e, f) \) the distance of two edges \( e, f \) in the graph \( G \) (i.e., the number of edges in the shortest path that connects \( e \) and \( f \)). A set \( F \subseteq E \) is called good if (i) for every edge \( f \in F \) its two endpoints have degree \( \Delta(G) \) in \( G \), and (ii) if for all \( f_i, f_j \in F \), \( d_G(f_i, f_j) \geq 4 \). We prove the following statement which is slightly stronger than Theorem 7.5.
**Lemma 7.9.** Let $b \geq 2$ and let $G = (V, E)$ be a graph with a good set $F \subseteq E$ such that

$$\frac{2\Delta(G) - 2}{b - 1} \leq \sum_{i=1}^{\lfloor |F| - 1 \rfloor} \frac{1}{i}.$$  

Then $\chi'_g(G, b) = 2\Delta(G) - 1$.

Let $G = (V, E)$ be a $\Delta$-regular graph with at least $2\Delta^3 \exp(\frac{2\Delta - 2}{b - 1})$ vertices. We then greedily find a good set $F \subseteq E$ as follows: choose an edge $f$ whose endpoints have both degree $\Delta$, put $f$ into $F$, delete every edge of $e \in E$ with $d_G(e, f) \leq 2$, and iterate as long as possible. Note that whenever we add an edge $f$ to $F$, so far no edge incident to $f$ has been removed because the endpoints of $f$ still have degree $\Delta$. So, for all edges $f'$ that are already included in $F$ we have $d_G(f, f') \geq 4$, and by induction, $F$ is a good set. Furthermore, whenever an edge $f$ is added to $F$ and edges of $E$ are deleted, we reduce the degree of at most $2\Delta^3$ vertices in $V$. In particular the number of vertices of degree $\Delta$ shrinks by at most $2\Delta^3$ per iteration and we obtain a set $F$ of size at least $\exp(\frac{2\Delta - 2}{b - 1})$. This implies

$$\frac{2\Delta - 2}{b - 1} \leq \log |F| \leq \sum_{i=1}^{\lfloor |F| - 1 \rfloor} \frac{1}{i},$$

where the last inequality holds for all integers $|F| > 0$. We see that indeed, Theorem 7.5 is a corollary of Lemma 7.9.

**Proof of Lemma 7.9.** Let $F = \{f_1, \ldots, f_s\} \subseteq E$ be a good set of a graph $G$ that satisfies (7.9), and consider the edge coloring game played with colors $\{1, \ldots, k\}$, where $k < 2\Delta(G) - 1$. We want to show that Breaker has a strategy such that at least one edge $f_i \in F$ runs out of available colors before it gets colored. Let $F' := \bigcup_{i=1}^{s} \Gamma(f_i)$, where $\Gamma(f_i)$ denotes the set of neighboring edges of $f_i$, i.e., the set of edges that share an endpoint with $f_i$. In the following we require Breaker to only color edges of $F'$, as long as possible. Moreover, we assume that whenever Breaker colors a neighbor of some $f_i \in F$, he uses a color that was so far not used at any neighboring edge of $f_i$, if possible.

The reduction from the coloring game to Box games now works as follows: for every edge $f_i \in F$ we introduce a box $A_i$, containing precisely $k$ elements. Whenever Breaker colors an edge of $F'$ that is a neighbor of some edge $f_i \in F$, in
the Box game we model this by Alice claiming an element of \( A_i \). Breaker’s right to color at most \( b \) edges per turn is mapped to the rule that in the Box game, Alice is allowed to claim up to \( q \) elements per turn. Furthermore, whenever Maker colors an edge \( e \in E \), we couple this by Bob playing an element of a box \( A_i \), where we choose \( i \) such that

\[
(7.10) \quad d_G(f_i, e) = \min_{1 \leq j \leq s} \{d_G(f_j, e)\}.
\]

Hence, as long as Breaker colors edges of \( F' \), for any strategy of Maker we can interpret the game process as Alice and Bob playing the Box game \( B(A_1, \ldots, A_s; b) \). Since \( F \) fulfills (7.9), we have

\[
\sum_{i=1}^{s} |A_i| \leq s(2\Delta(G) - 2) \leq s(b - 1) \sum_{i=1}^{s-1} \frac{1}{i}
\]

By Lemma 6.6 (ii) Alice has a winning strategy for this Box game, meaning that she is able to claim all \( k \) elements of at least one box \( A_i \) before Bob can claim one element of \( A_i \). Then our coupling implies that in the coloring game, Breaker has a strategy such that for at least one \( f_i \in F \), he can color \( k \) neighbors of \( f_i \) before Maker colors any edge \( e \) fulfilling (7.10) for \( f_i \). As \( F \) is a good set, this means that Maker colored no neighboring edge of \( f_i \), i.e., he never blocked a color for a neighboring edge of \( f_i \). But then, due to the applied strategy, Breaker was able to use all \( k \) colors exactly once when coloring the \( k \) neighbors of \( f_i \). Afterwards, for \( f_i \) clearly all colors are forbidden and Breaker wins the edge coloring game with bias \( b \) on the graph \( G \).

\[\blacksquare\]

### 7.4 Open Problems

With Theorem 7.3 we made a first step towards a proof of Conjecture 7.2. We verified the statement for all graphs \( G \) that satisfy \( \Delta(G) \geq C \log n \) by applying a random strategy for Maker. Our attempts to prove the full conjecture were not successful, neither by using the same strategy nor by analyzing more advanced and refined strategies. It is reasonable to believe that from Maker’s perspective, the game is harder to win in the case \( \Delta(G) \leq C \log n \), as indicated by Theorem 7.4
and Theorem 7.5 for the biased version of the game where the behavior of $\chi'_g(G, b)$ significantly changes around $\Delta(G) \approx \log n$.

In [Bev+08] it is also conjectured that there exist $c, d_0 > 0$ such that every graph $G$ with minimum degree $\delta(G) \geq d_0$ satisfies $\chi'_g(G) \geq (1 + c)\Delta(G)$. The interesting case of this statement is when $G$ is almost regular, i.e., $\Delta(G) \leq (1 + c)\delta(G)$. Note that so far, this conjecture is not even solved for examples like complete graphs. Another open question is to decide whether there exist $c, \Delta_0$ such that for any $\Delta \geq \Delta_0$, there are two $\Delta$-regular graphs $G_1$ and $G_2$ with $|\chi'_g(G_1) - \chi'_g(G_2)| \geq c\Delta$.

Finally, in order to gain a better understanding of the game process it would be desirable to determine the asymptotic expression of the game chromatic index at least for complete graphs, random graphs, or complete bipartite graphs.
The topic of this last chapter are colorability saturation games. Two players, called Mini and Maxi, play the saturation game with the rule that the chromatic number of the graph does not exceed $k$, for a given parameter $k \in \mathbb{N}$. The analysis of colorability saturation games has been initiated recently by Hefetz, Krivelevich, Naor, and Stojaković [Hef+16]. By using a randomized strategy for Maxi, they obtained a lower bound on the score of the game. In this chapter, we improve their results by providing almost matching lower and upper bounds on the score of the game for arbitrary choices of $k$ and $n > k$. In addition, we study the specific game with $k = 4$ in more details and prove that its score is $n^2/3 + O(n)$. This chapter is single-author work [Keu16].

### 8.1 Introduction

One of the most classic problems in extremal graph theory is to determine how many edges a graph on $n$ vertices can have without fulfilling a given monotone property $\mathcal{P}$. In this context, we say that a graph $G$ is saturated with respect to
if $G$ does not satisfy $\mathcal{P}$, but adding any additional edge $e \in \left(\binom{n}{2}\right) \setminus E$ to the graph results in $G \cup \{e\}$ satisfying $\mathcal{P}$. The Turán number $ex(n, \mathcal{P})$ is then the maximal number of edges that a $\mathcal{P}$-saturated graph on $n$ vertices can have. On the other hand, the saturation number $sat(n, \mathcal{P})$ denotes the minimal number of edges that a graph $G$ on $n$ vertices can have while being saturated w.r.t. $\mathcal{P}$. For a general survey on saturation numbers see [FFS11].

Saturation games are a class of combinatorial games that are closely related to saturated graphs. For a given monotone graph property $\mathcal{P}$, the saturation game is played as follows: two players Maxi and Mini start with the empty graph on $n$ vertices. They take turns, always extending the current graph $G$ with some additional edge $e$ such that $G \cup \{e\}$ does not satisfy $\mathcal{P}$. At some point, every free edge is forbidden, i.e., the obtained graph $G_{\text{end}}$ is saturated w.r.t. $\mathcal{P}$, and the game stops. Mini aims to minimize the number of edges in $G_{\text{end}}$ (that is, Mini wants that the game is over as soon as possible), while Maxi’s goal is to maximize the number of edges in $G_{\text{end}}$. The score of the game, denoted by $s(n, \mathcal{P})$, is the total number of edges in $G_{\text{end}}$ when both players apply optimal strategies. In the analysis of saturation games, we aim at finding significant lower and upper bounds on the score or ideally determining the score exactly.

In general, the score can depend on the identity of the first player. However, in this chapter we don’t specify who starts as all statements hold for both cases. Clearly, for every monotone property $\mathcal{P}$ we have

$$sat(n, \mathcal{P}) \leq s(n, \mathcal{P}) \leq ex(n, \mathcal{P}),$$

which connects saturation games to the well-studied saturation and Turán numbers of graphs. We see that if the saturation and the Turán numbers of the studied property are the same (for example if $\mathcal{P} =$“being non-planar” or $\mathcal{P} =$“having independence number at most $k$”), the score $s(n, \mathcal{P})$ is directly determined by the two numbers.

In the last quarter-century, it turned out that analyzing saturation games is both interesting and challenging. The two players not only want to follow their own strategy and play against the adversary at the same time, but moreover the two players also have opposing goals, which makes the game intriguing and
intricate. Note that the goal of a player is to create a certain graph structure which ensures a short (resp. long) game. The more extreme this structure is, the easier the opponent can play against it. But the weaker the structure is, the less our player gains. Often, it is much easier to destroy the opponent’s structure than to create the own desired structure. Hence a good strategy should (1) be resistant against attacks, and (2) make sure that $G_{\text{end}}$ will be sufficiently sparse (resp. dense). Therefore, finding optimal or almost optimal strategies for saturation games can be surprisingly hard and often requires tedious case distinctions. This is why the asymptotic value of the score is only known for a few particular games.

Let us briefly summarize the most important examples and existing results. Let $\mathcal{C}_k$ be the property of being $k$-connected and spanning, and let $\mathcal{PM}$ be the property of possessing a perfect matching. Carraher, Kinnersley, Reiniger, and West proved $s(n, \mathcal{C}_1) = \left(\frac{n^2}{2}\right) + 1$ for the connectivity game [Car+17] with $n \geq 6$. Hefetz et al. [Hef+16] generalized this result and asserted the bound $s(n, \mathcal{C}_k) \geq \left(\frac{n}{2}\right) - 5kn^{3/2}$. In the same paper, they proved $s(n, \mathcal{PM}) \geq \left(\frac{n^2}{2}\right)$ and further results on matching games. Additional saturation games have been studied by Lee and Riet [LR15], as well as variants on directed graphs [LR14].

Another typical example of saturation games is the triangle-free game. Here, the considered monotone property is $\mathcal{K}_3$, i.e., containing a triangle as a subgraph. It is well-known that $ex(n, \mathcal{K}_3) = \lceil n^2/4 \rceil$ and $sat(n, \mathcal{K}_3) = n - 1$ (see [EHM64], e.g.). In [FRS91] and [Ser92], Füredi, Reimer, and Seress proved a lower bound of $(\frac{1}{2} + o(1))n \log_2 n$ on the score of this game, and cite Erdős who has given an upper bound of $n^2/5$ in personal communication. However, the proof of this upper bound is lost and could not be retrieved until today. The currently best-known upper bound is $\frac{26}{121}n^2 + o(n^2)$ by Bíró, Horn and Wildstrom [BHW16], a small improvement compared to the trivial upper bound. Closing the large gap between $\Omega(n \log n)$ and $O(n^2)$ is still an open problem.

