ALGORITHMIC APPROACHES TO PROBLEMS IN PROBABILISTIC COMBINATORICS

A Dissertation
Presented to
The Academic Faculty

By

He Guo

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Algorithms, Combinatorics, and Optimization (ACO) Program

School of Mathematics
Georgia Institute of Technology

August 2021

ALGORITHMIC APPROACHES TO PROBLEMS IN PROBABILISTIC COMBINATORICS

Thesis committee:

Dr. Lutz Warnke (Advisor)
School of Mathematics
Georgia Institute of Technology

Dr. Eric Vigoda
School of Computer Science
Georgia Institute of Technology

Dr. Tom Bohman (Reader)
Department of Mathematical Sciences
Carnegie Mellon University

Dr. Xingxing Yu
School of Mathematics
Georgia Institute of Technology

Dr. Prasad Tetali
School of Mathematics and School of Computer Science
Georgia Institute of Technology

Date approved: June 10, 2021

江雪

千山鸟飞绝

万径人踪灭

孤舟蓑笠翁

独钓寒江雪

River-snow

From hill to hill no bird in flight;

From path to path no man in sight.

A little boat, a bamboo cloak,

An old man fishing in the cold river-snow.

柳宗元 *LIU Zongyuan*

回首向来萧瑟处，也无风雨也无晴

Turning my head, I find the dreary beaten track. Impervious to rain or shine, I’ll have my own will.

苏轼 *SU Shi*

Let everything happen to you

Beauty and terror

Just keep going

No feeling is final.

Rainer Maria Rilke

For my family

ACKNOWLEDGMENTS

Writing this emotional part of the dissertation is hard for me, as I cannot hold back my tears when I recall those happy moments and say goodbye to my wonderful six-year graduate life at Georgia Tech.

I would like to thank my advisor Lutz Warnke, who inspires me to dive into the research world of probabilistic combinatorics. I enjoy my PhD career working on various topics with him, many of which are covered in this dissertation. He patiently guides me from knowing nothing about research to knowing something. I cannot forget the happiness when I proved my first theorem. The happiness made all effort worth. Lutz encourages me and offers me lots of opportunities to improve my presentation and communication skills. Lutz also helps me a lot for improving my writing skills. I believe those skills are important parts of doing research. Lutz influences me by his optimism, which leaves me a peaceful mind to explore the unknown world with lots of uncertainties.

I would like to thank the ACO, School of Mathematics, and Center for the Enhancement of Teaching and Learning (CETL) community at Georgia Tech. I would like to thank Xingxing Yu who offers me lots of professional guidance. I would like to thank Prasad Tetali for his great role as ACO director. I would like to thank Tom Bohman to be the reader of this dissertation. I would like to thank Xu-Yan Chen, Larissa Serdukova, Martin Short, Plamen Iliev, and Thang Le for their professional assistance. I would like to thank Eric Vigoda, Santosh Vempala, Josephine Yu, Galyna Livshyts, Luca Dieci, John McCuan, Hayriye Ayhan, Christopher Heil, Martin Savelsbergh, Justin Romberg, Beryl C. Martinson, Jane Chisholm, and many other instructors for their teaching. I enjoy those courses. Especially, I would like to thank Robin Thomas, who did an excellent job as ACO director and helped me start my graduate life at Georgia Tech. As a world-class expert, Robin taught us graph theory (one of the most fundamental courses of the ACO program), even with very serious ALS disease. It was really a wonderful course and I found my academic

interest in his course. He still attended talks regularly in his last years. RIP, Robin!

I would like to thank many friends I know as a graduate student. Thank Shijie Xie, Yan Wang, Chi Ho Yuen, Laura Eslava, Xiaofan Yuan, Dantong Zhu, Renyi Chen, Haoyan Zhai, Xiao Liu, Samantha Petti, Saurabh Sawlani, Guanyi Wang, Zhiyu Wang, Hao Wu, Xiaonan Liu, Hongyi (Hugo) Zhou, Peng Zhang, Caleb Anderson, Anant Girdhar, Ruilin Li, Shu Liu, Jiaqi Yang, Xin Xing, Youngho Yoo, He Jia, Zongchen Chen, Jad Salem, Michael Wigal, Michael Anastos, Markus Scheprs, Zilin Jiang, Joseph Briggs, Xiao Chang, Molly and many others. I am not able to list all of them, but I treasure all the moments with them.

I would like to thank Kenya Payton, Lea Marzo, Mitchell Everett, Natache Sumter, Stephanie Cook, Iris Hamilton, Jeremy Terry, Chris Jankowski, Shuenn Siang Ng, and many others for their support about the departmental affairs. Thank Sarah Kegley for offering suggestions about the narrative parts in the first draft of this dissertation. Thank those who I know the faces but not names; thank dining staff, housing staff, CRC staff, library staff, IT group, Skiles workers, Green/Trolley/Grocery/Tech Express/Red/Blue lines and Stingerette drivers. Their daily hard work makes me enjoy the wonderful time at Georgia Tech. Thanks for players of the sports I joined at Georgia Tech. Thanks for the academic community from all over the world, which offers me chances to travel to many countries that I had never been to. Thank those teachers, instructors, and professors who have ever taught me in my life. Their teaching helps me build a solid foundation to pursue my PhD degree.

Last but not least, I would like to thank my parents. Their love is essential for my growth and supports me to get through tough times. Also thanks for my other family members and those old friends. Reunion with you is always my happiest time. I haven't been back to my hometown since January 2020 due to the COVID-19 pandemic and I do miss you. Hope the world will go back to normal soon!

All in all, thank the great life, which gives me lots of fresh experience! Be brave and keep going on!

TABLE OF CONTENTS

SUMMARY

The probabilistic method is one of the most powerful tools in combinatorics; it has been used to show the existence of many hard-to-construct objects with exciting properties. It also attracts broad interests in designing and analyzing algorithms to find and construct these objects in an efficient way. In this dissertation we obtain four results using algorithmic approaches in probabilistic method:

1. We study the structural properties of the triangle-free graphs generated by a semi-random variant of *triangle-free process* and obtain a packing extension of Kim's famous $R(3, t)$ results. This allows us to resolve a conjecture in Ramsey theory by Fox, Grinshpun, Liebenau, Person, and Szabó, and answer a problem in extremal graph theory by Esperet, Kang, and Thomassé.

2. We determine the order of magnitude of *Prague dimension*, which concerns efficient encoding and decomposition of graphs, of binomial random graph with high probability. We resolve conjectures by Füredi and Kantor. Along the way, we prove a Pippenger-Spencer type edge coloring result for random hypergraphs with edges of size $O(\log n)$.

3. We analyze the number set generated by r-*AP free process*, which answers a problem raised by Li and has connection with van der Waerden number in additive combinatorics and Ramsey theory.

4. We study a refined alteration approach to construct H-free graphs in binomial random graphs, which has applications in Ramsey games.

CHAPTER 1
INTRODUCTION

The work presented in this dissertation lies in the intersection of combinatorics and probability theory. Using tools from probabilistic combinatorics, we resolve several conjectures and problems in Ramsey theory, extremal combinatorics, and additive combinatorics.

A core methodology of this dissertation is a modern algorithmic approach to the probabilistic method. The probabilistic method, which is a powerful method in combinatorics and related areas, was pioneered by Paul Erdős. The classical approach is to show that a random object satisfies desired properties with positive probability. The algorithmic approach is to generate the random object using some suitable randomized algorithm, which, for example, generates the object step by step in a more sophisticated way. We apply this algorithmic approach in many areas, and the main mechanism is to keep track of various structural properties during the random processes, for example, to show certain pseudorandom properties are preserved. Using variations of this approach, the central contributions of this dissertation are based on the following three main paradigms:

- Using **semi-random algorithms**, we *resolve several conjectures and problems* in Ramsey theory and extremal combinatorics, which were raised by Fox, Grinshpun, Liebenau, Person, and Szabó [38], Füredi and Kantor [45], and Esperet, Kang, and Thomassé [36] (see Chapter 2, 3, 4 for details, which are based on [56] published in Combinatorica, [54], and [57], respectively). For example, we refine Kim's celebrated Ramsey $R(3,t)$ construction of triangle-free graphs and make it more robust, which allows us to determine the order of magnitude of the smallest minimum degree of r-Ramsey-minimal graphs $s_r(K_3)$, a parameter introduced by Burr, Erdős, and Lovász in 1976 (see Section 1.1 and Chapter 2 for details, which is based on [56] published in Combinatorica). For random graphs we determine the typical order of magnitude of Prague dimension, which was introduced by Nešetřil, Pultr, and Rödl in the 1970s (see Section 1.2 and Chapter 3 for details, which is based on [54]).

- Analyzing some **random greedy algorithms**, we answer a question by Li [77] related to van der Waerden numbers in additive combinatorics and Ramsey theory (see Section 1.3 and Chapter 4 for details, which is based on [57]), and we prove random variants of the influential Pippenger-Spencer hypergraph chromatic index result (see Section 1.2 and Chapter 3 for details, which is based on [54]). For example, in the random setting we are able to properly color an r-uniform n-vertex hypergraph with edges of very large size, i.e., of size $r = O(\log n)$.

- We also refine the related **alteration approach**. In particular, we prove that for suitable n and p, after removing *all* edges of H-copies in random graph $G_{n,p}$ for a fixed graph H, the resulting H-free graph is still pseudorandom. This contrasts with earlier approaches of Erdős [27] and Krivelevich [71], who constructed such H-free graph by removing *some* edges of H-copies in $G_{n,p}$. Our refined approach has applications in Ramsey games: we extend previous results by Conlon, Fox, Grinshpun, and He [20] and Fox, He, and Wigderson [39] (see Section 1.4 and Chapter 5 for details, which is based on [55]).

When applying the above paradigms, we use differential equation method, concentration inequalities, and martingale theory as the key ingredients of our analysis.

In the following subsections we further expand on the main results of this dissertation.

1.1 Triangle-free graphs and their applications in Ramsey theory

An interesting but mysterious phenomenon in mathematics is that for a sufficiently large structure, no matter how it is partitioned, there will always be some well-behaved substructure in one of the parts. The study of this phenomenon is called Ramsey theory. As a central parameter in Ramsey theory, *Ramsey number* $R(s, t)$ is the minimum number n

such that every red and blue edge coloring of complete graph K_n contains either a red K_s or a blue K_t. Understanding the behavior of $R(s,t)$ and other Ramsey-type parameters is notoriously difficult. A celebrated result in Ramsey theory is $R(3,t) = \Theta(t^2/\log t)$, the study of which has been very influential for the development of new tools and techniques in probabilistic combinatorics. Matching the upper bound by Ajtai, Komlós, and Szemerédi [2, 3] in 1980 where the fantastic semi-random approach was invented, in 1995 Kim [67] famously proved his Fulkerson Prize result $R(3,t) = \Omega(t^2/\log t)$ by analyzing so-called *semi-random triangle-free process*[1]. Kim's result improves the logarithmic factor in $R(3,t) = \Omega(t^2/(\log t)^2)$, which was first obtained by Erdős [27] in 1961 via a clear alteration method and was subsequently reproved over the following three decades via Lovász Local Lemma [107], a basic analysis of the triangle-free process [33], large deviation inequalities [71], and differential equations [109]. Kim's result was reproved by Bohman [10] in 2008.

Both Kim and Bohman proved that $R(3,t) = \Omega(t^2/\log t)$ by constructing an n-vertex triangle-free graph G in complete graph K_n with independence number $\alpha(G) = O(\sqrt{n \log n})$, which is optimal up to the constant factor and improves previous bound $O(\sqrt{n}\log n)$ by a logarithmic factor. By analyzing the semi-random triangle-free process and investigating more pseudorandom properties in the process, we prove a packing extension of Kim and Bohman's results: for any $\epsilon > 0$ we find an edge-disjoint collection $(G_i)_{i\in\mathcal{I}}$ of n-vertex graphs $G_i \subseteq K_n$ such that (a) each G_i is triangle-free and has independence number at most $C_\epsilon\sqrt{n \log n}$, and (b) the union of all the G_i contains at least $(1-\epsilon)\binom{n}{2}$ edges. As an application, we prove a conjecture in Ramsey theory by Fox, Grinshpun, Liebenau, Person, and Szabó [38] concerning a Ramsey-type parameter introduced by Burr, Erdős, and Lovász [16] in 1976. Namely, denoting by $s_r(H)$ the smallest minimum degree of

[1]The triangle-free process (proposed by Bollobás and Erdős) is defined by starting with an empty edge set on n vertices and then iteratively adding one edge at each step, chosen uniformly at random from all non-edges, subject to the constraint that adding it does not create a triangle. The semi-random version proceeds similarly to the triangle-free process, but at each step, instead of adding just one random edge, we add many random edges and if necessary do some correction on the random choices. See Chapter 2 for details.

r-Ramsey minimal graphs for H, we close the existing logarithmic gap for $H = K_3$ and establish that $s_r(K_3) = \Theta(r^2 \log r)$.

See Chapter 2 for details, which is based on joint work with Lutz Warnke [56] published in Combinatorica.

1.2 Prague dimension

Introduced by Nešetřil, Pultr, and Rödl [85, 84] in the 1970s, the *Prague dimension* (also called product dimension and there are many equivalent definitions, see [121, 59, 5]) $\dim_{\mathrm{P}}(G)$ of a graph G is the minimum number d such that G is an induced subgraph of the product of d complete graphs, which equals the minimum number of subgraphs of the complement $\overline{G}$ of G such that (i) each subgraph is a vertex-disjoint union of cliques, and (ii) each edge of $\overline{G}$ is contained in at least one of the subgraphs, but not all of them. Determining Prague dimension of many graphs remains to be difficult. Füredi and Kantor [45] noted that with high probability $\dim_{\mathrm{P}}(G_{n,p}) = \Omega(n/\log n)$ for constant edge-probabilities p and conjectured that their lower bound gives the correct order of magnitude; see [45, Conjecture 15] and [64]. We resolve this conjecture.

Along the way of the proof, we extend a previous chromatic index result by Kurauskas and Rybarczyk [74] and we obtain a Pippenger–Spencer type [93] edge coloring result for random hypergraphs with uniformity $O(\log n)$, which is a hypergraph extension of the famous Vizing's theorem for graphs.

See Chapter 3 for details, which is based on joint work with Kalen Patton and Lutz Warnke [54].

1.3 Some random greedy algorithms

1.3.1 Van der Waerden numbers

As an important parameter in additive combinatorics and Ramsey theory, *van der Waerden number* $W(r,k)$ is the minimum number N such that every red and blue coloring of numbers in $[N]=\{1,2,\ldots,N\}$ contains either a red r-term arithmetic progression (r-AP) or a blue k-AP. The celebrated van der Waerden's theorem guarantees that $W(r,k)$ is finite for all integers $r,k\geq 2$. It is natural, interesting, but difficult to determine the asymptotic behavior of $W(r,k)$ (see [47, 50]). Indeed, in the mid 2000s Graham conjectured that $W(3,k)\leq k^{O(1)}$ and mentioned that numerical evidence (for example, see [1] or the sequence `A007783` in [105]) suggests $W(3,k)=k^{2+o(1)}$; see [48, 49, 52]. Around 2015 Graham even started offering \$250 reward for his conjecture (see [49, p. 19]). The best known upper bound $W(3,k)\leq\exp(k^{1-\Omega(1)})$ was obtained by Schoen [100] in 2020. In terms of lower bounds, in 2008 Li and Shu [78] showed that $W(r,k)=\Omega\big((k/\log k)^{r-1}\big)$ for fixed $r\geq 3$, by applying the Lovász Local Lemma to a random subset of the integers $[n]$. Subsequently, Li raised in 2009 the natural question [77] whether this probabilistic lower bound can be improved via a randomized greedy algorithm that 'dynamically' constructs an r-AP free subset of the integers $[n]$. We answer Li's question affirmatively. See Chapter 4 for details, which is based on joint work with Lutz Warnke [57].

1.3.2 Induced bipartite subgraphs in triangle-free graphs

Studying induced bipartite subgraphs with large minimum degree in triangle-free graphs, recently Esperet, Kang, and Thomassé asked [36, Problem 4.1] to determine $f_\eta(n)$ that is the largest minimum degree of a bipartite induced subgraph over all n-vertex triangle-free graphs of minimum degree at least n^η, for fixed $\eta\in(0,1)$ as $n\to\infty$. They guess $f_\eta(n)$ has phase transition at $\eta=1/2$ due to the pseudorandom properties in triangle-free process as mentioned in Section 1.1. For η in some ranges, Esperet, Kang, and Thomassé [36], van

Batenburg, de Verclos, Kang, and Pirot [113], and Kwan, Letzter, Sudakov, and Tran [76] partially solve this problem up to constant factors, but for $\eta \in (1/2, 2/3]$ close to the critical value, there are logarithmic gaps in their work. Their proofs are based on alteration method or Lovász Local Lemma. Based on a more refined analysis of pseudorandom properties in the semi-random triangle-free process in [56] by the author and Lutz Warnke, which is mentioned in Section 1.1, we close the gaps and solve the problem up to constant. For $g(n, d)$ that is a natural generalization of $f_\eta(n)$ and can be viewed as replacing n^η in the definition of $f_\eta(n)$ by d, our refined analysis in [57] also closes the gaps in van Batenburg, de Verclos, Kang, and Pirot [113] and Kwan, Letzter, Sudakov, and Tran [76] so that we determine the order of magnitude of $g(n, d)$ for all $n^{\Omega(1)} \le d \le n/2$. See Chapter 4 for details, which is based on joint work with Lutz Warnke [57].

1.4 Refined alteration approach

To prove a lower bound on Ramsey number $R(H, K_k)$ that is the minimum integer n such that every red and blue edge coloring of complete graph K_n contains a red copy of H or a blue copy of K_k, one needs to show existence of an n-vertex H-free graph without large independent set. One way is in binomial random graph $G_{n,p}$ using *alteration method* to destroy all H-copies to make resulting graph H-free. By removing *some* edges of H-copies in $G_{n,p}$, Erdős [27] (for $H = K_3$) and Krivelevich [71] (for *strictly 2-balanced*[2] H) found a lower bound on $R(H, K_k)$ by showing existence of an H-free graph with no independent set of size k in $G_{n,p}$ for $n = \Theta((k/\log k)^{m_2(H)})$ and $p = O((\log k)/k)$. We consider a refined alteration approach and prove that removing *all* edges of H-copies does not significantly change the numbers of edges in all k-vertex sets so that the independence number of the remaining H-free graph is at most k.

One benefit of removing all edges of H-copies is that it can be applied in some on-

[2] $m_2(H) := \max_{F \subseteq H}\left(\mathbb{1}_{\{v_F \ge 3\}}\frac{e_F - 1}{v_F - 2} + \mathbb{1}_{\{F = K_2\}}\frac{1}{2}\right)$ and a graph H is strictly 2-balanced if $m_2(H) > m_2(F)$ for all $F \subsetneq H$.

line game settings. Firstly, we extend bounds of Conlon, Fox, Grinshpun, and He [20] regarding *online Ramsey game*. Secondly, we generalize the upper bound by Fox, He, and Wigderson [39] for *Ramsey, Paper, Scissors number*.

See Chapter 5 for details, which is based on joint work with Lutz Warnke [55].

1.5 Basic definitions and notations

In this final subsection we briefly introduce some basic definitions and notations used frequently in this dissertation.

An r-*uniform hypergraph* is an ordered pair of sets (V, E) with $E \subseteq \binom{V}{r} := \{A \subseteq V : |A| = r\}$. Especially, 2-uniform hypergraph is called *graph*. Given a (hyper)graph $H = (V, E)$, the set $V = V(H)$ is the *vertex set* consisting of *vertices* and the set $E = E(H)$ is the *(hyper)edge set* consisting of *(hyper)edges*. The *degree* of a vertex is the number of (hyper)edges containing the vertex. The *minimum (maximum) degree* of a (hyper)graph is the minimum (maximum) degree over all vertices of the (hyper)graph. An *independent set* I of a graph $G = (V, E)$ is a subset of the vertex set V that contains no edge of the graph, i.e., $\binom{I}{2} \cap E = \varnothing$. The *independence number* $\alpha(G)$ of a graph G is the maximum size over all the independent sets of G. The *complement* $\bar{G}$ of a graph $G = (V, E)$ is $(V, \binom{n}{2} \setminus E)$. Graph F is a *subgraph* of graph G if $V(F) \subseteq V(G)$ and $E(F) \subseteq E(G)$. Two graphs G_1 and G_2 are *isomorphic* if there exists a bijection $\phi : V(G_1) \to V(G_2)$ such that $\{u, v\} \in E(G_1)$ if and only if $\{\phi(u), \phi(v)\} \in E(G_2)$. Given a graph H, a graph G is called H-*free* if there is no subgraph of G that is isomorphic to H. The *complete graph* (or *clique* of n vertices) K_n is n-vertex graph with all of the $\binom{n}{2}$ pairs of vertices as edges. A graph G is *bipartite* if $V(G)$ is disjoint union of two independent sets.

An r-*term arithmetic progression (r-AP)* $A = \{a_1, a_2, \ldots, a_r\}$ in $[n] := \{1, 2, \ldots, n\}$ is a subset of $[n]$ satisfying $|A| = r$ and for some $d \neq 0$ and all $i = 1, 2, \ldots, r$, $a_i = a_1 + (i-1)d$.

The *binomial random graph* $G_{n,p}$ is the n-vertex graph where each of the $\binom{n}{2}$ pairs of vertices occurs independently as an edge with probability p. Given a sequence of events $(\mathcal{A}_n)_n$ in some probability space, it holds *with high probability (whp)* if $\lim_{n\to\infty} \mathbb{P}(\mathcal{A}_n) = 1$.

For asymptomatic notations in this dissertation, given two functions f and g, we write $f(n) = O(g(n))$, $f(n) = o(g(n))$, $f(n) = \omega(g(n))$, $f(n) = \Omega(g(n))$, and $f(n) \sim g(n)$ if $\limsup_{n\to\infty} \frac{|f(n)|}{g(n)} < \infty$, $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$, $\liminf_{n\to\infty} \frac{|f(n)|}{g(n)} = \infty$, $\liminf_{n\to\infty} \frac{|f(n)|}{g(n)} > 0$, and $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$, respectively. We write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$. We write $f(n) \gg g(n)$ or $g(n) \ll f(n)$ if $\liminf_{n\to\infty} \frac{f(n)}{g(n)} = \infty$.

CHAPTER 2

PACKING NEARLY OPTIMAL RAMSEY $R(3, T)$ GRAPHS

2.1 Background and main results

The 1947 paper of Erdős [26] on the diagonal Ramsey number $R(t, t)$ is often considered the start of the probabilistic method, where $R(s, t)$ is defined as the smallest integer $n \in \mathbb{N}$ such that every red-blue colouring of the edges of the complete n-vertex graph K_n contains either a red K_s or a blue K_t. In general, the estimation of $R(s, t)$ and other Ramsey-type parameters is known to be notoriously difficult.

One of the celebrated results in Ramsey theory is $R(3, t) = \Theta(t^2/\log t)$, and this special case has repeatedly served as a testbed for the development of new tools and techniques in probabilistic combinatorics. Indeed, complementing the basic bound $R(3, t) = O(t^2)$ of Erdős and Szekeres [34], in 1961 Erdős [27] used a sophisticated random greedy alteration argument to prove $R(3, t) = \Omega(t^2/(\log t)^2)$. This lower bound was subsequently reproved (or only slightly improved) using the Lovász Local Lemma [107], a basic analysis of the triangle-free process[1] [33], large deviation inequalities [71], and differential equations [109]. Furthermore, in 1980 Ajtai, Komlós, and Szemerédi [2, 3] invented the influential semi-random method (nowadays also called Rödl nibble approach) to prove the upper bound $R(3, t) = O(t^2/\log t)$. But it was not until 1995, when Kim [67] famously proved the matching lower bound $R(3, t) = \Omega(t^2/\log t)$ by analyzing a semi-random variation of the triangle-free process[2] (combining several of the aforementioned ideas with martingale concentration); for this major breakthrough he also received the Fulkerson Prize

[1] The triangle-free process (proposed by Bollobás and Erdős) proceeds as follows: starting with an empty n-vertex graph, in each step a single edge is added, chosen uniformly at random from all non-edges which do not create a triangle.

[2] Kim's semi-random variation proceeds similarly to the triangle-free process: it intuitively adds a large number of carefully chosen random-like edges in each step (instead of just a single edge); see Section 2.2 for more details.

in 1997. But the story does not end here: advancing the differential equation method, in 2008 Bohman [10] reproved $R(3,t) = \Omega(t^2/\log t)$ by analyzing the triangle-free process itself (and his analysis was recently further improved in [13, 37]).

In this chapter we refine the powerful techniques developed for $R(3,t) = \Theta(t^2/\log t)$ to determine the order of magnitude of another Ramsey-type parameter introduced in 1976 by Burr, Erdős, and Lovász [16], proving a conjecture of Fox, Grinshpun, Liebenau, Person, and Szabó [38] (in particular, analogous to Kim's $R(3,t)$-result, we again remove the last redundant logarithmic factor from existing bounds).

2.1.1 Main result: packing of nearly optimal Ramsey $R(3,t)$ graphs

Kim and Bohman both proved the Ramsey bound $R(3,t) = \Omega(t^2/\log t)$ by showing the existence of a triangle-free graph $G \subseteq K_n$ on n vertices with independence number $\alpha(G) = O(\sqrt{n \log n})$, which is best possible up to the value of the implicit constants. Our first theorem naturally extends their celebrated results, by approximately decomposing the complete graph K_n into a packing of such nearly optimal Ramsey $R(3,t)$ graphs.

Theorem 1. *For any $\epsilon > 0$ there exist $n_0, C, D > 0$ such that, for all $n \geq n_0$, there is an edge-disjoint collection $(G_i)_{i \in \mathcal{I}}$ of $|\mathcal{I}| = \lceil D\sqrt{n/\log n}\rceil$ triangle-free graphs $G_i \subseteq K_n$ on n vertices with $\max_{i \in \mathcal{I}} \alpha(G_i) \leq C\sqrt{n \log n}$ and $\sum_{i \in \mathcal{I}} e(G_i) \geq (1-\epsilon)\binom{n}{2}$.*

Our algorithmic proof proceeds by sequentially choosing the $|\mathcal{I}| = \Theta(\sqrt{n/\log n})$ edge-disjoint triangle-free subgraphs $G_i \subseteq K_n \setminus \bigcup_{0 \leq j < i} G_j$ with $\alpha(G_i) = O(\sqrt{n \log n})$ via a semi-random variation of the triangle-free process akin to Kim [67] (see Sections 2.1.3 and 2.2 for the details). In particular, we do not only show existence of the $(G_i)_{i \in \mathcal{I}}$, but also obtain a polynomial-time randomized algorithm which constructs these subgraphs.

Theorem 1 improves a construction of Fox et.al. [38, Lemma 4.2], who used the basic Lovász Local Lemma based $R(3,t)$-approach to sequentially choose $\Theta(\sqrt{n}/\log n)$ edge-disjoint triangle-free subgraphs with $\alpha(G_i) = O(\sqrt{n}\log n)$. It is natural to suspect that applying a more sophisticated $R(3,t)$-approach in each iteration ought to give an improved packing (with smaller independence number than the LLL approach), and here the usage of the triangle-free process was proposed by Fox et.al. [38, Section 5] as early as 2013 [79, 90]. One conceptual difficulty of this approach is to control various error terms over many iterations of the triangle-free process (so that these always stay small enough to carry out the next iteration), which in turn is the main technical reason why for Theorem 1 we instead iterate a semi-random variation.

It would be interesting to know if Theorem 1 also holds with $\epsilon = 0$, i.e., if one can completely decompose K_n into nearly optimal $R(3,t)$ graphs. Perhaps rashly, we conjecture that this is indeed possible (it might be insightful to first prove a variant of Theorem 1 where the constant C does *not* depend on ϵ).

2.1.2 Application in Ramsey theory: $s_r(K_3)$ has order of magnitude $r^2 \log r$

Turning to our main application, we say that a graph G is *r-Ramsey for H*, denoted by $G \to (H)_r$, if any r-colouring of the edges of G contains a monochromatic copy of H. Most fundamental questions and results in Ramsey theory can be formulated in terms of various parameters of the class

$$\mathcal{M}_r(H) := \big\{G : \ G \to (H)_r \text{ and } G' \nrightarrow (H)_r \text{ for all } G' \subsetneq G\big\}$$

of graphs which are *r-Ramsey minimal for H*. For example, Ramsey's theorem [95] states that $|\mathcal{M}_r(H)| > 0$ for all graphs H, which for cliques was strengthened to $|\mathcal{M}_r(K_t)| = \infty$ by Rödl and Siggers [98]. Furthermore, the archetypal problem of estimating various Ramsey-type parameters also corresponds to the study of certain extremal parameters of $\mathcal{M}_r(H)$, since, e.g., $R(t) = R(t,t) := \min_{G \in \mathcal{M}_2(K_t)} v(G)$ is the famous diagonal Ramsey number [34, 26, 22], $R_r(t) = R(t, \ldots, t) := \min_{G \in \mathcal{M}_r(K_t)} v(G)$ is the r-coloured Ramsey number [22], and $\hat{R}_r(H) := \min_{G \in \mathcal{M}_r(H)} e(G)$ is the widely-studied r-size-Ramsey number of H (see, e.g., [30, 7, 99, 22]).

In 1976 Burr, Erdős, and Lovász [16] initiated the systematic study of other extremal parameters of $\mathcal{M}_r(H)$, including the smallest minimum degree of all r-Ramsey minimal graphs for H, denoted by

$$s_r(H) := \min_{G \in \mathcal{M}_r(H)} \delta(G).$$

As usual, the clique-case $H = K_t$ is of particular interest, where $r(t-2) < s_r(K_t) < R_r(t)$ is easy to see (cf. [40, 112]). Perhaps surprisingly, for $r = 2$ colours Burr et.al. [16] were able to prove $s_2(K_t) = (t-1)^2$, showing that the simple exponential upper bound $R_2(t) = R(t) = 2^{\Theta(t)}$ is far from the truth. For $r \geq 2$ colours the behaviour of $s_r(K_t)$ was recently investigated in detail by Fox et.al. [38]: they proved super-quadratic bounds of form $s_r(K_t) = r^2 \cdot \text{polylog}\, r$ for fixed $t \geq 3$, and also determined $s_r(K_3)$ up to a logarithmic factor (by sharpening their general estimates). In particular, they showed $cr^2 \log r \leq s_r(K_3) \leq Cr^2(\log r)^2$, and conjectured that their lower bound gives the correct order of magnitude, see [38, Conjecture 5.4].

Our second theorem proves the aforementioned conjecture of Fox, Grinshpun, Liebenau, Person, and Szabó for $s_r(K_3)$, i.e., we close the logarithmic gap and establish $s_r(K_3) = \Theta(r^2 \log r)$.

Theorem 2. *There exists $C > 0$ such that $s_r(K_3) \leq Cr^2 \log r$ for all $r \geq 2$.*

Corollary 3. *We have $s_r(K_3) = \Theta(r^2 \log r)$ for $r \geq 2$.*

Using a reformulation of $s_r(K_3)$ from [38], Theorem 2 follows easily from our main packing result. Indeed, applying Theorem 1 with $\epsilon = 1/2$, say, it is routine to see that there is a constant $A > 0$ such that the following holds for each $r \geq 2$: there exists a collection of edge-disjoint triangle-free graphs $G_1, \ldots, G_r \subseteq K_{N_r}$ on $N_r := \lfloor Ar^2 \log r \rfloor$ vertices with independence number $\alpha(G_i) < N_r/r$ (as $N_r \geq n_0$, $D\sqrt{N_r/\log N_r} \geq r$ and $C\sqrt{N_r \log N_r} < N_r/r$ all hold for $A = A(n_0, C, D)$ large enough). By Theorem 1.5 and Lemma 4.1 in [38] (with $n = N_r$ and $k = 2$) this immediately implies $s_r(K_3) \leq N_r$, establishing Theorem 2.

Note that the above deduction of Theorem 2 did *not* use $\sum_{i\in\mathcal{I}} e(G_i) \geq (1-\epsilon)\binom{n}{2}$, i.e., that the nearly optimal $R(3,t)$ graphs $(G_i)_{i\in\mathcal{I}}$ approximately decompose the edge-set of K_n. It would be interesting to find applications (e.g., in Ramsey theory or extremal combinatorics) where this natural packing property is useful.

2.1.3 Main tool: pseudo-random triangle-free subgraphs

The $R(3,t)$-proofs of Kim and Bohman both in fact construct a triangle-free graph $G \subseteq K_n$ with pseudo-random properties (see also [109, 122, 13, 37]). Our third theorem extends their intriguing results to host graphs $H \subseteq K_n$ which are far from complete, by showing that one can again construct a triangle-free subgraph $G \subseteq H$ with pseudo-random properties. Here the crux is that Theorem 4 holds under very weak assumptions,[3] that G resembles a random subgraph of H with edge-probability $\rho = \Theta(\sqrt{(\log n)/n})$, and that the edge-estimate (2.1) implies $\alpha(G) = O(\sqrt{n \log n})$ for many well-behaved host graphs $H \subseteq K_n$.

Theorem 4. *There exist $\beta_0, D_0 > 0$ such that, for all $\gamma, \delta \in (0,1]$, $\beta \in (0,\beta_0)$ and $C \geq D_0/(\delta^2\sqrt{\beta}\gamma)$, the following holds for all $n \geq n_0(\gamma,\delta,\beta,C)$, with $\rho := \sqrt{\beta(\log n)/n}$. For any n-vertex graph H, there exists a triangle-free subgraph $G \subseteq H$ on the same vertex-set such that*

$$e_G(A,B) = (1 \pm \delta)\rho e_H(A,B) \tag{2.1}$$

for all disjoint vertex-sets $A, B \subseteq V(H)$ with $|A| = |B| = \lceil C\sqrt{n \log n}\rceil$ and $e_H(A,B) \geq \gamma|A||B|$.