We now turn to the topic of this chapter and focus on the property

$$\chi_{>k} = \text{“having chromatic number at least } k + 1\text{”}.$$
current graph $k$-colorable. Note that every graph that is saturated w.r.t. $\chi > k$ is a complete $k$-partite graph. Hence, the game is about settling the number of edges of this final, $k$-partite graph $G_{end}$. Clearly, the total number of edges in a complete $k$-partite graph $G$ with partition sizes $n_1, \ldots, n_k$ is $(\frac{n}{2}) - \sum_{i=1}^{k} \left(\frac{n_i}{2}\right)$. Then the well-known Turán number

$$ex(n, \chi > k) = (1 - 1/k + o(1))\left(\frac{n}{2}\right)$$

and the saturation number

$$sat(n, \chi > k) = (k - 1)(n - 1) - \left(\frac{k - 1}{2}\right)$$

give us first bounds on the score of the colorability saturation game.

Let us start by describing the case $k = 2$ where Maxi and Mini are forced to keep the graph bipartite. If $n$ is even, it is not difficult to observe that Maxi can play such that after each of her moves, every component of the current bipartite graph is balanced, except the isolated vertices, and $G_{end}$ will be perfectly balanced. For the general case, this argument implies

$$s(n, \chi > 2) = ex(n, \chi > 2) = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

A formal proof is provided in [Car+17]. We see that Mini has no power in this particular saturation game. However, things get more interesting and involved as soon as $k > 2$. Hefetz, Krivelevich, Naor, and Stojaković [Hef+16] proved

$$s(n, \chi > 3) \leq \frac{21}{64} n^2 + O(n),$$

revealing that Mini now has some influence on the game process. Furthermore, in the same paper they introduced a randomized strategy for Maxi that leads to a general lower bound on the score of colorability saturation games.

**Theorem 8.1** (Theorem 1.5 in [Hef+16]). *There exists a constant $C > 0$ such that for every $k \in \mathbb{N}$ and every $n \in \mathbb{N}$ that is sufficiently large compared to $k$ it holds*

$$s(n, \chi > k) \geq \left(\frac{n}{2}\right) \left(1 - \frac{C \log k}{k}\right).$$
Unfortunately, the proof of this result uses a relatively large constant $C$ (e.g., $C > 3000$), making the bound trivial for small choices of $k$.

As main contribution of this chapter, we provide almost matching lower and upper bounds on the score $s(n, \chi_{>k})$ which also demonstrate how $s(n, \chi_{>k})$ asymptotically depends on the parameter $k$. In contrast to Theorem 8.1, our results are non-trivial for small choices of $k$ as well and therefore enhance the intuitive understanding of the game process. The first result is a general lower bound on the score and improves Theorem 8.1.

**Theorem 8.2.** Let $k \in \mathbb{N}$ and $n > k$. Then

$$s(n, \chi_{>k}) \geq \binom{n}{2} \left(1 - \frac{1}{\lceil k/2 \rceil}\right) \geq \binom{n}{2} \left(1 - \frac{2}{k}\right).$$

Note that for $k = k(n)$ and $n \to \infty$, Theorem 8.2 and the Turán number $ex(n, \chi_{>k})$ together imply

$$s(n, \chi_{>k}) = \binom{n}{2} - O\left(\frac{n^2}{k}\right).$$

Next, we show a general upper bound which proves that if Mini follows an optimal strategy, the number of missing edges at the end of the game is by a constant factor larger than in a balanced complete $k$-partite graph.

**Theorem 8.3.** Let $k \in \mathbb{N}$ and $n > k$. Then

$$s(n, \chi_{>k}) \leq \binom{n}{2} \left(1 - \frac{1}{k - \lfloor (k-1)/3 \rfloor}\right) + n \leq \binom{n}{2} \left(1 - \frac{3}{2k + 3}\right) + n.$$

In particular, if $k$ is fixed and $n \to \infty$, then

$$s(n, \chi_{>k}) \leq \binom{n}{2} \left(1 - \frac{3}{2k + 3} + o(1)\right).$$

The provided lower and upper bounds on $s(n, \chi_{>k})$ are matching up to a small constant factor in the term that counts the missing edges of the final graph. It remains an interesting problem to determine the correct constant.

In addition, we investigate the specific game with parameter $k = 4$ where it turns out that the upper bound of Theorem 8.3 is tight.
Theorem 8.4. Let \( n \geq 5 \). Then \( s(n, \chi_{>4}) = n^2/3 + O(n) \).

We prove Theorem 8.2 and Theorem 8.3 by using carefully chosen potential functions that are closely related to the density of induced subgraphs. We then define the strategies in terms of these potentials in a general, abstract way such that we need to deal only with a reasonable number of case distinctions. Altogether, this is a novel approach for the analysis of saturation games.

We start by introducing notations and describing general aspects of our proofs strategies on a high level in Section 8.2. Then in Section 8.3 we provide a general strategy for Maxi and use it to prove Theorem 8.2. In Section 8.4 we investigate the game from Mini’s perspective and verify Theorem 8.3. Afterwards we study the special case \( k = 4 \) in Section 8.5 and prove Theorem 8.4 by using a more specific strategy for Maxi. Finally the last section contains some concluding remarks and open problems.

8.2 Preliminaries

Let \( k \) and \( n \) be two integers such that \( n > k \). We study the saturation game on a set \( V \) of \( n \) vertices w.r.t. the monotone property \( \chi_{>k} \). The game is considered as a process evolving in time where \( G(t) = (V, E_t) \) denotes the graph at the moment where Mini and Maxi have played \( t \) edges in total. Note that we start with the empty graph \( G(0) \). The game stops at time \( t_{end} \) and thus ends with a graph \( G_{end} := G(t_{end}) \) which is a complete \( k \)-partite graph. It turns out that in all our proofs, the game is partitioned into two phases. When analyzing the game from the perspective of one specific player, we always have a first phase in which our player wants to create a certain graph structure that is suitable for her goal of forcing \( G_{end} \) to be either sparse or dense. Once the desired graph structure is present, we enter the second phase where we always allow our player to play arbitrarily until the graph is saturated.

We will define the strategies via potential functions. Informally speaking, we measure the progress of “our” player by a function \( f : V \times \mathbb{N}_0 \rightarrow \mathbb{N}_0 \). Then the
progress of the player in a set $A \subseteq V$ at time $t$ is given by

$$f(A, t) := \sum_{v \in A} f(v, t).$$

We do not yet specify $f$ as the choice of the concrete function depends on the identity of the player and on further notations. In order to quantify the progress of the opponent, we introduce the following notation. Let $A$ and $B$ be two disjoint subsets of $V$ and let $t$ be an integer. We define

$$\phi(A, B, t) = |E_t(A)| + |E_t(A, B)|. \tag{8.1}$$

That is, we count the number of edges in the graph $G(t)$ that are either contained in the subgraph induced by $A$ or in the cut between $A$ and $B$. Suppose there exists a vertex set $B$ on which our player has already created her desired structure. Then it turns out that for a set $A \subseteq V \setminus B$, the function $\phi(A, B, t)$ is suitable for measuring the progress of the opponent on set $A$. The goal of our player is now the following: for every subset $A \subseteq V \setminus B$, her own pace should be at least as fast as the pace of her opponent. Hence, she aims to play such that at time $t$, $f(A, t) \geq \phi(A, B, t)$ holds for all $A \subseteq V \setminus B$.

After introducing the most important notations, we provide a criterion for a graph being $k$-colorable. Recall that the $k$-core of a graph $G$ denotes the largest induced subgraph where every vertex has degree at least $k$. Clearly, if the $k$-core is $k$-colorable, then $G$ itself is $k$-colorable because we can take a proper vertex coloring of the $k$-core and extend it vertex by vertex to the whole graph. The following easy lemma is a small modification of this fact. We use it in our proofs whenever we want to verify that an edge proposed by a strategy can indeed be inserted to the graph without violating the colorability constraint.

**Lemma 8.5.** Let $k \in \mathbb{N}$, let $G = (V, E)$ be a graph, and let $B \subseteq V$. Suppose that the induced subgraph $G[B]$ is $k$-colorable and that for every non-empty subset $A \subseteq V \setminus B$ it holds $2|E(A)| + |E(A, B)| < k \cdot |A|$. Then $G$ is $k$-colorable.

**Proof.** Let $A' := V \setminus B$. By assumption it holds

$$\sum_{v \in A'} \deg(v) = 2|E(A')| + |E(A', B)| < k \cdot |A'|.$$
Hence there exists a vertex \( v_1 \in A' \) with degree at most \( k - 1 \). Next, we apply the same argument for the set \( A'' := V \setminus (B \cup \{v_1\}) \), and afterwards we iterate the argument for all remaining vertices to find an ordering of \( V \) (with \( v_1 \) being the last vertex) where all vertices of \( A' \) have back-degree at most \( k - 1 \). Hence, if \( G[B] \) is \( k \)-colorable, we can take an arbitrary vertex coloring of \( G[B] \), use the ordering of the vertices, and extend the coloring vertex by vertex to the whole graph \( G \) since there is always at least one available color.

\[ \blacksquare \]

### 8.3 Lower Bound

In this section we prove Theorem 8.2. We thus provide a strategy for Maxi which ensures that the game process lasts sufficiently long and \( G_{\text{end}} \) becomes sufficiently dense. The main idea is the following: Maxi aims to create a collection of vertex-disjoint cliques that cover the entire vertex set \( V \). Then every independent set of \( G_{\text{end}} \) can contain at most one vertex per clique, yielding an upper bound on the independence number of \( G_{\text{end}} \) and thus a lower bound on the number of edges of the complete \( k \)-partite graph \( G_{\text{end}} \). Theorem 8.2 follows directly from the following lemma.

**Lemma 8.6.** Let \( n \in \mathbb{N} \), \( m \geq 2 \), and \( k \geq 2m - 1 \). Then in the colorability saturation game with \( k \) colors and \( n \) vertices, Maxi has a strategy such that in \( G_{\text{end}} \), the vertex set is covered by \( \lceil \frac{n}{m} \rceil \) vertex-disjoint cliques, all having size at most \( m \).

**Proof of Theorem 8.2.** First, we observe that for \( k \in \{1, 2\} \) the claimed lower bound equals zero and is trivial, so we can assume \( k \geq 3 \). Let \( n > k \), and put \( m := \lceil \frac{k}{2} \rceil \). \( G_{\text{end}} \) is a complete \( k \)-partite graph and contains all \( \binom{n}{2} \) possible edges except those where both incident vertices are contained in the same partition. By Lemma 8.6, Maxi has a strategy such that there are \( \lceil \frac{n}{m} \rceil \) vertex-disjoint cliques as induced subgraphs, covering \( V \) completely, where every clique contains at most \( m \) vertices. Clearly, all vertices of such a clique belong to different color classes of the saturated graph \( G_{\text{end}} \), therefore every independent set of \( G_{\text{end}} \) has size at most \( \lceil \frac{n}{m} \rceil \).
We claim that the number of missing edges is maximal if there are \( \left\lceil \frac{n}{m} \right\rceil \) color classes of size \( \left\lceil \frac{n}{m} \right\rceil \) and \( k - \left\lceil \frac{n}{m} \right\rceil - 1 \) classes of size 1. Indeed, for every other configuration we could move one vertex from a smaller color class to a larger color class and thereby increase the number of forbidden edges. Hence, the total number of missing edges in \( G_{end} \) is at most
\[
\frac{n-k}{\left\lceil \frac{n}{m} \right\rceil} \cdot \left( \frac{n}{m} \right) + \frac{(n-k)(n+m)}{2m} \leq \left( \frac{n}{2} \right) \frac{1}{m},
\]
proving the statement. ■

Before the game starts, Maxi partitions the vertex set \( V \) into disjoint sets \( V_1, \ldots, V_{\left\lceil \frac{n}{m} \right\rceil} \), where \( |V_i| = m \) holds for all \( i \leq \frac{n}{m} \). Her goal is to play such that in \( G_{end} \), each group \( V_i \) induces a clique of size \( |V_i| \). For every \( 1 \leq i \leq \left\lceil \frac{n}{m} \right\rceil \), every \( v \in V_i \), and every \( t \in \mathbb{N}_0 \) we define
\[
\alpha(v, t) := |V_i \cap \Gamma_t(v)|,
\]
where \( \Gamma_t(v) \) denotes the neighborhood of \( v \) in \( G(t) \). Furthermore, for every \( A \subseteq V \) and every \( t \in \mathbb{N}_0 \) we put \( \alpha(A, t) := \sum_{v \in A} \alpha(v, t) \). Clearly, for all \( v \in V \) we start with \( \alpha(v, 0) = 0 \), before the \( \alpha \)-values start to increase during the game. At time \( t \), we call a vertex \( v \in V_i \) full if \( \alpha(v, t) = |V_i| - 1 \), i.e., if \( v \) is connected to all other vertices of its set \( V_i \).

We want to use the function \( \phi(A, B, t) \) as defined in (8.1) to measure the progress of Mini in subsets \( A \subseteq V \setminus B \). The plan is to compare the \( \alpha \)-values with the \( \phi \)-values in order to determine where Maxi should insert her next edge. For all points in time \( t \in \mathbb{N}_0 \) and for all \( B \subseteq V \) we put
\[
\mathcal{D}(B, t) := \{ A \subseteq V \setminus B \mid \alpha(A, t) < \phi(A, B, t) \}.
\]
Informally speaking, the set \( \mathcal{D}(B, t) \) contains all dangerous subsets \( A \) where Mini made more progress than Maxi until time \( t \). Clearly, for all choices of \( B \) we start with \( \mathcal{D}(B, 0) = \emptyset \). Finally, we put
\[
A_0(B, t) := \bigcap_{A \in \mathcal{D}(B, t)} A.
\]
So far, we didn’t specify how we pick the set $B$. The concrete choice of $B$ depends on the game process, varies during the evolution of the game, and is quite subtle. We start with $B(0) = 0$. Afterwards, for all points in time $t$ where Mini is playing we let $B(t) = B(t - 1)$. Whenever Maxi is about to play at time $t$, before her turn we first define the set $B(t)$ according to the following “update rule”.

(U1) If all vertices of $V \setminus B(t - 1)$ are full in $G(t - 1)$, we put $B(t) := V$. If $V \setminus B(t - 1)$ contains non-full vertices, $\mathcal{D}(B(t - 1), t - 1)$ is non-empty, and all vertices of $A_0(B(t - 1), t - 1)$ are full in $G(t - 1)$, we put $B(t) := B(t - 1) \cup A_0(B(t - 1), t - 1)$. In all other cases, we let $B(t) := B(t - 1)$.

Clearly, for all points in time $t$ the set $B(t)$ only contains vertices that are full in $G(t - 1)$. (But not necessarily all of them!) As we will see later, it turns out that the rule (U1) guarantees that either the set $\mathcal{D}(B(t), t - 1)$ is empty or $A_0(B(t), t - 1)$ contains vertices that are not yet full. We now continue by providing Maxi’s strategy for playing her edge at time $t$.

(S1) If $B(t) = V$ but $G(t - 1)$ is not yet saturated, insert an arbitrary edge such that $G(t)$ is $k$-colorable.

(S2) If $\mathcal{D}(B(t), t - 1)$ is empty and $B(t) \neq V$, let $v \in V \setminus B(t)$ be a vertex which is not yet full in $G(t - 1)$. Insert a new edge $\{u, v\}$, where we require that $u$ is contained in the same set $V_i$ as $v$.

(S3) If $\mathcal{D}(B(t), t - 1)$ is non-empty, let $v \in A_0(B(t), t - 1)$ be a vertex which is not yet full in $G(t - 1)$. Insert a new edge $\{u, v\}$, where we require that $u$ is contained in the same set $V_i$ as $v$.

A priori, it is not evident that this strategy is well-defined. Amongst others, we have to verify that the rules (S1)-(S3) cover all cases and that the set $A_0(B(t), t - 1)$ considered in (S3) is non-empty. While proving Lemma 8.6 below we will also see that the proposed strategy is indeed well-defined. Before starting with the proof we first state one additional technical lemma whose proof is deferred to the end of this section.
Lemma 8.7. Let $f : V \to \mathbb{N}_0$ be a function, let $B \subset V$, and for all $A \subseteq V \setminus B$ and $t \in \mathbb{N}_0$ let $\phi(A, B, t)$ be defined as in (8.1). Let $t_0 \in \mathbb{N}_0$ be a point in time of the game process such that $G(t_0)$ is not saturated and such that for all $A \subseteq V \setminus B$ it holds $f(A) := \sum_{v \in A} f(v) \geq \phi(A, B, t_0)$. Then either the set

$$C := \{ A \subseteq V \setminus B \mid f(A) < \phi(A, B, t_0 + 1) \}$$

is empty, or the following statements are true.

(i) $A' := \bigcap_{A \notin C} A$ is non-empty and itself contained in $C$.

(ii) For all $A \in C$ it holds $\phi(A, B, t_0 + 1) = f(A) + 1$.

(iii) For all $A \subseteq V \setminus (B \cup A')$ we have

$$f(A) \geq \phi(A \cup A', B, t_0 + 1) - \phi(A', B, t_0 + 1).$$

Proof of Lemma 8.6. Let $n \in \mathbb{N}$, let $m \geq 2$, and let $k \geq 2m - 1$. Suppose Maxi applies the proposed strategy. We prove by induction that as long as $G(t − 1)$ is not saturated or there exists at least one vertex that is not full in $G(t − 1)$, for Maxi’s move at time $t$ the following invariants hold.

(I1) Exactly one rule of (S1)-(S3) can be applied.

(I2) Let Maxi insert the desired edge, regardless whether $G(t)$ is $k$-colorable or not. Then the set $\mathcal{D}(B(t), t)$ is empty.

(I3) Maxi can play her edge without violating the colorability constraint.

(I4) If there is a non-full vertex $v \in V \setminus B(t)$ in the graph $G(t)$, then $G(t)$ is not yet saturated.

Recall that we start with $B(0) = \emptyset$ and $\mathcal{D}(B(0), 0) = \emptyset$. Let $t$ be a point in time such that $G(t − 1)$ is not saturated or there exists a non-full vertex in $G(t − 1)$, and assume that either it is Maxi’s first move (providing the base case) or by induction (I1)-(I4) were true for all previous moves of Maxi. We first check property (I1). If every vertex is full in $G(t − 1)$, then by (U1) we have $B(t) = V$ and Maxi applies
(S1) for the remainder of the game. So let us assume that not every vertex is full in \(G(t - 1)\). If the set \(\mathcal{D}(B(t - 1), t - 1)\) is empty, we have \(B(t) = B(t - 1)\) and (S2) is matching. So we can assume that \(\mathcal{D}(B(t - 1), t - 1)\) is non-empty. Then \(t > 1\), and either by induction we have \(\mathcal{D}(B(t - 2), t - 2) = \emptyset\), or \(t = 2\) and \(\mathcal{D}(B(0), 0) = \emptyset\). Hence we can apply Lemma 8.7 with \(f(v) = \alpha(v, t - 2)\), \(B = B(t - 2)\), and \(t_0 = t - 2\). For these choices, the set \(\mathcal{C}\) as defined in (8.2) contains all subsets \(A \subseteq V \setminus B(t - 2)\) such that \(\alpha(A, t - 2) < \phi(A, B(t - 2), t - 1)\). Since \(\alpha(v, t - 1) \geq \alpha(v, t - 2)\) holds for every vertex \(v \in V\), we have

\[\emptyset \neq \mathcal{D}(B(t - 1), t - 1) \subseteq \mathcal{C}.\]

So \(\mathcal{C}\) is non-empty too. By statement (i) of Lemma 8.7, \(A' = \bigcap_{A \in \mathcal{C}} A\) is non-empty and itself a member of \(\mathcal{C}\), i.e., \(\alpha(A', t - 2) < \phi(A', B(t - 2), t - 1)\). Combining this fact with the assumption \(\mathcal{D}(B(t - 1), t - 1) \neq \emptyset\) and with statement (ii) of Lemma 8.7, we see that for all \(v \in A'\) it holds \(\alpha(v, t - 1) = \alpha(v, t - 2)\). Hence \(A' \in \mathcal{D}(B(t - 1), t - 1)\) and, moreover, \(A' = A_0(B(t - 1), t - 1)\). Then the update rule (U1) is well-defined.

When applying (U1), we define the set \(B(t)\). In case \(B(t) = B(t - 1)\), we have \(\mathcal{D}(B(t), t - 1) = \mathcal{D}(B(t - 1), t - 1)\) and \(A_0(B(t), t - 1) = A_0(B(t - 1), t - 1)\), so (S3) can be applied. It remains the case where \(\mathcal{D}(B(t - 1), t - 1)\) is non-empty and all vertices \(v \in A_0(B(t - 1), t - 1) = A'\) are full in \(G(t - 1)\) but there exist still non-full vertices in \(V\). By Lemma 8.7 (iii), for all \(A \subseteq V \setminus (B(t - 1) \cup A')\) we deduce

\[\alpha(A, t - 1) \geq \alpha(A, t - 2) \geq \phi(A \cup A', B(t - 1), t - 1) - \phi(A', B(t - 1), t - 1).\]

Hence, after setting \(B(t) = B(t - 1) \cup A_0(B(t - 1), t - 1)\), for every set \(A \subseteq V \setminus B(t)\) it holds \(\alpha(A, t - 1) \geq \phi(A, B(t), t - 1)\). So \(\mathcal{D}(B(t), t - 1)\) becomes empty and (S2) can be applied. This proves invariant (I1).

For invariant (I2), we only have to consider (S2) and (S3), because (S1) is applied when it already holds \(B(t) = V\). Maxi now inserts the edge \(e = \{u, v\}\). In case she applies (S2), we already know that \(\mathcal{D}(B(t), t - 1)\) is empty. Clearly we have \(u, v \in V \setminus B(t)\), implying that for a set \(A \subseteq V \setminus B(t)\), \(\phi(A, B(t), t) > \phi(A, B(t), t - 1)\) is only possible when \(u, v \in A\). But then, \(\phi(A, B(t), t) = \phi(A, B(t), t - 1) + 1\), \(\alpha(u, t) = \alpha(u, t - 1) + 1\) and \(\alpha(v, t) = \alpha(v, t - 1) + 1\), and we deduce \(A \notin \mathcal{D}(B(t), t)\).
Now suppose Maxi applies rule (S3). We have seen before that whenever \( B(t) \neq B(t - 1) \) it holds \( \mathcal{D}(B(t), t - 1) = \emptyset \), hence Maxi only uses (S3) in situations where \( B(t) = B(t - 1) = B(t - 2) \). For all sets \( A \notin \mathcal{D}(B(t), t - 1) \), by the same arguments as for rule (S2) it follows \( A \notin \mathcal{D}(B(t), t) \). So we only have to check the sets \( A \in \mathcal{D}(B(t), t - 1) \). Whenever Maxi uses rule (S3), we have \( t > 1 \), thus by induction \( \mathcal{D}(B(t - 2), t - 2) = \emptyset \). Let us apply again Lemma 8.7 with the same parameters as before. By statement (ii) of the lemma, for all \( A \notin \mathcal{D}(B(t), t - 1) \) we observe

\[
(8.3) \quad \alpha(A, t - 1) \geq \alpha(A, t - 2) \geq \phi(A, B(t - 2), t - 1) - 1 = \phi(A, B(t - 1), t - 1) - 1.
\]

Furthermore, recall that the set \( A' \) considered in Lemma 8.7 coincides with \( A_0(B(t), t - 1) \), and \( \mathcal{C} \) is a superset of \( \mathcal{D}(B(t - 1), t - 1) \). Let \( A \in \mathcal{D}(B(t), t - 1) \).