Our proof uses a semi-random variant of the triangle-free process to construct $G \subseteq H$, extending and simplifying Kim's $R(3,t)$-approach for the complete case $H = K_n$ (see Sections 2.2–2.3 and Theorem 9 for the details). In particular, besides handling the diffi-

[3] Note that Theorem 4 does *not* require the host graph H to be approximately degree or codegree regular. Furthermore, even if $G \subseteq H$ was a random subgraph with edge-probability ρ, then by standard calculations we would only expect the edge-estimate (2.1) to hold for vertex-sets $A, B \subseteq V(H)$ where the number of edges $e_H(A,B)$ is reasonably large (see Remark 11 for the details, which also indicates that the constant C in Theorem 4 has the correct dependence on γ, δ, β).

culties arising due to incomplete host graphs $H \subseteq K_n$ (by, e.g., exploiting a 'stabilization mechanism' to keep various parameters under control), the major technical difference lies in the way we analyze the properties of all large vertex-sets (by, e.g., focusing on bipartite subgraphs, applying a concentration inequality of Warnke [118], and showing concentration in (2.1) instead of just $e_G(A, B) \geq 1$). Together with some streamlining of Kim's arguments (by, e.g., using fewer variables, applying convenient bounded differences inequalities, and some changes to the semi-random construction), this leads to a shorter and hopefully more accessible proof even in the complete case $H = K_n$. As a by-product, we also obtain a randomized polynomial-time algorithm which constructs $G \subseteq H$ (see Remark 10).

Theorem 4 will be the main tool for establishing our main packing result Theorem 1. Let us briefly sketch the argument (deferring the details to Section 2.1.5). The idea is to sequentially choose the triangle-free subgraphs $G_i \subseteq H_i := K_n \setminus \bigcup_{0 \leq j < i} G_j$ via Theorem 4 with $\delta \in (0, 1)$, using the pseudo-random edge-estimate (2.1) to inductively control the number of remaining edges (between large sets) in H_i as

$$e_{H_i}(A, B) = (1 - (1 \pm \delta)\rho)^i \cdot |A||B| \tag{2.2}$$

for all disjoint $A, B \subseteq V(H)$ of size $s := \lceil C\sqrt{n \log n} \rceil$, stopping when the right hand side of (2.2) drops below $\epsilon|A||B|$ after $I = \Theta(\log(1/\epsilon)/\rho) = \Theta(\sqrt{n/\log n})$ steps. A double counting argument will then show that the leftover graph H_I contains at most $\epsilon\binom{n}{2}$ edges, so that $\sum_{0 \leq i < I} e(G_i) = e(K_n \setminus H_I) \geq (1 - \epsilon)\binom{n}{2}$. Furthermore, $e_{G_i}(A, B) = (1 \pm \delta)\rho e_{H_i}(A, B) > 0$ implies $\alpha(G_i) < 2s = O(\sqrt{n \log n})$, completing this rough proof sketch of Theorem 1 (assuming Theorem 4).

We believe that variants of Theorems 1 and 4 also hold for many other forbidden graphs (using semi-random variants of the H-free process [88, 12, 116, 115, 91]); we hope to return to this topic in a future work.

2.1.4 Organization

The remainder of this chapter is organized as follows. In Section 2.1.5 we use Theorem 4 to state and prove some extensions of our main packing result Theorem 1. In Section 2.2 we introduce a semi-random variation of the triangle-free process and state our main result for this Rödl nibble type construction (that implies our main tool Theorem 4, see Section 2.2.4), which is then subsequently proved in Section 2.3.

2.1.5 Further results

Our methods allow us to extend Theorem 1 to $R(3,t)$-packings of graphs which are far from complete. Our fourth theorem shows that if $H \subseteq K_n$ only satisfies certain uniformity conditions on its edge distribution (that resemble a weak form of pseudo-randomness, see (2.3) below), then we can still approximately decompose H into a packing of nearly optimal Ramsey $R(3,t)$ graphs (again by an efficient randomized algorithm).

Theorem 5. *For all $\epsilon, \xi, C_0 > 0$ there exist $n_0, C_1, D > 0$ such the following holds for all $n \geq n_0$. If H is an n-vertex graph satisfying*

$$\min_{\substack{\text{disjoint } A, B \subseteq V(H): \\ |A|=|B|=\lceil C_0\sqrt{n \log n}\rceil}} \frac{e_H(A,B)}{|A||B|} \geq \xi, \tag{2.3}$$

then there is an edge-disjoint collection $(G_i)_{i\in\mathcal{I}}$ of $|\mathcal{I}| = \lceil D\sqrt{n/\log n}\rceil$ triangle-free subgraphs $G_i \subseteq H$ with $V(G_i) = V(H)$, $\max_{i\in\mathcal{I}} \alpha(G_i) \leq C_1\sqrt{n \log n}$ and $\sum_{i\in\mathcal{I}} e(G_i) \geq (1-\epsilon)e(H)$.

Note that the case $H = K_n$ and $\xi = C_0 = 1$ implies Theorem 1. Furthermore, the case $H = G_{n,p}$, $\xi = p/2$ and $C_0 = 1$ routinely implies the following sparse analogue of Theorem 1 for binomial random graphs $G_{n,p}$.

Corollary 6. *For any $p \in (0,1]$ and $\epsilon > 0$ there exist $C, D > 0$ such that, with probability at least $1-o(1)$, the following event holds: there exists an edge-disjoint collection $(G_i)_{i\in\mathcal{I}}$*

of $|\mathcal{I}| = \lceil D\sqrt{n/\log n}\rceil$ *triangle-free graphs* $G_i \subseteq G_{n,p}$ *on* n *vertices with* $\max_{i\in\mathcal{I}} \alpha(G_i) \le C\sqrt{n\log n}$ *and* $\sum_{i\in\mathcal{I}} e(G_i) = (1\pm\epsilon)p\binom{n}{2}$.

We conjecture that Corollary 6 (with $|\mathcal{I}| = \lceil Dp\sqrt{n/\log n}\rceil$ and constants $C, D > 0$ depending only on ϵ) holds for much sparser random graphs $G_{n,p}$ with edge-probabilities of form $p = p(n) \ge n^{-1/2+o(1)}$, say.[4]

We conclude the introduction with the short proof of Theorem 5, which proceeds by sequentially choosing the graphs $G_i \subseteq H \setminus \bigcup_{0\le j<i} G_j$ via Theorem 4 (generalizing the argument sketched in Section 2.1.3). The reader mainly interested in the proof of Theorem 4 may perhaps wish to skip straight to Section 2.2.

Proof of Theorem 5 (assuming Theorem 4). We may assume $\epsilon < 1$ (as decreasing ϵ gives a stronger conclusion). For concreteness, set $\delta := 1/4$, $\gamma := \epsilon^2\xi$, $\beta := \beta_0/2$ and $C := \max\{C_0,\ D_0/(\delta^2\sqrt{\beta}\gamma)\}$, where β_0, D_0 are defined as in Theorem 4. Let $C_1 := 3C$, $s := \lceil C\sqrt{n\log n}\rceil$, $\rho := \sqrt{\beta(\log n)/n}$, and $I := \lceil \log(1/\epsilon)/(\rho(1-\delta))\rceil$.

Define $H_0 := H$. Let $\mathfrak{S}$ denote the set of all pairs (A, B) of disjoint vertex-sets $A, B \subseteq V(H)$ with $|A| = |B| = s$. Combining a 'handshaking lemma' like double counting argument with the assumed lower bound (2.3), writing $t := \lceil C_0\sqrt{n\log n}\rceil$ it follows that

$$\frac{e_{H_0}(A,B)}{|A||B|} = \frac{\sum_{\tilde{A}\subseteq A, \tilde{B}\subseteq B:\, |\tilde{A}|=|\tilde{B}|=t} e_H(\tilde{A},\tilde{B})}{s^2\cdot\binom{s-1}{t-1}\binom{s-1}{t-1}} \ge \frac{\binom{s}{t}\binom{s}{t}\cdot\xi t^2}{s^2\binom{s-1}{t-1}\binom{s-1}{t-1}} = \xi \tag{2.4}$$

for all $(A, B) \in \mathfrak{S}$.

The plan is to sequentially choose the graphs $(G_i)_{0\le i<I}$ with $G_i \subseteq H_i$ such that, setting $H_{i+1} := H_i \setminus G_i$ (which ensures that all the G_i are edge-disjoint), for all $0 \le i \le I$ we inductively have

$$\frac{e_{H_i}(A,B)}{e_{H_0}(A,B)} \in \Big[\big(1-(1+\delta)\rho\big)^i,\ \big(1-(1-\delta)\rho\big)^i\Big] \qquad \text{for all } (A,B)\in\mathfrak{S}. \tag{2.5}$$

[4]The range of $p = p(n)$ in this conjecture is essentially best possible, since it is well-known that typically $\alpha(G_{n,p}) \gg \sqrt{n\log n}$ for $p \ll \sqrt{(\log n)/n}$. Furthermore, although we have not checked all details, it seems that our proofs can be modified to verify the conjecture for $p \ge n^{-\delta}$, where $\delta > 0$ is some small constant; so the main question is whether $p \ge n^{-1/2+o(1)}$ suffices.

Turning to the details, note that inequality (2.5) holds trivially for $i = 0$. Given H_i with $0 \leq i \leq I-1$ satisfying (2.5), by combining the definition of I with $(1+2\delta)/(1-\delta) = 2$ and (2.4) it follows for $n \geq n_0(\beta)$ that, say,

$$\frac{e_{H_i}(A,B)}{|A||B|} \geq e^{-(1+2\delta)\rho(I-1)} \cdot \frac{e_{H_0}(A,B)}{|A||B|} \geq \epsilon^2 \cdot \xi = \gamma \qquad \text{for all } (A,B) \in \mathfrak{S}. \tag{2.6}$$

Using Theorem 4, for $n \geq n_0(\epsilon, \xi, \delta, \beta, C)$ we can thus find a triangle-free subgraph $G_i \subseteq H_i$ with $e_{G_i}(A,B) = (1 \pm \delta)\rho e_{H_i}(A,B) > 0$ for all $(A,B) \in \mathfrak{S}$. Hence $\alpha(G_i) < 2s \leq 3C\sqrt{n \log n}$, say. Furthermore, noting $e_{H_{i+1}}(A,B) = e_{H_i}(A,B) - e_{G_i}(A,B)$, it is immediate that $H_{i+1} = H_i \setminus G_i$ maintains (2.5).

Finally, for the number of edges of $\bigcup_{0 \leq i < I} G_i = H_0 \setminus H_I$, by (2.5) and definition of I it follows that

$$e_{H_0 \setminus H_I}(A,B) \geq \left(1 - e^{-(1-\delta)\rho I}\right) \cdot e_{H_0}(A,B) \geq (1-\epsilon) e_{H_0}(A,B) \tag{2.7}$$

for all $(A,B) \in \mathfrak{S}$. Using a double counting argument similar to (2.4), in view of (2.7) and $H_0 = H$ we infer

$$e(H_0 \setminus H_I) = \frac{\sum_{(A,B)\in\mathfrak{S}} e_{H_0 \setminus H_I}(A,B)}{2\binom{n-2}{s-1}\binom{n-2-(s-1)}{s-1}} \geq (1-\epsilon) \cdot \frac{\sum_{(A,B)\in\mathfrak{S}} e_H(A,B)}{2\binom{n-2}{s-1}\binom{n-2-(s-1)}{s-1}} = (1-\epsilon)e(H),$$

completing the proof of $\sum_{0 \leq i < I} e(G_i) = e(H_0 \setminus H_I) \geq (1-\epsilon)e(H)$. $\square$

2.2 The nibble: semi-random triangle-free process

The remainder of this chapter is devoted to the proof of our main tool Theorem 4. Given an n-vertex graph H with vertex-set $V = V(H)$ and edge-set $E(H)$, inspired by Kim [67] our strategy is to incrementally construct the triangle-free edge-set of $G \subseteq H$ using a semi-random variation of the triangle-free process (adding large chunks of random-like edges in each step; see also Footnotes 1–2 on page 1). One key difference to [67, 10] is that our approach only uses edges from the host graph H (and not the complete graph K_n). In particular, deferring the details to Section 2.2.1, the rough plan of our Rödl nibble type construction is to step-by-step build up a 'random' set of edges $E_i \subseteq E(H)$ and a triangle-free subset $T_i \subseteq E_i$; we also keep track of a set

$$O_i \subseteq \{e \in E(H) \setminus E_i : \ e \text{ does not form a triangle with any two edges of } E_i\} \tag{2.8}$$

of 'open' edges that can still be added. The idea of each step is to choose a small number of random edges $\Gamma_{i+1} \subseteq O_i$ so that only a few new triangles are created in $E_{i+1} = E_i \cup \Gamma_{i+1}$. This allows us to find an edge-subset $\Gamma'_{i+1} \subseteq \Gamma_{i+1}$, with $|\Gamma'_{i+1}| \approx |\Gamma_{i+1}|$, such that $T_{i+1} = T_i \cup \Gamma'_{i+1}$ remains triangle-free.[5] After

$$I := \lceil n^{\beta} \rceil \tag{2.9}$$

such alteration-method based steps, we eventually obtain a triangle-free graph $G = (V, T_I) \subseteq H$, which intuitively ought to be 'random enough' to resemble (many features of) a random subgraph of H.

2.2.1 Details of the nibble construction

Turning to the details of the nibble construction, consistent with (2.8) we start with

$$O_0 := E(H) \quad \text{and} \quad E_0 := T_0 := \Gamma_0 := \varnothing. \tag{2.10}$$

[5] For the construction of T_{i+1} it might seem overly complicated to define O_i with respect to E_i (and not T_i). However, this slightly wasteful definition actually simplifies the analysis: e.g., for the purpose of tracking various auxiliary variables, it intuitively is easier to understand the effect of adding the random edges Γ_{i+1} (rather than some subset $\Gamma'_{i+1} \subseteq \Gamma_{i+1}$). Using an inclusion in (2.8) might also seem overly complicated, but it again simplifies the analysis: by removing some extra edges it actually becomes easier to prove concentration (see the 'stabilization mechanism' discussion around (2.21) and Lemma 19).

In step $i+1 \geq 1$ we then set

$$E_{i+1} := E_i \cup \Gamma_{i+1}, \tag{2.11}$$

where each edge $e \in O_i$ is included in Γ_{i+1}, independently, with probability

$$p := \sigma/\sqrt{n}. \tag{2.12}$$

(The definition of the deterministic parameter $\sigma \ll 1$ is deferred to (2.35) in Section 2.2.3.) Note that $T_i \cup \Gamma_{i+1}$ is not necessarily triangle-free, since two or three edges of a triangle could enter via $\Gamma_{i+1} \subseteq O_i$ (one edge is not enough by (2.8) and $T_i \subseteq E_i$), i.e., via the following set of 'bad' pairs and triples of Γ_{i+1}–edges:

$$\begin{aligned} \mathcal{B}_{i+1} :=& \big\{\{wu, wv\} \subseteq \Gamma_{i+1} :\ uv \in T_i,\ |\{u,v,w\}| = 3\big\} \\ &\cup\ \big\{\{uv, vw, wu\} \subseteq \Gamma_{i+1} :\ |\{u,v,w\}| = 3\big\}, \end{aligned} \tag{2.13}$$

where we write $xy = \{x, y\}$ for brevity. To avoid triangles in T_{i+1} by alteration, we thus take $\mathcal{D}_{i+1}$ to be a maximal collection of pairwise edge-disjoint elements of $\mathcal{B}_{i+1}$ (say the first one in lexicographic order to resolve ties; any other deterministic choice also works, see Remark 7 and Section 2.3.5), and then set[6]

$$T_{i+1} := T_i \cup \big(\Gamma_{i+1} \setminus E(\mathcal{D}_{i+1})\big), \tag{2.14}$$

where we write $E(\mathcal{D}_{i+1}) := \bigcup_{\alpha \in \mathcal{D}_{i+1}} \alpha$ for the set of edges in the pairs and triples of $\mathcal{D}_{i+1}$. Note that T_{i+1} is indeed triangle-free by maximality of $\mathcal{D}_{i+1} \subseteq \mathcal{B}_{i+1}$. Defining

$$Y_{uv}(i) := \{uw \in O_i :\ vw \in E_i\}\ \cup\ \{vw \in O_i :\ uw \in E_i\}, \tag{2.15}$$

[6]The standard alteration approach of removing one edge from each element of $\mathcal{B}_{i+1}$ seems harder to analyze: e.g., removing the edges of a maximal edge-disjoint collection $\mathcal{D}_{i+1} \subseteq \mathcal{B}_{i+1}$ greatly facilitates the technical calculations in Section 2.3.5.

we now turn to the open edge-set $O_{i+1} \subseteq O_i \setminus \Gamma_{i+1}$: by (2.8) the set $C^{(1)}_{i+1} \cup C^{(2)}_{i+1} \subseteq O_i$ of newly 'closed' edges (that form a triangle with some two edges of E_{i+1}) is given by

$$C^{(1)}_{i+1} := \{f \in O_i :\ Y_f(i) \cap \Gamma_{i+1} \neq \varnothing\}, \tag{2.16}$$

$$C^{(2)}_{i+1} := \{uv \in O_i :\ \text{there is } w \text{ s.t. } uw \in \Gamma_{i+1},\ vw \in \Gamma_{i+1}\}. \tag{2.17}$$

Mimicking a technical idea of Alon, Kim and Spencer [6], we intuitively increase the set of closed edges (via the random set S_{i+1} below) in order to add a 'stabilization mechanism' to our construction,[7] and define

$$C_{i+1} := C^{(1)}_{i+1} \cup S_{i+1}, \tag{2.18}$$

$$O_{i+1} := O_i \setminus \big(\Gamma_{i+1} \cup C_{i+1} \cup C^{(2)}_{i+1}\big), \tag{2.19}$$

where each edge $e \in O_i$ is included in S_{i+1}, independently, with 'stabilization' probability

$$\hat{p}_{e,i} := 1 - (1-p)^{\max\{2q_i(\pi_i+\sqrt{\sigma})\sqrt{n} - |Y_e(i)|,\, 0\}}. \tag{2.20}$$

(The definition of the deterministic parameters q_i, π_i is deferred to (2.36)–(2.37) in Section 2.2.3.) Roughly put, the main point of the technical definitions of S_{i+1} and $\hat{p}_{e,i}$ will be that all the conditional probabilities

$$\begin{aligned} \mathbb{P}(e \notin C_{i+1} \mid O_i, E_i) &= \mathbb{P}(e \notin C^{(1)}_{i+1} \mid O_i, E_i) \cdot (1 - \hat{p}_{e,i}) \\ &= (1-p)^{\max\{2q_i(\pi_i+\sqrt{\sigma})\sqrt{n},\, |Y_e(i)|\}} \end{aligned} \tag{2.21}$$

can inductively be made equal and thus independent of the history (by only maintaining a weak upper bound on $\max_e |Y_e(i)|$; see (2.45), (2.62) and Lemma 19), which in turn helps

[7] Kim uses a different stabilization mechanism in [67, Section 5.1]: instead of introducing the random sets S_j, he deterministically modifies the underlying graphs in each step (by temporarily adding some extra edges and vertices), mimicking an earlier 'regularization' idea from [62]. We find our randomized approach more elegant, and easier to implement algorithmically.

to keep various error terms under control.

Remark 7. *Note that each step of our nibble construction requires only randomized polynomial time (since we can easily find a* maximal *edge-disjoint collection $\mathcal{D}_{i+1} \subseteq \mathcal{B}_{i+1}$ by a deterministic greedy algorithm).*

2.2.2 Pseudo-random intuition: trajectory equations

In this informal section we give a heuristic explanation of the differential equation that predicts the behaviour of (O_i, E_i) for $0 \leq i \leq I \approx n^{\beta}$. Inspired by [109, 67], our main non-rigorous ansatz is that the edge-sets (O_i, E_i) should resemble properties of a random subgraph of H with two types of edges, where

$$\mathbb{P}(e \in O_i) \approx q_i \quad \text{and} \quad \mathbb{P}(e \in E_i) \approx \pi_i/\sqrt{n} \tag{2.22}$$

are approximately independent. We now derive properties of q_i, π_i that are consistent with this ansatz. For example, combining $E_{i+1} = E_i \cup \Gamma_{i+1}$ with the random construction of $\Gamma_{i+1} \subseteq O_i$, we expect to have

$$\mathbb{P}(e \in E_{i+1}) - \mathbb{P}(e \in E_i) = \mathbb{P}(e \in \Gamma_{i+1} \mid e \in O_i)\,\mathbb{P}(e \in O_i) \approx p \cdot q_i = \sigma q_i/\sqrt{n}, \tag{2.23}$$

which together with (2.22) and $E_0 = \varnothing$ suggests that

$$\pi_{i+1} - \pi_i \approx \sigma q_i \quad \text{and} \quad \pi_0 \approx 0. \tag{2.24}$$

Furthermore, with lots of hand-waving, by (2.19) we intuitively have $O_i \setminus O_{i+1} = \Gamma_{i+1} \cup C_{i+1} \cup C_{i+1}^{(2)} \approx C_{i+1}$ (since each closed edge in $C_{i+1}^{(2)}$ requires the presence of at least two random edges from $\Gamma_{i+1} \subseteq O_i$). As (2.22) suggests $\mathbb{E}\,|Y_e(i)| \lesssim 2q_i\pi_i\sqrt{n}$, by the stabilization mechanism (2.21) and $p = \sigma/\sqrt{n}$ we thus loosely expect that

$$\mathbb{P}(e \in O_{i+1} \mid O_i, E_i) \approx \mathbb{P}(e \notin C_{i+1} \mid O_i, E_i) = (1-p)^{2q_i(\pi_i+\sqrt{\sigma})\sqrt{n}} \approx 1 - 2\sigma q_i \pi_i$$

for $e \in O_i$, where we bluntly ignored the $\sqrt{\sigma}$-term in the exponent. Similar to (2.23), using (2.22) we thus ought to have

$$q_{i+1} - q_i \approx \mathbb{P}(e \in O_{i+1}) - \mathbb{P}(e \in O_i) \approx -2\sigma q_i \pi_i \cdot \mathbb{P}(e \in O_i) \approx -2\sigma q_i^2 \pi_i. \tag{2.25}$$

To extract the behaviour of π_I from (2.24) and (2.25), we further assume that $\pi_i \approx \Psi(i\sigma)$ holds for some smooth function $\Psi(x)$, where $\sigma \ll 1$ is tiny. Using Taylor series, in view of (2.24) and $O_0 = E(H)$ this suggests that

$$q_i \approx \Psi'(i\sigma) \quad \text{and} \quad q_0 \approx 1. \tag{2.26}$$

Together with (2.25) and the initial values from (2.24) and (2.26), this leads to the second order differential equation $\Psi''(x) = -2\Psi'(x)^2\Psi(x)$ with $\Psi'(0) = 1$ and $\Psi(0) = 0$, which in turn reduces to the simple ODE

$$\Psi'(x) = e^{-\Psi^2(x)} \quad \text{and} \quad \Psi(0) = 0. \tag{2.27}$$

Noting the implicit solution $x = \int_0^{\Psi(x)} e^{t^2} dt$, it now is easy to derive that $\Psi(x) \approx \sqrt{\log x}$ as $x \to \infty$ (see, e.g., the proof of (2.57) in Section 2.4). Since $I \approx n^\beta$ is sufficiency large compared to σ (which will be of form $\sigma = (\log n)^{-\Theta(1)}$, see (2.35) in Section 2.2.3), this makes it plausible that

$$\pi_I \approx \Psi(I\sigma) \approx \sqrt{\log(I\sigma)} \approx \sqrt{\beta \log n}. \tag{2.28}$$

Finally, since by construction we expect $|E_{i+1} \setminus E_i| \approx |T_{i+1} \setminus T_i|$ to hold for all $0 \leq i <$

I, the edge-sets E_I and T_I ought to share many properties. Together with (2.22) and (2.28) this intuitively suggests

$$\mathbb{P}(e \in T_I) \approx \mathbb{P}(e \in E_I) \approx \sqrt{\beta(\log n)/n}, \tag{2.29}$$

making the pseudo-random edge-estimate (2.1) plausible for $G = (V, T_I)$ with $T_I \subseteq E_I \subseteq E(H)$.

2.2.3 Definitions and parameters

In this section we formally define several variables and parameters used in our analysis of the nibble construction. We start with two standard notions from graph theory: for any edge-subset $S \subseteq \binom{V}{2}$ we write

$$S(A, B) := \{ab \in S : \ a \in A, \ b \in B\}, \tag{2.30}$$

$$N_S(v) := \{w \in V : \ vw \in S\}, \tag{2.31}$$

where $A, B \subseteq V$ are vertex-disjoint. For all pairs of distinct vertices $u, v \in V$ we then define

$$X_{uv}(i) := N_{O_i}(u) \cap N_{O_i}(v), \tag{2.32}$$

$$Z_{uv}(i) := N_{E_i}(u) \cap N_{E_i}(v), \tag{2.33}$$

where $|X_{uv}(i)|$ and $|Z_{uv}(i)|$ intuitively correspond to an 'open codegree' and the usual codegree, respectively (note that $|Y_{uv}(i)|$ defined in (2.15) corresponds to a 'mixed codegree').

Guided by Section 2.2.2, we define $\Psi(x)$ as the unique solution to the differential equation

$$\Psi'(x) = e^{-\Psi^2(x)} \quad \text{and} \quad \Psi(0) = 0, \tag{2.34}$$

as suggested by (2.27). With the heuristics (2.22) in mind, we then introduce the parameters

$$\sigma := (\log n)^{-2}, \tag{2.35}$$

$$q_i := \Psi'(i\sigma) = e^{-\Psi^2(i\sigma)}, \tag{2.36}$$

$$\pi_i := \sigma + \sum_{j=0}^{i-1} \sigma q_j = \pi_{i-1} + \sigma q_{i-1} \mathbb{1}_{\{i \geq 1\}}, \tag{2.37}$$

making (2.24) and (2.26) rigorous (starting with $\pi_0 = \sigma > 0$ leads to cleaner formulae later on). With foresight, for $i \leq I$ we also introduce the 'relative error' parameter

$$\tau_i := 1 - \frac{\delta \pi_i}{2\pi_I} = \tau_{i-1} - \frac{\delta \sigma q_{i-1}}{2\pi_I} \mathbb{1}_{\{i \geq 1\}}, \tag{2.38}$$

which slowly degrades from $\tau_0 = 1 - o(\delta)$ to $\tau_I = 1 - \delta/2$.

With an eye on Theorem 4, for concreteness we introduce the absolute constants[8]

$$D_0 := 108 \quad \text{and} \quad \beta_0 := 1/14, \tag{2.39}$$

as well as the set-sizes (with $s_0 \ll s$) and idealized edge-probability

$$s := \lceil C\sqrt{n \log n} \rceil, \qquad s_0 := \lfloor \sigma^4 q_I^2 s \rfloor, \quad \text{and} \quad \rho := \sqrt{\beta(\log n)/n}, \tag{2.40}$$

and, recalling $O_0 = E(H)$, the collection of 'relevant' pairs of large vertex-sets

$$\begin{aligned} \mathfrak{S}_{s,\gamma} := \{(A, B) : \text{disjoint } A, B \subseteq V \text{ with } |A| = |B| = s \\ \text{and } |O_0(A, B)| \geq \gamma |A||B|\}. \end{aligned} \tag{2.41}$$

[8]To make this chapter easier to read, we have made no attempt to optimize the constants D_0, β_0 in (2.39).

2.2.4 Main nibble result: pseudo-random properties

In this section we state our main nibble result Theorem 9, which implies our main tool Theorem 4 and establishes various pseudo-random properties of $(O_i, E_i, T_i, \Gamma_i)_{0 \le i \le I}$. The following event is of core interest:

$$\mathcal{T}_I := \big\{ |T_I(A,B)| = (1 \pm \delta)\rho |O_0(A,B)| \text{ for all } (A,B) \in \mathfrak{S}_{s,\gamma} \big\}. \tag{2.42}$$

Indeed, it implies the conclusion of Theorem 4 with $G = (V, T_I)$ since the edge-set $T_I \subseteq E_I \subseteq E(H) = O_0$ is triangle-free. To get a handle on $\mathcal{T}_I$, in view of Section 2.2.1 it is natural that we also require some control over the other edge-sets $(E_i, O_i, \Gamma_i)_{0 \le i \le I}$. To this end we introduce the 'good' events

$$\mathfrak{X}_i := \mathcal{N}_i \cap \mathcal{P}_i \cap \mathcal{Q}_i^+ \cap \mathcal{Q}_i \quad \text{and} \quad \mathfrak{X}_{\le i} := \bigcap_{0 \le j \le i} \mathfrak{X}_j, \tag{2.43}$$

where the following auxiliary events encapsulate various pseudo-random properties:

$$\mathcal{N}_i := \big\{ |N_{O_i}(v)| \le q_i n \text{ and } |N_{\Gamma_i}(v)| \le 2\sigma q_{i-1}\sqrt{n} \text{ for all } v \in V \big\}, \tag{2.44}$$

$$\begin{aligned} \mathcal{P}_i := \big\{ &|X_{uv}(i)| \le q_i^2 n,\ |Y_{uv}(i)| \le 2 q_i \pi_i \sqrt{n}, \\ &\text{and } |Z_{uv}(i)| \le i (\log n)^9 \text{ for all } u, v \in V \text{ with } u \neq v \big\}, \end{aligned} \tag{2.45}$$

$$\mathcal{Q}_i^+ := \big\{ |O_i(A,B)| \le q_i |A||B| \text{ for all disjoint } A, B \subseteq V \text{ with } |A|, |B| \ge s_0 \big\}, \tag{2.46}$$

$$\mathcal{Q}_i := \big\{ \tau_i q_i |O_0(A,B)| \le |O_i(A,B)| \le q_i |O_0(A,B)| \text{ for all } (A,B) \in \mathfrak{S}_{s,\gamma} \big\}. \tag{2.47}$$

In words, the above events give bounds for degree-like variables ($\mathcal{N}_i$), codegree-like variables ($\mathcal{P}_i$), and the number of open edges ($\mathcal{Q}_i^+$ and $\mathcal{Q}_i$). A subtle but important point is that $\mathcal{N}_i$, $\mathcal{P}_i$ and $\mathcal{Q}_i^+$ only guarantee one-sided concentration, i.e., ensure upper bounds but no matching lower bounds (which can fail badly, for example, $|Y_{uv}(i)| = 0$ holds when $uv \in E_i$). Merely $\mathcal{Q}_i$ guarantees two-sided concentration, which is harder to prove, but crucial for establishing the edge-estimate from $\mathcal{T}_I$ (see the heuristic below Theorem 9).

With $\tau_i \approx 1$ and $O_0 = E(H) \subseteq E(K_n)$ in mind, most of the bounds in (2.42) and (2.44)–(2.47) can easily be guessed by the pseudo-random heuristics (2.22) and (2.29) from Section 2.2.2 (the $|N_{\Gamma_i}(v)|$-bound is one exception: based on $\mathbb{E}\,|N_{\Gamma_i}(v)| = p \cdot \mathbb{E}\,|N_{O_{i-1}}(v)|$, it contains an extra factor of 2 to avoid additive error terms; another exception is the $|Z_{uv}(i)|$-bound: it relaxes the prediction $\mathbb{E}\,|Z_{uv}(i)| \lesssim \pi_i^2 = O(\log n)$ for technical reasons).

Inspecting (2.44)–(2.47) in the special case $i = 0$, it is not difficult to see that the good event $\mathfrak{X}_0 = \mathfrak{X}_{\le 0}$ always holds (by combining $q_0 = 1 \ge \tau_0$ and $\sigma, q_{-1}, \pi_0 \ge 0$ with $E_0 = T_0 = \Gamma_0 = \varnothing$).

Remark 8. *The event $\mathfrak{X}_0$ holds deterministically for any n-vertex host graph H.*

Our main nibble result (which is at the heart of this chapter) states that, under fairly natural constraints, the pseudo-random events $\mathcal{T}_I$ and $\mathfrak{X}_{\le I}$ both hold with very high probability. Recall that $I \approx n^{\beta}$, and that all pairs $(A, B) \in \mathfrak{S}_{s,\gamma}$ of vertex-sets satisfy $|O_0(A, B)| \ge \gamma s^2$ and $|A| = |B| = s \approx C\sqrt{n \log n}$.