By definition of (S3), Maxi plays such that at least one vertex of the edge \( e \) is contained in \( A_0(B(t), t - 1) \subseteq A \). We now distinguish two cases. If both \( u, v \in A \), then

\[
\alpha(u, t) + \alpha(v, t) = \alpha(u, t - 1) + \alpha(v, t - 1) + 2.
\]

Since \( \phi(A, B(t), t) \leq \phi(A, B(t - 1), t - 1) + 1 \), together with (8.3) we deduce that \( A \notin \mathcal{D}(B(t), t) \). On the other hand, if \( u \in A \) but \( v \notin A \), the \( \phi \)-value of the set \( A \) does not increase with the new edge \( e \), because \( v \notin B(t) \) as every vertex of \( B(t) \) is full in \( G(t - 1) \). At the same time, the \( \alpha \)-value of \( u \) increases by one. Again it follows \( A \notin \mathcal{D}(B(t), t) \). All together we see that indeed, \( \mathcal{D}(B(t), t) \) is empty.

We proceed with invariant (I3). For the rule (S1) it is obvious that \( G(t) \) is \( k \)-colorable. Otherwise, we observe that since Maxi’s edge \( e \) uses two vertices of the same set \( V_i \), we have \( e \in E(V \setminus B(t)) \) because every vertex of \( B(t) \) was already full in \( G(t - 1) \). Thus in \( G(t) \), at least the subgraph induced by \( B(t) \) is \( k \)-colorable. Now, since \( \mathcal{D}(B(t), t) \) is empty by (I2), for every non-empty set \( A \subseteq V \setminus B(t) \) we have

\[
(8.4) \quad 2|E_t(A)| + |E_t(A, B(t))| \leq 2\phi(A, B(t), t) \leq 2\alpha(A, t) \\
\leq 2(m - 1)|A| < (2m - 1)|A|.
\]

By Lemma 8.5 and by \( k \geq 2m - 1 \), we see that \( G(t) \) is \( k \)-colorable and indeed, Maxi is allowed to play the desired edge.
It remains the last invariant (I4). Suppose there exists a vertex \( v \in V \setminus B(t) \) which is not full in \( G(t) \). Then \( v \in V_i \) for some index \( i \), and there exists another vertex \( u \in V_i \) such that \( \{u, v\} \notin E_t \). We now fictitiously assume that Mini plays the edge \( f = \{u, v\} \) in her move at time \( t + 1 \) and then verify that \( G(t + 1) \) would be \( k \)-colorable, implying in turn that \( G(t) \) could not be saturated.

By (I2) it holds \( D(B(t), t) = \emptyset \). We see that if Mini inserts edge \( f \), she in fact applies herself rule (S2)! Using similar arguments as above when analyzing rule (S2), we see that for all sets \( A \subseteq V \setminus B(t) \) where the \( \phi \)-value increases by one due to the edge \( f \), the \( \alpha \)-value increases too. Therefore, \( D(B(t + 1), t + 1) \) is empty, given that Mini plays \( f \), where we used \( B(t + 1) = B(t) \). Then (8.4) for \( t + 1 \) instead of \( t \) and Lemma 8.5 together imply that \( G(t + 1) \) would be \( k \)-colorable. Hence (I4) is also true.

We are now able to finish the proof as follows. Suppose Maxi follows the proposed strategy and Mini plays the last edge of the game process at time \( t_{\text{end}} \). By (I1)-(I4), for every move of Mini at a point in time \( t - 1 \), the following is true: if there exists at least one vertex \( v \in V \) that is not full in \( G(t - 1) \), Maxi has a valid answer. Since \( B(t) \) only contains full vertices, Maxi applies either (S2) and (S3) and thus inserts a new edge at time \( t \). In this case, \( G(t - 1) \) could not be saturated, and we see that if the game stops directly after a move of Mini, indeed every vertex is full in \( G(t - 1) \) which is fine. On the other hand, by invariant (I4) it is only possible for Maxi to play the last edge of the game at a point in time \( t \) if every vertex of \( V \setminus B(t) \) is full in \( G(t) \). Since all vertices of \( B(t) \) are full as well, we see that if the game stops after a move of Maxi, again all vertices are full in \( G(t) \) and indeed, in the graph \( G_{\text{end}} \) every subset \( V_i \) induces a clique of size \( |V_i| \).

**Proof of Lemma 8.7.** Let \( t_0 \) be a point in time of the game process that satisfies the two preconditions of the statement. Suppose that the set \( \mathcal{C} \) is non-empty. We first observe that the empty set is not an element of \( \mathcal{C} \) as \( \phi(\emptyset, B, t_0 + 1) = 0 \). Next, assume that \( \mathcal{C} \) is non-empty and let \( A_1, A_2 \in \mathcal{C} \). We claim that \( A_1 \cap A_2 \in \mathcal{C} \). By assumption, we have \( \sum_{v \in A_1} f(v) \geq \phi(A_1, B, t_0) \). \( G(t_0) \) is not saturated, so there is a player who inserts a new edge \( e = \{x, y\} \) at time \( t_0 + 1 \). Since \( A_1 \in \mathcal{C} \), we have

\[
f(A_1) \leq \phi(A_1, B, t_0 + 1) - 1 \leq \phi(A_1, B, t_0) \leq f(A_1),
\]

implying

\[(8.5) \quad f(A_1) = \phi(A_1, B, t_0).\]

Obviously, \(A_2\) achieves the same property. It follows

\[
\phi(A_1, B, t_0) + \phi(A_2, B, t_0) = f(A_1) + f(A_2) = f(A_1 \cup A_2) + f(A_1 \cap A_2)
\]

\[
\geq \phi(A_1 \cup A_2, B, t_0) + \phi(A_1 \cap A_2, B, t_0)
\]

\[
\geq \phi(A_1, B, t_0) + \phi(A_2, B, t_0),
\]

where the first inequality follows by the first assumption on \(G(t_0)\) and second inequality follows from the fact that every edge of the graph \(G(t_0)\) is counted in \(\phi(A_1 \cup A_2, B, t_0) + \phi(A_1 \cap A_2, B, t_0)\) at least as often as in \(\phi(A_1, B, t_0) + \phi(A_2, B, t_0)\), which can be observed by a simple case distinction. However, the above inequality chain implies that we have equality everywhere, and in particular

\[(8.6) \quad f(A_1 \cap A_2) = \phi(A_1 \cap A_2, B, t_0).\]

From (8.5) we know that \(\phi(A_1, B, t_0 + 1) = \phi(A_1, B, t_0) + 1\), and the same is true for \(A_2\). Then either both vertices \(x, y\) of the new edge are contained in \(A_1 \cap A_2\), or one endpoint is in \(A_1 \cap A_2\) and the other in \(B\). We see that in both cases, the new edge \(e\) contributes to \(\phi(A_1 \cap A_2, B, t_0 + 1)\). Together with (8.6) we deduce \(A_1 \cap A_2 \in \mathcal{C}\). Now the whole argument can be repeated for any two sets \(A_1, A_2 \in \mathcal{C}\), and we conclude that \(A' = \bigcap_{A \in \mathcal{C}} A\) is itself in the family \(\mathcal{C}\). This proves (i).

Next we observe that statement (ii) follows directly from the property \(\phi(A, B, t_0 + 1) \leq \phi(A, B, t_0) + 1\) and from (8.5). Regarding (iii), the same observation implies \(\phi(A', B, t_0) = f(A')\). Let \(A \subseteq V \setminus (B \cup A')\) and recall that by assumption we have \(f(A \cup A') \geq \phi(A \cup A', B, t_0)\). Putting things together, we then arrive at

\[(8.7) \quad \phi(A \cup A', B, t_0) - \phi(A', B, t_0) \leq f(A \cup A') - f(A') = f(A).\]

However, we observe that

\[
\phi(A \cup A', B, t_0) - \phi(A', B, t_0) = |E_{t_0}(A)| + |E_{t_0}(A, C \cup A')|.
\]
Since the edge $e$ uses at least one vertex of $A'$ but no vertex of $A$, it is neither contained in $E_{t_0+1}(A)$ nor in $E_{t_0+1}(A, C \cup A')$, implying

$$\phi(A \cup A', B, t_0) - \phi(A', B, t_0) = \phi(A \cup A', B, t_0 + 1) - \phi(A', B, t_0 + 1).$$

Combining this equality with inequality (8.7) then proves (iii). □

8.4 Upper Bound

We prove Theorem 8.3 by describing and analyzing a strategy for Mini which significantly shortens the game. Here, the main idea is to play such that in $G_{end}$ there are many vertices of degree $n-1$ (“star vertices”). Then, the color classes of the complete $k$-partite graph $G_{end}$ are rather unbalanced, making $G_{end}$ sufficiently sparse. Our general upper bound on the score of colorability saturation games is a consequence of the following lemma.

Lemma 8.8. Let $n, \ell \in \mathbb{N}$ and let $k \geq 3\ell + 1$. Then in the colorability saturation game with $k$ colors and $n$ vertices, Mini has a strategy such that there are at least $\ell$ vertices of degree $n-1$ in $G_{end}$.

Proof of Theorem 8.3. We first verify the claimed bound

$$s(n, \chi_{>k}) \leq \left(\frac{n}{2}\right)\left(1 - \frac{1}{k - \lfloor(k-1)/3\rfloor}\right) + n$$

for a few special cases. Obviously, $s(n, \chi_{>1}) = 0$ which agrees with (8.8). For $k = 2$ our upper bound simplifies to $n^2/4 + 3n/4$, which is higher than the Turán number and thus fine again. Finally, for the case $k = 3$ the proof of Hefetz et al. yields $s(n, \chi_{>3}) \leq 21n^2/64$ [Hef+16] while with (8.8), we only claim an upper bound of $n^2/3 + n$. So we can assume $k \geq 4$ for the remainder of this section.

Let $n > k$, and put $\ell := \lfloor \frac{k-1}{3} \rfloor$. No matter how Mini and Maxi play, the graph $G_{end}$ is a complete $k$-partite graph with partition sizes $n_1, \ldots, n_k$, containing $\binom{n}{2} - \sum_{i=1}^{k} \binom{n_i}{2}$ edges. W.l.o.g. assume $n_1 \geq \ldots \geq n_k$. By Lemma 8.6, Mini has a strategy such that there are $\ell$ vertices of degree $n-1$ in $G_{end}$. So we can assume that $n_{k-\ell+1} = \ldots = n_k = 1$. 
We want to lower-bound the number of missing edges in $G_{\text{end}}$. Aside from the $\ell$ color classes of size 1, there are $k - \ell$ color classes left over among which we have to distribute the $n - \ell$ remaining vertices. Then the number of missing edges becomes minimal if for all $1 \leq i \leq k - \ell$ it holds $\lceil \frac{n - \ell}{k - \ell} \rceil \leq n_i \leq \lfloor \frac{n - \ell}{k - \ell} \rfloor$. Indeed, for any other distribution we would have $n_1 > \lceil \frac{n - \ell}{k - \ell} \rceil$ and $n_{k - \ell} < \lfloor \frac{n - \ell}{k - \ell} \rfloor$, and transferring one vertex from the first class to class $k - \ell$ would decrease the total number of missing edges. Hence, it follows that the number of edges that are missing in $G_{\text{end}}$ is at least

$$\binom{k - \ell}{\frac{n - \ell}{k - \ell}} = \binom{n - \ell}{n - k} \geq \frac{1}{k - \ell} \binom{n}{2} - \frac{n(k + \ell)}{2(k - \ell)}.$$}

Furthermore, $\frac{k}{3} \geq \ell \geq \frac{k - 3}{3}$ by our choice of $\ell$, therefore $\frac{1}{k - \ell} \geq \frac{3}{2k + 3}$ and $\frac{k + \ell}{k - \ell} \leq 2$. We see that the number of missing edges is at least

$$\frac{1}{k - \ell} \binom{n}{2} - n \geq \frac{3}{2k + 3} \binom{n}{2} - n,$$

which proves the theorem.

In Section 8.3 we analyzed a strategy where Maxi creates a collection of disjoint cliques. Even though Mini’s strategy will be different, we use similar proof techniques to verify Lemma 8.8. There will be one major difference: while Maxi defined before the start of the game which vertex sets should become cliques, the strategy that we propose for Mini is more adaptive in the sense that she does not announce at the start which nodes should become stars, but she chooses these distinguished vertices carefully at specific moments during the game process. In the analysis, we therefore use a set $S(t)$, containing all vertices that have been designated until time $t$ to become star vertices until $t_{\text{end}}$. The set $S(t)$ will be increasing in $t$.

For every vertex $v \in V$, every set $S \subseteq V$, and all $t \in \mathbb{N}_0$ we define

$$\alpha(v, S, t) := |S \cap \Gamma_t(v)|$$

for measuring the progress of Mini at vertex $v$. Moreover, for all $A \subseteq V$, $S \subseteq V$, and $t \in \mathbb{N}_0$ let $\alpha(A, S, t) := \sum_{v \in A} \alpha(v, S, t)$. For all choices of $v$ and $S$, we start with $\alpha(v, S, 0) = 0$. Clearly, the $\alpha$-values are increasing as the game evolves.
Similarly as in Section 8.3, we use the function \( \phi(A, B, t) \) as defined in (8.1) to measure the progress of the opponent, but this time the opponent is Maxi. The plan is to compare the \( \alpha \)-values with the \( \phi \)-values, for the purpose of finding a suitable edge for Mini’s moves. Again, for all points in time \( t \in \mathbb{N}_0 \) and all \( B, S \subseteq V \) we put
\[
D(B, S, t) := \{ A \subseteq V \setminus B \mid \alpha(A, S, t) < \phi(A, B \setminus S, t) \}.
\]
Notice that for all sets \( B \) and \( S \) we start with \( D(B, S, 0) = \emptyset \). Finally, let
\[
A_0(B, S, t) := \bigcap_{A \in D(B, S, t)} A.
\]
For all \( t \in \mathbb{N}_0 \) we have to define which sets \( B = B(t) \) and \( S = S(t) \) we want to use when comparing \( \alpha \)- and \( \phi \)-values. We choose \( B(t) \) and \( S(t) \) in such a way that \( S(t) = B(t) \) as long as \( |S(t)| < \ell \). As soon as \( |S(t)| = \ell \), only \( B(t) \) will further grow. We start with \( S(0) = B(0) = \emptyset \). Whenever it is Maxi’s turn at some point in time \( t \), we put \( S(t) = S(t-1) \) and \( B(t) = B(t-1) \). Whenever it is Mini’s turn at round \( t \), we first define the sets \( S(t) \) and \( B(t) \) and afterwards specify which edge she should play. It turns out that finding suitable sets \( S(t) \) and \( B(t) \) is quite tricky. Suppose at some point in time \( t \) it is Mini’s turn. We then advise Mini to run Algorithm 8.1 for appropriately defining \( S(t) \) and \( B(t) \).

The idea behind Algorithm 8.1 is the following. Mini wants to shape the sets \( B(t) \) and \( S(t) \) such that by playing an edge \( e = \{u, v\} \) with \( u \in V \setminus B(t) \) and \( v \in S(t) \), she can make the set \( D(B(t), S(t), t) \) empty. In situations where it is not possible to find sets \( B(t) \) and \( S(t) \) such that already \( D(B(t), S(t), t-1) \) is empty, Mini needs to choose \( u \in A_0(B(t), S(t), t-1) \). Hence, the set \( Z \) used in the algorithm indicates where Mini should choose \( u \) from. While running the algorithm, we store the current choices for \( B(t) \) and \( S(t) \) in \( B \) and \( S \). We observe that if all edges between \( Z \) and \( S \) are already present in \( G(t-1) \), there is no suitable edge \( \{u, v\} \) available for Mini. We overcome this problem by repeatedly updating the sets \( S \) and \( B \) with the rules (U1)-(U3) until the desired properties are present. When applying rule (U1) or rule (U3), a new vertex \( v \) is added to \( S \). Here, \( v \) is already connected to all other vertices of \( S \), therefore for all points in time \( t \) the vertices of \( S(t) \) induce a complete graph in \( G(t-1) \). In case \( S \) already contains
Algorithm 8.1 Algorithm for defining the sets $S(t)$ and $B(t)$

1: $S := S(t-1), B := B(t-1)$
2: \textbf{if} $\emptyset(B, S, t-1) = \emptyset$ \textbf{then} $Z := V \setminus B$
3: \textbf{else} $Z := A_0(B, S, t-1)$
4: \textbf{while} $B \neq V$ \textbf{and} $\alpha(Z, S, t-1) = |Z| \cdot |S|$ \textbf{do}
5: \hspace{1em} \textbf{if} $Z = V \setminus B$ \textbf{and} $|S| = \ell$ \textbf{then} $B := V$
6: \hspace{1em} \textbf{else if} $\emptyset(B, S, t-1) = \emptyset$ \textbf{then}
7: \hspace{2em} let $v \in V \setminus B$
8: \hspace{1em} $S := S \cup \{v\}, B := B \cup \{v\}$ \hspace{1em} \textgreater{} Rule (U1)
9: \hspace{1em} \textbf{else}
10: \hspace{2em} \textbf{if} $|S| = \ell$ \textbf{then}
11: \hspace{3em} $B := B \cup A_0(B, S, t-1)$ \hspace{1em} \textgreater{} Rule (U2)
12: \hspace{2em} \textbf{else}
13: \hspace{3em} let $v \in A_0(B, S, t-1)$
14: \hspace{3em} $S := S \cup \{v\}, B := B \cup \{v\}$ \hspace{1em} \textgreater{} Rule (U3)
15: \hspace{2em} \textbf{if} $\emptyset(B, S, t-1) = \emptyset$ \textbf{then} $Z := V \setminus B$
16: \hspace{2em} \textbf{else} $Z := A_0(B, S, t-1)$
17: \textbf{return} $S(t) := S, B(t) := B$

$\ell$ vertices, instead of adding a new vertex to $S$ we apply rule (U2) and transfer some nodes to the set $B$. We see that whenever the sets $B(t)$ and $S(t)$ returned by Algorithm 8.1 satisfy $S(t) \neq B(t)$, we have $|S(t)| = \ell$ and all edges between $S(t)$ and $B(t) \setminus S(t)$ are present in $G(t-1)$. Note that a priori it is not clear whether the set $A_0(B, S, t-1)$ considered in lines 11, 13, and 16 is non-empty and whether the algorithm terminates or not.

We now continue by describing Mini’s strategy for playing her edge at time $t$ immediately after executing Algorithm 8.1.

(S1) If $B(t) = V$ but $G(t-1)$ is not yet saturated, insert an arbitrary edge such that $G(t)$ is $k$-colorable.

(S2) If $\emptyset(B(t), S(t), t-1)$ is empty and $B(t) \neq V$, let $u \in V \setminus B(t)$ and $v \in S(t)$ be two vertices such that the edge \{u, v\} is not yet contained in $G(t-1)$. Insert \{u, v\}.

(S3) If $\emptyset(B(t), S(t), t-1)$ is non-empty, let $u \in A_0(B(t), S(t), t-1)$ and $v \in S(t)$
be two vertices such that the edge \( \{u, v\} \) is not yet contained in \( G(t - 1) \). Insert \( \{u, v\} \).

This finishes the definition of Mini’s strategy. We now start proving Lemma 8.8. Thereby, we also verify that Algorithm 8.1 terminates and that the proposed strategy is well-defined and covers all possible cases.

**Proof of Lemma 8.8.** Let \( n, \ell \in \mathbb{N} \) and let \( k \geq 3\ell + 1 \). Suppose Mini applies the proposed strategy. We prove by induction that as long as \( G(t - 1) \) is not saturated or there exists a vertex \( u \in V \setminus B(t - 1) \) with \( \alpha(u, S(t - 1), t - 1) < \ell \), for Mini’s move at time \( t \) the following invariants hold.

(I1) Algorithm 8.1 always terminates and afterwards, exactly one rule of (S1)-(S3) can be applied.

(I2) Let Mini insert the desired edge, regardless whether \( G(t) \) is \( k \)-colorable or not. Then the set \( \mathcal{D}(B(t), S(t), t) \) is empty.

(I3) Mini can play her edge without violating the colorability constraint.

(I4) If there is a vertex \( u \in V \setminus B(t) \) with \( \alpha(u, S(t), t) < \ell \), then \( G(t) \) is not yet saturated.

We start with \( B(0) = \emptyset, S(0) = \emptyset, \) and \( \mathcal{D}(B(0), S(0), 0) = \emptyset \). Let \( t \) be a point in time such that \( G(t - 1) \) is not saturated or such that there exists a vertex \( u \in V \setminus B(t - 1) \) with \( \alpha(u, S(t - 1), t - 1) < \ell \). Assume that either it is Mini’s first move of the game (providing the base case) or by induction, (I1)-(I4) were true for all previous moves of Mini. We start with (I1). Recall that the set \( Z \) used in Algorithm 8.1 indicates from which set Mini wants to pick a vertex \( u \) and connect with a vertex \( v \in S(t) \). With the criterion \( \alpha(Z, S, t - 1) = |Z| \cdot |S| \) in line 4 we test whether all these edges are already present in the graph \( G(t - 1) \).

First we assume that the set \( \mathcal{D}(B(t - 1), S(t - 1), t - 1) \) is empty. If this is the case, we put \( B = B(t - 1), S = S(t - 1), \) and in line 2 we set \( Z = V \setminus B \). In the special case \( V = B \) Algorithm 8.1 immediately stops and Mini can apply rule (S1). If \( B \neq V \), we have three subcases. First, if \( \alpha(Z, S, t - 1) < |Z| \cdot |S| \), Algorithm 8.1
terminates, and rule (S2) can be applied. Next, if $\alpha(Z, S, t - 1) = |Z| \cdot |S|$ and $|S| = \ell$ and we put $B = V$ in line 5, again the algorithm terminates, and Mini uses (S1) until the end of the game. It remains the subcase $\alpha(Z, S, t - 1) = |Z| \cdot |S|$ but $|S| < \ell$. In this situation, when running Algorithm 8.1 we apply rule (U1) and add one vertex $v$ to the sets $S$ and $B$. We claim that after applying (U1), the set $\mathcal{D}(B, S, t - 1)$ is still empty. Let $A \subseteq V \setminus B$ be any non-empty set (where we take the freshly updated set $B$ including $v$). Since $\mathcal{D}(B(t - 1), S(t - 1), t - 1)$ was empty by assumption and $B(t - 1) = S(t - 1)$, we have

$$\alpha(A, S(t - 1), t - 1) \geq \phi(A, B(t - 1) \setminus S(t - 1), t - 1) = |E_{t - 1}(A)|.$$  

However, when adding $v$ to $S$ and $B$, clearly the $\alpha$-value of $A$ is non-decreasing while the $\phi$-value remains the same, so $A \notin \mathcal{D}(B, S, t - 1)$. Hence after using rule (U1), $\mathcal{D}(B, S, t - 1)$ is empty as claimed, $Z$ is set to $V \setminus B$, and we can repeat the whole argument for the case, potentially update the sets $S$ and $B$ several times, until either $\alpha(Z, S, t - 1) < |Z| \cdot |S|$ or $\alpha(Z, S, t - 1) = |Z| \cdot \ell$. Then indeed the algorithm terminates, and as discussed above Mini can apply either (S1) or (S2).