Theorem 9 (Main nibble result)**.** *For all $\gamma, \delta \in (0, 1]$, $\beta \in (0, \beta_0)$ and $C \ge D_0/(\delta^2\sqrt{\beta}\gamma)$ the following holds for $n \ge n_0(\gamma, \delta, \beta, C)$: we have $\mathbb{P}(\mathcal{T}_I \cap \mathfrak{X}_{\le I}) \ge 1 - n^{-\omega(1)}$ for any n-vertex host graph H.*

Proof of Theorem 4. If the event $\mathcal{T}_I$ holds, then the triangle-free graph $G := (V, T_I)$ has the claimed properties by (2.42), $V = V(H)$ and $T_I \subseteq E_I \subseteq E(H) = O_0$, so Theorem 9 completes the proof. $\square$

Remark 10. *In view of $I = O(n^{\beta_0})$ and Remark 7, the nibble thus yields a randomized polynomial time algorithm (with error probability $\le n^{-\omega(1)}$) for constructing the triangle-free $G \subseteq H$ from Theorem 4.*

Remark 11. *The heuristic edge-estimate* (2.29) *suggests that in Theorem 9 the dependence of the constant C on δ, β, γ is qualitatively best possible, since it would also naturally arise* *if $G = (V, T_I) \subseteq H$ was a random subgraph with edge-probability $\rho \approx \sqrt{\beta(\log n)/n}$. Indeed, for all $(A, B) \in \mathfrak{S}_{s,\gamma}$ the expected number of edges between A and B would then be at least $\lambda_{A,B} := \mathbb{E}\,|T_I(A, B)| = \rho|O_0(A, B)| \geq \rho \cdot \gamma s^2 \geq \sqrt{\beta}\gamma C \cdot s \log n$, and the probability that the event $\mathcal{T}_I$ from* (2.42) *fails would therefore be (using a union bound and standard Chernoff bounds) at most $\sum_{(A,B)\in\mathfrak{S}_{s,\gamma}} e^{-\Theta(\delta^2\lambda_{A,B})} \leq n^{2s-\Omega(\delta^2\sqrt{\beta}\gamma C s)} = o(1)$ for $C = \Omega(1/(\delta^2\sqrt{\beta}\gamma))$ large enough.*

We defer the proof of Theorem 9 to Section 2.3, and now just outline a brief heuristic argument that illustrates how the event $\mathfrak{X}_{\leq I} \subseteq \bigcap_{0\leq i\leq I} \mathcal{Q}_i$ is instrumental for establishing the edge-estimate from $\mathcal{T}_I$ (which seems informative). Similar to (2.29), in view of Section 2.2.1 we expect that in each step only few edges are removed due to the creation of triangles, which intuitively suggests

$$|T_{i+1}(A, B) \setminus T_i| \approx |E_{i+1}(A, B) \setminus E_i|.$$

Combining the construction of $E_{i+1} \setminus E_i = \Gamma_{i+1} \subseteq O_i$ with the event $\mathcal{Q}_i$ and $\tau_i \approx 1$, we also expect that

$$|E_{i+1}(A, B) \setminus E_i| = |\Gamma_{i+1}(A, B)| \approx p \cdot |O_i(A, B)| \approx p \cdot q_i|O_0(A, B)|.$$

Recalling $p = \sigma/\sqrt{n}$ and $\rho = \sqrt{\beta(\log n)/n}$, using the definition (2.37) of π_I and the approximation $\pi_I \approx \sqrt{\beta \log n}$ from (2.28) it now becomes plausible that

$$\begin{aligned}
|T_I(A, B)| &= \sum_{0\leq i<I} |T_{i+1}(A, B) \setminus T_i| \\
&\approx \frac{\sum_{0\leq i<I} \sigma q_i}{\sqrt{n}} \cdot |O_0(A, B)| \\
&\approx \frac{\pi_I}{\sqrt{n}} \cdot |O_0(A, B)| \\
&\approx \rho|O_0(A, B)|,
\end{aligned}$$

as suggested by $\mathcal{T}_I$ (Section 2.3.5 contains a rigorous version of this heuristic argument).

2.2.5 Tools and auxiliary estimates

In this preparatory section we gather, for later reference, some results that will be used throughout the proof of Theorem 9 (mostly probabilistic and combinatorial tools, and ending with some auxiliary estimates). On a first reading the reader may perhaps wish to skip straight to Section 2.3.

We start with a convenient version of the bounded differences inequality [81, 82, 117] for Bernoulli variables. Note that the upper tail estimate (2.48) for decreasing functions does *not* have an extra Ct term in the exponent like (2.49). Remarks 13–14 are well-known, see, e.g., [82, Theorem 2.3, 3.8, and 3.9] or [117, Corollary 1.4]. Inequality (2.48) can be deduced from the arguments in [81, Lemma 7.14], but this monotone version does not seem to be widely known; in Section 2.4 we thus include a simple proof for completeness.

Theorem 12. *Let $(\xi_\alpha)_{\alpha\in\mathcal{I}}$ be a finite family of independent random variables with $\xi_\alpha \in \{0,1\}$. Let $f : \{0,1\}^{|\mathcal{I}|} \to \mathbb{R}$ be a function, and assume that there exist numbers $(c_\alpha)_{\alpha\in\mathcal{I}}$ such that the following holds for all $z = (z_\alpha)_{\alpha\in\mathcal{I}} \in \{0,1\}^{|\mathcal{I}|}$ and $z' = (z'_\alpha)_{\alpha\in\mathcal{I}} \in \{0,1\}^{|\mathcal{I}|}$: $|f(z) - f(z')| \le c_\beta$ if $z_\alpha = z'_\alpha$ for all $\alpha \neq \beta$. Define $X := f\big((\xi_\alpha)_{\alpha\in\mathcal{I}}\big)$ and $\lambda := \sum_{\alpha\in\mathcal{I}} c_\alpha^2\, \mathbb{P}(\xi_\alpha = 1)$. Then, for all $t \ge 0$,*

$$\mathbb{P}(X \ge \mathbb{E}\, X + t) \le \exp\left(-\frac{t^2}{2\lambda}\right) \tag{2.48}$$

if the function f is decreasing (i.e., that $f(z) \le f(z')$ whenever $z_\alpha \ge z'_\alpha$ for all $\alpha \in \mathcal{I}$).

Remark 13. *Define $C := \max_{\alpha\in\mathcal{I}} c_\alpha$. If we drop the assumption that f is decreasing, then*

$$\mathbb{P}(X \le \mathbb{E}\, X - t) \le \exp\left(-\frac{t^2}{2(\lambda + Ct)}\right). \tag{2.49}$$

Remark 14. *In the special case $X = \sum_{\alpha\in\mathcal{I}} \xi_\alpha$ we have $C = c_\alpha = 1$ and $\lambda = \mathbb{E}\, X$. Standard Chernoff bounds (or applying* (2.48)-(2.49) *to the decreasing function $-X$) then* *show that in this case $\mathbb{P}(X \leq \mathbb{E}\, X - t)$ and $\mathbb{P}(X \geq \mathbb{E}\, X + t)$ are at most the right hand side of* (2.48) *and* (2.49)*, respectively.*

For random variables with a special combinatorial form (based on the occurrence of events with 'limited overlaps') we shall use the following Chernoff-type upper tail inequality, which is a convenient corollary of a more general result by Warnke [118, Theorem 9]. Note that the exponent of (2.50) scales with $1/C$.

Theorem 15. *Let $(\xi_i)_{i \in \mathfrak{S}}$ be a finite family of independent random variables with $\xi_i \in \{0, 1\}$. Let $(Y_\alpha)_{\alpha \in \mathcal{I}}$ be a finite family of variables $Y_\alpha := \mathbb{1}_{\{\xi_i = 1 \text{ for all } i \in \alpha\}}$ with $\sum_{\alpha \in \mathcal{I}} \mathbb{E}\, Y_\alpha \leq \mu$. Define $Z_C := \max \sum_{\alpha \in \mathcal{J}} Y_\alpha$, where the maximum is taken over all $\mathcal{J} \subseteq \mathcal{I}$ with $\max_{\beta \in \mathcal{J}} |\{\alpha \in \mathcal{J} : \alpha \cap \beta \neq \varnothing\}| \leq C$. Then, for all $C, t > 0$,*

$$\mathbb{P}(Z_C \geq \mu + t) \leq \min\left\{ \left(\frac{e\mu}{\mu + t}\right)^{(\mu+t)/C}, \ \exp\left(-\frac{t^2}{2C(\mu + t)}\right) \right\}. \tag{2.50}$$

The following simple combinatorial lemma formalizes the intuition that we expect $\sum_i |U_i| = O(|U|)$ whenever the subsets $U_i \subseteq U$ are nearly disjoint (i.e., have small pairwise intersections).

Lemma 16. *Suppose that $(U_i)_{i \in \mathcal{I}}$ is a family of subsets $U_i \subseteq U$ with $|U_i| \geq z > 0$ and $|U_i \cap U_j| \leq y$ for all $i \neq j$. Then $z \geq \sqrt{4|U|y}$ implies $|\mathcal{I}| \leq 2|U|/z$ and $\sum_{i \in \mathcal{I}} |U_i| \leq 2|U|$.*

Proof. Aiming at a contradiction, suppose that $|\mathcal{I}| > 2|U|/z$. Then there is $\mathcal{J} \subseteq \mathcal{I}$ with $|\mathcal{J}| = \lfloor 2|U|/z \rfloor + 1$. Observe that, for any $i \in \mathcal{J}$,

$$\sum_{j \in \mathcal{J} : i \neq j} |U_j \cap U_i| \leq (|\mathcal{J}| - 1)y \leq 2|U|y/z \leq z/2 \leq |U_i|/2. \tag{2.51}$$

So we obtain a contradiction by noting that

$$|U| \geq \Big|\bigcup_{i \in \mathcal{J}} U_i\Big| \geq \sum_{i \in \mathcal{J}} \Big(|U_i| - \sum_{j \in \mathcal{J} : i \neq j} |U_j \cap U_i|\Big) \geq \sum_{i \in \mathcal{J}} |U_i|/2 \geq |\mathcal{J}| z/2 > |U|. \tag{2.52}$$

With $|\mathcal{I}| \le 2|U|/z$ in hand, after replacing $\mathcal{J}$ with $\mathcal{I}$, note that (2.51) and the first three inequalities of (2.52) remain valid, completing the proof of $\sum_{i \in \mathcal{I}} |U_i| \le 2|U|$. $\square$

Our final auxiliary result contains a number of convenient estimates involving the parameters q_i, π_i, σ, I defined in Section 2.2.3 and (2.9). Roughly put, (2.55)–(2.57) state that $q_i \approx q_{i+1}$, $1 - 2\sigma q_i \pi_i \approx q_{i+1}/q_i$ and $\pi_I \approx \sqrt{\log(I\sigma)}$, as predicted by (2.25) and (2.28). The technical estimates (2.53)–(2.54) can safely be ignored on a first reading. The proof of Lemma 17 is based on elementary calculus and thus deferred to Section 2.4 (it also establishes $q_i \ge q_I = n^{-\beta+o(1)}$, which together with $I \approx n^\beta$ and (2.54) motivates our choice of $\beta_0 = 1/14$).

Lemma 17. *If $\beta \in (0, \beta_0)$, then there exists $\tau, n_0 > 0$ such that, for all $n \ge n_0$ and $0 \le i \le I$,*

$$\max\Big\{ \max_{j \in \{0,1,2\}} \{q_i \pi_i^j\},\ \sqrt{\sigma}\pi_i \Big\} \le 1, \tag{2.53}$$

$$\min\Big\{ \min_{j \in [4]} \{q_i^j \sqrt{n}\},\ q_i^2 \sqrt{n}/I,\ q_i^3 \sqrt[4]{n}/\sqrt{I} \Big\} \ge n^\tau, \tag{2.54}$$

$$0 \le q_i - q_{i+1} \le 4\sigma \cdot \min\{q_i,\ q_{i+1},\ q_i \pi_i\}, \tag{2.55}$$

$$\big|(1 - 2\sigma q_i \pi_i) - q_{i+1}/q_i\big| \le 8\sigma^2 q_i, \tag{2.56}$$

$$\big|\pi_I - \sqrt{\log(I\sigma)}\big| \le 2. \tag{2.57}$$

As a simple application, for $0 \le i \le I$ we now bound the stabilization probability $\hat{p}_{e,i}$ defined in (2.20). Since (2.54) implies $q_i \sqrt{\sigma}\sqrt{n} \gg 1$, by applying $(1-p)^r \ge 1 - pr = 1 - \sigma r/\sqrt{n}$ (valid for $r \ge 1$) we infer

$$\hat{p}_{e,i} \le 1 - (1-p)^{2q_i(\pi_i + \sqrt{\sigma})\sqrt{n}} \le 2\sigma q_i(\pi_i + \sqrt{\sigma}) \le q_i, \tag{2.58}$$

where we used $\sqrt{\sigma}\pi_i \le 1$ and $\sigma \ll 1$ (see (2.53) and (2.35)) for the last inequality.

2.3 Analyzing the nibble

In this section we prove our main nibble result Theorem 9 (which in turn implies our main tool Theorem 4, see Section 2.2.4) as a corollary of the following auxiliary lemma.

Lemma 18. *Under the assumptions of Theorem 9, for* $n \geq n_0(\gamma, \delta, \beta, C)$ *we have*

$$\mathbb{P}(\neg \mathfrak{X}_{i+1} \mid \mathfrak{X}_{\leq i}) \leq n^{-\omega(1)} \qquad \textit{for all } 0 \leq i \leq I-1, \tag{2.59}$$

$$\mathbb{P}(\neg \mathcal{T}_I \cap \mathfrak{X}_{\leq I}) \leq n^{-\omega(1)}. \tag{2.60}$$

Proof of Theorem 9. Recalling $I \leq \lceil n^{\beta_0} \rceil = n^{O(1)}$ and $\mathfrak{X}_{\leq i} = \bigcap_{0 \leq j \leq i} \mathfrak{X}_j$, note that $\mathbb{P}(\neg \mathfrak{X}_0) = 0$ (see Remark 8) and (2.59) readily imply $\mathbb{P}(\neg \mathfrak{X}_{\leq I}) \leq n^{-\omega(1)}$, which together with (2.60) completes the proof. □

The remainder of this section is devoted to the proof of Lemma 18: the proof of (2.59) with $\neg \mathfrak{X}_{i+1} = \neg \mathcal{N}_{i+1} \cup \neg \mathcal{P}_{i+1} \cup \neg \mathcal{Q}^+_{i+1} \cup \neg \mathcal{Q}_{i+1}$ is spread across Sections 2.3.2–2.3.4, and the proof of (2.60) is given in Section 2.3.5.

2.3.1 Preliminaries: setup and conventions

To avoid clutter, up to (and including) Section 2.3.4 we shall suppress the conditioning in the notation: we will write $\mathbb{P}(\cdot)$ and $\mathbb{E}(\cdot)$ as shorthand for $\mathbb{P}(\cdot \mid \mathcal{F}_i)$ and $\mathbb{E}(\cdot \mid \mathcal{F}_i)$, where $(\mathcal{F}_i)_{0 \leq i \leq I}$ denotes the natural filtration associated with $(O_i, E_i, T_i, \Gamma_i, S_i)_{0 \leq i \leq I}$, as usual. We will also tacitly assume that the $\mathcal{F}_i$-measurable event $\mathfrak{X}_{\leq i}$ holds. Conditional on $\mathcal{F}_i$, note that by construction of the random edge-sets Γ_{i+1} and S_{i+1}, the (conditional) probability space formally consists of the $2|O_i|$ independent Bernoulli random variables $(\mathbb{1}_{\{e \in \Gamma_{i+1}\}}, \mathbb{1}_{\{e \in S_{i+1}\}})_{e \in O_i}$, with $\mathbb{P}(e \in \Gamma_{i+1}) = p = \sigma/\sqrt{n}$ and $\mathbb{P}(e \in S_{i+1}) = \hat{p}_{e,i} \leq q_i$, see (2.58).

Using the above setup and conventions, we shall repeatedly consider random variables of form

$$X = f\Big(\big(\mathbb{1}_{\{e\in\Gamma_{i+1}\}}, \mathbb{1}_{\{e\in S_{i+1}\}}\big)_{e\in O_i}\Big). \tag{2.61}$$

To get a handle on the (conditional) expectation $\mathbb{E}\, X$ we will often use $O_{i+1} \subseteq O_i \setminus C_{i+1}$ together with the following key lemma, which hinges on the stabilization mechanism to equalize all (conditional) probabilities $\mathbb{P}(e \notin C_{i+1})$, see (2.62) below. (The extra $\sqrt{\sigma}$ term in (2.20) ensures that $\mathbb{P}(e \notin C_{i+1}) < q_{i+1}/q_i$ holds with plenty of elbow room, which is convenient for avoiding ugly error terms in the upper bounds of (2.44)–(2.47).)

Lemma 19. *We have* $\mathbb{P}(e \notin C_{i+1}) - q_{i+1}/q_i \in [-3\sigma^{3/2}q_i,\ -\sigma^{3/2}q_i]$ *for all* $e \in O_i$.

Proof. For any $e \in O_i$, since $|Y_e(i)| \le 2q_i\pi_i\sqrt{n}$ by $\mathfrak{X}_{\le i} \subseteq \mathcal{P}_i$, by definition of $C_{i+1} = C^{(1)}_{i+1} \cup S_{i+1}$ we have

$$\begin{aligned} \mathbb{P}(e \notin C_{i+1}) &= \mathbb{P}(e \notin C^{(1)}_{i+1}) \cdot \mathbb{P}(e \notin S_{i+1}) = (1-p)^{|Y_e(i)|} \cdot (1 - \hat{p}_{e,i}) \\ &= (1-p)^{2q_i(\pi_i+\sqrt{\sigma})\sqrt{n}}. \end{aligned} \tag{2.62}$$

It is well-known (and easy to check) that $1-rx \le (1-x)^r \le 1-rx+\binom{r}{2}x^2$ for all $x \in [0,1]$ and $r \ge 2$. Using $\sqrt{n}p = \sigma \ll 1$ and $\max\{q_i, q_i\pi_i, q_i\pi_i^2\} \le 1$ (see (2.53)), it follows that

$$\Big|\mathbb{P}(e \notin C_{i+1}) - \big[1 - 2\sigma q_i(\pi_i + \sqrt{\sigma})\big]\Big| \le 2\sigma^2 q_i^2(\pi_i + \sqrt{\sigma})^2 = O(\sigma^2 q_i) = o(\sigma^{3/2}q_i).$$

This completes the proof since $1 - 2\sigma q_i\pi_i = q_{i+1}/q_i + o(\sigma^{3/2}q_i)$ by (2.56). $\square$

To deduce concentration properties of such random variables X we shall frequently rely on the bounded differences inequality (Theorem 12). In those cases we will bound the associated parameter λ via

$$\lambda = \sum_{e\in O_i} c_e^2\,\mathbb{P}(e \in \Gamma_{i+1}) + \sum_{e\in O_i} \hat{c}_e^2\,\mathbb{P}(e \in S_{i+1}) \le p\sum_{e\in O_i} c_e^2 + q_i \sum_{e\in O_i} \hat{c}_e^2, \tag{2.63}$$

where the *edge-effect* c_e is an upper bound on how much X can change if we modify the indicator $\mathbb{1}_{\{e\in\Gamma_{i+1}\}}$ (alter whether e is in Γ_{i+1} or not), and the *stabilization-effect* $\hat{c}_e$ is an upper bound on how much X can change if we modify the indicator $\mathbb{1}_{\{e\in S_{i+1}\}}$ (alter whether e is in S_{i+1} or not). Moreover, the following simple observation will later allow us to control the above sum (2.63) of these effects.

Lemma 20. *If $\mathfrak{X}_{\le i}$ holds, then $\sum_{e\in O_i}|Y_e(i)\cap J| \le 2q_i\pi_i\sqrt{n}\cdot|J|$ for any edge-subset $J\subseteq\binom{V}{2}$.*

Proof. For any $e\in O_i$, note that $f\in Y_e(i)$ implies $e\in Y_f(i)$. As $Y_f(i)\subseteq O_i$, we infer

$$\sum_{e\in O_i}|Y_e(i)\cap J| = \sum_{f\in J}\sum_{e\in O_i}\mathbb{1}_{\{f\in Y_e(i)\}} \le \sum_{f\in J}\sum_{e\in O_i}\mathbb{1}_{\{e\in Y_f(i)\}} = \sum_{f\in J}|Y_f(i)|.$$

This completes the proof since $\mathfrak{X}_{\le i}\subseteq\mathcal{P}_i$ implies $|Y_f(i)|\le 2q_i\pi_i\sqrt{n}$. □

2.3.2 Event $\mathcal{N}_{i+1}$: degree-like variables $|N_{O_{i+1}}(v)|$ and $|N_{\Gamma_{i+1}}(v)|$

Recall that the event $\mathcal{N}_{i+1}$ defined in (2.44) concerns degree-like variables, ensuring that $|N_{O_{i+1}}(v)|\le q_{i+1}n$ and $|N_{\Gamma_{i+1}}(v)|\le 2\sigma q_i\sqrt{n}$ for all vertices v; see (2.31) for the definition of $N_S(v)$.

Lemma 21. *We have $\mathbb{P}(\neg\mathcal{N}_{i+1})\le n^{-\omega(1)}$.*

Proof. We start with $|N_{O_{i+1}}(v)|$. Note that $O_{i+1}\subseteq O_i\setminus C_{i+1}$ implies

$$|N_{O_{i+1}}(v)| \le \sum_{w\in N_{O_i}(v)}\mathbb{1}_{\{vw\notin C_{i+1}\}} =: X. \tag{2.64}$$

Since $\mathfrak{X}_{\le i}\subseteq\mathcal{N}_i$ implies $|N_{O_i}(v)|\le q_in$, using Lemma 19 we obtain

$$\mathbb{E}\,X = \sum_{w\in N_{O_i}(v)}\mathbb{P}(vw\notin C_{i+1}) \le |N_{O_i}(v)|\cdot(q_{i+1}/q_i-\sigma^{3/2}q_i) \le q_{i+1}n-\sigma^{3/2}q_i^2n. \tag{2.65}$$

Gearing up to apply Theorem 12 to X, we now bound the associated parameter $\lambda\le p\sum_{e\in O_i}c_e^2+q_i\sum_{e\in O_i}\hat{c}_e^2$ from (2.63). Set $\mathcal{X}_v:=\{v\}\times N_{O_i}(v)\subseteq O_i$, and recall that $C_{i+1}=$

$C_{i+1}^{(1)} \cup S_{i+1}$, where $C_{i+1}^{(1)}$ depends only on Γ_{i+1} and thus is independent of S_{i+1}. The edge-effect c_e (an upper bound on how much X changes if we alter whether $e \in \Gamma_{i+1}$ or $e \notin \Gamma_{i+1}$) is thus at most the number of changes to $C_{i+1}^{(1)} \cap \mathcal{X}_v = \{vw \in \mathcal{X}_v : Y_{vw}(i) \cap \Gamma_{i+1} \neq \varnothing\}$. Since $e \in Y_{vw}(i)$ implies $vw \in Y_e(i)$ when $vw \in \mathcal{X}_v$, we infer $c_e \leq |Y_e(i) \cap \mathcal{X}_v| \leq |Y_e(i)| \leq 2q_i\pi_i\sqrt{n}$ by $\mathfrak{X}_{\leq i} \subseteq \mathcal{P}_i$. Using Lemma 20, $|\mathcal{X}_v| = |N_{O_i}(v)| \leq q_i n$, and $q_i\pi_i^2 \leq 1$ (see (2.53)), it follows that

$$\begin{aligned} p\sum_{e\in O_i} c_e^2 &\leq p \cdot 2q_i\pi_i\sqrt{n} \cdot \sum_{e\in O_i} |Y_e(i) \cap \mathcal{X}_v| \\ &\leq \sigma/\sqrt{n} \cdot (2q_i\pi_i\sqrt{n})^2 \cdot |\mathcal{X}_v| \\ &\leq 4\sigma q_i^3\pi_i^2 n^{3/2} \\ &\leq 4\sigma q_i^2 n^{3/2}. \end{aligned} \tag{2.66}$$

By our above discussion, the stabilization-effect $\hat{c}_e$ (an upper bound on how much X changes if we alter whether $e \in S_{i+1}$ or $e \notin S_{i+1}$) is at most the number of changes to $S_{i+1} \cap \mathcal{X}_v$. Hence $\hat{c}_e \leq \mathbb{1}_{\{e\in\mathcal{X}_v\}}$, so that

$$q_i \sum_{e\in O_i} \hat{c}_e^2 \leq q_i \cdot |\mathcal{X}_v| \leq q_i^2 n \ll \sigma q_i^2 n^{3/2}.$$

Noting that X is a decreasing function of the edge-indicators $(\mathbb{1}_{\{e\in\Gamma_{i+1}\}}, \mathbb{1}_{\{e\in S_{i+1}\}})_{e\in O_i}$, using Theorem 12 together with the λ–bound (2.63) and $q_i^2 n^{1/2} \geq n^{\tau}$ (see (2.54)) it follows that

$$\mathbb{P}(|N_{O_{i+1}}(v)| \geq q_{i+1}n) \leq \mathbb{P}(X \geq \mathbb{E}\, X + \sigma^{3/2}q_i^2 n) \leq \exp\left(-\frac{\sigma^3 q_i^4 n^2}{2\cdot 5\sigma q_i^2 n^{3/2}}\right) \leq n^{-\omega(1)}.$$

Taking a union bound over all vertices v completes the proof for the $|N_{O_{i+1}}(v)|$ variables.

Finally, note that $|N_{\Gamma_{i+1}}(v)|$ is a sum of independent Bernoulli random variables with

$$\mathbb{E}\,|N_{\Gamma_{i+1}}(v)| = |N_{O_i}(v)| \cdot p \leq q_i n \cdot \sigma/\sqrt{n} = \sigma q_i\sqrt{n} =: \mu,$$

where we used $\mathfrak{X}_{\le i} \subseteq \mathcal{N}_i$ to bound $|N_{O_i}(v)| \le q_i n$. Applying standard Chernoff bounds (see, e.g., Remark 14), using $q_i\sqrt{n} \ge n^{\tau}$ (see (2.54)) it is routine to deduce that $\mu \gg \log n$ and

$$\mathbb{P}(|N_{\Gamma_{i+1}}(v)| \ge 2\sigma q_i\sqrt{n}) = \mathbb{P}(|N_{\Gamma_{i+1}}(v)| \ge 2\mu) \le \exp\left(-\frac{\mu^2}{2\cdot 2\mu}\right) = \exp\left(-\frac{\mu}{4}\right) \le n^{-\omega(1)}.$$

Taking a union bound over all vertices v completes the proof for the $|N_{\Gamma_{i+1}}(v)|$ variables. □

2.3.3 Event $\mathcal{P}_{i+1}$: codegree-like variables $|X_{uv}(i+1)|$, $|Y_{uv}(i+1)|$ and $|Z_{uv}(i+1)|$

Recall that the event $\mathcal{P}_{i+1}$ defined in (2.45) concerns codegree-like variables, ensuring that $|X_{uv}(i+1)| \le q_{i+1}^2 n$, $|Y_{uv}(i+1)| \le 2q_{i+1}\pi_{i+1}\sqrt{n}$, and $|Z_{uv}(i+1)| \le (i+1)(\log n)^9$ for all pairs uv of vertices.

Lemma 22. *We have* $\mathbb{P}(\neg\mathcal{P}_{i+1}) \le n^{-\omega(1)}$.

Proof. We start with $|X_{uv}(i+1)|$. Recalling $O_{i+1} \subseteq O_i \setminus C_{i+1}$, note that by construction

$$|X_{uv}(i+1)| \le \sum_{w\in X_{uv}(i)} \mathbb{1}_{\{uw\notin C_{i+1} \text{ and } vw\notin C_{i+1}\}} =: X. \tag{2.67}$$

To estimate $\mathbb{E}\,X$, note that the event $f \notin C_{i+1}^{(1)} = \{f \in O_i : Y_f(i) \cap \Gamma_{i+1} \neq \varnothing\}$ is determined by the values of the independent indicator variables $(\mathbb{1}_{\{e\in\Gamma_{i+1}\}})_{e\in Y_f(i)}$. In view of the reasoning (2.62) for the value of $\mathbb{P}(e \notin C_{i+1})$, it follows by construction of $C_{i+1} = C_{i+1}^{(1)} \cup S_{i+1}$ that

$$\begin{aligned} &\mathbb{P}(uw \notin C_{i+1} \text{ and } vw \notin C_{i+1}) \\ &= \mathbb{P}(uw \notin C_{i+1})\,\mathbb{P}(vw \notin C_{i+1}) \cdot (1-p)^{-|Y_{uw}(i)\cap Y_{vw}(i)|}. \end{aligned} \tag{2.68}$$

Since $\mathfrak{X}_{\le i} \subseteq \mathcal{P}_i$ implies $|Y_{uw}(i) \cap Y_{vw}(i)| \le |Z_{uv}(i)| \le I(\log n)^9$ and $|X_{uv}(i)| \le q_i^2 n$, by combining (2.68) with Lemma 19 it follows that

$$\mathbb{E}\, X \leq |X_{uv}(i)| \cdot (q_{i+1}/q_i - \sigma^{3/2} q_i)^2 \cdot (1-p)^{-I(\log n)^9} \leq q_{i+1}^2 n - \sigma^{3/2} q_i^3 n, \qquad (2.69)$$

where for the last inequality we used $pI(\log n)^9 \ll \sigma^{3/2} q_i^3 \ll 1$ (since $q_i^3 \sqrt{n}/I \geq n^{\tau}$ by (2.54)) and $\sigma^3 q_i^4 \ll \sigma^{3/2} q_i^3 \sim \sigma^{3/2} q_{i+1} q_i^2$ (see (2.53)–(2.55)). With an eye on Theorem 12, we now bound the parameter $\lambda \leq p \sum_{e \in O_i} c_e^2 + q_i \sum_{e \in O_i} \hat{c}_e^2$ from (2.63). Set $\mathcal{X}_{uv} := \{u, v\} \times X_{uv}(i) \subseteq O_i$. Analogous to the proof of Lemma 21 for $|N_{O_{i+1}}(v)|$, here we have edge-effect $c_e \leq |Y_e(i) \cap \mathcal{X}_{uv}| \leq |Y_e(i)| \leq 2q_i \pi_i \sqrt{n}$ and stabilization-effect $\hat{c}_e \leq \mathbb{1}_{\{e \in \mathcal{X}_{uv}\}}$. Similar to (2.66), using Lemma 20, $|\mathcal{X}_{uv}| = 2 \cdot |X_{uv}(i)| \leq 2q_i^2 n$ and $q_i \pi_i^2 \leq 1$ it follows that

$$p \sum_{e \in O_i} c_e^2 \leq \sigma/\sqrt{n} \cdot (2q_i \pi_i \sqrt{n})^2 \cdot |\mathcal{X}_{uv}| \leq 8\sigma q_i^4 \pi_i^2 n^{3/2} \leq 8\sigma q_i^3 n^{3/2}. \qquad (2.70)$$

Furthermore, $q_i \sum \hat{c}_e^2 \leq q_i |\mathcal{X}_{uv}| \leq 2q_i^3 n \ll \sigma q_i^3 n^{3/2}$. Noting that X is a decreasing function of the edge-indicators $(\mathbb{1}_{\{e \in \Gamma_{i+1}\}}, \mathbb{1}_{\{e \in S_{i+1}\}})_{e \in O_i}$, using Theorem 12 and $q_i^3 n^{1/2} \geq n^{\tau}$ (see (2.54)) it follows that

$$\mathbb{P}(|X_{uv}(i+1)| \geq q_{i+1}^2 n) \leq \mathbb{P}(X \geq \mathbb{E}\, X + \sigma^{3/2} q_i^3 n) \leq \exp\left(-\frac{\sigma^3 q_i^6 n^2}{2 \cdot 9\sigma q_i^3 n^{3/2}}\right) \leq n^{-\omega(1)}.$$

Taking a union bound over all pairs of vertices u, v completes the proof for the $|X_{uv}(i+1)|$ variables.