Let us now assume that the set $\mathcal{D}(B(t - 1), S(t - 1), t - 1)$ is non-empty. Then in line 3 we put $Z = A_0(B(t - 1), S(t - 1), t - 1) = \cap_{A \in \mathcal{D}(B(t - 1), S(t - 1), t - 1)} A$. We have to verify that this set is non-empty. Note that this case can only happen when $t > 1$. By induction we have $\mathcal{D}(B(t - 2), S(t - 2), t - 2) = \emptyset$, so we can apply Lemma 8.7 with $f(v) = \alpha(v, S(t - 2), t - 2)$, $B = B(t - 2) \setminus S(t - 2)$, and $t_0 = t - 2$. Then the set $\mathcal{C}$ given by (8.2) contains all subsets $A \subseteq V \setminus B(t - 2)$ where $\alpha(A, S(t - 2), t - 2) < \phi(A, B(t - 2) \setminus S(t - 2), t - 1)$. Since $S(t - 1) = S(t - 2)$, $B(t - 1) = B(t - 2)$, and $\alpha(v, S(t - 1), t - 1) \geq \alpha(v, S(t - 2), t - 2)$ holds for every vertex $v \in V$, it follows

$$\emptyset \neq \mathcal{D}(B(t - 1), S(t - 1), t - 1) \subseteq \mathcal{C}.$$  

By Lemma 8.7 (i), $A' = \cap_{A \in \mathcal{C}}$ is non-empty and itself a member of $\mathcal{C}$, i.e., $\alpha(A', S(t - 2), t - 2) < \phi(A', B(t - 2) \setminus S(t - 2), t - 1)$. Together with the assumption $\mathcal{D}(B(t - 1), S(t - 1), t - 1) \neq \emptyset$ and statement (ii) of Lemma 8.7, we conclude that
for all $v \in A'$ it holds

$$\alpha(v, S(t-1), t-1) = \alpha(v, S(t-1), t-2).$$

Hence $A' \in \mathcal{D}(B(t-1), S(t-1), t-1)$, $A' = A_0(B(t-1), S(t-1), t-1)$, and therefore $A_0(B(t-1), S(t-1), t-1)$ is non-empty.

If at least one edge between $A_0(B(t-1), S(t-1), t-1)$ and $S(t-1)$ is not present in $G(t-1)$, Algorithm 8.1 immediately terminates, returns $S(t) = S(t-1)$ resp. $B(t) = B(t-1)$, and Mini can play her edge according to rule (S3). If all edges between $A_0(B(t-1), S(t-1), t-1)$ and $S(t-1)$ are already contained in $E_{t-1}$, there are two subcases.

Let us first investigate the subcase $|S(t-1)| = \ell$ where we are going to apply rule (U2) in line 11 and thus put $B = B(t-1) \cup A_0(B(t-1), S(t-1), t-1)$. (Note that we are still in the first iteration of the while-loop, so the “old” $B$ is the same as $B(t-1)$. Also note that here, we have $S = S(t-1)$.) We now apply Lemma 8.7 (iii) and using $S(t-1) = S(t-2)$ and $B(t-1) = B(t-2)$, we see that for all $A \subseteq V \setminus (B(t-1) \cup A')$ it holds

$$\alpha(A, S(t-1), t-1) \geq \alpha(A, S(t-2), t-2)$$

$$\geq \phi(A \cup A', B(t-1) \setminus S(t-1), t-1) - \phi(A', B(t-1) \setminus S(t-1), t-1).$$

Hence, after applying rule (U2), for each such set $A$ we have

$$\alpha(A, S, t-1) \geq \alpha(A, S(t-1), t-1) \geq \phi(A, B \setminus S(t-1), t-1) = \phi(A, B \setminus S, t-1),$$

and $\mathcal{D}(B, S, t-1)$ becomes empty. Consequently we put $Z = V \setminus B$ in line 15. Now either $\alpha(Z, S, t-1) < |Z| \cdot \ell$, Algorithm 8.1 terminates, and rule (S2) can be applied, or $\alpha(Z, S, t-1) = |Z| \cdot \ell$, we put $B = V$ in line 5 (during the second iteration of the while-loop), the algorithm terminates, and Mini uses (S1) until the end of the game.

In the subcase $|S(t-1)| < \ell$, we apply rule (U3), pick one vertex $v$ of the set $A_0(B(t-1), S(t-1), t-1)$, designate it as future “star-vertex” and add it to the set $S$. We claim that after executing rule (U3) in line 14, the set $\mathcal{D}(B, S, t-1)$ is empty. Indeed, by construction all sets $A \in \mathcal{D}(B(t-1), S(t-1), t-1)$ contain $v$, ...
thus for every set \( A \subseteq V \setminus (B(t - 1) \cup \{v\}) \) we have
\[
\alpha(A, S(t - 1), t - 1) \geq \phi(A, B(t - 1) \setminus S(t - 1), t - 1).
\]
However, here it holds \( B(t - 1) = S(t - 1) \) and \( B = S \), thus
\[
\phi(A, B \setminus S, t - 1) = \phi(A, B(t - 1) \setminus S(t - 1), t - 1).
\]
It follows that \( A \notin \mathcal{D}(B, S, t - 1) \) as the \( \alpha \)-value of \( A \) is non-decreasing. Hence \( \mathcal{D}(B, S, t - 1) \) is empty as claimed, and we consequently set \( Z = V \setminus B \) in line 15.

Now we are in the same situation as discussed above: either \( \alpha(Z, S, t - 1) < |Z| \cdot |S| \), or \( \alpha(Z, S, t - 1) = |Z| \cdot \ell \), or \( \alpha(Z, S, t - 1) = |Z| \cdot |S| \) but \( |S| < \ell \). We have already seen that in all three subcases, at some point the algorithm stops and afterwards, either (S1) or (S2) serves as a matching rule. This proves invariant (I1).

We continue with the second invariant (I2). Suppose Mini plays the edge \( e = \{u, v\} \). If Mini applied rule (S1), then \( B(t) = V \) and the invariant is trivial. If Mini uses (S2), then \( \mathcal{D}(B(t), S(t), t - 1) \) is empty. W.l.o.g. we assume \( u \in V \setminus B(t) \) and \( v \in S(t) \). We observe that for no set \( A \subseteq V \setminus B(t) \) its \( \phi \)-value can increase due to Mini’s edge, so \( \mathcal{D}(B(t), S(t), t) \) is empty as well.

It remains to check rule (S3). As we have seen before when analyzing Algorithm 8.1, whenever Mini applies rule (S3) it holds \( S(t) = S(t - 1) = S(t - 2) \) and \( B(t) = B(t - 1) = B(t - 2) \), because the set \( B(t) \) can differ only from \( B(t - 1) \) if at least one of the update rules (U1)-(U3) was applied, but then the set \( \mathcal{D}(B(t), S(t), t - 1) \) would be empty (and the same is true for \( S(t) \) and \( S(t - 1) \)). Similarly as for rule (S2) we see that each set \( A \subseteq V \setminus B(t) \) that is not contained in \( \mathcal{D}(B(t - 1), S(t - 1), t - 1) \) is also not contained in \( \mathcal{D}(B(t), S(t), t) \). Regarding the sets in \( \mathcal{D}(B(t - 1), S(t - 1), t - 1) \), we notice that whenever (S3) is used we have \( t > 1 \) and thus by induction \( \mathcal{D}(B(t - 2), S(t - 2), t - 2) = \emptyset \). We apply Lemma 8.7 with the same parameters as before and argue that for every set \( A \in \mathcal{D}(B(t - 1), S(t - 1), t - 1) \) we have
\[
\alpha(A, S(t - 1), t - 1) \geq \alpha(A, S(t - 2), t - 2) \geq \phi(A, B(t - 2) \setminus S(t - 2), t - 1) - 1
\]
\[
= \phi(A, B(t - 1) \setminus S(t - 1), t - 1) - 1.
\]
Recall that the set \( A' \) given by Lemma 8.7 is the same set as \( A_0(B(t), S(t), t - 1) \) and \( \mathcal{D}(B(t - 1), S(t - 1), t - 1) \subseteq \mathcal{C} \). Assume w.l.o.g. that Mini plays edge \( e = \{u, v\} \)
where \( u \in A_0(B(t), S(t), t-1) \) and \( v \in S(t) \). Then for all \( A \in \emptyset(B(t-1), S(t-1), t-1) \) we have \( u \in A \). We see that with the new edge \( e \), the \( \phi \)-value of \( A \) does not increase whereas its \( \alpha \)-value increases by one, implying that \( A \notin \emptyset(B(t), S(t), t) \) as required. We deduce that \( \emptyset(B(t), S(t), t) \) must be empty, which verifies (I2).

Regarding (I3), we first detect that for rule (S1) it is obvious that \( G(t) \) is \( k \)-colorable. For rules (S2) and (S3), we argue that since \( |S(t)| \leq \ell \) by construction, it is sufficient to prove that the subgraph of \( G(t) \) induced by \( V \setminus S(t) \) is \((k - \ell)\)-colorable. Recall that whenever \( S(t) \neq B(t) \), it holds \( |S(t)| = \ell \) in \( G(t-1) \), every vertex of \( B(t) \setminus S(t) \) is connected with each star vertex of \( S(t) \), and the vertices of \( S(t) \) induce a complete graph in \( G(t-1) \). From these facts we infer that since the subgraph of \( G(t-1) \) induced by \( B(t) \) is \( k \)-colorable, the subgraph of \( G(t-1) \) induced by \( B(t) \setminus S(t) \) must be \((k - \ell)\)-colorable. In the other case \( S(t) = B(t) \), this is trivial. Further, we know that when Mini plays the edge \( e = \{u, v\} \), both nodes \( u \) and \( v \) are contained in the set \( (V \setminus B(t)) \cup S(t) \). Hence, in \( G(t) \) the subgraph induced by \( B(t) \setminus S(t) \) is \((k - \ell)\)-colorable too. We now want to extend this property from \( B(t) \setminus S(t) \) to the set \( V \setminus S(t) \). By (I2), the set \( \emptyset(B(t), S(t), t) \) is empty, so for every non-empty set \( A \subseteq V \setminus B(t) \) it holds

\[
2|E_t(A)| + |E_t(A, B(t) \setminus S(t))| \leq 2\phi(A, B(t) \setminus S(t), t) \leq 2\alpha(A, S(t), t)
\]

\[
\leq 2\ell|A| < (k - \ell)|A|.
\]

Using Lemma 8.5 we conclude that in \( G(t) \), the subgraph induced by \( V \setminus S(t) \) is \((k - \ell)\)-colorable and thus \( G(t) \) itself must be \( k \)-colorable.

We turn to the last invariant (I4) and assume that there exists \( u \in V \setminus B(t) \) with \( a(u, S(t), t) < \ell \). Above when verifying invariant (I3) we realized that the subgraph of \( G(t) \) induced by \( V \setminus S(t) \) is \((k - \ell)\)-colorable. Hence if \( S(t) \) contains less than \( \ell \) vertices, it is obvious that \( G(t) \) can not be saturated. So suppose \( |S(t)| = \ell \). Then there exists a vertex \( v \in S(t) \) such that \( \{u, v\} \notin E_t \). We fictitiously assume that Maxi inserts the edge \( f = \{u, v\} \) in her turn at time \( t + 1 \). By invariant (I2) it holds \( \emptyset(B(t), S(t), t) = \emptyset \). Therefore Maxi playing edge \( f \) corresponds to applying rule (S2) herself. Using similar arguments as above when analyzing rule (S2), we see that for no set \( A \in V \setminus B(t) \) its \( \phi \)-value increases due to edge \( f \). Then \( \emptyset(B(t + 1), S(t + 1), t + 1) \) is empty, given that Mini plays \( f \), and from (8.9) for
t + 1 instead of t and Lemma 8.5 we deduce that G(t + 1) would be k-colorable. Therefore, G(t) cannot be saturated and (I4) is true as well.