Turning to the more involved $|Y_{uv}(i+1)|$ variables, note that by construction

$$|Y_{uv}(i+1)| \leq \sum_{w \in X_{uv}(i)} \mathbb{1}_{\{uw \in \Gamma_{i+1} \text{ or } vw \in \Gamma_{i+1}\}} + \sum_{f \in Y_{uv}(i)} \mathbb{1}_{\{f \notin C_{i+1}\}} =: Y_{uv}^{+} + Y_{uv}^{*}. \qquad (2.71)$$

(To clarify: Y_{uv}^{+} and Y_{uv}^{*} are defined by the first and second sum in (2.71), respectively.) Using Lemma 19 together with $\sigma q_i^2 = \sigma q_i q_{i+1} + o(\sigma^{3/2} q_i^2 \pi_i)$ (see (2.55)) and $\pi_i q_{i+1} =$

$q_{i+1}\pi_{i+1} - \sigma q_i q_{i+1}$ (as $\pi_{i+1} = \pi_i + \sigma q_i$ by (2.37)), it follows that

$$\begin{aligned} \mathbb{E}(Y_{uv}^+ + Y_{uv}^*) &\le |X_{uv}(i)| \cdot 2p \ + \ |Y_{uv}(i)| \cdot (q_{i+1}/q_i - \sigma^{3/2} q_i) \\ &\le 2\sigma q_i^2 \sqrt{n} + 2\pi_i \sqrt{n}(q_{i+1} - \sigma^{3/2} q_i^2) \\ &\le 2q_{i+1}\pi_{i+1}\sqrt{n} - \sigma^{3/2} q_i^2 \pi_i \sqrt{n}. \end{aligned} \tag{2.72}$$

We now estimate Y_{uv}^+ and Y_{uv}^* separately. Noting $\mathbb{E}\, Y_{uv}^+ \le 2\sigma q_i^2 \sqrt{n}$ and $\sigma^2 q_i^2 \pi_i \sqrt{n} = o(\sigma q_i^2 \sqrt{n})$ (see (2.53)), using standard Chernoff bounds together with $\pi_i^2 \ge \pi_0^2 = \sigma^2$ and $q_i^2 \sqrt{n} \ge n^\tau$ (see (2.54)) it follows that

$$\begin{aligned} \mathbb{P}(Y_{uv}^+ \ge \mathbb{E}\, Y_{uv}^+ + \sigma^2 q_i^2 \pi_i \sqrt{n}) &\le \exp\left(-\frac{\left(\sigma^2 q_i^2 \pi_i \sqrt{n}\right)^2}{4 \cdot 2\sigma q_i^2 \sqrt{n}}\right) \\ &\le \exp\left(-\frac{\sigma^5 q_i^2 \sqrt{n}}{8}\right) \le n^{-\omega(1)}. \end{aligned} \tag{2.73}$$

For Y_{uv}^* we shall apply Theorem 12, and we thus now bound $\lambda \le p \sum_{e \in O_i} c_e^2 + q_i \sum_{e \in O_i} \hat{c}_e^2$ from (2.63). As usual, we have edge-effect $c_e \le |Y_e(i) \cap Y_{uv}(i)| \le |Y_e(i)| \le 2q_i \pi_i \sqrt{n}$ and stabilization-effect $\hat{c}_e \le \mathbb{1}_{\{e \in Y_{uv}(i)\}}$. Here we can significantly improve the simple worst case estimate $c_e \le |Y_e(i)|$ when $e \neq uv$. Indeed, if $e = w_1 w_2$ does not intersect uv, then $c_e \le 4$ since $Y_e(i) \cap Y_{uv}(i) \subseteq \{u, v\} \times \{w_1, w_2\}$, say. Furthermore, if $e = w_1 w_2$ intersects uv in one vertex, say $u = w_1$, then $c_e \le \max_f |Z_f(i)| \le I(\log n)^9$ since $Y_e(i) \cap Y_{uv}(i) \subseteq \{u\} \times [N_{E_i}(w_2) \cap N_{E_i}(v)]$. To sum up, for $e \neq uv$ we have $c_e \le \max\{4, I(\log n)^9\} \le \sigma^{-5} I$, say. Similar to (2.66) and (2.70), using Lemma 20 and $|Y_{uv}(i)| \le 2q_i \pi_i \sqrt{n}$ it follows that

$$p \sum_{e \in O_i} c_e^2 \le \sigma/\sqrt{n} \cdot \Big((2q_i \pi_i \sqrt{n})^2 + \sigma^{-5} I \cdot 2q_i \pi_i \sqrt{n} \cdot |Y_{uv}(i)|\Big) \le 8\sigma^{-4} q_i^2 \pi_i^2 I \sqrt{n}.$$

Furthermore, using $\pi_i \ge \sigma$ and $I \ge 1$ we obtain $q_i \sum_{e \in O_i} \hat{c}_e^2 \le q_i |Y_{uv}(i)| \le 2q_i^2 \pi_i \sqrt{n} \ll \sigma^{-4} q_i^2 \pi_i^2 I \sqrt{n}$. Noting that Y_{uv}^* is decreasing, using Theorem 12 and $q_i^2 \sqrt{n}/I \ge n^\tau$ (see (2.54)) it follows that

$$\mathbb{P}(Y^*_{uv} \ge \mathbb{E}\, Y^*_{uv} + \sigma^2 q_i^2 \pi_i \sqrt{n}) \le \exp\left(-\frac{\sigma^4 q_i^4 \pi_i^2 n}{2 \cdot 9\sigma^{-4} q_i^2 \pi_i^2 I \sqrt{n}}\right) \le n^{-\omega(1)}. \tag{2.74}$$

Combining the probability estimates (2.73) and (2.74) with inequalities (2.71)–(2.72) and $\sigma^2 \ll \sigma^{3/2}$, now a union bound argument (to account for all pairs of vertices u, v) completes the proof for the $|Y_{uv}(i+1)|$ variables.

Finally, for $|Z_{uv}(i+1)|$ note that the one-step difference

$$\Delta Z := |Z_{uv}(i+1)| - |Z_{uv}(i)| = \sum_{w \in X_{uv}(i)} \mathbb{1}_{\{uw \in \Gamma_{i+1} \text{ and } vw \in \Gamma_{i+1}\}} + \sum_{f \in Y_{uv}(i)} \mathbb{1}_{\{f \in \Gamma_{i+1}\}} \tag{2.75}$$

is a sum of independent Bernoulli random variables with

$$\mathbb{E}(\Delta Z) = |X_{uv}(i)| \cdot p^2 + |Y_{uv}(i)| \cdot p \le \sigma^2 q_i^2 + 2\sigma q_i \pi_i \le 3\sigma \ll 1, \tag{2.76}$$

where we used $|X_{uv}(i)| \le q_i^2 n$ and $|Y_{uv}(i)| \le 2q_i\pi_i\sqrt{n}$ for the first inequality, and $\max\{q_i^2, q_i\pi_i\} \le 1$ (see (2.53)) and $\sigma \ll 1$ for the last two inequalities. Inspecting (2.75), note that $\mathfrak{X}_{\le i} \subseteq \mathcal{P}_i$ implies $|Z_{uv}(i+1)| \le \Delta Z + i(\log n)^9$. Applying standard Chernoff bounds, using $\mathbb{E}(\Delta Z) \ll 1$ it readily follows that, say,

$$\mathbb{P}\big(|Z_{uv}(i+1)| \ge (i+1)(\log n)^9\big) \le \mathbb{P}\big(\Delta Z \ge (\log n)^9\big) \le n^{-\omega(1)}.$$

Taking a union bound over all pairs of vertices u, v completes the proof for the $|Z_{uv}(i+1)|$ variables. □

Remark 23. *If desired, it would not be difficult to establish the better upper bound* $|Z_{uv}(i)| \le (\log n)^2$*, say (using the stochastic domination arguments leading to* (2.95) *in Section 2.3.5; in view* (2.75)–(2.76) *the main point is that, for* $0 \le i \le I$*, the event* $\mathfrak{X}_{\le i}$ *implies* $\sum_{0 \le j \le i}(|X_{uv}(j)|p^2 + |Y_{uv}(j)|p) = O(\log n)$*). This in turn could, e.g., be used to increase the constant* β_0 *slightly* *(as we could then remove $I = \lceil n^{\beta} \rceil$ from constraint* (2.54)*).*

2.3.4 Event $\mathcal{Q}^{+}_{i+1} \cap \mathcal{Q}_{i+1}$: number $|O_{i+1}(A,B)|$ of open edges between large sets

Recall that the events $\mathcal{Q}^{+}_{i+1}$, $\mathcal{Q}_{i+1}$ defined in (2.46)–(2.47) concern the open edge-set $O_{i+1} \subseteq E(H) = O_0$, ensuring that $|O_{i+1}(A,B)| \le q_{i+1}|A||B|$ for all disjoint $A, B \subseteq V$ with $|A|, |B| \ge s_0$, and $\tau_{i+1} q_{i+1} |O_0(A,B)| \le |O_{i+1}(A,B)| \le q_{i+1}|O_0(A,B)|$ for all $(A,B) \in \mathfrak{S}_{s,\gamma}$; see (2.40)–(2.41) for the definition of s_0 and $\mathfrak{S}_{s,\gamma}$.

Turning to $|O_{i+1}(A,B)|$, note that one edge $e \in \Gamma_{i+1}$ can add up to $|Y_e(i) \cap O_i(A,B)| \le \sum_{w \in e} |N_{E_i}(w) \cap (A \cup B)|$ edges to $C^{(1)}_{i+1}(A,B) \subseteq O_i(A,B) \setminus O_{i+1}(A,B)$, which can potentially lead to large edge-effects c_e. To sidestep such technical difficulties, we now introduce the following auxiliary variables for vertex-sets $A, B \subseteq V$ with $|A| = |B|$ (to avoid clutter we suppress the dependence on A, B, i in parts of our notation):

$$
\begin{aligned}
z &:= \sigma^4 q_i^2 |A|,\\
W_1 &:= \{w \in V : \ |N_{E_i}(w) \cap (A \cup B)| \ge z\},\\
W_2 &:= \{w \in V : \ |N_{\Gamma_{i+1}}(w) \cap (A \cup B)| \ge z\},\\
\hat{C}^{(1)}_{i+1} &:= \{uv \in O_i : \text{ there is } w \notin W_1 \text{ s.t. } |\{uw, vw\} \cap \Gamma_{i+1}| = |\{uw, vw\} \cap E_i| = 1\},\\
\hat{C}^{(2)}_{i+1} &:= \{uv \in O_i : \text{ there is } w \notin W_2 \text{ s.t. } uw \in \Gamma_{i+1}, \ vw \in \Gamma_{i+1}\},\\
\hat{C}_{i+1} &:= \hat{C}^{(1)}_{i+1} \cup S_{i+1}.
\end{aligned}
$$

Note that $\hat{C}^{(j)}_{i+1} \subseteq C^{(j)}_{i+1}$ for $j \in \{1,2\}$, and that $\hat{C}_{i+1} \subseteq C_{i+1}$. Furthermore, recalling $q_i \ge q_I$ (see (2.55)), using inequality (2.54) it is routine to check that $s_0 \gg 1$ holds, that $|A| \ge s_0$ implies $z \gg 1$, and moreover that

$$\min_{|A| \ge s_0} z/\sqrt{|A| I} \ge \sigma^4 q_i^2 \sqrt{s_0}/\sqrt{I} \gg \sigma^6 q_I^3 \sqrt[4]{n}/\sqrt{I} \gg n^{\tau/2}. \tag{2.77}$$

Lemma 24. *We have $\mathbb{P}(\neg \mathcal{Q}^{+}_{i+1}) \le n^{-\omega(1)}$.*

Proof. Mimicking the double counting argument from (2.4), it follows that the special case $|A| = |B|$ of $\mathcal{Q}^+_{i+1}$ implies the event $\mathcal{Q}^+_{i+1}$ in full. Hence $\neg\mathcal{Q}^+_{i+1}$ implies that $|O_{i+1}(A,B)| \leq q_{i+1}|A||B|$ fails for some disjoint vertex-sets $A, B \subseteq V$ with $|A| = |B| \geq s_0$, and we shall below estimate the probability of this special case.

Recalling $\hat{C}_{i+1} \subseteq C_{i+1}$, noting $O_{i+1} \subseteq O_i \setminus C_{i+1} \subseteq O_i \setminus \hat{C}_{i+1}$ we obtain

$$|O_{i+1}(A,B)| \leq |O_i(A,B) \setminus \hat{C}_{i+1}| = \sum_{f \in O_i(A,B)} \mathbb{1}_{\{f \notin \hat{C}_{i+1}\}} =: X. \tag{2.78}$$

To estimate $\mathbb{E}\,X$, recall that $C^{(1)}_{i+1} = \{f \in O_i : Y_f(i) \cap \Gamma_{i+1} \neq \varnothing\}$. Note that if the event $\mathcal{Q}_f := \{(f \times W_1) \cap \Gamma_{i+1} = \varnothing\}$ holds, then $f \notin \hat{C}^{(1)}_{i+1}$ implies $f \notin C^{(1)}_{i+1}$, so that $f \notin \hat{C}_{i+1}$ implies $f \notin C_{i+1} = C^{(1)}_{i+1} \cup S_{i+1}$. Since $f \notin C^{(1)}_{i+1}$ and $\mathcal{Q}_f$ are both monotone decreasing functions of the edge-indicators $(\mathbb{1}_{\{e \in \Gamma_{i+1}\}}, \mathbb{1}_{\{e \in S_{i+1}\}})_{e \in O_i}$, using Harris's inequality [58] and $\mathbb{P}(\mathcal{Q}_f) \geq (1-p)^{2|W_1|}$ it follows that

$$\mathbb{P}(f \notin C_{i+1}) \geq \mathbb{P}(f \notin \hat{C}_{i+1} \text{ and } \mathcal{Q}_f) \geq \mathbb{P}(f \notin \hat{C}_{i+1})\,\mathbb{P}(\mathcal{Q}_f) \geq \mathbb{P}(f \notin \hat{C}_{i+1}) \cdot (1-p)^{2|W_1|}.$$

Note that $\mathfrak{X}_{\leq i}$ and $i < I$ imply $|N_{E_i}(u) \cap N_{E_i}(v)| = |Z_{uv}(i)| \leq I(\log n)^9 =: y$ when $u \neq v$, and that (2.77) implies $z \gg \sqrt{|A \cup B| y}$. Using the definition of W_1 and Lemma 16 (with $\mathcal{I} = W_1$, $U = A \cup B$ and $U_w = N_{E_i}(w) \cap U$), we infer $|W_1| \leq 2|A \cup B|/z = 4/(\sigma^4 q_i^2) \leq q_i \sigma \sqrt{n}$ by (2.54), say. Similar to (2.69), using Lemma 19, $|O_i(A,B)| \leq q_i|A||B|$, $p|W_1| \leq q_i\sigma^2 \ll 1$ and $q_i q_{i+1} \sim q_i^2$ (see (2.55)) it is routine to deduce that

$$\begin{aligned} \mathbb{E}\,X &\leq |O_i(A,B)| \cdot (q_{i+1}/q_i - \sigma^{3/2} q_i) \cdot (1-p)^{-2|W_1|} \\ &\leq |A||B| \cdot (q_{i+1} - \sigma^{3/2} q_i^2/2). \end{aligned} \tag{2.79}$$

Gearing up to apply Theorem 12, we now bound $\lambda \leq p \sum_{e \in O_i} c_e^2 + q_i \sum_{e \in O_i} \hat{c}_e^2$. Noting $\hat{C}_{i+1} \subseteq C_{i+1}$, as usual we have edge-effect $c_e \leq |Y_e(i) \cap O_i(A,B)|$ and stabilization-effect $\hat{c}_e \leq \mathbb{1}_{\{e \in O_i(A,B)\}}$. Here the definition of $\hat{C}_{i+1}$ allows us to improve the simple worst case estimate $c_e \le |Y_e(i)|$. Indeed, inspecting the corresponding argument for $|N_{O_{i+1}}(v)|$ from Lemma 21, we see that the edge-effect c_e (an upper bound on how much X changes if we alter whether $e \in \Gamma_{i+1}$ or $e \notin \Gamma_{i+1}$) is at most the number of changes to

$$\begin{aligned} \hat{C}^{(1)}_{i+1} \cap O_i(A,B) = \big\{uv \in O_i(A,B) : \text{there is } w \notin W_1 \text{ s.t.} \\ \text{either } uw \in \Gamma_{i+1},\ vw \in E_i \\ \text{or } vw \in \Gamma_{i+1},\ uw \in E_i\big\}. \end{aligned} \tag{2.80}$$

Since any $w \notin W_1$ has at most z neighbours in $A \cup B$ via E_i–edges, we infer that $c_e \le 2z$ (the factor of two takes into account that each vertex of e could potentially play the role of w in (2.80) above). Similar to (2.66) and (2.70), using Lemma 20, $\sigma\pi_i \le \sqrt{\sigma} \ll 1$ (see (2.53)), and $|O_i(A,B)| \le q_i|A||B|$ it follows that

$$p\sum_{e \in O_i} c_e^2 \le \sigma/\sqrt{n} \cdot 2z \cdot 2q_i\pi_i\sqrt{n} \cdot |O_i(A,B)| \ll zq_i|O_i(A,B)| \le zq_i^2|A||B|.$$

Furthermore, using $z \ge 1$ we obtain $q_i \sum \hat{c}_e^2 \le q_i|O_i(A,B)| \le zq_i|O_i(A,B)| \le zq_i^2|A||B|$. Noting that X is decreasing, using Theorem 12 and the λ–bound (2.63) it follows that

$$\begin{aligned} \mathbb{P}(|O_{i+1}(A,B)| \ge q_{i+1}|A||B|) &\le \mathbb{P}(X \ge \mathbb{E}\,X + \sigma^{3/2}q_i^2|A||B|/2) \\ &\le \exp\left(-\frac{\left(\sigma^{3/2}q_i^2|A||B|/2\right)^2}{2 \cdot 2zq_i^2|A||B|}\right) \\ &= \exp\left(-\frac{\sigma^3 q_i^2|A||B|}{16z}\right) \le n^{-\omega(|B|)}, \end{aligned} \tag{2.81}$$

where for the last inequality we used $z = \sigma^4 q_i^2|A|$ and $\sigma^{-1} \gg \log n$. Finally, taking a union bound over all disjoint vertex-sets $A, B \subseteq V$ with $|A| = |B| \ge s_0$ completes the proof (as discussed). □

For the 'relative error' τ_i used in the event $\mathcal{Q}_i$, see (2.38), we now record the following convenient bounds:

$$1 \geq \tau_i \geq \tau_I = 1 - \delta/2 \geq 1/2 \qquad \text{for all } 0 \leq i \leq I. \tag{2.82}$$

Lemma 25. *We have* $\mathbb{P}(\neg \mathcal{Q}_{i+1} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}) \leq n^{-\omega(1)}$.

The proof strategy is to estimate the different contributions to $O_{i+1} = O_i \setminus (\Gamma_{i+1} \cup C_{i+1} \cup C^{(2)}_{i+1})$ separately (here $\mathcal{Q}^+_i$ will be crucial for bounding some of the large edge-effects ignored in Lemma 24).

Claim 26. *Let* $\mathcal{Q}_{A,B}$ *be the event that the following bounds hold:*

$$\begin{aligned}
X_1 &:= \big|O_i(A,B) \setminus \hat{C}_{i+1}\big| \in \big[|O_i(A,B)| \cdot (q_{i+1}/q_i - 4\sigma^{3/2} q_i),\ |O_i(A,B)| \cdot q_{i+1}/q_i\big],\\
X_2 &:= \big|O_i(A,B) \cap \hat{C}^{(2)}_{i+1}\big| \leq |O_i(A,B)| \cdot 2\sigma^2 q_i,\\
X_3 &:= |O_i(A,B) \cap \Gamma_{i+1}| \leq |O_i(A,B)| \cdot 2\sigma^2 q_i,\\
X_4 &:= \big|O_i(A,B) \cap (C_{i+1} \cup C^{(2)}_{i+1}) \setminus (\hat{C}_{i+1} \cup \hat{C}^{(2)}_{i+1})\big| \leq 36\sigma q_i^2 \sqrt{n}|A|.
\end{aligned}$$

Then $\mathbb{P}(\neg \mathcal{Q}_{A,B} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}) \leq n^{-\omega(s)}$ *for all vertex-sets* $(A,B) \in \mathfrak{S}_{s,\gamma}$.

Before giving the proof, we first show that Claim 26 implies Lemma 25. Using a union bound argument (to account for the $|\mathfrak{S}_{s,\gamma}| \leq n^{2s}$ vertex-sets $(A,B) \in \mathfrak{S}_{s,\gamma}$), it is enough to show that $\mathcal{Q}_{A,B} \cap \mathfrak{X}_{\leq i}$ implies $\tau_{i+1} q_{i+1}|O_0(A,B)| \leq |O_{i+1}(A,B)| \leq q_{i+1}|O_0(A,B)|$. By definition of $O_{i+1}(A,B)$ we have

$$X_1 - X_2 - X_3 - X_4 \leq |O_{i+1}(A,B)| \leq X_1.$$

Combining $\mathcal{Q}_{A,B}$ with the fact that $|O_i(A,B)| \leq q_i|O_0(A,B)|$ by $\mathfrak{X}_{\leq i} \subseteq \mathcal{Q}_i$, we readily infer the upper bound $|O_{i+1}(A,B)| \leq q_{i+1}|O_0(A,B)|$. Turning to the lower bound, using $\mathcal{Q}_{A,B}$ it follows that

$$
\begin{aligned}
X_1 - X_2 - X_3 - X_4 &\geq |O_i(A,B)| \cdot \big(q_{i+1}/q_i - 8\sigma^{3/2} q_i\big) - 36\sigma q_i^2 \sqrt{n} |A| \\
&\geq \left(\tau_i q_i \big(q_{i+1}/q_i - 8\sigma^{3/2} q_i\big) - \frac{36\sigma q_i^2}{\gamma C \sqrt{\log n}}\right) \cdot |O_0(A,B)| \\
&\geq \left(\tau_i - \frac{45\sigma q_i}{\gamma C \sqrt{\log n}}\right) \cdot q_{i+1} |O_0(A,B)| \geq \tau_{i+1} \cdot q_{i+1} |O_0(A,B)|,
\end{aligned}
$$

where for the second inequality we used $|O_i(A,B)| \geq \tau_i q_i |O_0(A,B)|$ (by $\mathfrak{X}_{\leq i} \subseteq \mathcal{Q}_i$) and $|O_0(A,B)| \geq \gamma |A||B| \geq \gamma C \sqrt{\log n} \cdot \sqrt{n} |A|$, for the third inequality we used $\tau_i \leq 1$ (see (2.82)), $\sigma^{1/2} \ll 1/\sqrt{\log n}$, and $q_i \sim q_{i+1}$ (see (2.55)), and for the last inequality we used $\sqrt{\log n} \sim \sqrt{\log(I\sigma)/\beta} \sim \pi_I/\sqrt{\beta}$ (see (2.57)), $\gamma C/\sqrt{\beta} \geq D_0/\delta^2 \geq 91/\delta$ (by assumption and (2.39)) and $\tau_i - \delta\sigma q_i/\pi_I = \tau_{i+1}$ (see (2.38)). This completes the proof of Lemma 25 (assuming Claim 26).

Proof of Claim 26. We start with $X_1 = |O_i(A,B) \setminus \hat{C}_{i+1}|$. Since $s \geq s_0$, the upper tail argument for $X = X_1$ defined in (2.78) carries over from Lemma 24, with $\mathbb{E} X_1 \leq |O_i(A,B)|(q_{i+1}/q_i - \sigma^{3/2} q_i/2)$ and $\lambda \leq 2z q_i |O_i(A,B)|$, say. In particular, noting that here $|O_i(A,B)| \geq \tau_i q_i |O_0(A,B)| \geq \gamma \tau_i q_i |A||B|$, an application of Theorem 12 along the lines of (2.81) gives

$$
\begin{aligned}
\mathbb{P}(X_1 \geq |O_i(A,B)| q_{i+1}/q_i) &\leq \exp\left(-\frac{\big(\sigma^{3/2} q_i |O_i(A,B)|/2\big)^2}{2 \cdot 2z q_i |O_i(A,B)|}\right) \\
&\leq \exp\left(-\frac{\gamma \tau_i \sigma^3 q_i^2 |A||B|}{16z}\right) \leq n^{-\omega(s)},
\end{aligned}
\tag{2.83}
$$

where for the last inequality we used $z = \sigma^4 q_i^2 |A|$, $\tau_i \geq 1/2$ (see (2.82)), $\gamma\sigma^{-1} \gg \log n$ and $|B| = s$. For the lower tail of X_1 we proceed similarly. Since $\hat{C}_{i+1} \subseteq C_{i+1}$, using Lemma 19 we obtain

$$
\mathbb{E} X_1 = \sum_{e \in O_i(A,B)} \mathbb{P}(e \notin \hat{C}_{i+1}) \geq \sum_{e \in O_i(A,B)} \mathbb{P}(e \notin C_{i+1}) \geq |O_i(A,B)| \cdot (q_{i+1}/q_i - 3\sigma^{3/2} q_i).
$$

Furthermore, the edge-effect and stabilization-effect estimates from the proof of Lemma 24 again carry over, giving $\lambda \le 2zq_i|O_i(A,B)|$ and $\max_{e\in O_i}\max\{c_e,\hat{c}_e\} \le 2z$, say. Applying inequality (2.49) of Remark 13 (with $C = 2z$), it follows similarly to (2.83) that

$$\begin{aligned}
&\mathbb{P}\big(X_1 \le |O_i(A,B)|(q_{i+1}/q_i - 4\sigma^{3/2}q_i)\big)\\
&\le \mathbb{P}\big(X_1 \le \mathbb{E}\,X_1 - \sigma^{3/2}q_i|O_i(A,B)|\big)\\
&\le \exp\left(-\frac{\big(\sigma^{3/2}q_i|O_i(A,B)|\big)^2}{2\big(2zq_i|O_i(A,B)| + 2z\cdot\sigma^{3/2}q_i|O_i(A,B)|\big)}\right)\\
&\le \exp\left(-\frac{\gamma\tau_i\sigma^3q_i^2|A||B|}{8z}\right) \le n^{-\omega(s)}.
\end{aligned} \tag{2.84}$$

Turning to $X_2 = |O_i(A,B)\cap \hat{C}^{(2)}_{i+1}|$, note that by construction of $\hat{C}^{(2)}_{i+1}$ we have

$$X_2 = \sum_{e\in O_i(A,B)} \mathbb{1}_{\{e\in \hat{C}^{(2)}_{i+1}\}} \le \sum_{ab\in O_i(A,B)}\sum_{w\in V\setminus W_2} \mathbb{1}_{\{\{wa,wb\}\subseteq \Gamma_{i+1}\}} =: X_2^+. \tag{2.85}$$

Gearing up to apply Theorem 15 to X_2^+, in view of $\Gamma_{i+1}\subseteq O_i$ we define

$$\begin{aligned}
\mathcal{I} &:= \big\{\{wa,wb\}\subseteq O_i:\ ab\in O_i(A,B),\ w\in V,\ |\{a,b,w\}| = 3\big\},\\
\mathcal{K} &:= \{\{wa,wb\}\in\mathcal{I}:\ w\notin W_2,\ \{wa,wb\}\subseteq\Gamma_{i+1}\}.
\end{aligned}$$

Since $p^2\cdot|X_{ab}(i)| \le \sigma^2q_i^2 \le \sigma^2q_i$ by $\mathfrak{X}_{\le i}\subseteq\mathcal{P}_i$ and $q_i\le 1$ (see (2.53)), we obtain

$$\begin{aligned}
\sum_{\alpha\in\mathcal{I}}\mathbb{E}\,\mathbb{1}_{\{\alpha\subseteq\Gamma_{i+1}\}} &= p^2\sum_{ab\in O_i(A,B)}\sum_{v\in V}\mathbb{1}_{\{\{va,vb\}\subseteq O_i\}}\\
&= p^2\sum_{ab\in O_i(A,B)}|X_{ab}(i)| \le \sigma^2q_i\cdot|O_i(A,B)| =: \mu.
\end{aligned}$$

Furthermore, since $\mathcal{K}$ only contains edge-pairs $\{wa,wb\}$ with $\{a,b\}\subseteq N_{\Gamma_{i+1}}(w)\cap(A\cup B)$ where the 'central vertex' w satisfies $w\notin W_2$ and thus $|N_{\Gamma_{i+1}}(w)\cap(A\cup B)|\le z$, for all

$\beta \in \mathcal{K}$ we see that

$$|\{\alpha \in \mathcal{K} : \alpha \cap \beta \neq \varnothing\}| \leq \sum_{f \in \beta} |\{\alpha \in \mathcal{K} : f \in \alpha\}| \leq \sum_{f \in \beta} \sum_{v \in f \setminus W_2} |N_{\Gamma_{i+1}}(v) \cap (A \cup B)| \leq 2 \cdot 2 \cdot z.$$

It follows that $X_2^+ = \sum_{\alpha \in \mathcal{K}} \mathbb{1}_{\{\alpha \subseteq \Gamma_{i+1}\}} \leq Z_{4z}$, where Z_{4z} is defined as in Theorem 15. Applying first (2.85) and then inequality (2.50) with $C = 4z$, using $|O_i(A, B)| \geq \gamma \tau_i q_i |A||B|$ it follows similarly to (2.83) that

$$\begin{aligned} \mathbb{P}(X_2 \geq 2\sigma^2 q_i |O_i(A, B)|) &\leq \mathbb{P}(Z_{4z} \geq 2\mu) \leq \exp\Big(-\frac{\mu^2}{2 \cdot 4z \cdot 2\mu}\Big) \\ &\leq \exp\Big(-\frac{\gamma \tau_i \sigma^2 q_i^2 |A||B|}{16z}\Big) \leq n^{-\omega(s)}. \end{aligned} \tag{2.86}$$

We next turn to $X_3 = |O_i(A, B) \cap \Gamma_{i+1}|$, which is a sum of independent Bernoulli random variables with $\mathbb{E} X_3 = |O_i(A, B)| \cdot p \ll \sigma^2 q_i |O_i(A, B)| =: t$, as $q_i \sqrt{n} \geq n^{\tau}$ by (2.54). Applying standard Chernoff bounds, using $|O_i(A, B)| \geq \gamma \tau_i q_i |A||B|$ and $z \geq 1$ it follows by comparison with the last inequality of (2.86) that

$$\begin{aligned} \mathbb{P}(X_3 \geq 2\sigma^2 q_i |O_i(A, B)|) &\leq \mathbb{P}(X_3 \geq \mathbb{E} X_3 + t) \leq \exp\Big(-\frac{t^2}{2 \cdot 2t}\Big) \\ &\leq \exp\Big(-\frac{\gamma \tau_i \sigma^2 q_i^2 |A||B|}{4}\Big) \leq n^{-\omega(s)}. \end{aligned} \tag{2.87}$$

Finally, X_4 is a more difficult variable: assuming that $\mathcal{N}_{i+1} \cap \mathcal{P}_{i+1} \cap \mathfrak{X}_{\leq i}$ holds, we shall bound X_4 by *deterministic* counting arguments (here the edge-effects can potentially be fairly large, so concentration inequalities seem less effective). Noting $C_{i+1} \setminus \hat{C}_{i+1} =$

$C^{(1)}_{i+1} \setminus \hat{C}^{(1)}_{i+1}$, similarly to (2.85) we obtain

$$\begin{aligned} X_4 &\le \sum_{e \in O_i(A,B)} \mathbb{1}_{\{e \in C^{(1)}_{i+1} \setminus \hat{C}^{(1)}_{i+1}\}} + \sum_{e \in O_i(A,B)} \mathbb{1}_{\{e \in C^{(2)}_{i+1} \setminus \hat{C}^{(2)}_{i+1}\}} \\ &\le \sum_{w \in W_1} \Big(|O_i(N_{\Gamma_{i+1}}(w) \cap A,\ N_{E_i}(w) \cap B)| \\ &\qquad\qquad + |O_i(N_{\Gamma_{i+1}}(w) \cap B,\ N_{E_i}(w) \cap A)| \Big) \\ &\qquad + \sum_{w \in W_2} |O_i(N_{\Gamma_{i+1}}(w) \cap A,\ N_{\Gamma_{i+1}}(w) \cap B)|. \end{aligned} \tag{2.88}$$

Using the upper bound estimate from $\mathfrak{X}_{\le i} \subseteq \mathcal{Q}_i^+$ when $\min\{|N_{\Gamma_{i+1}}(v) \cap A|, |N_{E_i}(v) \cap B|\} \ge z$ holds (note that $z = \sigma^4 q_i^2 s \ge s_0$), and a trivial estimate otherwise, it follows that

$$\begin{aligned} &|O_i(N_{\Gamma_{i+1}}(w) \cap A,\ N_{E_i}(w) \cap B)| \\ &\qquad \le q_i |N_{\Gamma_{i+1}}(w) \cap A||N_{E_i}(w) \cap B| + z \max\{|N_{\Gamma_{i+1}}(w) \cap A|, |N_{E_i}(w) \cap B|\} \\ &\qquad \le \big(q_i |N_{\Gamma_{i+1}}(w)| + z\big) \cdot |N_{E_i \cup \Gamma_{i+1}}(w) \cap (A \cup B)|. \end{aligned} \tag{2.89}$$

With an eye on (2.88), we note that an analogous estimate also holds when we reverse the role of A and B in (2.89). Furthermore, $q_i |N_{\Gamma_{i+1}}(w)| \le 2\sigma q_i^2 \sqrt{n}$ by $\mathcal{N}_{i+1}$, and $z = \sigma^4 q_i^2 s = O(\sigma^3 q_i^2 \sqrt{n}) \ll \sigma q_i^2 \sqrt{n}$. Recalling $E_i \cup \Gamma_{i+1} = E_{i+1}$, observe that $\mathcal{P}_{i+1}$ and $i+1 \le I$ imply $|N_{E_i \cup \Gamma_{i+1}}(u) \cap N_{E_i \cup \Gamma_{i+1}}(v)| = |Z_{uv}(i+1)| \le I (\log n)^9 =: y$ when $u \ne v$, and that (2.77) implies $z \gg \sqrt{|A \cup B| y}$ (as $|A| = s \ge s_0$). Using the definition of W_1 and Lemma 16 (with $\mathcal{I} = W_1$, $U = A \cup B$ and $U_w = N_{E_i \cup \Gamma_{i+1}}(w) \cap U$), it follows that

$$\begin{aligned} &\sum_{w \in W_1} \Big(|O_i(N_{\Gamma_{i+1}}(w) \cap A,\ N_{E_i}(w) \cap B)| + |O_i(N_{\Gamma_{i+1}}(w) \cap B,\ N_{E_i}(w) \cap A)| \Big) \\ &\qquad \le 2 \cdot 3\sigma q_i^2 \sqrt{n} \cdot \sum_{w \in W_1} |N_{E_i \cup \Gamma_{i+1}}(w) \cap (A \cup B)| \\ &\qquad \le 2 \cdot 3\sigma q_i^2 \sqrt{n} \cdot 2|A \cup B| \le 24 \sigma q_i^2 \sqrt{n} |A|. \end{aligned} \tag{2.90}$$

Proceeding analogously to (2.89)–(2.90), using the definition of W_2 and Lemma 16 we similarly obtain

$$
\begin{aligned}
\sum_{w \in W_2} & |O_i(N_{\Gamma_{i+1}}(w) \cap A,\ N_{\Gamma_{i+1}}(w) \cap B)| \\
&\le 3\sigma q_i^2 \sqrt{n} \cdot \sum_{w \in W_2} |N_{\Gamma_{i+1}}(w) \cap (A \cup B)| \\
&\le 3\sigma q_i^2 \sqrt{n} \cdot 2|A \cup B| \le 12\sigma q_i^2 \sqrt{n}|A|.
\end{aligned} \tag{2.91}
$$

To sum up, inserting the bounds (2.90)–(2.91) into (2.88), we showed that $\mathcal{N}_{i+1} \cap \mathcal{P}_{i+1} \cap \mathfrak{X}_{\le i}$ implies $X_4 \le 36\sigma q_i^2 \sqrt{n}|A|$. This completes the proof together with the probability estimates (2.83), (2.84), (2.86), and (2.87). $\square$

Remark 27. *If desired, it would not be difficult to extend the event $\mathcal{Q}_i$ to larger vertex-sets $(A, B) \in \mathfrak{S}_{\ge s,\gamma} := \bigcup_{s \le r \le n} \mathfrak{S}_{r,\gamma}$ (the above arguments all carry over, except for the modified bound $X_4 \le 3 \cdot \max_w (q_i |N_{\Gamma_{i+1}}(w)| + z) \cdot 2|A \cup B| \le 36\sigma q_i^2 \max\{\sqrt{n}, \sigma^3 |B|\}|A|$, which is still strong enough to deduce Lemma 25). This in turn could, e.g., be used to also extend the event $\mathcal{T}_I$ to $(A, B) \in \mathfrak{S}_{\ge s,\gamma}$ (the proofs in Section 2.3.5 then carry over).*

Remark 28. *Under a mild extra assumption such as $|O_0| \ge \sigma n$, say, it would not be difficult to add two-sided bounds for the total number of open edges $|O_i|$ and edges $|T_I|$ to the events $\mathcal{Q}_i$ and $\mathcal{T}_I$. For example, much simpler variants of the above arguments then imply $\tau_i q_i |O_0| \le |O_i| \le q_i |O_0|$ (by directly estimating $|O_i \setminus C_{i+1}| - |\Gamma_{i+1}| - |C_{i+1}^{(2)}| \le |O_{i+1}| \le |O_i \setminus C_{i+1}|$, without using $\hat{C}_{i+1}$ or $\hat{C}_{i+1}^{(2)}$, nor a union bound over all vertex-sets), which in turn gives $|T_I| = (1 \pm \delta)\rho|O_0|$ by straightforward variants of the proofs in Section 2.3.5.*

2.3.5 Event $\mathcal{T}_I$: number $|T_I(A, B)|$ of edges between large sets

Recall that the event $\mathcal{T}_I$ defined in (2.42) concerns the triangle-free edge-set $T_I \subseteq E(H) = O_0$, ensuring that $|T_I(A, B)| = (1 \pm \delta)\rho|O_0(A, B)|$ for all $(A, B) \in \mathfrak{S}_{s,\gamma}$; see (2.41) for the definition of $\mathfrak{S}_{s,\gamma}$.

For $|T_I(A,B)|$ it is convenient to think of the entire nibble construction as one evolving random process. Thus, in contrast to previous sections, in Lemma 29 and Claim 30 below we shall *not* tacitly condition on $\mathcal{F}_i$.

Lemma 29. *We have* $\mathbb{P}(\neg\mathcal{T}_I \cap \mathfrak{X}_{\leq I}) \leq n^{-\omega(1)}$.

Since $T_I = \bigcup_{0\leq i<I}(T_{i+1}\setminus T_i)$ forms a partition, the proof strategy is to estimate the two contributions to $T_{i+1}\setminus T_i = \Gamma_{i+1}\setminus E(\mathcal{D}_{i+1})$ separately (here the deleted edges $E(\mathcal{D}_{i+1})$ will have negligible impact).

Claim 30. *Let $\mathcal{T}_{A,B}$ be the event that the following bounds hold:*

$$\begin{aligned} X &:= \sum_{0\leq i<I} |O_i(A,B)\cap\Gamma_{i+1}| \in \big[(1-\delta/2)\mu^-,\ (1+\delta/2)\mu^+\big], \\ Y &:= \sum_{0\leq i<I} \big|O_i(A,B)\cap E(\mathcal{D}_{i+1})\big| \leq \delta^2\mu^-/9, \end{aligned}$$

where $\mu^+ := \sum_{0\leq i<I}\lfloor q_i|O_0(A,B)|\rfloor p$ and $\mu^- := \sum_{0\leq i<I}\lceil \tau_i q_i|O_0(A,B)|\rceil p$. Then $\mathbb{P}(\neg\mathcal{T}_{A,B}\cap \mathfrak{X}_{\leq I}) \leq 3n^{-3s}$ for all vertex-sets $(A,B)\in\mathfrak{S}_{s,\gamma}$.

Before giving the proof, we first show that Claim 30 implies Lemma 29. Using a union bound argument (to account for the $|\mathfrak{S}_{s,\gamma}| \leq n^{2s}$ vertex-sets $(A,B)\in\mathfrak{S}_{s,\gamma}$), it is enough to show that $\mathcal{T}_{A,B}$ implies $|T_I(A,B)| = (1\pm\delta)\rho|O_0(A,B)|$. Since all the $(\Gamma_{i+1})_{0\leq i<I}$ are edge-disjoint, by the recursive definition (2.14) of T_I we have

$$X - Y \leq |T_I(A,B)| \leq X. \tag{2.92}$$

Noting $\mu^- \geq \tau_I\mu^+ = (1-\delta/2)\mu^+$ (see (2.82)), it follows that $\mathcal{T}_{A,B}$ implies $X \leq (1+\delta/2)\mu^+$ and

$$X - Y \geq \big(1-\delta/2-\delta^2/9\big)\cdot\mu^- \geq (1-\delta+\delta^2/8)\mu^+.$$

It thus suffices to show that $\mu^+ \sim \rho|O_0(A,B)|$, where $\rho = \sqrt{\beta(\log n)/n}$. But this is routine: indeed, since $q_i|O_0(A,B)| \geq q_i\cdot\gamma s^2 \gg q_i n \gg \sqrt{n}$ by (2.54), and $\pi_I \sim \sqrt{\log(I\sigma)} \sim$

$\sqrt{\beta \log n}$ by (2.57), using the definition (2.37) of π_I we readily infer

$$\begin{aligned} \mu^+ &= \sum_{0 \le i < I} (q_i |O_0(A,B)| \pm 1) p \sim \sum_{0 \le i < I} \sigma q_i / \sqrt{n} \cdot |O_0(A,B)| \\ &= (\pi_I - \sigma)/\sqrt{n} \cdot |O_0(A,B)| \sim \rho |O_0(A,B)|, \end{aligned} \tag{2.93}$$

completing the proof of Lemma 29 (assuming Claim 30).

Proof of Claim 30. We start with $X = \sum_{0 \le i < I} |O_i(A,B) \cap \Gamma_{i+1}|$. Define

$$X^+_{i+1} := \mathbb{1}_{\{\mathfrak{X}_i\}} \sum_{e \in O_i(A,B)} \mathbb{1}_{\{e \in \Gamma_{i+1}\}} \qquad \text{and} \qquad X^+ := \sum_{0 \le i < I} X^+_{i+1}.$$

Note that $X = X^+$ when $\mathfrak{X}_{\le I} = \bigcap_{0 \le i \le I} \mathfrak{X}_i$ holds. Let $Z^+_{i+1} \overset{\mathrm{d}}{=} \mathrm{Bin}(\lfloor q_i |O_0(A,B)| \rfloor,\ p)$ be independent random variables (where $\overset{\mathrm{d}}{=}$ means equality in distribution, as usual). Since the $\mathcal{F}_i$-measurable event $\mathfrak{X}_i \subseteq \mathcal{Q}_i$ implies $|O_i(A,B)| \le q_i |O_0(A,B)|$, it is easy to see that $\mathbb{P}(X^+_{i+1} \ge t \mid \mathcal{F}_i) \le \mathbb{P}(Z^+_{i+1} \ge t)$ for $t \in \mathbb{R}$. Setting

$$Z^+ := \sum_{0 \le i < I} Z^+_{i+1} \overset{\mathrm{d}}{=} \mathrm{Bin}\Big(\sum_{0 \le i < I} \lfloor q_i |O_0(A,B)| \rfloor,\ p \Big), \tag{2.94}$$

a standard stochastic domination argument then shows $\mathbb{P}(X^+ \ge t) \le \mathbb{P}(Z^+ \ge t)$ for $t \in \mathbb{R}$, so that

$$\mathbb{P}(X \ge t \text{ and } \mathfrak{X}_{\le I}) \le \mathbb{P}(X^+ \ge t) \le \mathbb{P}(Z^+ \ge t). \tag{2.95}$$

Since $\mathfrak{X}_i$ also implies $|O_i(A,B)| \ge \tau_i q_i |O_0(A,B)|$, an analogous argument gives

$$\begin{aligned} &\mathbb{P}(X \le t \text{ and } \mathfrak{X}_{\le I}) \le \mathbb{P}(Z^- \le t) \qquad \text{with} \\ &\qquad Z^- \overset{\mathrm{d}}{=} \mathrm{Bin}\Big(\sum_{0 \le i < I} \lceil \tau_i q_i |O_0(A,B)| \rceil,\ p \Big). \end{aligned} \tag{2.96}$$

Combining $\mu^- \ge \tau_I \mu^+ \ge \mu^+/2$ (see (2.82)) and (2.93) with $|O_0(A,B)| \ge \gamma s^2$, using

$\delta^2\sqrt{\beta}\gamma \cdot C \geq D_0 = 108$ (by assumption and (2.39)) we have

$$\begin{aligned} \delta^2 \min\{\mu^-, \mu^+\} &\geq \tfrac{\delta^2}{2}\mu^+ \geq \tfrac{\delta^2}{3}\rho|O_0(A,B)| \\ &\geq \tfrac{\delta^2}{3}\sqrt{\beta(\log n)/n} \cdot \gamma C\sqrt{n\log n} \cdot s \geq 36 s \log n. \end{aligned} \tag{2.97}$$

Using (2.94)–(2.96) and $\mathbb{E}\, Z^{\pm} = \mu^{\pm}$, by standard Chernoff bounds (see, e.g., Remark 14) we obtain, say,

$$\begin{aligned} &\mathbb{P}\big(X \notin \big[(1-\delta/2)\mu^-,\ (1+\delta/2)\mu^+\big] \text{ and } \mathfrak{X}_{\leq I}\big) \\ &\leq \mathbb{P}\big(Z^- \leq (1-\delta/2)\mu^-\big) + \mathbb{P}\big(Z^+ \geq (1+\delta/2)\mu^+\big) \\ &\leq \exp\big(-\delta^2\mu^-/8\big) + \exp\big(-\delta^2\mu^+/12\big) \leq 2n^{-3s}. \end{aligned} \tag{2.98}$$

Finally, turning to $Y = \sum_{0 \leq i < I} |O_i(A,B) \cap E(\mathcal{D}_{i+1})|$, for brevity we define

$$Y_{i+1} := |O_i(A,B) \cap E(\mathcal{D}_{i+1})| \qquad \text{and} \qquad y := \delta^2\mu^-/9.$$

Note that $Y = \sum_{0 \leq i < I} Y_{i+1}$ and $Y_{i+1} \in \mathbb{N}$. Since $\mathfrak{X}_{\leq i} = \bigcap_{0 \leq j \leq i} \mathfrak{X}_j$, a union bound argument gives

$$\begin{aligned} &\mathbb{P}\big(Y \geq \delta^2\mu^-/9 \text{ and } \mathfrak{X}_{\leq I}\big) \\ &\leq \sum_{\substack{(y_1,\ldots,y_I) \in \mathbb{N}^I \\ \sum_{1 \leq i \leq I} y_i = \lceil y \rceil}} \mathbb{P}\Big(\bigcap_{0 \leq i < I} \big(Y_{i+1} \geq y_{i+1} \text{ and } \mathfrak{X}_{\leq i+1}\big)\Big) \\ &\leq \sum_{\substack{(y_1,\ldots,y_I) \in \mathbb{N}^I \\ \sum_{0 \leq i < I} y_{i+1} = \lceil y \rceil}} \prod_{0 \leq i < I} \mathbb{P}\Big(Y_{i+1} \geq y_{i+1} \,\Big|\, \bigcap_{0 \leq j < i} \big(Y_{j+1} \geq y_{j+1} \text{ and } \mathfrak{X}_{\leq j+1}\big)\Big). \end{aligned} \tag{2.99}$$

Gearing up to apply Theorem 15 to Y_{i+1}, with an eye on $\mathcal{D}_{i+1} \subseteq \mathcal{B}_{i+1}$ and $T_i \subseteq E_i$ (see

Section 2.2.1) we define

$$\mathcal{I} := \big\{\{wu, wv\} \subseteq O_i : \ uv \in E_i, \ |\{u,v,w\}| = 3, \ \{wu, wv\} \cap O_i(A,B) \neq \varnothing\big\}$$
$$\cup \ \big\{\{uv, vw, wu\} \subseteq O_i : \ |\{u,v,w\}| = 3, \ \{uv, vw, wu\} \cap O_i(A,B) \neq \varnothing\big\}.$$

Since each edge-set $\alpha \in \mathcal{I}$ contains at least one edge from $O_i(A,B)$, when the $\mathcal{F}_i$-measurable event $\mathfrak{X}_{\leq i}$ holds we infer by the usual reasoning (using, e.g., $\mathcal{P}_i \cap \mathcal{Q}_i$ and $\max\{\pi_i q_i, q_i^2\} \leq 1$) that

$$\sum_{\alpha \in \mathcal{I}} \mathbb{E}(\mathbb{1}_{\{\alpha \subseteq \Gamma_{i+1}\}} \mid \mathcal{F}_i) \leq \sum_{e \in O_i(A,B)} \sum_{\alpha \in \mathcal{I} : e \in \alpha} p^{|\alpha|} \leq \sum_{e \in O_i(A,B)} \Big(|Y_e(i)| \cdot p^2 + |X_e(i)| \cdot p^3\Big)$$
$$\leq q_i |O_0(A,B)| \cdot \big(2\pi_i q_i \sqrt{n} \cdot p^2 + q_i^2 n \cdot p^3\big) \leq 3\sigma \cdot q_i |O_0(A,B)| p =: \mu_{i+1}^*.$$

Since $\mathcal{D}_{i+1}$ is a collection of edge-disjoint elements of $\mathcal{B}_{i+1}$ (and thus $\{\alpha \in \mathcal{D}_{i+1} : \alpha \cap \beta \neq \varnothing\} = \{\beta\}$ for all $\beta \in \mathcal{D}_{i+1}$), using $E(\mathcal{D}_{i+1}) = \bigcup_{\alpha \in \mathcal{D}_{i+1}} \alpha \subseteq \Gamma_{i+1} \subseteq O_i$, $|\alpha| \leq 3$ and $T_i \subseteq E_i$ it is not difficult to check that

$$Y_{i+1} = \sum_{\alpha \in \mathcal{D}_{i+1}} |\alpha \cap O_i(A,B)| \leq 3 \cdot \sum_{\alpha \in \mathcal{I} \cap \mathcal{D}_{i+1}} \mathbb{1}_{\{\alpha \in \Gamma_{i+1}\}} \leq 3Z_1,$$

where Z_1 is defined as in Theorem 15. Applying inequality (2.50) with $C = 1$ and $\mu = \mu_{i+1}^*$ (in the probability space conditional on $\mathcal{F}_i$; cf. the beginning of Section 2.3.1), when $\mathfrak{X}_{\leq i}$ holds it follows that, say,

$$\mathbb{P}(Y_{i+1} \geq y_{i+1} \mid \mathcal{F}_i) \leq \mathbb{P}(Z_1 \geq y_{i+1}/3 \mid \mathcal{F}_i) \leq \begin{cases} \left(\frac{e\mu_{i+1}^*}{y_{i+1}/3}\right)^{y_{i+1}/3} \leq \sigma^{y_{i+1}/6} & \text{if } y_{i+1} \geq 9\mu_{i+1}^*/\sqrt{\sigma}, \\ 1 & \text{otherwise.} \end{cases} \tag{2.100}$$

Comparing the definition of $\sum_{0 \leq i < I} \mu_{i+1}^*$ with μ^-, using $\tau_i \geq \tau_I \geq 1/2$ (see (2.82)) and

$\sigma \ll 1$ we see that

$$\sum_{\substack{0 \le i < I:\\ y_{i+1} \le 9\mu^*_{i+1}/\sqrt{\sigma}}} y_{i+1} \le 9/\sqrt{\sigma} \cdot \sum_{0 \le i < I} \mu^*_{i+1} \le 9/\sqrt{\sigma} \cdot 6\sigma\mu^- \ll \delta^2\mu^-/9 = y.$$

So, inserting (2.100) into (2.99), using (2.97) and the definition of s it follows that $y/\log y = \Omega(\sqrt{n}) \gg I$ and

$$\mathbb{P}\big(Y \ge \delta^2\mu^-/9 \text{ and } \mathfrak{X}_{\le I}\big) \le \sum_{\substack{(y_1,\dots,y_I) \in \mathbb{N}^I\\ \sum_{0 \le i < I} y_{i+1} = \lceil y \rceil}} \sigma^{\lceil y \rceil/6 - o(y)} \le (y+2)^I \cdot \sigma^{y/7} \le e^{-\omega(\delta^2\mu^-)} \le n^{-\omega(s)},$$

completing the proof together with the probability estimate (2.98). $\square$

2.4 Appendix

Proof of Theorem 12. We may assume that $\mathcal{I} = \{1, \dots, |\mathcal{I}|\}$. Recalling $X = f\big((\xi_i)_{i \in \mathcal{I}}\big)$, we define

$$D_i := \mathbb{E}(X \mid \xi_1, \dots, \xi_{i-1},\ \xi_i = 1) - \mathbb{E}(X \mid \xi_1, \dots, \xi_{i-1},\ \xi_i = 0) \in [-c_i, 0],$$

where $D_i \le 0$ follows from the assumption that f is decreasing, and $|D_i| \le c_i$ follows, as usual, from the assumed discrete Lipschitz property of f. Analogous to, e.g., the proof of [117, Theorem 1.3], writing $p_i = \mathbb{P}(\xi_i = 1)$ it is routine to check that

$$\Delta_i := \mathbb{E}(X \mid \xi_1, \dots, \xi_i) - \mathbb{E}(X \mid \xi_1, \dots, \xi_{i-1}) = D_i(1-p_i)\mathbb{1}_{\{\xi_i = 1\}} - D_i p_i \mathbb{1}_{\{\xi_i = 0\}}.$$

Since $1+x \le e^x$ for $x \in \mathbb{R}$ and $e^x \le 1+x+x^2/2$ for $x \le 0$, for $\theta \ge 0$ it follows easily that

$$\begin{aligned}\mathbb{E}\big(e^{\theta\Delta_i} \mid \xi_1,\dots,\xi_{i-1}\big) &= (1-p_i)\cdot e^{-\theta D_i p_i} + p_i \cdot e^{\theta D_i(1-p_i)} = e^{-\theta D_i p_i}(1-p_i+p_i e^{\theta D_i}) \\ &\le e^{-\theta D_i p_i + p_i(e^{\theta D_i}-1)} \le e^{\theta^2 D_i^2 p_i/2} \le e^{\theta^2 c_i^2 p_i/2}.\end{aligned}$$

Hence $\mathbb{E}\big(e^{\theta\sum_{i\in\mathcal{I}}\Delta_i}\big) \le e^{\theta^2\lambda/2}$, where $\lambda = \sum_{i\in\mathcal{I}} c_i^2 p_i$. Noting $X - \mathbb{E}\,X = \sum_{i\in\mathcal{I}}\Delta_i$, we deduce

$$\mathbb{P}(X \ge \mathbb{E}\,X + t) = \mathbb{P}\big(e^{\theta\sum_{i\in\mathcal{I}}\Delta_i} \ge e^{\theta t}\big) \le \mathbb{E}\big(e^{\theta\sum_{i\in\mathcal{I}}\Delta_i}\big)e^{-\theta t} \le e^{\theta^2\lambda/2-\theta t} = e^{-t^2/(2\lambda)}$$

by choosing $\theta = t/\lambda$, completing the proof of (2.48). $\square$

Proof of Lemma 17. Note that the ODE $\Psi'(x) = e^{-\Psi^2(x)}$ and $\Psi(0) = 0$ has the implicit solution

$$x = \int_0^{\Psi(x)} e^{t^2}\,dt. \tag{2.101}$$

For $x \ge 0$ it follows that $\Psi(x)$ is strictly increasing, so that $\Psi'(x) \ge 0$ is strictly decreasing. Recalling $q_i = \Psi'(i\sigma)$, we deduce $q_i \ge q_{i+1}$ and $0 \le q_i \le q_0 = 1$ for all $i \ge 0$.

To facilitate our upcoming calculations, we first prove the auxiliary claim that, for all $i \ge 0$,

$$\pi_i - \Psi(i\sigma) \in [\sigma, 2\sigma]. \tag{2.102}$$

Indeed, using $\Psi(0) = 0$ and monotonicity of Ψ' (for the first two inequalities) together with $\Psi'(0) = 1$ and $\Psi' \ge 0$ (for the last inequality) it follows that

$$0 \le \left(\sum_{0\le j\le i-1} \sigma\Psi'(j\sigma)\right) - \Psi(i\sigma) \le \sigma(\Psi'(0) - \Psi'(i\sigma)) \le \sigma,$$

which establishes (2.102) by the definition (2.37) of π_i and $\Psi'(j\sigma) = q_j$.

For (2.57), note that by (2.102) and $I = \lceil n^\beta \rceil \gg 1$ it suffices to show $\sqrt{\log x} - 1 \le$

$\Psi(x) \leq \sqrt{\log x} + 1$ for $x \geq e$ (with room to spare). The upper bound follows from $\int_0^{\sqrt{\log x}+1} e^{t^2} dt \geq x$ and (2.101). Using the inequality $(y-1)e^{-2y+1} \leq 1$ with $y = \sqrt{\log x}$, the lower bound follows from $\int_0^{\sqrt{\log x}-1} e^{t^2} dt \leq x$ and (2.101).

Turning to (2.54), note that the above calculations for (2.57) imply $\Psi'(x) = e^{-\Psi^2(x)} = x^{-1+o(1)}$ as $x \to \infty$, so that $q_I = n^{-\beta+o(1)}$. Together with $q_i \geq q_I$, it then is routine to see that (2.54) holds for $\beta < \beta_0 = 1/14$.

Now we focus on (2.53). As a warm-up, note that $\pi_i \leq \pi_I$ for $0 \leq i \leq I$ by the definition (2.37) of π_i, and that $\pi_I \leq \sqrt{\log(I\sigma)} + 2 \ll \log n = \sigma^{-1/2}$ by (2.57), so that $\sqrt{\sigma}\pi_i \leq 1$. Next, using (2.102) together with the simple inequalities $e^{-x^2}x \leq 1/2$ and $e^{-x^2}x^2 \leq 1/2$, we also infer that

$$q_i\pi_i \leq e^{-\Psi^2(i\sigma)}\big(\Psi(i\sigma) + 2\sigma\big) \leq 1, \tag{2.103}$$

$$q_i\pi_i^2 \leq e^{-\Psi^2(i\sigma)}\big(\Psi^2(i\sigma) + 4\sigma\Psi(i\sigma) + 4\sigma^2\big) \leq 1. \tag{2.104}$$

Combined with $q_i \leq 1$ this implies $q_i\pi_i^j \leq 1$ for all $j \in \{0,1,2\}$, completing the proof of (2.53).

Turning to (2.55), note that $\Psi((i+1)\sigma) \leq \pi_{i+1} - \sigma \leq \pi_i$ by (2.102), (2.37) and $q_i \leq 1$. Since $\Psi \geq 0$ is increasing and $\Psi' \geq 0$ is decreasing, using $q_j = \Psi'(j\sigma)$ together with $\Psi''(x) = -2\Psi'(x)^2\Psi(x)$ and (2.103) it follows that

$$|q_i - q_{i+1}| \leq \sigma \max_{i\sigma \leq \xi \leq (i+1)\sigma} |\Psi''(\xi)| \leq \sigma \cdot 2\Psi'(i\sigma)^2 \cdot \Psi((i+1)\sigma) \leq \sigma \cdot 2q_i^2\pi_i \leq \sigma \cdot 2\min\{q_i, q_i\pi_i\}. \tag{2.105}$$

Noting that (2.105) also implies $q_i \sim q_{i+1}$, this completes the proof of (2.55) since $q_i \geq q_{i+1}$.

Finally, for (2.56) it suffices to show $|q_i - q_{i+1} - 2\sigma q_i^2\pi_i| \leq 8\sigma^2 q_i^2$. Since $q_i = \Psi'(i\sigma)$, it follows that

$$\big|q_i - q_{i+1} + \sigma\Psi''(i\sigma)\big| \leq \tfrac{\sigma^2}{2} \max_{i\sigma \leq \xi \leq (i+1)\sigma} |\Psi'''(\xi)|.$$

As $\Psi'(x) = e^{-\Psi^2(x)}$, it is routine to check that $\Psi'''(x) = 2\Psi'(x)^3\big(4\Psi^2(x)-1\big)$. Since $\Psi \geq 0$ is increasing and $\Psi' \geq 0$ is decreasing, using $\Psi((i+1)\sigma) \leq \pi_i$ (as above), (2.104) and $q_i \leq 1$ we infer

$$\max_{i\sigma\leq\xi\leq(i+1)\sigma} |\Psi'''(\xi)| \leq 2\Psi'(i\sigma)^3 \cdot \max\big\{4\Psi^2((i+1)\sigma),\ 1\big\} \leq 2q_i^3 \max\big\{4\pi_i^2,\ 1\big\} \leq 8q_i^2.$$

Furthermore, since $\Psi''(x) = -2\Psi'(x)^2\Psi(x)$, using (2.102) we deduce

$$\big|\Psi''(i\sigma) - (-2q_i^2\pi_i)\big| = \big|-2q_i^2\Psi(i\sigma) + 2q_i^2\pi_i\big| \leq 4\sigma q_i^2,$$

which completes the proof of (2.56). $\square$

CHAPTER 3

PRAGUE DIMENSION OF RANDOM GRAPHS

3.1 Background and main results

Various notions of dimension are important in many areas of mathematics, as a measure for the complexity of objects. For graphs, one interesting notion of dimension was introduced by Nešetřil, Pultr and Rödl [85, 84] in the 1970s. The Prague dimension $\dim_{\mathrm{P}}(G)$ of a graph G (also called product dimension) is the minimum number d such that G is an induced subgraph of the product of d complete graphs. There are many equivalent definitions of $\dim_{\mathrm{P}}(G)$, see [121, 59, 5], indicating that this is a natural combinatorial notion of dimension [80, 59, 104], which in fact has appealing connections with efficient representations of graphs [121, 65, 45].

Despite receiving considerable attention during the last 40 years (including combinatorial [24, 44], information theoretic [69, 68] and algebraic [85, 80, 4, 5] approaches), the Prague dimension is still not well understood, i.e., its determination usually remains a notoriously[1] difficult task [24, 45]. To gain further insight into the behavior of this intriguing graph parameter, it thus is natural and instructive to investigate the Prague dimension of random graphs, as initiated by Nešetřil and Rödl [86] already in the 1980s. For the binomial random graph $G_{n,p}$, Füredi and Kantor conjectured that with high probability[2] (whp) the order is $\dim_{\mathrm{P}}(G_{n,p}) = \Theta(n/\log n)$ for constant edge-probabilities p, see [45, Conjecture 15] and [64].

In this chapter we prove the aforementioned Füredi–Kantor Prague dimension conjecture, by showing that the binomial random graph whp satisfies $\dim_{\mathrm{P}}(G_{n,p}) = \Theta(n/\log n)$

[1] The decision problem of whether $\dim_{\mathrm{P}}(G) \leq k$ holds is also known to be NP-complete for $k \geq 3$, see [84].

[2] As usual, we say that an event holds *whp* (with high probability) if it holds with probability tending to 1 as $n \to \infty$.

for constant edge-probabilities p.

Theorem 31 (Prague dimension of random graphs)**.** *For any fixed edge-probability $p \in (0,1)$ there are constants $c, C > 0$ so that the Prague dimension of the random graph $G_{n,p}$ satisfies with high probability*

$$c\frac{n}{\log n} \leq \dim_{\mathrm{P}}(G_{n,p}) \leq C\frac{n}{\log n}. \tag{3.1}$$

The Prague dimension of n-vertex graphs can be as large as $n-1$, see [80, 121], so an important consequence of Theorem 31 is that almost all n-vertex graphs have a significantly smaller Prague dimension of order $n/\log n$ (this follows since the random graph $G_{n,1/2}$ is uniformly distributed over all n-vertex graphs).

For our purposes it will be useful to view the Prague dimension as a clique covering and coloring problem. This convenient perspective hinges on the following equivalent definition [121, 5]: that $\dim_{\mathrm{P}}(G)$ equals the minimum number of subgraphs of the complement $\overline{G}$ of G such that (i) each subgraph is a vertex-disjoint union of cliques, and (ii) each edge of $\overline{G}$ is contained in at least one of the subgraphs, but not all of them.

Our main contribution is the upper bound on the Prague dimension in (3.1), whose proof carefully combines two different random greedy approaches: firstly, a semi-random 'nibble-type' algorithm to iteratively decompose the edges of $\overline{G_{n,p}}$ into edge-disjoint cliques of size $O(\log n)$, and, secondly, a random greedy coloring algorithm to regroup these cliques into $O(n/\log n)$ subgraphs of $\overline{G_{n,p}}$ consisting of vertex-disjoint cliques, which together eventually gives $\dim_{\mathrm{P}}(G_{n,p}) = O(n/\log n)$; see Section 3.1.3 for more details. Interestingly, this combination allows us to exploit the best features of both greedy approaches: the semi-random approach makes it easier to guarantee certain pseudo-random properties in the first decomposition step, and the random greedy approach makes it easier to guarantee that all cliques are efficiently colored in the second regrouping step (which in fact requires the pseudo-random properties established in the first step).

One major obstacle for this natural proof approach is that the cliques have size $O(\log n)$, which makes many standard tools and techniques unavailable, as they are usually restricted to objects of constant size. Notably, in order to overcome this technical difficulty in the second regrouping step, in this chapter we develop a new Pippenger–Spencer type coloring result for random hypergraphs with edges of size $O(\log n)$, which we believe to be of independent interest; see Section 3.1.1. Beyond Prague dimension and hypergraph coloring results, further contributions of this chapter include the proof of a related conjecture of Füredi and Kantor [45], and a strengthening of an old edge-covering result of Frieze and Reed [43]; see Section 3.1.2.

3.1.1 Chromatic index of random subhypergraphs

Coloring problems play an important role in much of combinatorics, and in our Prague dimension proof one key ingredient also corresponds to a hypergraph coloring result. The chromatic index $\chi'(\mathcal{H})$ of a hypergraph $\mathcal{H}$ is the smallest number of colors needed to properly color its edges, i.e., so that no two intersecting edges receive the same color. Writing $\Delta(\mathcal{H})$ for the maximum degree, it is of fundamental interest to understand when the trivial lower bound $\chi'(\mathcal{H}) \geq \Delta(\mathcal{H})$ is close to the truth. Vizing's theorem from the 1960s states that $\chi'(G) \leq \Delta(G) + 1$ for any graph G. Influential work of Pippenger and Spencer [93] from the 1980s gives a partial answer for r-uniform hypergraphs $\mathcal{H}$ with edges of size $r = \Theta(1)$: for any $\delta > 0$ they showed that $\chi'(\mathcal{H}) \leq (1+\delta)\Delta(\mathcal{H})$ for any nearly regular $\mathcal{H}$ with small codegrees, effectively removing the edge-size dependence from the trivial greedy upper bound $\chi'(\mathcal{H}) \leq r(\Delta(\mathcal{H}) - 1) + 1$.

It is challenging to extend the Pippenger–Spencer coloring arguments to edges of size $r = O(\log n)$, which is what we desire in our main Prague dimension proof (where cliques correspond to edges of an auxiliary hypergraph). Our Theorem 32 overcomes this size obstacle in the random setting, i.e., for coloring random edges of any nearly regular hypergraph $\mathcal{H}$ with small codegrees. As we shall see in Sections 3.1.3 and 3.2.2.1, this probabilistic Pippenger–Spencer type coloring result indeed suffices for our purposes. Here $\deg_{\mathcal{H}}(v) := |\{e \in E(\mathcal{H}) : v \in e\}|$ and $\deg_{\mathcal{H}}(u,v) := |\{e \in E(\mathcal{H}) : \{u,v\} \subseteq e\}|$ denote the degree and codegree, as usual.

Theorem 32 (Chromatic index of random subhypergraphs)**.** *For all reals $\delta, \sigma, b > 0$ with $b \le \delta\sigma/30$ there is $n_0 = n_0(\delta, \sigma, b) > 0$ such that, for all integers $n \ge n_0$, $2 \le r \le b \log n$, $n^{1+\sigma} \le m \le n^{rn^{\sigma/5}}$ and all reals $D > 0$, the following holds for every n-vertex r-uniform hypergraph $\mathcal{H}$ satisfying*

$$\max_{v \in V(\mathcal{H})} |\deg_{\mathcal{H}}(v) - D| \le n^{-\sigma} D \qquad \text{and} \qquad \max_{u \neq v \in V(\mathcal{H})} \deg_{\mathcal{H}}(u,v) \le n^{-\sigma} D. \tag{3.2}$$

We have $\mathbb{P}(\chi'(\mathcal{H}_m) \le (1+\delta) rm/n) \ge 1 - m^{-\omega(r)}$, where $\mathcal{H}_m$ denotes the random subhypergraph of $\mathcal{H}$ with edges $e_1, \dots, e_m$, where each edge e_i is independently chosen uniformly at random from $\mathcal{H}$.