After verifying all four invariants by induction, we now finish the proof of the lemma. We assume that Mini follows our strategy during the whole game. Her goal is to ensure that G_{end} contains \ell vertices of degree n − 1. By (I1)-(I4), for every move of Maxi at a point in time t − 1 the following is true: if there exists at least one vertex \( u \in V \setminus B(t − 1) \) with \( \alpha(u, S(t − 1), t − 1) < \ell \), then Mini has a valid answer and runs in particular Algorithm 8.1 for finding \( S(t) \) and \( B(t) \). Unless \( B(t) = V \), Mini adds a new edge to the graph and thus \( G(t − 1) \) was not saturated. However, if \( B(t) = V \), then \( |S(t)| = \ell \) and all vertices of \( S(t) \) have degree \( n − 1 \) in \( G(t − 1) \) by construction. Finally, exactly the same holds if every vertex of \( u \in V \setminus B(t − 1) \) satisfies \( \alpha(u, S(t − 1), t − 1) \geq \ell \) or if \( V = B(t − 1) \). We conclude that if the game stops after the move of Maxi at time \( t − 1 \), then the desired graph structure is indeed present in \( G(t − 1) \). On the other hand, if Mini makes \( G(t) \) saturated with her edge played at time \( t \), by invariant (I4) we have \( \alpha(V \setminus B(t), S(t), t) = |V \setminus B(t)| \cdot \ell \), and the \ell vertices of \( S(t) \) have degree \( n − 1 \) in \( G(t) \). We see that no matter which player terminates the game, there are at least \( \ell \) vertices of degree \( n − 1 \) in \( G_{end} \). ■

8.5 The Four Color Game

In this section we study a specific saturation game by considering the special case \( k = 4 \). By Theorem 8.2 and Theorem 8.3 we already have the bounds

\[
\frac{n^2}{4} \leq s(n, \chi_{>4}) \leq \frac{n^2}{3}(1 + o(1)).
\]

In particular, the upper bound follows because Mini has a strategy such that \( G_{end} \) contains a vertex \( s \) of degree \( n − 1 \) (see Lemma 8.8). The goal of this section is to prove Theorem 8.4 and thus close the gap in (8.10). We improve the lower bound by providing an alternative strategy for this specific game, being more effective than the general strategy proposed in Section 8.3. In principle, the idea is the same as before: building up a collection of vertex-disjoint cliques ensures that no color class becomes too large. But in contrast to our previous strategy,
we advise Maxi to proceed “greedily” and play the cliques within successive moves such that they cover the vertices that have been used most recently by Mini. Depending on Mini’s strategy, Maxi answers by drawing cliques of size 4, 3, or 2 (that is, single edges).

After this informal description, let us proceed by introducing some specific notation. On the one hand, our refined strategy for Maxi defines which edges she should play in her turns. On the other hand, it also describes how the collection of vertex-disjoint cliques grows during the process. At several points in time \( t_i \), after the edge of round \( t_i \) has been played Maxi defines a vertex set \( V_i \subseteq V \) which then yields the \( i \)-th clique of the desired collection. More precisely, we require that the vertices of \( V_i \) induce a complete graph in \( G(t_i) \) and that all sets \( V_i \) are vertex-disjoint. We then denote \( t_i \) as the birth time of the \( i \)-th clique. The sequence \( t_i \) is non-decreasing, but subsequent cliques are allowed to have the same birth time. It turns out that it is convenient to also require that all birth times \( t_i \) are chosen such that Mini plays at time \( t_i \). During the game process, Maxi only marks a finite number of cliques, and this number depends on Mini’s strategy. Consequently, \( V_i \) and \( t_i \) are only defined for indices \( i \) where the \( i \)-th clique actually exists. Finally, for each clique \( V_i \) we put

\[
W(i) := V \setminus \left( \bigcap_{j \leq i} V_j \right).
\]

Similarly as in Section 8.3 and Section 8.4, we use the function \( \phi \) as defined in (8.1) to measure the progress of the opponent. For all \( i \in \mathbb{N} \) where the \( i \)-th clique exists, we put

\[
\phi(i) := \phi(W(i), V \setminus W(i), t_i) = |E_{t_i}(W(i))| + |E_{t_i}(W(i), V \setminus W(i))|.
\]

One of Maxi’s goals is to play such that the values \( \phi(i) \) are globally bounded by a small constant, which means that we are always in the situation that her collection of vertex-disjoint cliques covers almost all non-isolated vertices of the current graph.

We start by describing under which conditions it is reasonable for Maxi to create a triangle as next clique.
Lemma 8.9. Let $i \in \mathbb{N}$ such that $\phi(i) \leq 3$, $|W(i)| \geq 4$, and at most one edge of $E_{t_i}(W(i))$ is isolated in $G(t_i)$. Then Maxi has a strategy for playing her edges and defining $V_{i+1}$ and $t_{i+1}$ such that

(i) $t_{i+1} \leq t_i + 6$,

(ii) $|V_{i+1}| = 3$, and

(iii) $\phi(i + 1) \leq 3$.

Proof. We prove the lemma by a case distinction. First suppose that in $G(t_i)$, the vertex set $W(i)$ already induces a triangle with vertices $x, y, z$. Then obviously we can put $V_{i+1} := \{x, y, z\}$ and $t_{i+1} := t_i$. In particular, it then holds $\phi(i + 1) = 0$.

Next assume that in $G(t-1)$ there exists a path $\{x, y, z\}$ where all three vertices are contained in $W(i)$. Then Maxi can play the edge $\{x, z\}$ at time $t_i + 1$. After Mini’s move at time $t_i + 2$, we put $V_{i+1} := \{x, y, z\}$ and $t_{i+1} := t_i + 2$. We also have $\phi(i + 1) \leq 2$ since the two edges $\{x, y\}$ and $\{y, z\}$ that were contributing to $\phi(i)$ are not counted for $\phi(i + 1)$.

Now we investigate the case where in $G(t_i)$, the graph induced by $W(i)$ is non-empty but neither contains a triangle nor a path of 3 vertices. Let $e = \{x, y\}$ be an arbitrary edge of $E_{t_i}(W(i))$, and let $z \in W(i) \setminus \{x, y\}$ with maximal degree in $G(t_i)$. We then advise Maxi to complete the triangle $\{x, y, z\}$ within her next two moves at times $t_i + 1$ and $t_i + 3$. Note that in $G(t + 3)$, there will be at most three edges between $V_{i+1} := \{x, y, z\}$ and $V \setminus V_{i+1}$: by testing all possible cases, we see that at most two edges of this type were already present in $G(t_i)$, and we have to add the edge that Mini eventually played at time $t_i + 2$. Then it is easy to see that $G(t + 3)$ is indeed 4-colorable. Furthermore, $G(t + 3)$ is not saturated due to the assumption $|W(i)| \geq 4$. So we put $t_{i+1} := t_i + 4$. Regarding property (iii), we observe that due to our choice of $z$, the assumption on $\phi(i)$ and the assumption that $E_{t_i}(W(i))$ contains at most one isolated edge in $G(t_i)$, at most one edge that was counted in $\phi(i)$ can now contribute to $\phi(i + 1)$. Adding Mini’s edges at times $t_i + 2$ and $t_i + 4$ we arrive at $\phi(i + 1) \leq 3$. Note that here, we assumed that Mini does not help us in creating the triangle on the vertices $\{x, y, z\}$. In case Mini
would play one of these edges on her own, it is easy to see that we obtain the same invariants already by picking \( t_{i+1} := t_i + 2 \).

Finally, we consider the remaining case where \( E_{t_i}(W(i)) \) is empty. We pick \( x, y \in W(i) \) such that \( \deg(x) + \deg(y) \) is maximal in the graph \( G(t_i) \). We first advise Maxi to play the edge \( \{x, y\} \) at time \( t_i + 1 \). After Mini’s subsequent answer, at time \( t_i + 3 \) Maxi picks \( z \in W(i) \setminus \{x, y\} \) with maximal degree in \( G(t_i + 2) \). Then in her next two moves at times \( t_i + 3 \) and \( t_i + 5 \), Maxi completes the triangle \( V_{i+1} := \{x, y, z\} \). Again, we assume for simplicity that Mini does not help Maxi in creating this triangle and plays different edges. We first have to verify that the proposed strategy is possible without violating the colorability constraint. Assume w.l.o.g. that in \( G(t_i + 5) \) we have \( \deg(x) \geq \deg(y) \geq \deg(z) \). We observe that in this graph there can be at most five edges between \( V_{i+1} \) and \( V \setminus V_{i+1} \): at most three that were counted with \( \phi(i) \), and at most two additional edges played by Mini at times \( t_i + 2 \) and \( t_i + 4 \). Therefore in \( G(t_i + 5) \) we have \( \deg(y) \leq 4 \) and \( \deg(z) \leq 3 \). The subgraph of \( G(t_i + 5) \) induced by \( V \setminus \{y, z\} \) must be 4-colorable because Maxi is not playing in this subgraph and Mini is also forced to maintain the colorability property. Since \( |\Gamma_{t_i+5}(y)| \leq 4 \), \( |\Gamma_{t_i+5}(z)| \leq 3 \), and \( \{y, z\} \in E_{t_i+5} \), we can take an arbitrary coloring of \( V \setminus \{y, z\} \) and extend it to the entire vertex set. Due to the assumption \( |W(i)| \geq 4 \), we also infer that \( G(t_i + 5) \) is not yet saturated. So we can take \( t_{i+1} := t_i + 6 \). It remains to bound \( \phi(i + 1) \). Let

\[
E' := E_{t_i+2}(W(i)) \cup E_{t_i+2}(W(i), V \setminus W(i)).
\]

The assumption \( \phi(i) \leq 3 \) implies \( |E'| \leq 5 \), but the edge \( \{x, y\} \) is clearly not counted in \( \phi(i + 1) \). Because we are in the case \( E_{t_i}(W(i)) = \emptyset \), the strategy now ensures that at most one edge of the set \( E' \) can contribute to \( \phi(i + 1) \). At last, we take into account the two edges that Mini plays at times \( t_i + 4 \) and \( t_i + 6 \) and conclude that invariant (iii) is indeed satisfied. \( \blacksquare \)

Next, we give a similar lemma for the special situation where Maxi is able to create a clique of size four during her next four moves.

**Lemma 8.10.** Let \( i \in \mathbb{N} \) such that \( |W(i)| \geq 6 \) and such that there exist at least two edges \( e, e' \in E_{t_i}(W(i)) \) that are isolated in \( G(t_i) \). Then Maxi has a strategy for playing her edges and defining \( V_{i+1} \) and \( t_{i+1} \) such that
8.5. THE FOUR COLOR GAME 271

(i) \( t_{i+1} \leq t_i + 8, \)

(ii) \( |V_{i+1}| = 4, \) and

(iii) \( \phi(i + 1) \leq \phi(i) + 2. \)

Proof. Suppose that the described situation occurs at time \( t_i. \) Then Maxi takes the two edges \( e = \{x, y\} \) and \( e' = \{x', y'\}. \) During her next four moves, she claims the remaining edges that complete the \( K_4 \) on the vertex set \( V_{i+1} := \{x, x', y, y'\}. \) Since Mini can insert at most three additional edges meanwhile, it is not difficult to check that Mini has no possibility to forbid any of these four edges, so Maxi indeed succeeds in creating this \( K_4 \) until time \( t_i + 7. \) Because we are assuming that \( |W(i)| \geq 6 \) and \( e, e' \) have been isolated in \( G(t_i) \), the graph \( G(t_i + 7) \) is not yet saturated. We then put \( t_{i+1} := t_i + 8. \) For property (iii), we infer that \( e \) and \( e' \) contributed to \( \phi(i) \) but don't contribute to \( \phi(i + 1). \) On the other hand, there are at most four edges that Mini played in meantime, therefore \( \phi(i + 1) - \phi(i) \leq 2. \) Note that we assumed again that Mini does not help Maxi in creating the \( K_4. \) Otherwise, it is not difficult to see that (i)-(iii) can be already achieved with a smaller choice of \( t_{i+1}. \) ■

Finally, in some situations it is the best for Maxi to only play cliques of size 2, i.e., simple edges. The purpose of such a strategy is to reduce the \( \phi \)-value and ensure that it is globally bounded during the game process.