Remark 33. *Noting $D \cdot m/|E(\mathcal{H})| = (1+o(1)) rm/n \gg r \log m$, for any real $\epsilon > 0$ it is straightforward to see that the maximum degree satisfies $\Delta(\mathcal{H}_m) = (1 \pm \epsilon) rm/n$ with probability at least $1 - m^{-\omega(r)}$, say.*

As discussed, for this chapter the key point is that Theorem 32 permits edges of size $r = O(\log n)$; we have made no attempt to optimize the ad-hoc assumptions on the number of edges m or the $n^{-\sigma}$ approximation terms in (3.2). The explicit technical assumption $b \le \delta\sigma/30$ allows for some flexibility in applications: setting $b = \delta\sigma/30$ and $\delta = 30b/\sigma$, respectively, using Remark 33 we readily infer that whp

$$\chi'(\mathcal{H}_m) \le \begin{cases} \big(1 + 2\delta\big) \cdot \Delta(\mathcal{H}_m) & \text{if } r = o(\log n), \\ O(1) \cdot \Delta(\mathcal{H}_m) & \text{if } r = O(\log n), \end{cases} \tag{3.3}$$

which gives Pippenger–Spencer like chromatic index bounds for many non-constant edge-sizes r; we believe that these bounds are of independent interest (see also Corollary 40).

We prove Theorem 32 by showing that a simple random greedy algorithm (that differs from the one used by Pippenger and Spencer [93]) whp produces the desired coloring of the random edges $e_1, \dots, e_m$ from $\mathcal{H}$. The algorithm we use sequentially assigns each edge e_i a random color in $\{1, \dots, \lfloor (1+\delta)rm/n \rfloor\}$ that does not appear on some adjacent edge e_j with $j < i$; see Section 3.3. This random greedy edge coloring algorithm is very natural: Kurauskas and Rybarczyk [74] analyzed it when $\mathcal{H}$ is the complete n-vertex r-uniform hypergraph, and its idea also underpins earlier work that extends the Pippenger–Spencer result to list-colorings [62, 83]. Taking advantage of the random setting, our proof of Theorem 32 uses differential equation method [123, 10, 119] based martingale arguments to show that this greedy algorithm whp properly colors the first m out of $(1+\delta)m$ random edges. This 'more random edges' twist enables us to sidestep some of the 'last few edges' complications[3] that usually arise in the deterministic setting [93, 62, 83], which is one of the reasons why our analysis can allow for edges of size $O(\log n)$; see Section 3.3 for the details.

3.1.2 Partitioning the edges of a random graph into cliques

Further motivation for studying the Prague dimension comes from its close connection to the covering and decomposition problems that pervade combinatorics, one interesting non-standard feature being that Theorem 31 requires usage of cliques with $O(\log n)$ vertices, rather than just subgraphs of constant size. The clique covering number $\mathrm{cc}(G)$ of a graph G (also called intersection number) is the minimum number of cliques in G that cover the edge-set of G. Similarly, the clique partition number $\mathrm{cp}(G)$ is the minimum number of cliques in G that partition the edge-set of G. The question of estimating these natural graph parameters was raised by Erdős, Goodman and Pósa [31] in 1966. Motivated in part

[3]Many deterministic approaches such as [93, 62] first efficiently color most of the edges of $\mathcal{H}$ using $(1+\delta/2)\Delta(\mathcal{H})$ colors, say, so that the remaining uncolored 'last few edges' yield a hypergraph with maximum degree at most $\epsilon\Delta(\mathcal{H})$, say. By choosing the constant $\epsilon = \epsilon(r,\delta) > 0$ sufficiently small, these 'last few edges' can then trivially be colored using $r \cdot \epsilon\Delta(\mathcal{H}) \leq \delta/2 \cdot \Delta(\mathcal{H})$ additional colors, which clearly becomes harder to implement when $r = r(n) \to \infty$ (as now the dependence of ϵ on r matters).

by applications such as keyword conflicts, traffic phasing and competition graph analysis [87, 70, 96, 23], both $\mathrm{cc}(G)$ and $\mathrm{cp}(G)$ have since been extensively studied for many interesting graph classes, see e.g. [114, 4, 17, 32, 18, 21] and the many references therein.

For random graphs, the study of the clique covering number was initiated in the 1980s by Poljak, Rödl and Turzík [94] and Bollobás, Erdős, Spencer and West [15]. In 1995, Frieze and Reed [43] showed that whp $\mathrm{cc}(G_{n,p}) = \Theta(n^2/(\log n)^2)$ for constant edge-probabilities p. Constructing a clique covering is certainly easier than constructing a clique partition, since it does not have to satisfy such a rigid edge constraint. Indeed, while obviously $\mathrm{cc}(G) \le \mathrm{cp}(G)$, the ratio $\mathrm{cp}(G)/\mathrm{cc}(G)$ can in fact be arbitrarily large, see [28]. However, our Theorem 34 demonstrates that for most graphs the clique partition number and clique covering number have the same order of magnitude.

Theorem 34 (Clique covering and partition number of random graphs)**.** *For every fixed real $\gamma \in (0,1)$ there are constants $c > 0$ and $C = C(\gamma) > 0$ so that if the edge-probability $p = p(n)$ satisfies $n^{-2} \ll p \le 1-\gamma$, then with high probability*

$$c\frac{n^2p}{(\log_{1/p} n)^2} \le cc(G_{n,p}) \le cp(G_{n,p}) \le C\frac{n^2p}{(\log_{1/p} n)^2}. \tag{3.4}$$

The main contribution of (3.4) is the upper bound, which strengthens the main result of Frieze and Reed [43] from clique coverings to clique partitions, and also allows for $p = p(n) \to 0$. Here the mild assumption $p \le 1-\gamma$ turns out to be necessary, since Lemma 50 implies that whp $\mathrm{cc}(G_{n,p})/(n^2p/(\log_{1/p} n)^2) \to \infty$ as $p \to 1$. The lower bound in (3.4) is straightforward: it is well-known that $G_{n,p}$ whp has $m = \Theta(n^2p)$ edges and largest clique of size $\omega = O(\log_{1/p} n)$, which gives $\mathrm{cc}(G_{n,p}) \ge m/\binom{\omega}{2} = \Omega(n^2p/(\log_{1/p} n)^2)$.

To gain a better combinatorial understanding of clique coverings, it is instructive to study and optimize other properties besides the size, such as their thickness $\mathrm{cc}_\Delta(G) := \min_{\mathcal{C}} \Delta(\mathcal{C})$ and chromatic index $\mathrm{cc}'(G) := \min_{\mathcal{C}} \chi'(\mathcal{C})$, where the minimum is taken over all clique coverings $\mathcal{C}$ of the edges of G (formally thinking of $\mathcal{C}$ as a hypergraph with vertex-set $V(G)$

and edge-set $\mathcal{C}$). Notably, the parameters $\mathrm{cc}'(\overline{G})$ and $\mathrm{cc}_\Delta(\overline{G})$ approximate the Prague dimension and the so-called Kneser rank of G, see [45]. In particular, we have

$$\mathrm{cc}'(\overline{G}) \le \dim_{\mathrm{P}}(G) \le \mathrm{cc}'(\overline{G}) + 1, \tag{3.5}$$

which follows by noting that the color classes of a properly colored collection $\mathcal{C}$ of cliques naturally correspond to subgraphs consisting of vertex-disjoint unions of cliques (the $+1$ in the upper bound is only needed to handle boundary cases where an edge is contained in cliques from all color classes); see [85, 45].

For random graphs, Füredi and Kantor [45] showed that the clique covering thickness is whp $\mathrm{cc}_\Delta(G_{n,p}) = \Theta(n/\log n)$ for constant edge-probabilities p. Supported by $\mathrm{cc}_\Delta(G) \le \mathrm{cc}'(G)$ and further evidence, they conjectured that the clique covering chromatic index is whp also $\mathrm{cc}'(G_{n,p}) = \Theta(n/\log n)$ for constant p, see [45, Conjecture 17]. The following theorem proves their chromatic index conjecture in a strong form, allowing for $p = p(n) \to 0$. More importantly, Theorem 35 and inequality (3.5) together imply our main Prague dimension result Theorem 31, since the complement $\overline{G_{n,p}}$ of $G_{n,p}$ has the same distribution as $G_{n,1-p}$.

Theorem 35 (Thickness and chromatic index of clique coverings of random graphs)**.** *For every fixed real $\gamma \in (0,1)$ there are constants $c > 0$ and $C = C(\gamma) > 0$ so that if the edge-probability $p = p(n)$ satisfies $n^{-1}\log n \ll p \le 1-\gamma$, then with high probability*

$$c\frac{np}{\log_{1/p} n} \le cc_\Delta(G_{n,p}) \le cc'(G_{n,p}) \le C\frac{np}{\log_{1/p} n}. \tag{3.6}$$

Remark 36. *Our proof shows that the upper bound in* (3.6) *remains valid when the definition of $cc'(G)$ is restricted to clique partitions of the edges (instead of clique coverings); see Sections 3.1.3.1 and 3.2.*

The main contribution of (3.6) is the upper bound, where the mild assumption $p \le 1-\gamma$

again turns out to be necessary, since Lemma 50 implies that whp $\mathrm{cc}_\Delta(G_{n,p})/(np/\log_{1/p} n) \to \infty$ as $p \to 1$. The lower bound in (3.6) is straightforward: it is well-known that $G_{n,p}$ whp has maximum degree $\Delta = \Theta(np)$ and largest clique of size $\omega = O(\log_{1/p} n)$, which gives $\mathrm{cc}_\Delta(G_{n,p}) \geq \Delta/(\omega - 1) = \Omega(np/\log_{1/p} n)$.

3.1.3 Proof strategy: finding a clique partition of a random graph

We now comment on the proofs of Theorems 34–35, for which it remains to establish the upper bounds in inequalities (3.4) and (3.6). In Section 3.2 we shall establish these upper bounds using the following proof strategy, which finds a clique partition $\mathcal{P}$ of $G_{n,p}$ with the desired properties, i.e., size and chromatic index bounds.

Step 1: Decomposing the edges of $G_{n,p}$ into a clique partition $\mathcal{P}$. We first use a semi-random 'nibble-type' algorithm to incrementally construct a decreasing sequence of n-vertex graphs

$$G_{n,p} = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_I, \tag{3.7}$$

inspired by the semi-random approaches of Frieze and Reed [43] and Guo and Warnke [56]. Omitting some technicalities, the main idea is to obtain G_{i+1} from G_i by removing the edges of a random collection $\mathcal{K}_i$ of cliques of size $k_i = O(\log n)$ from G_i. We iterate this until G_I is sufficiently sparse, i.e., has maximum degree $\Delta(G_I) = o(np/(\log_{1/p} n)^2)$, say, and then put all remaining edges of G_I into $\mathcal{K}_I$, to ensure that

$$\mathcal{P} \;=\; \mathcal{K}_0 \cup \cdots \cup \mathcal{K}_I \tag{3.8}$$

covers all edges of $G_{n,p}$. Here we exploit the flexibility of the semi-random approach, which allows us to add extra wrinkles to the algorithm. In particular, using concentration inequalities, these extra wrinkles enable us to show that whp all graphs G_i stay pseudo-random, i.e., that G_i 'looks like' a random graph G_{n,p_i} with suitably decaying edge-probabilities p_i; see Section 3.2.1 and Theorem 39 for the details.

Step 2: Coloring the clique partition $\mathcal{P}$. We then use the basic observation

$$\chi'(\mathcal{P}) \;\le\; \sum_{0\le i\le I} \chi'(\mathcal{K}_i) \;\le\; \sum_{0\le i<I} \chi'(\mathcal{K}_i) \;+\; 2\Delta(G_I), \tag{3.9}$$

where the last inequality $\chi'(\mathcal{K}_i) \le 2\Delta(G_I)$ follows from Vizing's theorem, since $\mathcal{K}_I = E(G_I)$ simply contains all edges of G_I. Thinking of $\mathcal{K}_i$ as a hypergraph with vertex-set $V(G_i)$ and edge-set $\mathcal{K}_i$, we would like to similarly bound $\chi'(\mathcal{K}_i) = O(\Delta(\mathcal{K}_i))$, but there is a major obstacle here. Namely, as discussed, such Pippenger–Spencer type coloring results only apply to hypergraphs with edges of constant size, and their proofs are hard to extend to hypergraphs with edges of size $O(\log n)$ such as $\mathcal{K}_i$. We overcome this technical obstacle by exploiting that $\mathcal{K}_i$ is a random collection of cliques from G_i. Crucially, this enables us to bound $\chi'(\mathcal{K}_i)$ using our new probabilistic Pippenger–Spencer type result Theorem 32, which efficiently colors such random hypergraphs with large edges. In view of (3.3), it thus becomes plausible that whp

$$\chi'(\mathcal{P}) \;\le\; \sum_{0\le i<I} O\big(\Delta(\mathcal{K}_i)\big) \;+\; O\big(\Delta(G_I)\big), \tag{3.10}$$

where the pseudo-random properties are key for verifying the technical assumptions of Theorem 32. Using again pseudo-randomness to estimate $\Delta(\mathcal{K}_i)$ and $\Delta(G_I)$, it turns out[4] that whp

$$\chi'(\mathcal{P}) \;\le\; \sum_{0\le i<I} O\bigg(\frac{np_i}{\log_{1/p_i} n}\bigg) \;+\; O\big(np_I\big) \;\le\; \cdots \;\le\; O\bigg(\frac{np}{\log_{1/p} n}\bigg), \tag{3.11}$$

where the exponentially decaying edge-probabilities p_i will ensure that in estimate (3.11) the bulk of the contribution comes from the case $i = 0$ with $p_0 = p$; see Section 3.2.2.1 for the details. Finally, the whp size estimate $|\mathcal{P}| = O\big(n^2p/(\log_{1/p} n)^2\big)$ can be obtained in a

[4]Heuristically, the form of the upper bound (3.11) can be motivated as follows: (3.7) and $G_i \approx G_{n,p_i}$ loosely suggest $\mathrm{cc}'(G_{n,p}) \le \sum_{0\le i\le I} \mathrm{cc}'(G_{n,p_i})$, which together with (3.6) and $\mathrm{cc}'(G_{n,p_I}) \le 2\Delta(G_{n,p_I}) = O(np_I)$ makes (3.11) a natural target bound.

similar but simpler way; see Section 3.2.2.2.

3.1.3.1 Technical result: weakly pseudo-random clique partition

As we shall see in Section 3.2, the outlined proof strategy gives the following technical result, which for large edge-probabilities $p = p(n)$ intuitively guarantees that the random graph $G_{n,p}$ has a weakly pseudo-random clique partition $\mathcal{P}$, i.e., which simultaneously has small size, thickness and chromatic index.

Theorem 37. *There is a constant $\alpha > 0$ so that, for every fixed real $\gamma \in (0,1)$, there are constants $B, C > 0$ such that the following holds. If the edge-probability $p = p(n)$ satisfies $n^{-\alpha} \le p \le 1-\gamma$, then whp there exists a clique partition $\mathcal{P}$ of the edges of $G_{n,p}$ satisfying $\max_{K\in\mathcal{P}} |K| \le \log_{1/p} n$, $|\mathcal{P}| \le Bn^2p/(\log_{1/p} n)^2$ and $\Delta(\mathcal{P}) \le \chi'(\mathcal{P}) \le Cnp/\log_{1/p} n$.*

Remark 38. *The proof shows that the whp conclusion in fact holds with probability at least $1 - n^{-\omega(1)}$.*

After potentially increasing the constants $B, C > 0$, this theorem readily implies the upper bounds in (3.4) and (3.6) of Theorems 34–35, since for smaller edge-probabilities $p = p(n) \le n^{-\alpha}$ the trivial clique partition $\mathcal{P} := E(G_{n,p})$ consisting of all edges of $G_{n,p}$ easily[5] gives the desired bounds due to $1 \le 1/(\alpha \log_{1/p} n)$.

3.1.4 Organization

In Section 3.2 we prove our main technical clique partition result Theorem 37 (which as discussed implies Theorems 31, 34 and 35), by analyzing a semi-random greedy clique partition algorithm using concentration inequalities and our new chromatic index result Theorem 32. We then prove our key tool Theorem 32 in Section 3.3, by analyzing a natural random greedy edge coloring algorithm using the differential equation method. The final

[5]Using well-known estimates, it is easy to see that whp $|\mathcal{P}| = |E(G_{n,p})| \sim \binom{n}{2}p < 2\alpha^{-2} \cdot \binom{n}{2}p/(\log_{1/p} n)^2$ for $n^{-2} \ll p \le n^{-\alpha}$ and whp $\chi'(\mathcal{P}) = \chi'(G_{n,p}) \le \Delta(G_{n,p}) + 1 \sim np < 2\alpha^{-1} \cdot np/(\log_{1/p} n)$ for $n^{-1}\log n \ll p \le n^{-\alpha}$.

Section 3.4 discusses some open problems, sharpens the lower bounds of Theorems 34–35 for constant edge-probabilities p, and also records strengthenings of Theorems 34–35 for many small $p = p(n) \to 0$.

3.2 Semi-random greedy clique partition algorithm

In this section we prove Theorem 37 (and thus Theorems 31, 34 and 35, see Sections 3.1.2–3.1.3) by showing that a certain semi-random greedy algorithm is likely to find the desired clique partition $\mathcal{P}$ of the binomial random graph $G_{n,p}$. This algorithm iteratively adds cliques to $\mathcal{P}$, and the main idea is roughly as follows. Writing $G_i \subseteq G_{n,p}$ for the subgraph containing all edges of $G_{n,p}$ which are edge-disjoint from the cliques added to $\mathcal{P}$ during the first i iterations, we randomly sample a collection $\mathcal{K}_i$ of cliques from G_i (of suitable size k_i). We then alter this collection to ensure that there are no edge-overlaps between the cliques, and add the resulting edge-disjoint collection $\mathcal{K}_i^* \cup D_i$ of cliques to $\mathcal{P}$. Finally, after a sufficiently large number of I iterations, we add all remaining so-far uncovered edges of $G_I \subseteq G_{n,p}$ to $\mathcal{P}$ (as cliques of size two).

In fact, we shall use an additional wrinkle for technical reasons: in each iteration of the algorithm we add an extra set S_i of random edges to $\mathcal{P}$, which helps us to ensure that the graphs $G_i = ([n], E_i)$ stay pseudo-random, i.e., resemble a random graph G_{n,p_i} with suitable decaying edge-probabilities p_i.

3.2.1 Details of the semi-random 'nibble' algorithm

Turning to the technical details of our clique partition algorithm, let

$$k := \lceil \sigma \log_{1/p} n \rceil, \quad I := \lceil \tau k^{\tau} \log k \rceil, \quad p_i := p e^{-i/k^{\tau}}, \quad k_i := \lceil \sigma \log_{1/p_i} n \rceil, \quad \epsilon := n^{-\sigma}, \tag{3.12}$$

where we fix the absolute constants $\sigma := 1/9$ and $\tau := 9$ for concreteness (we have made no attempt to optimize these constants, and the reader looses little by simply assuming that σ and τ are always sufficiently small and large, respectively, whenever needed). For any vertex-subset $U \subseteq [n]$ with $|U| \le j$ we define

$$\mathcal{C}_{U,j,i} := \left\{ J \subseteq [n] : \ U \subseteq J, \ |J| = j, \ \tbinom{J}{2} \setminus \tbinom{U}{2} \subseteq E_i \right\}. \tag{3.13}$$

In words, if U forms a clique in the graph $G_i = ([n], E_i)$, then $\mathcal{C}_{U,j,i}$ corresponds to the set of all j-vertex cliques of G_i that contain U. Furthermore, if G_i indeed heuristically resembles the random graph G_{n,p_i} (as suggested above, and later made precise by Theorem 39), then we expect that $|\mathcal{C}_{U,j,i}| \approx \mu_{|U|,j,i}$, where

$$\mu_{s,j,i} := \binom{n-s}{j-s} p_i^{\binom{j}{2} - \binom{s}{2}}. \tag{3.14}$$

Writing $E(\mathcal{C}) := \bigcup_{K \in \mathcal{C}} E(K)$ for the edges covered by a family $\mathcal{C}$ of cliques, after defining

$$q_i := \frac{1}{(1+\epsilon) k^{\tau} \mu_{2,k_i,i}} \quad \text{and} \quad \zeta_{e,i} := 1 - (1 - q_i)^{\max\{(1+\epsilon)\mu_{2,k_i,i} - |\mathcal{C}_{e,k_i,i}|, \, 0\}} \tag{3.15}$$

we now formally state the algorithm that finds the desired clique partition $\mathcal{P}$ of $G_{n,p}$.

Algorithm: Semi-random greedy clique partition

Set $\mathcal{P}_0 := \varnothing$ and $G_0 := ([n], E_0)$, where $E_0 := E(G_{n,p})$.
for $i = 0$ **to** $I - 1$ **do**
  Let $\mathcal{C}_i := \mathcal{C}_{\varnothing,k_i,i}$ contain all k_i-vertex cliques of G_i.
  Generate $\mathcal{K}_i \subseteq \mathcal{C}_i$: independently include each clique $K \in \mathcal{C}_i$ with probability q_i.
  Generate $S_i \subseteq E_i$: independently include each edge $e \in E_i$ with probability $\zeta_{e,i}$.
  Let $\mathcal{K}_i^*$ be a size-maximal collection of edge-disjoint k_i-vertex cliques in $\mathcal{K}_i$.
  Set $\mathcal{P}_{i+1} := \mathcal{P}_i \cup \mathcal{K}_i^* \cup D_i \cup \big(S_i \setminus E(\mathcal{K}_i)\big)$, where $D_i := E(\mathcal{K}_i) \setminus E(\mathcal{K}_i^*)$.
  Set $G_{i+1} := ([n], E_{i+1})$, where $E_{i+1} := E_i \setminus \big(E(\mathcal{K}_i) \cup S_i\big)$.
end for
Return $\mathcal{P} := \mathcal{P}_I \cup E_I$.

One may heuristically motivate the technical definitions (3.15) of q_i and $\zeta_{e,i}$ as follows. The 'inclusion' probability q_i will intuitively ensure that, for any fixed edge $e \in E_i$, the expected number of cliques in $\mathcal{K}_i$ containing e is roughly $|\mathcal{C}_{e,k_i,i}| \cdot q_i \approx \mu_{2,k_i,i} \cdot q_i \approx 1/k^{\tau}$. This makes it plausible that the cliques in $\mathcal{K}_i$ are largely edge-disjoint, i.e., that $|\mathcal{K}_i^*| \approx |\mathcal{K}_i|$. The 'stabilization' probability $\zeta_{e,i}$ will intuitively ensure that

$$\mathbb{P}(e \in E_{i+1} \mid e \in E_i) = (1 - q_i)^{|\mathcal{C}_{e,k_i,i}|} \cdot (1 - \zeta_{e,i}) \approx (1 - q_i)^{(1+\epsilon)\mu_{2,k_i,i}} \approx e^{-1/k^{\tau}}.$$

Since all edges $e \in E_i$ of G_i have roughly the same probability of appearing in E_{i+1}, it then inductively becomes plausible that G_{i+1} resembles a random graph $G_{n,p_{i+1}}$ with edge-probability $p_{i+1} \approx p_i \cdot e^{-1/k^{\tau}}$.

3.2.2 The clique partition $\mathcal{P}$: proof of Theorem 37

In this section we prove Theorem 37 by analyzing the clique partition $\mathcal{P}$ produced by the semi-random greedy algorithm. Recalling the definitions (3.13)–(3.14) of $|\mathcal{C}_{U,j,i}|$ and $\mu_{s,j,i}$, Theorem 39 confirms our heuristic that G_i stays pseudo-random, i.e., resembles the random graph G_{n,p_i} with respect to various clique statistics.

Theorem 39 (Pseudo-randomness of the graphs G_i)**.** *Let $p = p(n)$ satisfy $n^{-\sigma/\tau} \le p \le 1-\gamma$, where $\gamma \in (0,1)$ is a constant. Then, with probability at least $1-n^{-\omega(1)}$, for all $0 \le i \le I$ the following event $\mathcal{R}_i$ holds: for all $U \subseteq [n]$ and j with $0 \le |U| \le j \le k_i$, we have*

$$|\mathcal{C}_{U,j,i}| \;=\; (1 \pm \epsilon) \cdot \mu_{|U|,j,i}. \tag{3.16}$$

We defer the proof of this important technical auxiliary result to Section 3.2.3, and first use it (together with our new edge-coloring result Theorem 32) to prove Theorem 37 with $\alpha := \sigma/\tau$. To this end, we henceforth tacitly assume $n^{-\sigma/\tau} \le p \le 1-\gamma$. In particular, for $0 \le i \le I$ it then is routine to check that

$$8 < \tau - o(1) \le \frac{\sigma \log n}{\log(k^{2\tau}/p)} \le k_i \le k \le n^{o(1)} \qquad \text{and} \qquad \min_{0 \le s \le k_i - 1} p_i^s \ge p_i^{k_i-1} \ge n^{-\sigma}. \tag{3.17}$$

By construction of $\mathcal{P}$, it also follows that $\max_{K \in \mathcal{P}} |K| \le k \le \log_{1/p} n$. To complete the proof of Theorem 37, it thus remains to bound the size and chromatic index of the clique partition $\mathcal{P}$.

3.2.2.1 Chromatic index of $\mathcal{P}$

We first focus on the chromatic index of the clique partition $\mathcal{P}$, which is easily seen to be (by separately coloring different subsets of the cliques, using disjoint sets of colors) at most

$$\chi'(\mathcal{P}) \le \sum_{0 \le i \le I-1} \Big(\chi'(\mathcal{K}_i) + \chi'(D_i) + \chi'(S_i)\Big) + \chi'(E_I). \tag{3.18}$$

For $\mathcal{S} \in \{\mathcal{K}_i, D_i, S_i, E_I\}$, let $\mathcal{S}^{(v)} \subseteq \mathcal{S}$ denote the subset of cliques that contain the vertex v. Since the cliques in D_i, S_i, E_I are all simply edges, using Vizing's theorem it follows that

$$\chi'(\mathcal{P}) \leq \sum_{0\leq i\leq I-1} \Big(\chi'(\mathcal{K}_i) + \max_{v\in[n]}\big|D_i^{(v)}\big| + \max_{v\in[n]}\big|S_i^{(v)}\big|\Big) + \max_{v\in[n]}\big|E_I^{(v)}\big| + (2I+1). \quad (3.19)$$

In the following we bound the contributions of each of these terms. We start with the main term $\chi'(\mathcal{K}_i)$, where the trivial upper bound $\chi'(\mathcal{K}_i) \leq O(k_i)\cdot\Delta(\mathcal{K}_i)$ would be too weak for our purposes. Gearing up to instead apply our stronger Pippenger–Spencer type chromatic index result Theorem 32 to the random set $\mathcal{K}_i \subseteq \mathcal{C}_i$ of cliques, let $\mathcal{H} := ([n], \mathcal{C}_i)$ denote the k_i-uniform auxiliary hypergraph consisting of all k_i-vertex cliques in G_i. Note that $\mathcal{K}_i$ has the same distribution as the edge-set of $\mathcal{H}_{q_i}$, where the random subhypergraph $\mathcal{H}_{q_i} \subseteq \mathcal{H}$ is defined as in Corollary 40 with $r = k_i$ and $q = q_i$ (we defer the proof of Corollary 40 to Section 3.6, since this standard reduction to Theorem 32 is rather tangential to the main argument here).

Corollary 40 (Convenient variant of Theorem 32)**.** *There is $\xi = \xi(\delta) > 0$ such that if the assumptions of Theorem 32 hold for a given n-vertex r-uniform hypergraph $\mathcal{H}$, with assumption $m \leq n^{rn^{\sigma/5}}$ replaced by $m \leq \xi e(\mathcal{H})$, then we have $\mathbb{P}(\chi'(\mathcal{H}_q) \leq (1+2\delta)rm/n) \geq 1 - n^{-\omega(r)}$, where $\mathcal{H}_q$ denotes the random subhypergraph of $\mathcal{H}$ where each edge $e \in \mathcal{H}$ is independently included with probability $q := m/|E(\mathcal{H})|$.*

Conditional on $\mathcal{R}_i$, we will apply this corollary to $\mathcal{H} = ([n], \mathcal{C}_i)$ with $r := k_i$, $m := |E(\mathcal{H})|q_i$, $D := \mu_{1,k_i,i}$, $q := q_i$, as well as

$$b := 2\sigma/\log(1/(1-\gamma)) \quad \text{and} \quad \delta := 30b/\sigma. \quad (3.20)$$

We now verify the technical assumptions of Corollary 40 (and thus Theorem 32). Using the definition (3.12) of k_i and inequality (3.17) together with $p_i \leq p \leq 1-\gamma$, we obtain $2 < k_i \leq 2\sigma(\log n)/\log(1/p_i) \leq b\log n$. Using the estimate (3.16) of $\mathcal{R}_i$ together with the definition (3.15) of q_i, it follows that

$$m = |\mathcal{C}_{\varnothing,k_i,i}| \cdot q_i = \frac{(1 \pm \epsilon)\mu_{0,k_i,i}}{(1+\epsilon)k^\tau \mu_{2,k_i,i}} = \frac{1 \pm \epsilon}{1+\epsilon} \cdot \frac{n(n-1)p_i}{k_i(k_i-1)k^\tau}, \tag{3.21}$$

so that $m \geq n^{2-\sigma-o(1)} \gg n^{1+\sigma}$ by (3.17) and choice of σ. Recalling that $\epsilon = n^{-\sigma}$, estimate (3.16) implies that $\mathcal{H} = ([n], \mathcal{C}_i)$ satisfies the degree condition in (3.2). We also have $\mu_{2,k_i,i}/\mu_{1,k_i,i} \leq \big(\Omega(n/k_i) \cdot p_i\big)^{-1} \leq n^{-1+\sigma+o(1)} \ll n^\sigma$, which in view of (3.16) and $D = \mu_{1,k_i,i}$ implies that $\mathcal{H}$ also satisfies the codegree condition in (3.2). We similarly infer $D = \big(\Omega(n/k_i) \cdot p_i^{k_i/2}\big)^{k_i-1} \geq (n^{1-\sigma-o(1)})^4 \gg n^3$, so that $m = O(n^2/k_i) \ll D/r \ll e(\mathcal{H})$. We thus may apply Corollary 40 to $\mathcal{H}$, which together with our above discussion gives

$$\begin{aligned} &\mathbb{P}\big(\chi'(\mathcal{K}_i) \geq (1+2\delta)k_i m/n \mid \mathcal{R}_i\big) \\ =&\mathbb{P}\big(\chi'(\mathcal{H}_q) \geq (1+2\delta)k_i m/n \mid \mathcal{R}_i\big) \leq n^{-\omega(1)}. \end{aligned} \tag{3.22}$$

In the following we fix a vertex $v \in [n]$, and bound $|S_i^{(v)}|$ and $|D_i^{(v)}|$ separately. For these terms we will have some elbow-room, and we can thus be more generous in our upcoming estimates. Using (3.16) together with $1-(1-q_i)^{2\epsilon\mu_{2,k_i,i}} \leq 2\epsilon\mu_{2,k_i,i}q_i \leq 2\epsilon k^{-\tau}$ and $\epsilon = n^{-\sigma} \ll k^{-2} \leq k_i^{-2}$, it follows that

$$\mathbb{E}\big(|S_i^{(v)}| \mid \mathcal{R}_i\big) \leq |\mathcal{C}_{\{v\},2,i}| \cdot \big(1-(1-q_i)^{2\epsilon\mu_{2,k_i,i}}\big) \leq \frac{2\epsilon n p_i}{k^\tau} \ll \frac{np_i}{k_i^2 k^\tau} =: \lambda. \tag{3.23}$$

Note that $\lambda \geq n^{1-\sigma-o(1)} \gg \log n$ by inequality (3.17) and choice of σ. Furthermore, since $|S_i^{(v)}|$ is a sum of independent indicator random variables, standard Chernoff bounds (such as [60, Theorem 2.1]) imply

$$\mathbb{P}\big(|S_i^{(v)}| \geq 2\lambda \mid \mathcal{R}_i\big) \leq \exp\big(-\Theta(\lambda)\big) \leq n^{-\omega(1)}. \tag{3.24}$$

Turning to $|D_i^{(v)}|$, let X denote the number of unordered pairs $\{K', K''\} \in \binom{\mathcal{K}_i}{2}$ with

$|\{K', K''\} \cap \mathcal{K}_i^{(v)}| \geq 1$ and $|E(K') \cap E(K'')| \geq 1$. Since each of these edge-overlapping clique pairs contributes at most $k_i \leq k$ edges to $|D_i^{(v)}|$, we infer $|D_i^{(v)}| \leq kX$. Furthermore, using (3.16) and (3.15), it follows similarly to (3.21) that

$$\begin{aligned}\mathbb{E}(X \mid \mathcal{R}_i) &\leq \sum_{K' \in \mathcal{C}_{\{v\},k_i,i}} \sum_{e \in \binom{K'}{2}} \sum_{K'' \in \mathcal{C}_{e,k_i,i}} q_i \\ &\leq |\mathcal{C}_{\{v\},k_i,i}| \cdot \binom{k_i}{2} \cdot (1+\epsilon)\mu_{2,k_i,i} \cdot q_i^2 \leq \frac{k_i n p_i}{k^{2\tau}} =: \mu.\end{aligned} \tag{3.25}$$

Conditioning on the event $\mathcal{R}_i$, we shall bound X using the following upper tail inequality for combinatorial random variables, which is a convenient corollary of [118, Theorem 9].

Lemma 41. *Let $(\xi_j)_{j \in \Lambda}$ be a finite family of independent random variables with $\xi_j \in \{0,1\}$. Let $(Y_\alpha)_{\alpha \in \mathcal{I}}$ be a finite family of random variables with $Y_\alpha := \mathbb{1}_{\{\xi_j = 1 \text{ for all } j \in \alpha\}}$. Defining $\mathcal{I}^+ := \{\alpha \in \mathcal{I} : Y_\alpha = 1\}$, let $\mathcal{G}$ be an event that implies $\max_{\alpha \in \mathcal{I}^+} |\{\beta \in \mathcal{I}^+ : \beta \cap \alpha \neq \varnothing\}| \leq C$. Set $X := \sum_{\alpha \in \mathcal{I}} Y_\alpha$, and assume that $\mathbb{E}\, X \leq \mu$. Then, for all $x > \mu$,*

$$\mathbb{P}(X \geq x \text{ and } \mathcal{G}) \;\leq\; \big(e\mu/x\big)^{x/C}. \tag{3.26}$$

We will apply Lemma 41 to X with $\Lambda = \mathcal{C}_i$, the independent random variables $\xi_K := \mathbb{1}_{\{K \in \mathcal{K}_i\}}$, and $\mathcal{I}$ equal to the set of unordered pairs $\{K', K''\} \in \binom{\mathcal{C}_i}{2}$ with $|\{K', K''\} \cap \mathcal{C}_{\{v\},k_i,i}| \geq 1$ and $|E(K') \cap E(K'')| \geq 1$. Let $\mathcal{G}$ denote that the event that each edge $e \in E_i$ is contained in at most $z := \lceil \log n \rceil$ cliques in $\mathcal{K}_i$. Clearly, $\mathcal{G}$ implies that each clique $K' \in \mathcal{K}_i$ has edge-overlaps with a total of at most $\binom{k_i}{2} \cdot z$ cliques $K'' \in \mathcal{K}_i$, so that the parameter $C := 2 \cdot \binom{k_i}{2} z \leq k^2 z$ works in Lemma 41. Recalling $|D_i^{(v)}| \leq kX$, by invoking inequality (3.26) with $x := \lambda/k \geq k^{\tau-4}\mu > e^2\mu$ it follows that

$$\begin{aligned}\mathbb{P}\big(|D_i^{(v)}| \geq \lambda \text{ and } \mathcal{G} \mid \mathcal{R}_i\big) &\leq \mathbb{P}\big(X \geq \lambda/k \text{ and } \mathcal{G} \mid \mathcal{R}_i\big) \\ &\leq \exp\big(-\Theta(\lambda/(k^3 z))\big) \leq n^{-\omega(1)},\end{aligned} \tag{3.27}$$

where the last inequality uses $\lambda/(k^3 z) \geq \lambda n^{-o(1)} \gg \log n$ analogous to (3.24). With an eye on the event $\mathcal{G}$, note that conditional on $\mathcal{R}_i$ we have $|\mathcal{C}_{e,k_i,i}|q_i \le k^{-\tau} \le 1$ for each edge $e \in E_i$. Recalling $z = \lceil \log n \rceil$, by taking a union bound over all edges $e \in E_i$ it now is routine to see that

$$\mathbb{P}(\neg \mathcal{G} \mid \mathcal{R}_i) \le \sum_{e \in E_i} \binom{|\mathcal{C}_{e,k_i,i}|}{z} q_i^z \le |E_i| \cdot \Big(|\mathcal{C}_{e,k_i,i}| q_i e/z\Big)^z \le n^2 \cdot (e/z)^z \le n^{-\omega(1)}. \quad (3.28)$$

To sum up, by combining the above inequalities (3.22), (3.24), and (3.27)–(3.28) for $0 \le i \le I-1$ with the degree estimate $|E_I^{(v)}| = |\mathcal{C}_{\{v\},k_I,I}| = (1 \pm \epsilon)(n-1)p_I$ from (3.16), using $I = n^{o(1)}$ and Theorem 39 it follows (by a standard union bound argument) that the chromatic index (3.19) of $\mathcal{P}$ is whp at most

$$\chi'(\mathcal{P}) \le \sum_{0 \le i \le I-1} \left(\frac{(1+2\delta)2np_i}{k_i k^\tau} + \frac{3np_i}{k_i^2 k^\tau} \right) \; + \; (1+\epsilon)np_I + n^{o(1)}, \qquad (3.29)$$

where the $k_i^2 > k_i$ term will be useful in Section 3.2.2.2. Let $\pi := \log(1/p)$ and $f(x) := e^{-x}(1 + x/\pi)$. Using $p_i = p \cdot e^{-i/k^\tau}$ and $k_i \ge \sigma \log_{1/p}(n)/(1 + i/(k^\tau \pi))$ as well as $n^\sigma \ll n^{1-\sigma} \le np_I \le np/k^\tau$, it follows that

$$\chi'(\mathcal{P}) \le \frac{(5+4\delta)np}{\sigma \log_{1/p} n} \sum_{0 \le i \le I-1} \frac{f(i/k^\tau)}{k^\tau} \; + \; \frac{3np}{\big(\sigma \log_{1/p} n\big) k^{\tau-1}}. \qquad (3.30)$$

On $[0, \infty)$ the function $f(x)$ first increases and then decreases, with a maximum at $x^* := \max\{0, 1-\pi\}$. By comparing the sum with an integral, it then is standard to see that

$$\sum_{0 \le i \le I-1} \frac{f(i/k^\tau)}{k^\tau} \le \int_0^\infty f(x)dx + 2f(x^*)/k^\tau \; \le \; 1 + O\big(\pi^{-1} + k^{-\tau}\big). \qquad (3.31)$$

Combining inequalities (3.30)–(3.31) with the definition (3.20) of δ, after noting $\pi \ge \log(1/(1-\gamma)) > 0$ and $\min\{k^\tau, k^{\tau-1}\} > 1$ it follows that there is a constant $C = C(\sigma, \gamma) > 0$ such that whp $\chi'(\mathcal{P}) \le Cnp/\log_{1/p} n$.

3.2.2.2 Size of $\mathcal{P}$

It remains to bound the size of the clique partition $\mathcal{P}$, which by construction is at most

$$|\mathcal{P}| \leq \sum_{0 \leq i \leq I-1} \Big(|\mathcal{K}_i| + |D_i| + |S_i|\Big) + |E_I|. \tag{3.32}$$

Rather than estimating each of these terms (which is conceptually straightforward), we shall instead reuse known estimates from Section 3.2.2.1. A routine double-counting argument gives $|\mathcal{K}_i| \cdot k_i \leq \sum_{K \in \mathcal{K}_i} |K| = \sum_{v \in [n]} |\mathcal{K}_i^{(v)}| \leq n \cdot \chi'(\mathcal{K}_i)$. Recalling that D_i, S_i, E_I are simply sets of edges, it follows that

$$|\mathcal{P}| \leq \sum_{0 \leq i \leq I-1} \Big(n/k_i \cdot \chi'(\mathcal{K}_i) + n \cdot \max_{v \in [n]} \big|D_i^{(v)}\big| + n \cdot \max_{v \in [n]} \big|S_i^{(v)}\big|\Big) + n \cdot \max_{v \in [n]} \big|E_I^{(v)}\big|. \tag{3.33}$$

After comparing the above upper bound for $|\mathcal{P}|$ with (3.19), we see that the proof of (3.29) implies the following estimate: the size (3.32) of $\mathcal{P}$ is whp at most

$$|\mathcal{P}| \leq \sum_{0 \leq i \leq I-1} \left(\frac{(1+2\delta)2n^2 p_i}{k_i^2 k^\tau} + \frac{3n^2 p_i}{k_i^2 k^\tau}\right) \; + \; (1+\epsilon) n^2 p_I. \tag{3.34}$$

Recalling $\pi = \log(1/p)$, set $g(x) := e^{-x}(1 + x/\pi)^2$. Proceeding similarly to (3.29)–(3.31), using $\int_0^\infty g(x)dx = 1 + O(\pi^{-1} + \pi^{-2})$ it follows that there is a constant $B = B(\sigma, \gamma) > 0$ such that whp

$$|\mathcal{P}| \leq \frac{(5+4\delta)n^2 p}{\big(\sigma \log_{1/p} n\big)^2} \sum_{0 \leq i \leq I-1} \frac{g(i/k^\tau)}{k^\tau} \; + \; \frac{2n^2 p}{\big(\sigma \log_{1/p} n\big)^2 k^{\tau-2}} \; \leq \; B \frac{n^2 p}{(\log_{1/p} n)^2}, \tag{3.35}$$

which completes the proof Theorem 37 (modulo the deferred proof of Theorem 39). $\square$

3.2.3 Pseudo-randomness of the graphs G_i: proof of Theorem 39

In this section we give the deferred proof of Theorem 39. For technical reasons, we will establish concentration of the $|\mathcal{C}_{U,j,i}|$ variables in a somewhat indirect way, by focusing on auxiliary random variables that are more amenable to concentration inequalities. Turning to the details, for any vertex-subset $U \subseteq [n]$ we define

$$N_{U,i} := \Big|\big\{w \in [n] \setminus U : \ U \times \{w\} \subseteq E_i\big\}\Big|. \tag{3.36}$$

In words, $N_{U,i}$ denotes the number of common neighbors of U in $G_i = ([n], E_i)$. Recalling that G_i heuristically resembles the random graph G_{n,p_i}, we expect that $N_{U,i} \approx (n-|U|)p_i^{|U|}$; so to avoid clutter we set

$$\lambda_{s,i} := (n-s)p_i^s. \tag{3.37}$$

The following pseudo-random result establishes Theorem 39 by confirming this heuristic prediction. Our proof of Theorem 42 exploits the technical definition of $E_{i+1} = E_i \setminus \big(E(\mathcal{K}_i) \cup S_i\big)$: the extra 'stabilization' set S_i will intuitively ensure that edges of E_i remain in E_{i+1} with roughly the correct probability, see (3.44)–(3.45).

Theorem 42 (Strengthening of Theorem 39)**.** *Let $p = p(n)$ satisfy $n^{-\sigma/\tau} \le p \le 1-\gamma$, where $\gamma \in (0,1)$ is a constant. Then, with probability at least $1 - n^{-\omega(1)}$, for all $0 \le i \le I$ the following event $\mathcal{N}_i$ holds: for all $U \subseteq [n]$ with $0 \le |U| \le k_i - 1$,*

$$N_{U,i} \;=\; \big(1 \pm (i+1)\epsilon^2\big) \cdot \lambda_{|U|,i}. \tag{3.38}$$

Furthermore, $\mathcal{N}_i$ implies the event $\mathcal{R}_i$ from Theorem 39 for $0 \le i \le I$ and $n \ge n_0(\sigma,\tau)$.

Proof. Noting $kI\epsilon^2 \le n^{o(1)-\sigma}\epsilon \ll \epsilon$ it is routine to see that $\mathcal{N}_i$ implies $\mathcal{R}_i$, but we include the proof for completeness. Fixing $U \subseteq [n]$ and $0 \le i \le I$ with $|U| \le j \le k_i$, we shall double-count the number of vertex-sequences $x_{|U|+1}, \ldots, x_j \in [n] \setminus U$ with the property that $U \cup \{x_{|U|+1}, \ldots, x_j\} \in \mathcal{C}_{U,j,i}$. Using (3.38) to sequentially estimate the number of common neighbors of $U \cup \{x_{|U|+1}, \ldots, x_s\}$, noting $j \cdot I\epsilon^2 \le kI\epsilon^2 \ll \epsilon$ it follows that

$$(j-|U|)! \cdot |\mathcal{C}_{U,j,i}| = \prod_{|U| \le s \le j-1} \Big(\big(1 + O(I\epsilon^2)\big) \cdot (n-s)p_i^s\Big) = (1+o(\epsilon)) \cdot \mu_{|U|,j,i} \cdot (j-|U|)!,$$

which readily gives (3.16) for $n \geq n_0(\sigma, \tau)$, establishing the claim that $\mathcal{N}_i$ implies $\mathcal{R}_i$. With this implication and $I = n^{o(1)}$ in mind, the below auxiliary Lemmas 43–44 then complete the proof of Theorem 42. $\square$

Lemma 43. *We have* $\mathbb{P}(\neg \mathcal{N}_0) \leq n^{-\omega(1)}$.

Lemma 44. *We have* $\mathbb{P}(\neg \mathcal{N}_{i+1} \mid \mathcal{N}_i) \leq n^{-\omega(1)}$ *for all* $0 \leq i < I$.

Proof of Lemma 43. Fix $U \subseteq [n]$ with $|U| \leq k-1$, where $k = k_0$. Note that $N_{U,0}$ has a Binomial distribution with $\mathbb{E}\, N_{U,0} = (n-|U|)p^{|U|} = \lambda_{|U|,0}$, where $p = p_0$. Since $\epsilon^4 \lambda_{|U|,0} = \Theta\big(n^{1-4\sigma} p_0^{|U|}\big) \geq \Omega(n^{1-5\sigma}) \gg k \log n$ by inequality (3.17) and choice of σ, standard Chernoff bounds (such as [60, Theorem 2.1]) imply that

$$\mathbb{P}\big(|N_{U,0} - \lambda_{|U|,0}| \geq \epsilon^2 \lambda_{|U|,0}|\big) \leq 2 \cdot \exp\big(-\Theta\left(\epsilon^4 \lambda_{|U|,0}\right)\big) \leq n^{-\omega(k)}, \tag{3.39}$$

which completes the proof by taking a union bound over all $n^{O(k)}$ choices of the sets U. $\square$

Conditioning on the event $\mathcal{N}_i$, in the proof of Lemma 44 we shall estimate $N_{U,i+1}$ using the following bounded differences inequality for Bernoulli variables, see [117, Corollary 1.4] and [82, Theorem 3.8].

Lemma 45. *Let* $(\xi_\alpha)_{\alpha \in \mathcal{I}}$ *be a finite family of independent random variables with* $\xi_\alpha \in \{0,1\}$. *Let* $f : \{0,1\}^{|\mathcal{I}|} \to \mathbb{R}$ *be a function, and assume that there exist numbers* $(c_\alpha)_{\alpha \in \mathcal{I}}$ *such that the following holds for all* $z = (z_\alpha)_{\alpha \in \mathcal{I}} \in \{0,1\}^{|\mathcal{I}|}$ *and* $z' = (z'_\alpha)_{\alpha \in \mathcal{I}} \in \{0,1\}^{|\mathcal{I}|}$*:* $|f(z) - f(z')| \leq c_\beta$ *if* $z_\alpha = z'_\alpha$ *for all* $\alpha \neq \beta$. *Define* $X := f\big((\xi_\alpha)_{\alpha \in \mathcal{I}}\big)$, $V := \sum_{\alpha \in \mathcal{I}} c_\alpha^2\, \mathbb{P}(\xi_\alpha = 1)$*, and* $C := \max_{\alpha \in \mathcal{I}} c_\alpha$. *Then, for all* $t \geq 0$,

$$\mathbb{P}(|X - \mathbb{E}\, X| \geq t) \;\leq\; 2 \cdot \exp\bigg(-\frac{t^2}{2(V + Ct)}\bigg). \tag{3.40}$$

Proof of Lemma 44. To avoid clutter, we henceforth omit the conditioning on $\mathcal{N}_i$ from our notation. Fix $U \subseteq [n]$ with $|U| \leq k_i - 1$. Gearing up to apply Lemma 45 to $N_{U,i+1}$, note that the associated parameter V is given by

$$V = \sum_{K \in \mathcal{C}_i} c_K^2 \cdot q_i + \sum_{e \in E_i} \hat{c}_e^2 \cdot \zeta_{e,i}, \tag{3.41}$$

where c_K is an upper bound on how much $N_{U,i+1}$ can change if we alter whether the clique K is in $\mathcal{K}_i$ or not, and $\hat{c}_e$ is an upper bound on how much $N_{U,i+1}$ can change if we alter whether the edge e is in S_i or not. To estimate c_K and $\hat{c}_e$, note that any edge in $U \times \{w\}$ uniquely determines w. By definition (3.36) of $N_{U,i+1}$, it follows that $\hat{c}_e \le 1$ and $c_K \le \binom{k_i}{2} \le k^2$, say. In addition, the number of edges $e \in E_i$ with $\hat{c}_e \neq 0$ is at most $N_{U,i} \cdot |U|$. Similarly, the number of cliques $K \in \mathcal{C}_i$ with $c_K \neq 0$ is at most $N_{U,i} \cdot |U| \cdot \max_e |\mathcal{C}_{e,k_i,i}| \le N_{U,i}|U| \cdot (1+\epsilon)\mu_{2,k_i,i}$, where we used that $\mathcal{N}_i$ implies $\mathcal{R}_i$ (as established above) to bound $|\mathcal{C}_{e,k_i,i}|$ via (3.16). Since $(1+\epsilon)\mu_{2,k_i,i} \cdot q_i = k^{-\tau}$ by definition (3.15) of q_i, and $\zeta_{e,i} \ll k^{-\tau}$ by the calculation above (3.23), using $|U| \le k$ and $\tau \ge 5$ we infer that

$$V \le N_{U,i}|U| \cdot k^{-\tau} \cdot k^4 + N_{U,i}|U| \cdot k^{-\tau} \le 2N_{U,i} \le 4\lambda_{|U|,i} = \Theta(\lambda_{|U|,i+1}),$$

where we used (3.38) and $i\epsilon^2 \le I\epsilon^2 \le n^{o(1)-2\sigma} \ll 1$ to bound $N_{U,i}$. Invoking inequality (3.40) of Lemma 45 with $C = k^2$, noting $C\epsilon^2 \le n^{o(1)-2\sigma} \ll 1$ it follows that

$$\mathbb{P}\big(|N_{U,i+1} - \mathbb{E}\, N_{U,i+1}| \ge 0.5\epsilon^2\lambda_{|U|,i+1}\big) \le 2 \cdot \exp\big(-\Theta(\epsilon^4\lambda_{|U|,i+1})\big) \le n^{-\omega(k)},$$

where the last estimate is analogous to (3.39). To complete the proof it thus suffices to show that

$$\big|\mathbb{E}\, N_{U,i+1} - \lambda_{|U|,i+1}\big| \le (i+1.5)\epsilon^2\lambda_{|U|,i+1}. \tag{3.42}$$

Indeed, $\mathbb{P}(\neg\mathcal{N}_{i+1} \mid \mathcal{N}_i) \le n^{-\omega(1)}$ then follows by taking a union bound over all $n^{O(k)}$ sets U.

Turning to the remaining proof of (3.42), note that by construction

$$\mathbb{E}\, N_{U,i+1} = \sum_{\substack{w \in V \setminus U: \\ U \times \{w\} \subseteq E_i}} \mathbb{P}\big(U \times \{w\} \subseteq E_{i+1}\big). \tag{3.43}$$

Let us henceforth tacitly assume $U \times \{w\} \subseteq E_i$. Since $\mathcal{N}_i$ implies $\mathcal{R}_i$ we obtain $(1+\epsilon)\mu_{2,k_i,i} \geq |\mathcal{C}_{e,k_i,i}|$ via (3.16), so recalling $E_{i+1} = E_i \setminus \big(E(\mathcal{K}_i) \cup S_i\big)$ and the definition (3.15) of $\zeta_{e,i}$ it follows that

$$\begin{aligned} &\mathbb{P}\left(U \times \{w\} \subseteq E_{i+1}\right) \\ =&(1-q_i)^{|\bigcup_{e \in U \times \{w\}} \mathcal{C}_{e,k_i,i}|} \cdot \prod_{e \in U \times \{w\}} (1-\zeta_{e,i}) \\ =&(1-q_i)^{|\bigcup_{e \in U \times \{w\}} \mathcal{C}_{e,k_i,i}| - \sum_{e \in U \times \{w\}} |\mathcal{C}_{e,k_i,i}|} \cdot (1-q_i)^{|U|(1+\epsilon)\mu_{2,k_i,i}}. \end{aligned} \tag{3.44}$$

Recalling the definition (3.15) of q_i, using estimates (3.16)–(3.17) we infer that

$$\begin{aligned} q_i \cdot \left| \Big| \bigcup_{e \in U \times \{w\}} \mathcal{C}_{e,k_i,i} \Big| - \sum_{e \in U \times \{w\}} |\mathcal{C}_{e,k_i,i}| \right| &\leq q_i \sum_{u \neq v \in U} |\mathcal{C}_{\{u,v,w\},k_i,i}| \leq k^2 \mu_{3,k_i,i} / \mu_{2,k_i,i} \\ &\leq k^2 / \big(\Omega(n/k_i) p_i^2\big) \leq n^{-1+\sigma+o(1)} \ll n^{-2\sigma} = \epsilon^2. \end{aligned}$$

We similarly obtain $q_i \ll \epsilon^2$ and $q_i \cdot |U|(1+\epsilon)\mu_{2,k_i,i} = |U|/k^\tau \leq 1$. Inserting $1 - q_i = e^{-(1+O(q_i))q_i}$ into (3.44), using $e^{o(\epsilon^2)} = 1 + o(\epsilon^2)$ it routinely follows that

$$\mathbb{P}\left(U \times \{w\} \subseteq E_{i+1}\right) = \big(1 + o(\epsilon^2)\big) \cdot e^{-|U|/k^\tau}. \tag{3.45}$$

Recalling (3.43) and $(i+1)\epsilon^2 \leq I\epsilon^2 \ll \epsilon \ll 1$, using (3.38) and $\lambda_{|U|,i} \cdot e^{-|U|/k^\tau} = \lambda_{|U|,i+1}$ we infer that

$$\mathbb{E}\, N_{U,i+1} = N_{U,i} \cdot \big(1 + o(\epsilon^2)\big) e^{-|U|/k^\tau} = \Big(1 \pm \big(i + 1 + o(1)\big)\epsilon^2\Big) \cdot \lambda_{|U|,i+1},$$

which establishes (3.42) with room to spare, completing the proof of Lemma 44. $\square$

3.3 Random greedy edge coloring algorithm

In this section we prove Theorem 32 by showing that the following simple random greedy algorithm is likely to produce the desired proper edge coloring of the random edges from the hypergraph $\mathcal{H}$ (allowing for repeated edges), using the colors $[q] = \{1, \ldots, q\}$ for suitable $q \geq 1$. For $i \geq 0$, we sequentially choose an edge $e_{i+1} \in E(\mathcal{H})$ uniformly at random, and then assign e_{i+1} a color c chosen uniformly at random from all colors in $[q]$ that are still available at e_{i+1}, i.e., which have not been assigned to an edge e_j with $e_j \cap e_{i+1} \neq \varnothing$ and $j \leq i$ (this also ensures the usage of different colors for each occurrence of the same edge). This random greedy coloring algorithm terminates when no more colors are available at some edge $e \in E(\mathcal{H})$.

3.3.1 Dynamic concentration of key variables: proof of Theorem 32

Our main goal is to understand the evolution of the colors available for each edge $e \in E(\mathcal{H})$, i.e., the size of $Q_e(i)$, where for any set of vertices $S \subseteq V(\mathcal{H})$ we more generally define

$$Q_S(i) := \bigl\{c \in [q] : \text{color } c \text{ not assigned to any edge } f \in \{e_j : 1 \leq j \leq i\} \text{ with } f \cap S \neq \varnothing\bigr\}. \tag{3.46}$$

At the beginning of the algorithm we have $|Q_e(0)| = q$. In order to keep track of the number of available colors $|Q_e(i)|$, we need to understand changes in the colors assigned to edges adjacent to the vertices of e. To take such changes into account, for all vertices $v \in V(\mathcal{H})$ and colors $c \in [q]$ we introduce

$$Y_{v,c}(i) := \bigl\{f \in E(\mathcal{H}) : \ v \in f \text{ and } c \in Q_{f \setminus \{v\}}(i)\bigr\}, \tag{3.47}$$

which in case of $c \in Q_{\{v\}}(i)$ denotes the set of all edges adjacent to v that could still be colored by c (since for any $f \in Y_{v,c}(i)$ then $c \in Q_{f \setminus \{v\}}(i) \cap Q_{\{v\}}(i) = Q_f(i)$ holds). Note that initially $|Y_{v,c}(0)| = \deg_{\mathcal{H}}(v)$.

Our main technical result for the random greedy algorithm shows that, when $q \approx rm/n$ colors are used, then the above-mentioned key random variables closely follow the trajectories $|Q_e(i)| \approx \hat{q}(t)$ and $|Y_{v,c}(i)| \approx \hat{y}(t)$ during the first $m_0 \approx (1-\gamma)m$ steps, tacitly using the continuous time scaling

$$t = t(i, m) := i/m. \tag{3.48}$$

In particular, $\min_{e \in E(\mathcal{H})} |Q_e(m_0)| > 0$ ensures that the algorithm properly colors the first m_0 edges using at most q colors, as no edge has run out of available colors. The form of the trajectories (3.50)–(3.52) can easily be predicted via modern (pseudo-random or expected one-step changes based) heuristics, see Appendix 3.7.

Theorem 46 (Dynamic concentration of the variables)**.** *For all reals $\gamma \in (0,1)$ and $\sigma, b > 0$ with*

$$b \log(1/\gamma) \leq \sigma/30 \tag{3.49}$$

there is $n_0 = n_0(\sigma, b) > 0$ such that, for all integers $n \geq n_0$, $2 \leq r \leq b \log n$ and all reals $n^{1+\sigma} \leq m \leq n^{rn^{\sigma/4}}$, $D > 0$, the following holds for every n-vertex r-uniform hypergraph $\mathcal{H}$ satisfying the degree and codegree assumptions (3.2)*. With probability at least $1 - m^{-\omega(r)}$, we have $\min_{e \in E(\mathcal{H})} |Q_e(i)| > 0$ and*

$$|Q_e(i)| = \big(1 \pm \hat{e}(t)\big) \cdot \hat{q}(t) \text{ for all } e \in E(\mathcal{H}), \tag{3.50}$$

$$|Y_{v,c}(i)| = \big(1 \pm \hat{e}(t)\big) \cdot \hat{y}(t) \text{ for all } v \in V(\mathcal{H}) \text{ and } c \in [q], \tag{3.51}$$

for all $0 \leq i \leq m_0 := \lfloor (1-\gamma)m \rfloor$, where $q := \lfloor rm/n \rfloor$ and

$$\hat{q}(s) := (1-s)^r q, \qquad \hat{y}(s) := (1-s)^{r-1} D \quad \text{and} \quad \hat{e}(s) := (1-s)^{-9r} n^{-\sigma/3}. \tag{3.52}$$

Remark 47. *The assumption* (3.49) *simply ensures that $\hat{e}(t) = (1-t)^{-9r} n^{-\sigma/3} \leq n^{9b\log(1/\gamma) - \sigma/3} \leq n^{-\sigma/30} = o(1)$ for all $0 \leq i \leq m_0$, so that estimates* (3.50)–(3.51) *imply $|Q_e(i)| \sim \hat{q}(t)$*

and $|Y_{v,c}(i)| \sim \hat{y}(t)$.

Remark 48. *The proof carries over to the case* $\gamma = \gamma(n) \to 0$, *provided that the assumption* (3.49) *is replaced by* $r \log(1/\gamma)/\log n \le \sigma/30$ *(to again ensure that* $\hat{e}(t) \le n^{-\sigma/30} = o(1)$ *holds).*

Before giving the differential equation method based proof of this result, we first show how it implies Theorem 32 by slightly increasing the number of edges from m to m', to ensure that the greedy algorithm properly colors the first $\lfloor (1-\gamma)m' \rfloor \ge m$ random edges using at most $\lfloor rm'/n \rfloor \le (1+\epsilon)rm/n$ colors.

Proof of Theorem 32. Set $\gamma := 1 - 1/(1+\delta)$, so that $b \log(1/\gamma) = b \log(1 + 1/\delta) \le b/\delta \le \sigma/30$ implies (3.49). Invoking Theorem 46 with m set to $m' := (1+\delta)m = o(n^{rn^{\sigma/4}})$ it follows that, with probability at least $1 - m^{-\omega(r)}$, the greedy algorithm properly colors the first $m_0 := \lfloor (1-\gamma)m' \rfloor = \lfloor m \rfloor = m$ random edges $e_1, \ldots, e_m$ using at most $q := \lfloor rm'/n \rfloor \le (1+\delta)rm/n$ colors, completing the proof. $\square$

3.3.2 Differential equation method: proof of Theorem 46

In this subsection we prove Theorem 46 by showing $\mathbb{P}(\neg \mathcal{G}_{m_0}) \le m^{-\omega(r)}$, where $\mathcal{G}_j$ denotes the event that $\min_{e \in E(\mathcal{H})} |Q_e(i)| > 0$ and estimates (3.50)–(3.51) hold for all $0 \le i \le j$. We henceforth tacitly assume $0 \le i \le m_0$, and also that $n \ge n_0(\sigma, b)$ is sufficiently large (whenever necessary). In particular, estimate (3.50) implies $\min_{e \in E(\mathcal{H})} |Q_e(i)| \ge \hat{q}(t)/2 > 0$ by Remark 47. To establish (3.50)–(3.51) using the differential equation method approach to dynamic concentration, we introduce the following sequences of auxiliary random variables:

$$Q_e^{\pm}(i) := \pm\big[|Q_e(i)| - \hat{q}(t)\big] - \hat{e}(t)\hat{q}(t) \text{ for all } e \in E(\mathcal{H}), \tag{3.53}$$

$$Y_{v,c}^{\pm}(i) := \pm\big[|Y_{v,c}(i)| - \hat{y}(t)\big] - \hat{e}(t)\hat{y}(t) \text{ for all } v \in V(\mathcal{H}) \text{ and } c \in [q]. \tag{3.54}$$

Note that the desired estimates (3.50)–(3.51) follow when the four inequalities $Q_e^{\pm}(i) \leq 0$ and $Y_{v,c}^{\pm}(i) \leq 0$ all hold. To establish these inequalities, in Section 3.3.2.1 we first estimate the expected one-step changes of $|Q_e(i)|$ and $|Y_{v,c}(i)|$, which in Section 3.3.2.2 then enables us to show that the sequences $Q_e^{\pm}(i)$ and $Y_{v,c}^{\pm}(i)$ are supermartingales. Next, in Section 3.3.2.3 we bound the one-step changes of the variables, which in Section 3.3.2.4 then enables us to invoke a supermartingale inequality (that is optimized for the differential equation method, see Lemma 49) in order to show that $Q_e^{\pm}(i) \geq 0$ or $Y_{v,c}^{\pm}(i) \geq 0$ are extremely unlikely events.

3.3.2.1 Expected one-step changes

We first derive estimates for the expected one-step changes of the available colors variables $|Q_e(i)|$ and the available edges variables $|Y_{v,c}(i)|$, tacitly assuming that $0 \leq i \leq m_0$ and $\mathcal{G}_i$ hold. As we shall see, the expected changes (3.57) and (3.59) will be consistent with the deterministic approximations $|Q_e(i+1)| - |Q_e(i)| \approx \hat{q}(t+1/m) - \hat{q}(t) \approx \hat{q}'(t)/m = -r(1-t)^{r-1}q/m$ and $|Y_{v,c}(i+1)| - |Y_{v,c}(i)| \approx \hat{y}'(t)/m = -(r-1)(1-t)^{r-2}D/m$, which is one motivation for the choice of $\hat{q}(t)$ and $\hat{y}(t)$; see also (3.82)–(3.84) in Appendix 3.7.