Lemma 8.11. Let \( i \in \mathbb{N} \) such that \( \phi(i) \leq 5 \) and \( |W(i)| \geq 3. \) Then Maxi has a strategy for playing her edges and defining \( V_{i+1} \) and \( t_{i+1} \) such that

(i) \( t_{i+1} \leq t_i + 2, \)

(ii) \( |V_{i+1}| = 2, \) and

(iii) \( \phi(i + 1) \leq \max\{\phi(i) - 1, 1\}. \)

Proof. First suppose that the edge set \( E_t(W(i)) \) is non-empty. Then we can take an arbitrary edge \( e = \{x, y\} \) of this edge set, put \( V_{i+1} := \{x, y\} \) and \( t_{i+1} = t_i. \)
Clearly, we then have $|E_{t_i+1}(W(i+1))| < |E_{t_i}(W(i))|$, which shows (iii) in this case.

If the set $E_{t_i}(W(i))$ is empty, we take $x, y \in W(i)$ such that $\deg(x) + \deg(y)$ is maximal in $G(t_i)$, and tell Maxi to play the edge $e = \{x, y\}$ at time $t_i + 1$. Assume w.l.o.g. that $\deg(x) \geq \deg(y)$. Then $\deg(y) \leq 2$ by our assumption on $\phi(i)$. After inserting the edge $e$, clearly the subgraph of $G(t_i+1)$ induced by $V \setminus \{y\}$ is 4-colorable, and since $\deg(y) \leq 3$ in $G(t_i+1)$, every proper coloring of $V \setminus \{y\}$ can be easily extended to $y$. Since $|W(i)| \geq 3$ and $E_{t_i}(W(i)) = \emptyset$, the strategy also ensures that $G(t_i+1)$ is not saturated. So we can put $V_{i+1} := \{x, y\}$ and $t_{i+1} := t_i + 2$. It remains to prove (iii). We observe that if $\phi(i) \leq 2$, the edge $e$ covers all edges that contributed to $\phi(i)$, hence the only edge that potentially counts for $\phi(i+1)$ is the edge that Mini plays at time $t_i + 2$, thus $\phi(i+1) \leq 1$. On the other hand, if $\phi(i) \geq 2$, the edge $e$ covers at least two edges that counted for $\phi(i)$, and we obtain $\phi(i+1) = \phi(i) - 1$.

After these preparations we can start proving Theorem 8.4. Recall that we only have to provide a matching lower bound on $s(n, \chi_{>4})$.

**Proof of Theorem 8.4.** Let $n \geq 5$. The proof strategy is to verify by induction that given the game process for the first $i$ cliques, we can apply one of the three auxiliary lemmas of this section to see that there exists a strategy for Maxi to create the next clique sufficiently fast. For the base case, we distinguish two cases regarding the identity of the starting player. If Maxi starts the game, she can draw an arbitrary triangle with vertex set $V_1$ during her first three turns, and we put $t_1 := 6$. No matter how Mini plays, we have $\phi(1) \leq 3$. If Mini starts with the game, Maxi extend the first edge played by Mini to a triangle with vertex set $V_1$ during her first two moves. We then put $t_1 := 5$ and observe that $\phi(1) \leq 2$. In addition, we observe here, the assumption $n \geq 5$ implies that $G(t_1)$ is not saturated as the smallest saturated graph requires at least seven edges.

We now claim that whenever Maxi has defined the $i$-th clique for some $i \in \mathbb{N}$ and $|W(i)| \geq 5$, then she also has a strategy to build the next clique such that one of the following four conditions is satisfied.

1. $\phi(i+1) \leq 3$ and $|V_{i+1}| = 3$,
(ii) $\phi(i + 1) \leq 5$ and $|V_{i+1}| = 4$,

(iii) $\phi(i + 1) \leq 4$ and $|V_i| = 4$, or

(iv) $\phi(i + 1) \leq 3$ and $|V_{i-1}| = 4$.

We prove this claim by induction over $i$, so we assume by induction that $V_i$ and $\phi(i)$ satisfied one of the four conditions. Recall that $\phi(1) \leq 3$ and $|V_1| = 3$, so indeed $V_1$ and $\phi(1)$ serve as base case. The induction step now follows directly from Lemma 8.9, Lemma 8.10, and Lemma 8.11. Let $i \geq 1$ and first suppose that $V_i$ and $t_i$ satisfied (i) or (iv). Then we distinguish two cases. If $E_{t_i}(W(i))$ contains at least two edges which are isolated in $G(t_i)$, by Lemma 8.10 Maxi has a fast strategy to create the next clique such that $|V_{i+1}| = 4$ and $\phi(i + 1) \leq 5$, so $V_{i+1}$ and $\phi(i + 1)$ fulfill (ii). In the other case where we don’t have this pair of isolated edges, by Lemma 8.9 Maxi can play such that $|V_{i+1}| = 3$, $\phi(i + 1) \leq 3$, and (i) is satisfied. Next suppose $V_i$ and $\phi(i)$ satisfy (ii). Here we advise Maxi to play such that the next clique is only a single edge (i.e. $|V_{i+1}| = 2$). By Lemma 8.11, Maxi is able to do so such that (iii) is true. Finally, in the case (iii), by assumption it holds $|V_{i-2}| = 4$. Again, we require Maxi to play such that $|V_i| = 2$, and then Lemma 8.11 establishes (iv).

We see that as long as $|W(i)| \geq 5$, the game does not stop and Maxi has a strategy to create at least one additional clique with vertex set $V_{i+1}$. Let $j$ be the unique index where $|W(j)| < 5$ and the procedure stops. At this point in time, we have $n' := \sum_{i=1}^{j} |V_i| > n - 5$. For the remainder of the game, we let Maxi play arbitrarily until the graph is saturated and the game ends. We now prove that with the given strategy, Maxi ensures that $G_{\text{end}}$ contains at least $\frac{n^2}{3} + O(n)$ edges.

At time $t_j$, we have $V \setminus W(j) = \bigcup_{i=1}^{j} V_i$. For $k \in \{2, 3, 4\}$, we denote by $a_k$ the number of sets $V_i$ in this collection with size $k$. By definition we have $a_2 + a_3 + a_4 = j$. Moreover it holds

$$2a_2 + 3a_3 + 4a_4 = n' > n - 5. \quad (8.11)$$

Next we observe that for every clique where $|V_i| = 2$, $V_i$ and $\phi(i)$ either satisfy (iii) or (iv). Hence either $|V_{i-1}| = 4$ or $|V_{i-2}| = 4$, and we deduce

$$a_2 \leq 2a_4.$$
No matter how Mini and Maxi play, $G_{end}$ is a complete 4-partite graph and up to permutations, there exists exactly one proper 4-coloring. Denote by $C_1, \ldots, C_4$ the four color classes. Clearly every $K_4$ of the collection contributes one vertex to each class. Next, every triangle of the collections spends one vertex to three different color classes. Let $m := |E_{end}|$ be the total number of edges in $G_{end}$. We observe that $m$ is minimal if the triangles always contribute to the same three color classes, say $C_1, C_2,$ and $C_3$. Indeed, otherwise we could move one vertex from a smaller class to a higher class and forbid more edges. Finally, the same argument yields that from Maxi’s perspective, in the worst case all $K_2$ of the collection account to the two heaviest color classes, say $C_1$ and $C_2$. Finally, we assume that the remaining $n - n'$ vertices that are not covered by the cliques all account for $C_1$. We summarize that $m$ becomes minimal if we have $|C_4| = a_4$, $|C_3| = a_3 + a_4$, $|C_2| = a_2 + a_3 + a_4$, and $|C_1| = |C_2| + n - n'$. Using $n - n' = O(1)$, we obtain

$$m \geq |C_2|^2 + 2 \cdot |C_2| \cdot (|C_3| + |C_4|) + |C_3| \cdot |C_4| + O(n). \tag{8.12}$$

Minimizing this number subject to the boundary conditions on $a_2, a_3,$ and $a_4$ is a standard optimization problem. For every fixed value $a_3$, the number of edges will be minimized when $a_2$ is maximal with respect to $a_4$ because this makes the color classes of $G_{end}$ as unbalanced as possible. Therefore, in the extremal case we have $a_2 = 2a_4$, which eliminates one variable. From (8.12) we then deduce

$$m \geq 3a_3^2 + 17a_3a_4 + 22a_4^2 + O(n). \tag{8.13}$$

On the other hand, from (8.11) it follows

$$a_2 = \frac{n' - 3a_3}{4} \quad \text{and} \quad a_4 = \frac{n' - 3a_3}{8}, \tag{8.14}$$

and combining (8.13) and (8.14) yields

$$m \geq \frac{1}{32} \left(11n'^2 + 2a_3n' - 9a_3^2\right) + O(n).$$

The term $11n'^2 + 2a_3n' - 9a_3^2$ is concave in $a_3$ and obtains its minimum at the boundary, i.e., when $a_3 \in [0, n'/3]$. Let us shortly compare the two cases. If
We see that with the choice $a_3 = n'/3$, the game stops earlier in the worst-case, and we argue that $a_3 = n'/3$ minimizes the value of $m$, up to error term $O(n)$. We conclude that $m \geq n^2/3 + O(n)$ and indeed, the score of this particular saturation game is $n^2/3 + O(n)$.

**Remark 8.12.** Using the suggested strategies for Mini and Maxi we determined $s(n, \chi > 4) = n^2/3 + O(n)$. Having the analysis on hand we can now precisely describe the game process, given that both players follow an optimal strategy. During a first period of the game, by applying Lemma 8.8 Mini always uses the same vertex $s$ for her edges until $\deg(s) = n - 1$. Hence $s$ becomes a star vertex. Meanwhile, Maxi uses the strategy provided with Lemma 8.9 and covers the leaves of this star greedily with a collection of triangles. The pace of both players is equal: both insert in total $n + O(1)$ edges for completing their tasks. This first phase ends at the moment where $\deg(s) = n - 1$ and thus $s$ reserves one color class on its own. On the other hand, at the same time the vertex-disjoint triangles guarantee that the three other color classes are equally large. The score of the game is then determined, and the players spend the remaining time by filling the graph arbitrarily with edges until it becomes saturated and the game ends.

### 8.6 Concluding Remarks

We described strategies for both Maxi and Mini that work for all parameters $k$ and turned out to be almost optimal and sufficiently strong for proving Theorem 8.2 and Theorem 8.3. In Section 8.5 we have seen that at least in the case $k = 4$ it is possible to improve and refine Maxi’s strategy such that the lower and upper bounds are matching. We think that also in the general case, Maxi’s strategy can be further improved and that the bound given by Theorem 8.2 is not optimal, but that it requires more advanced strategies to improve the lower bound and that it becomes technical to determine the score of the colorability saturation game precisely.

As discussed in the introduction of this chapter, very little is known about the saturation game with respect to the property “$G$ contains a copy of $H$” where
$H$ is a fixed subgraph, even for the choice $H = K_3$. One natural and very interesting example is the Hamiltonian saturation game where we pick $H = C_n$. It is conjectured that the score of the Hamiltonian game is $\Theta(n^2)$ [Hef+16]. We hope that in near future, the understanding of this fascinating family of games can be improved and some of the aforementioned specific games can be solved.


BIBLIOGRAPHY


CURRICULUM VITAE

RALPH KEUSCH

born on June 27, 1986 in Muri AG, Switzerland

2002 - 2006  Aargauische Kantonsschule, Baden
2006 - 2015  Classical Concerto Pianist, including performance of
             10 disjoint recital programs (solo and chamber music)
2006 - 2010  Teaching Diploma for Piano, SAMP,
             degree with distinction
2007 - 2014  Piano Teacher, private and as substitute at
             Aargauische Kantonsschule, Baden
2009 - 2012  Bachelor of Sciences in Mathematics, ETH Zurich
2010 - 2014  Concerto Diploma for Classical Piano, SAMP
2012 - 2014  Master of Sciences in Mathematics, ETH Zurich,
             degree with distinction
since 2014  PhD Candidate in Theoretical Computer Science,
            ETH Zurich