To calculate the expectation of the one-step changes $\Delta Q_e(i) := |Q_e(i+1)| - |Q_e(i)|$, we consider a color $c \in Q_e(i)$ and the event that $c \notin Q_e(i+1)$. By definition (3.46) of $Q_e(i)$ this only occurs if the algorithm chooses an edge f with $f \cap e \neq \varnothing$, and then assigns the color c to f. By definition (3.47) of $Y_{v,c}(i)$ this color assignment is only possible if $f \in \bigcup_{v \in e} Y_{v,c}(i)$, as $c \in Q_e(i) \subseteq Q_{\{v\}}(i)$ for any $v \in e$. Since the algorithm chooses both the edge $e_{i+1} \in E(\mathcal{H})$ and the color $c \in Q_{e_{i+1}}(i)$ uniformly at random, it follows that

$$\mathbb{E}(\Delta Q_e(i) \mid \mathcal{F}_i) = -\sum_{c \in Q_e(i)} \sum_{f \in \bigcup_{v \in e} Y_{v,c}(i)} \frac{1}{|E(\mathcal{H})| \cdot |Q_f(i)|}, \tag{3.55}$$

where $\mathcal{F}_i$ denotes, as usual, the natural filtration associated with the algorithm after i steps (which intuitively keeps track of the history algorithm, i.e., contains all the information available up to step i). Recalling the codegree assumption (3.2) and $r = O(\log n)$, note that the cardinality of the union $\bigcup_{v\in e} Y_{v,c}(i)$ differs from the sum $\sum_{v\in e} |Y_{v,c}(i)|$ by at most $\sum_{v\neq w\in e} \deg_{\mathcal{H}}(v,w) < n^{-\sigma/2}D < \hat{e}(t)\hat{y}(t)$. The degree assumption (3.2) also implies $r\cdot|E(\mathcal{H})| = \sum_{v\in V(H)} \deg_{\mathcal{H}}(v) = n\cdot(1\pm n^{-\sigma})D$. Using estimates (3.50)–(3.51), it follows that

$$\mathbb{E}(\Delta Q_e(i) \mid \mathcal{F}_i) = -\frac{(1\pm\hat{e})\hat{q}\cdot r\cdot(1\pm 2\hat{e})\hat{y}}{(1\pm n^{-\sigma})nD/r\cdot(1\pm\hat{e})\hat{q}}, \tag{3.56}$$

where we suppressed the dependence on t to avoid clutter in the notation. Noting $|rm/n - q| \leq 1 < n^{-\sigma}q$ and $n^{-\sigma} < \hat{e}(t) = o(1)$, using $\hat{y}(t) = (1-t)^{r-1}D$ we routinely arrive at

$$\mathbb{E}(\Delta Q_e(i) \mid \mathcal{F}_i) = -\big(1\pm 7\hat{e}(t)\big)\cdot r(1-t)^{r-1}q/m. \tag{3.57}$$

To calculate the expectation of the one-step changes $\Delta Y_{v,c}(i) := |Y_{v,c}(i+1)| - |Y_{v,c}(i)|$, we consider an edge $f\in Y_{v,c}(i)$ and the event that $f\notin Y_{v,c}(i+1)$. By definition (3.47) of $Y_{v,c}(i)$ this only occurs if the algorithm chooses an edge e with $e\cap(f\setminus\{v\})\neq\varnothing$, and then assigns the color c to e, which in turn is only possible if $e\in\bigcup_{w\in f\setminus\{v\}} Y_{w,c}(i)$. Proceeding similarly to (3.55), it follows that

$$\mathbb{E}(\Delta Y_{v,c}(i) \mid \mathcal{F}_i) = -\sum_{f\in Y_{v,c}(i)} \sum_{e\in\bigcup_{w\in f\setminus\{v\}} Y_{w,c}(i)} \frac{1}{|E(\mathcal{H})|\cdot|Q_e(i)|}, \tag{3.58}$$

where $|\bigcup_{w\in f\setminus\{v\}} Y_{w,c}(i)|$ differs from $\sum_{w\in f\setminus\{v\}} |Y_{w,c}(i)|$ by at most $\sum_{u\neq w\in f} \deg_{\mathcal{H}}(u,w) < \hat{e}(t)\hat{y}(t)$. Proceeding similarly to (3.56)–(3.57), using $|q - rm/n| \leq 1 < n^{-\sigma}rm/n$ and $n^{-\sigma} < \hat{e}(t) = o(1)$ it follows that

$$\begin{aligned}\mathbb{E}(\Delta Y_{v,c}(i) \mid \mathcal{F}_i) &= -\frac{(1\pm\hat{e})\hat{y}\cdot(r-1)\cdot(1\pm 2\hat{e})\hat{y}}{(1\pm n^{-\sigma})nD/r\cdot(1\pm\hat{e})\hat{q}} \\ &= -\big(1\pm 7\hat{e}(t)\big)\cdot(r-1)(1-t)^{r-2}D/m.\end{aligned} \tag{3.59}$$

3.3.2.2 Supermartingale conditions

We now show that the expected one-step changes of the auxiliary variables $Q_e^{\pm}(i)$ and $Y_{v,c}^{\pm}(i)$ are negative (as required for supermartingales), tacitly assuming that $0 \le i \le m_0 - 1$ and $\mathcal{G}_i$ hold. As we shall see, the main terms in the expected changes (3.60) and (3.63) will cancel due to the estimates of Section 3.3.2.1, and the careful choice of $\hat{e}(t)$ then ensures that the resulting expected changes (3.62) and (3.64) are indeed negative (by ensuring that the ratios $e'_X(t)/e_X(t)$ of the below-defined error functions $e_X(t)$ are sufficiently large).

For the one-step changes $\Delta Q_e^{\pm}(i) := Q_e^{\pm}(i+1) - Q_e^{\pm}(i)$, set $e_Q(s) := \hat{e}(s)\hat{q}(s) = (1-s)^{-8r}n^{-\sigma/3}q$. Recalling $t = i/m$, by applying Taylor's theorem with remainder it follows that

$$\begin{aligned}\mathbb{E}(\Delta Q_e^{\pm}(i) \mid \mathcal{F}_i) &= \pm\Big[\mathbb{E}(\Delta Q_e(i) \mid \mathcal{F}_i) - \Big[\hat{q}\big(t+1/m\big) - \hat{q}(t)\Big]\Big] - \Big[e_Q\big(t+1/m\big) - e_Q(t)\Big] \\ &= \pm\Big[\mathbb{E}(\Delta Q_e(i) \mid \mathcal{F}_i) - \frac{\hat{q}'(t)}{m}\Big] - \frac{e'_Q(t)}{m} + O\bigg(\max_{s\in[0,m_0/m]} \frac{|\hat{q}''(s)| + |e''_Q(s)|}{m^2}\bigg).\end{aligned} \tag{3.60}$$

The key point is that the derivative $\hat{q}'(t)/m = -r(1-t)^{r-1}q/m$ equals the main term in (3.57), and that the other term in (3.57) satisfies $7\hat{e}(t) \cdot r(1-t)^{r-1}q/m = 7r(1-t)^{-1}e_Q(t)/m$. Furthermore, using the estimate from Remark 47 together with $m \ge n^{1+\sigma}$ and $r = O(\log n)$, for all $s \in [0, m_0/m]$ it is routine to see that

$$\frac{|\hat{q}''(s)| + |e''_Q(s)|}{m} \le O\bigg(\frac{r^2 q + r^2(1-s)^{-8r-2}n^{-\sigma/3}q}{m}\bigg) = o(n^{-\sigma/3}q). \tag{3.61}$$

Putting things together, now the crux is that $e'_Q(t) = 8r(1-t)^{-1}e_Q(t) = \Omega(n^{-\sigma/3}q)$ implies

$$\mathbb{E}(\Delta Q_e^{\pm}(i) \mid \mathcal{F}_i) \le \frac{7r(1-t)^{-1}e_Q(t) - e'_Q(t) + o(n^{-\sigma/3}q)}{m} < 0. \tag{3.62}$$

For the one-step changes $\Delta Y^{\pm}_{v,c}(i) := Y^{\pm}_{v,c}(i+1) - Y^{\pm}_{v,c}(i)$, set $e_Y(s) := \hat{e}(s)\hat{y}(s) = (1-s)^{-8r-1}n^{-\sigma/3}D$. Proceeding similarly to (3.60), we obtain

$$\begin{aligned} \mathbb{E}(\Delta Y^{\pm}_{v,c}(i) \mid \mathcal{F}_i) = \pm\biggl[\mathbb{E}(\Delta Y_{v,c}(i) \mid \mathcal{F}_i) - \frac{\hat{y}'(t)}{m}\biggr] - \frac{e'_Y(t)}{m} \\ + O\biggl(\max_{s\in[0,m_0/m]} \frac{|\hat{y}''(s)| + |e''_Y(s)|}{m^2}\biggr). \end{aligned} \tag{3.63}$$

The key point is that the derivative $\hat{y}'(t)/m = -(r-1)(1-t)^{r-2}D/m$ equals the main term in (3.59), and that the other term in (3.59) satisfies $7\hat{e}(t) \cdot (r-1)(1-t)^{r-2}D/m = 7(r-1)(1-t)^{-1}e_Y(t)/m$. Analogously to (3.61), it is routine to see that $|\hat{y}''(s)|+|e''_Y(s)| = o(n^{-\sigma/3}Dm)$ for all $s \in [0, m_0/m]$. Putting things together similarly to (3.62), here the crux is that $e'_Y(t) = (8r+1)(1-t)^{-1}e_Y(t) = \Omega(n^{-\sigma/3}D)$ implies

$$\mathbb{E}(\Delta Y^{\pm}_{v,c}(i) \mid \mathcal{F}_i) \le \frac{7(r-1)(1-t)^{-1}e_Y(t) - e'_Y(t) + o(n^{-\sigma/3}D)}{m} < 0. \tag{3.64}$$

3.3.2.3 Bounds on one-step changes

We next derive bounds on the one-step changes of the variables $|Q_e(i)|$ and $|Y_{v,c}(i)|$ (as required by the supermartingale inequality in Section 3.3.2.4), tacitly assuming that $0 \le i \le m_0$ and $\mathcal{G}_i$ hold. As we shall see, the expected changes (3.66) and (3.68) are easy to bound due to step-wise monotonicity of the variables.

The one-step changes $\Delta Q_e(i) = |Q_e(i+1)| - |Q_e(i)|$ of the available colors satisfy

$$|\Delta Q_e(i)| \le 1. \tag{3.65}$$

Since $|Q_e(i)|$ is step-wise decreasing, by inserting $\hat{e}(t) = o(1)$ and $r(1-t)^{r-1} \le r$ into (3.57) we obtain

$$\mathbb{E}(|\Delta Q_e(i)| \mid \mathcal{F}_i) = -\,\mathbb{E}(\Delta Q_e(i) \mid \mathcal{F}_i) \le 2rq/m. \tag{3.66}$$

The one-step changes $\Delta Y_{v,c}(i) = |Y_{v,c}(i+1)| - |Y_{v,c}(i)|$ of the available edges satisfy

$$|\Delta Y_{v,c}(i)| \le \sum_{w \in e_{i+1} \setminus \{v\}} \deg_{\mathcal{H}}(v,w) \le r \cdot n^{-\sigma} D \tag{3.67}$$

due to the codegree assumption (3.2). Since $|\Delta Y_{v,c}(i)|$ is step-wise decreasing, by inserting $\hat{e}(t) = o(1)$ and $(r-1)(1-t)^{r-2} \le r$ into (3.59) we also obtain

$$\mathbb{E}(|\Delta Y_{v,c}(i)| \mid \mathcal{F}_i) = -\,\mathbb{E}(\Delta Y_{v,c}(i) \mid \mathcal{F}_i) \le 2rD/m. \tag{3.68}$$

3.3.2.4 Supermartingale estimates

We finally bound $\mathbb{P}(\neg \mathcal{G}_{m_0})$ by focusing on the first step where the estimates (3.50)–(3.51) are violated, which by the discussion below (3.53)–(3.54) can only happen if $Q_e^{\pm}(i) \le 0$ or $Y_{v,c}^{\pm}(i) \le 0$ is violated. Our main tool for bounding the probabilities of these 'first bad events' will be the following Freedman type supermartingale inequality: it is optimized for the differential equation method approach to dynamic concentration, where supermartingales S_i are constructed by adding a deterministic quantity to a random variable X_i, cf. the definition of $Q_e^{\pm}(i)$ and $Y_{v,c}^{\pm}(i)$ in (3.53)–(3.54). Here the convenient point is that Lemma 49 only assumes upper bounds on the one-step changes of X_i (and not of S_i, as usual, cf. [14, Lemma 3.4]).

Lemma 49. *Let $(S_i)_{i\ge 0}$ be a supermartingale adapted to the filtration $(\mathcal{F}_i)_{i \ge 0}$. Assume that $S_i = X_i + D_i$, where X_i is $\mathcal{F}_i$-measurable and D_i is $\mathcal{F}_{\max\{i-1,0\}}$-measurable. Writing $\Delta X_i := X_{i+1} - X_i$, assume that* $\max_{i \ge 0} |\Delta X_i| \le C$ *and* $\sum_{i \ge 0} \mathbb{E}(|\Delta X_i| \mid \mathcal{F}_i) \le V$. *Then, for all $z > 0$,*

$$\mathbb{P}\big(S_i \ge S_0 + z \textit{ for some } i \ge 0\big) \;\le\; \exp\left(-\frac{z^2}{2C(V+z)}\right). \tag{3.69}$$

Proof. Writing $\Delta S_i := S_{i+1} - S_i$, set $M_i := S_i - \sum_{0 \le j < i} \mathbb{E}(\Delta S_j \mid \mathcal{F}_j)$. Note that

$S_i = X_i + D_i$ implies

$$\Delta M_i := M_{i+1} - M_i = \Delta S_i - \mathbb{E}(\Delta S_i \mid \mathcal{F}_i) = \Delta X_i - \mathbb{E}(\Delta X_i \mid \mathcal{F}_i),$$

which readily gives $\mathbb{E}(\Delta M_i \mid \mathcal{F}_i) = 0$ and $\max_{i \geq 0} |\Delta M_i| \leq 2 \cdot C$. Note that we also have

$$\mathrm{Var}(\Delta M_i \mid \mathcal{F}_i) = \mathrm{Var}(\Delta X_i \mid \mathcal{F}_i) \leq \mathbb{E}(\Delta X_i^2 \mid \mathcal{F}_i) \leq C \cdot \mathbb{E}(|\Delta X_i| \mid \mathcal{F}_i), \tag{3.70}$$

so that $\sum_{i \geq 0} \mathrm{Var}(\Delta M_i \mid \mathcal{F}_i) \leq C \cdot V$. Clearly $M_0 = S_0$. Also $M_i \geq S_i$, since $(S_i)_{i \geq 0}$ is a supermartingale. Hence a standard application of Freedman's martingale inequality (see [41] or [117, Lemma 2.2]) yields

$$\begin{aligned} \mathbb{P}\big(S_i \geq S_0 + z \text{ for some } i \geq 0\big) &\leq \mathbb{P}\big(M_i \geq M_0 + z \text{ for some } i \geq 0\big) \\ &\leq \exp\left(-\frac{z^2}{2(CV + 2C \cdot z/3)}\right), \end{aligned} \tag{3.71}$$

which completes the proof of inequality (3.69). □

Turning to the details, we define the stopping time I as the minimum of m_0 and the first step $i \geq 0$ where $\mathcal{G}_i$ fails. Writing $i \wedge I := \min\{i, I\}$, as usual, by our above discussion it follows that

$$\begin{aligned} \mathbb{P}\big(\neg \mathcal{G}_{m_0}\big) \leq &\sum_{e \in E(\mathcal{H})} \sum_{\tau \in \{+,-\}} \mathbb{P}\big(Q_e^\tau(i \wedge I) \geq 0 \text{ for some } i \geq 0\big) \\ &+ \sum_{v \in V(\mathcal{H})} \sum_{c \in [q]} \sum_{\tau \in \{+,-\}} \mathbb{P}\big(Y_{v,c}^\tau(i \wedge I) \geq 0 \text{ for some } i \geq 0\big). \end{aligned} \tag{3.72}$$

Note that initially $|Q_e(i)| = q$ and $|Y_{v,c}(0)| = \deg_{\mathcal{H}}(v)$, which in view of the degree assumption (3.2) and the definitions (3.53)–(3.54) of $Q_e^\tau(0)$ and $Y_{v,c}^\tau(0)$ gives the initial value estimates

$$
\begin{aligned}
Q^{\tau}_e(0 \wedge I) &= Q^{\tau}_e(0) = -\hat{e}(0)q = -n^{-\sigma/3}q,\\
Y^{\tau}_{v,c}(0 \wedge I) &= Y^{\tau}_{v,c}(0) = O(n^{-\sigma}D) - \hat{e}(0)D \le -n^{-\sigma/3}D/2.
\end{aligned}
$$

Noting that the estimates from Sections 3.3.2.2–3.3.2.3 apply for $0 \le i \le I-1$ (since then $0 \le i \le m_0 - 1$ and $\mathcal{G}_i$ hold), the point is that the stopped sequence $S_i := Q^{\tau}_e(i \wedge I)$ is a supermartingale with $S_0 = -n^{-\sigma/3}q$, to which Lemma 49 can be applied with $X_i = \tau|Q_e(i \wedge I)|$, $C = 1$ and $V = m_0 \cdot 2rq/m = O(rq)$. Invoking inequality (3.69) with $z = n^{-\sigma/3}q$, using $q = \Omega(rn^{\sigma})$ together with $m^r \le n^{r^2 n^{\sigma/4}}$ and $r = O(\log n)$ it follows that

$$
\begin{aligned}
\mathbb{P}\big(Q^{\tau}_e(i \wedge I) \ge 0 \text{ for some } i \ge 0\big) &\le \exp\Big\{-\Theta(n^{-2\sigma/3}q/r)\Big\}\\
&\le \exp\Big\{-\Theta(n^{\sigma/3})\Big\} \le m^{-\omega(r)}.
\end{aligned}
\tag{3.73}
$$

Similarly, the sequence $S_i := Y^{\tau}_{v,c}(i \wedge I)$ is a supermartingale with $S_0 \le -n^{-\sigma/3}D/2$, to which Lemma 49 can be applied with $X_i = \tau|Y_{v,c}(i \wedge I)|$, $C = rn^{-\sigma}D$ and $V = m_0 \cdot 2rD/m = O(rD)$. Invoking inequality (3.69) with $z = n^{-\sigma/3}D/2$, it follows analogously to (3.73) that

$$
\mathbb{P}\big(Y^{\tau}_{v,c}(i \wedge I) \ge 0 \text{ for some } i \ge 0\big) \le \exp\Big\{-\Theta(n^{\sigma/3}/r^2)\Big\} \le m^{-\omega(r)}. \tag{3.74}
$$

Inserting (3.73)–(3.74) into inequality (3.72), noting $|V(H)| = n \le m$, $|E(\mathcal{H})| \le n^r \le m^r$ and $q \le m$ it then follows that $\mathbb{P}(\neg\mathcal{G}_{m_0}) \le m^{-\omega(r)}$, which completes the proof of Theorem 46. □

3.4 Concluding remarks

The main remaining open problem is to determine the typical asymptotic behavior of the Prague dimension $\dim_{\mathrm{P}}(G_{n,p}) \approx \mathrm{cc}'(G_{n,1-p})$ as well as the clique covering and partition numbers $\mathrm{cc}(G_{n,p})$ and $\mathrm{cp}(G_{n,p})$, i.e., to refine the estimates from Theorems 31, 34 and 35. Here edge-probability $p = 1/2$ is of special interest, since this would reveal the asymptotics of these intriguing parameters for almost all n-vertex graphs.

Problem 1. *Determine the whp asymptotics of the parameters $cc(G_{n,p})$, $cp(G_{n,p})$, $cc_\Delta(G_{n,p})$, and $cc'(G_{n,p})$ for constant edge-probabilities $p \in (0,1)$.*

3.4.1 Non-trivial lower bounds for dense random graphs

For constant edge-probabilities $p \in (0,1)$ our understanding of the asymptotics remains unsatisfactory, even on a heuristic level. Indeed, it is well-known that the largest clique of $G_{n,p}$ whp has size $s \sim 2\log_{1/p} n$, which together with the simple lower bound reasoning for Theorem 34 makes it tempting to speculate that perhaps $\mathrm{cc}(G_{n,p}) \sim \binom{n}{2}p/\binom{s}{2}$ holds whp. However, Lemma 50 shows that this natural guess is false, by further improving the simple lower bound (which for $p = 1/2$ was already noted in [15]). The analogous speculation $\mathrm{cc}_\Delta(G_{n,p}) \sim np/(s-1)$ is also refuted by Lemma 50, whose proof we defer to Section 3.5.

Lemma 50. *If $p = p(n)$ satisfies $n^{-o(1)} \le p \le 1 - n^{-o(1)}$, then for any $\epsilon \in (0,1)$ whp*

$$cc(G_{n,p}) \ \ge\ (1-\epsilon)\cdot\big(1+\varphi(p)\big)\tbinom{n}{2}p/\tbinom{s}{2}, \tag{3.75}$$

$$cc_\Delta(G_{n,p}) \ \ge\ (1-\epsilon)\cdot\big(1+\varphi(p)\big)np/(s-1), \tag{3.76}$$

where $s := \lceil 2\log_{1/p} n\rceil$ and $\varphi(p) := (1-p)\log(1-p)/(p\log p)$. The function $\varphi : (0,1) \to (0,\infty)$ is increasing, with $\lim_{p\searrow 0}\varphi(p) = 0$, $\varphi(1/2) = 1$, and $\lim_{p\nearrow 1}\varphi(p) = \infty$.

For Problem 1 the main conceptual message of Lemma 50 is as follows: it simply is not enough to mainly use cliques of near maximal size, which in turn indicates that the correct asymptotics are somewhat tricky. Perhaps rashly, we speculate that the lower bounds in (3.75)–(3.76) are asymptotically best possible.

3.4.2 Asymptotics for sparse random graphs

We now record strengthenings of Theorems 34–35 for many small edge-probabilities $p = p(n) \to 0$, where the asymptotics follow from Pippenger–Spencer type hypergraph results. As we shall see, here the crux is that when all cliques have size $O(1)$, then it suffices to simply cover a $1 - o(1)$ fraction of the relevant edges.

Theorem 51. *If $p = p(n)$ satisfies $n^{-2/(s+1)} \ll p \ll n^{-2/(s+2)}$ for some fixed integer $s \geq 3$, then cc$(G_{n,p})$ and cp$(G_{n,p})$ are whp both asymptotic to $\binom{n}{2}p/\binom{s}{2}$.*

We leave it as an open problem to determine the whp asymptotics for $p = \Theta(n^{-2/(s+1)})$, and now outline the proof of Theorem 51, which uses $\mathrm{cc}(G_{n,p}) \leq \mathrm{cp}(G_{n,p})$. The lower bound on $\mathrm{cc}(G_{n,p})$ is routine: the expected number of edges in cliques of size at least $s+1$ is at most $\sum_{k \geq s+1} \binom{k}{2}\binom{n}{k}p^{\binom{k}{2}} \ll \binom{n}{2}p$, which makes it easy to see that whp $\mathrm{cc}(G_{n,p}) \geq (1 - o(1))\binom{n}{2}p/\binom{s}{2}$. For the upper bound on $\mathrm{cp}(G_{n,p})$ we shall mimic the natural strategy of Kahn and Park [63] for $s = 3$: using Kahn's fractional version of Pippenger's hypergraph packing result [63, Theorem 7.1] it is not difficult[6] to see that $G_{n,p}$ whp contains a collection $\mathcal{C}$ of $|\mathcal{C}| = (1 - o(1))\binom{n}{2}p/\binom{s}{2}$ edge-disjoint cliques K_s. Writing $\mathcal{U}$ for the edges of $G_{n,p}$ not covered by the cliques in $\mathcal{C}$, it then easily follows that whp $\mathrm{cp}(G_{n,p}) \leq |\mathcal{C}| + |\mathcal{U}| \leq (1 + o(1))\binom{n}{2}p/\binom{s}{2}$, as desired.

Theorem 52. *If $p = p(n)$ satisfies $(\log n)^{\omega(1)}n^{-2/(s+1)} \leq p \ll n^{-2/(s+2)}$ for some fixed integer $s \geq 3$, then $cc_{\Delta}(G_{n,p})$ and $cc'(G_{n,p})$ are whp both asymptotic to $np/(s-1)$.*

Remark 53. *These asymptotics remain valid when the definitions of $cc_{\Delta}(G_{n,p})$ and $cc'(G_{n,p})$ are restricted to clique partitions of the edges (instead of clique coverings).*

We leave it as an open problem to determine the whp asymptotics for $p = (\log n)^{O(1)}n^{-2/(s+1)}$, and now outline the proof of Theorem 52, which uses $\mathrm{cc}_{\Delta}(G_{n,p}) \leq$

[6] We consider the auxiliary hypergraph $\mathcal{H}$, where the vertices correspond to the edges of $G_{n,p}$ and the edges correspond to the edge-sets of the cliques K_s of $G_{n,p}$. The technical conditions of [63, Theorem 7.1] required for mimicking [63, Section 7] can then be verified using (careful applications of) standard tail bounds such as Lemma 41 and [120, Theorems 30 and 32].

$\mathrm{cc}'(G_{n,p})$. The lower bound on $\mathrm{cc}_\Delta(G_{n,p})$ is routine: the expected number of edges in cliques of size at least $s+1$ containing a fixed vertex v is at most $\sum_{k \geq s+1} \binom{k}{2}\binom{n-1}{k-1}p^{\binom{k}{2}} \ll np$, which makes it easy to see that whp $\mathrm{cc}_\Delta(G_{n,p}) \geq (1-o(1))np/(s-1)$. Turning to the upper bound on $\mathrm{cc}'(G_{n,p})$, using a pseudo-random variant of Pippenger's packing result due to Ehard, Glock and Joos [25], it is not difficult[7] to see that $G_{n,p}$ whp contains a collection $\mathcal{C}$ of edge-disjoint cliques K_s where each vertex is contained in $(1-o(1))np/(s-1)$ cliques of $\mathcal{C}$. Writing $\mathcal{U}$ for the edges of $G_{n,p}$ not covered by the cliques in $\mathcal{C}$, using Pippenger and Spencer's chromatic index result [93] and Vizing's theorem it then is not difficult to see that whp $\mathrm{cc}'(G_{n,p}) \leq \chi'(\mathcal{C}) + \chi'(\mathcal{U}) \leq (1+o(1))np/(s-1)$, as desired.

Acknowledgements. We would like to thank Annika Heckel for valuable discussions about Problem 1.

3.5 Appendix: Lower bounds: proof of Lemma 50

Proof of Lemma 50. Writing $\mathcal{S}$ for the event that the largest clique of $G_{n,p}$ has size at most $s = \lceil 2\log_{1/p} n\rceil$, it well-known that $\mathcal{S}$ holds whp (by a straightforward first moment argument). Writing $\mathcal{E}$ for the event that $G_{n,p}$ contains $(1 \pm \epsilon)\binom{n}{2}p$ edges for $\epsilon := n^{-1/2}$, say, it is easy to see that $\mathcal{E}$ holds whp (using Chebychev's inequality). Furthermore, recalling $\varphi(p) = (1-p)\log(1-p)/(p\log p)$, the probability that $G_{n,p}$ equals any fixed spanning subgraph $G \subseteq K_n$ with $e(G) = (1 \pm \epsilon)\binom{n}{2}p$ edges is routinely seen to be at most

$$\Pi := \max_{m \in (1\pm\epsilon)\binom{n}{2}p} p^m(1-p)^{\binom{n}{2}-m} \;\leq\; \exp\Big(-(1-o(1))\cdot\tbinom{n}{2}p\big(1+\varphi(p)\big)\cdot\log(1/p)\Big). \quad (3.77)$$

For the clique covering number $\mathrm{cc}(G_{n,p})$, the crux is that there are at most

$$\binom{n+s}{s}^T \;\leq\; o(n^{sT})$$

[7] For the same auxiliary hypergraph $\mathcal{H}$ as considered before, the required technical conditions of [25, Theorem 1.2] with $\Delta \approx \binom{n-2}{s-2}p^{\binom{s}{2}-1} \geq \Omega\big((\log n)^{\omega(1)}\big)$ and $\log e(\mathcal{H}) \leq s\log n \ll \Delta^{\Theta(1)}$ can be verified using Lemma 41 and [102, Theorem 1].

many collections $\{C_1, \ldots, C_t\}$ with $t \le T$ that are a clique covering for some graph $G \subseteq K_n$ with largest clique of size at most s. Hence, since each clique covering uniquely determines the entire edge-set and thus the underlying spanning subgraph $G \subseteq K_n$, it follows by a union bound argument that

$$\mathbb{P}(\mathrm{cc}(G_{n,p}) \le T) \;\le\; \mathbb{P}(\neg\mathcal{S} \text{ or } \neg\mathcal{E}) \;+\; o(n^{sT}) \cdot \Pi. \tag{3.78}$$

Note that $\mathbb{P}(\neg\mathcal{S} \text{ or } \neg\mathcal{E}) = o(1)$ and $s \log n \sim \binom{s}{2} \cdot \log(1/p)$. In view of inequality (3.77), for any $\epsilon \in (0,1)$ it follows that (3.78) is at most $o(1)$ when $T \le (1-\epsilon) \cdot (1+\varphi(p))\binom{n}{2}p/\binom{s}{2}$, establishing (3.75).

Turning to the thickness $\mathrm{cc}_\Delta(G_{n,p})$, we associate each clique covering $\mathcal{C}$ of some graph $G \subseteq K_n$ with an auxiliary bipartite graph $\mathcal{B}$ on vertex-set $[n] \cup \mathcal{C}$, where $v \in [n]$ and $C_i \in \mathcal{C}$ are connected by an edge whenever $v \in V(C_i)$. If the thickness of $\mathcal{C}$ is at most T, then in $\mathcal{B}$ the degree of each $v \in [n]$ is at most $\lfloor T \rfloor$, which also gives $|\mathcal{C}| \le n\lfloor T \rfloor$. Since the structure of the auxiliary bipartite graph $\mathcal{B}$ uniquely determines $\mathcal{C}$ (as the neighbors of C_i in $\mathcal{B}$ determine the clique vertex-set $V(C_i)$), it follows that there are at most

$$\binom{n\lfloor T \rfloor + \lfloor T \rfloor}{\lfloor T \rfloor}^n \;\le\; O\big((6n)^{nT}\big)$$

many collections $\mathcal{C}$ with thickness at most T that are a clique covering of some graph $G \subseteq K_n$. Since each such $\mathcal{C}$ uniquely determines the underlying spanning subgraph $G \subseteq K_n$, we obtain similarly to (3.78) that

$$\mathbb{P}(\mathrm{cc}_\Delta(G_{n,p}) \le T) \;\le\; \mathbb{P}(\neg\mathcal{E}) \;+\; O\big((6n)^{nT}\big) \cdot \Pi. \tag{3.79}$$

Note that $\mathbb{P}(\neg\mathcal{E}) = o(1)$ and $n \log(6n) \sim \binom{n}{2} \log(1/p) \cdot (s-1)/n$. In view of inequality (3.77), for any $\epsilon \in (0,1)$ it follows that (3.79) is at most $o(1)$ when $T \le (1-\epsilon) \cdot (1 + \varphi(p))np/(s-1)$, completing the proof of (3.76). $\square$

3.6 Appendix: Variant of Theorem 32: proof of Corollary 40

Proof of Corollary 40. Choosing $\xi = \xi(\delta) \in (0, 1/16]$ such that $(1+\delta)(1+\xi)/(1-4\xi)^2 \leq 1+2\delta$, set $m_0 := \lfloor(1+\xi)m\rfloor$, $m_1 := \lfloor m_0/(1-4\xi)^2\rfloor$, and $c := (1+\delta)rm_1/n$. Let $\mathcal{H}_i^*$ be chosen uniformly at random from all $\binom{|E(\mathcal{H})|}{i}$ subhypergraphs of $\mathcal{H}$ with exactly i edges. Since $\mathcal{H}_q$ conditioned on having exactly i edges has the same distribution as $\mathcal{H}_i^*$, by the law of total probability and monotonicity it follows that

$$\begin{aligned} \mathbb{P}(\chi'(\mathcal{H}_q) \geq c) \ &\leq \ \mathbb{P}(|E(\mathcal{H}_q)| > m_0) + \sum_{0 \leq i \leq m_0} \mathbb{P}(\chi'(\mathcal{H}_i^*) \geq c)\, \mathbb{P}(|E(\mathcal{H}_q)| = i) \\ &\leq \ n^{-\omega(r)} + \mathbb{P}(\chi'(\mathcal{H}_{m_0}^*) \geq c), \end{aligned} \tag{3.80}$$

where we used standard Chernoff bounds (such as [60, Theorem 2.1]) and $\mathbb{E}\,|E(\mathcal{H}_q)| = |E(\mathcal{H})|q = m \geq n^{1+\sigma} \gg r \log n$. Sequentially choosing the random edges $e_1, \ldots, e_{m_1} \in E(\mathcal{H})$ of $\mathcal{H}_{m_1}$ as defined in Theorem 32, note that $e_{i+1} \in E(\mathcal{H}) \setminus \{e_1, \ldots, e_i\}$ holds with probability at least $1 - m_1/e(\mathcal{H}) > 1 - 4\xi$, as $m_1 < 4m \leq 4\xi e(\mathcal{H})$. Since we can equivalently construct the edge-set $\{f_1, \ldots, f_{m_0}\}$ of $\mathcal{H}_{m_0}^*$ by sequentially choosing $f_{i+1} \in E(\mathcal{H}) \setminus \{f_1, \ldots, f_i\}$ uniformly at random, a natural coupling of $\mathcal{H}_{m_1}$ and $\mathcal{H}_{m_0}^*$ thus satisfies

$$\mathbb{P}(\mathcal{H}_{m_0}^* \subseteq \mathcal{H}_{m_1}) \ \geq \ \mathbb{P}(\mathrm{Bin}(m_1, 1-4\xi) \geq m_0) \ \geq \ 1 - n^{-\omega(r)},$$

where we used standard Chernoff bounds and that $m_1(1-4\xi) > m_0/(1-\xi)$ for $n \geq n_0(\xi)$. Hence

$$\mathbb{P}(\chi'(\mathcal{H}_{m_0}^*) \geq c) \ \leq \ \mathbb{P}(\chi'(\mathcal{H}_{m_1}) \geq c) + n^{-\omega(r)} \ \leq \ n^{-\omega(r)}, \tag{3.81}$$

where we invoked Theorem 32 with m set to m_1 (which applies since $n^{1+\sigma} \leq m \leq m_1 < 4\xi e(\mathcal{H}) < n^r$). This completes the proof by combining (3.80) and (3.81) with $c \leq (1+$

$2\delta)rm/n$. $\square$

3.7 Appendix: Heuristics: random greedy edge coloring algorithm

In this appendix we give, for the greedy coloring algorithm from Section 3.3, two heuristic explanations for the trajectories $|Q_e(i)| \approx \hat{q}(t)$ and $|Y_{v,c}(i)| \approx \hat{y}(t)$ that these random variables follow, where $t = t(i,m) = i/m$.

For our first *pseudo-random heuristic*, we write $E_i = \{e_1, \dots, e_i\}$ for the multi-set of edges appearing during the first i steps of the algorithm. Ignoring that edges can appear multiple times, our pseudo-random ansatz is that the edges in E_i and their assigned colors are approximately independent with

$$\mathbb{P}\big(e \text{ in } E_i \text{ and colored } c\big) \approx \frac{|E_i|}{|E(\mathcal{H})|} \cdot \frac{1}{q} \approx \frac{i}{nD/r} \cdot \frac{1}{rm/n} = \frac{t}{D} =: p(t,D) = p,$$

where independence only holds with respect to colorings that are proper, i.e., possible in the algorithm. Using this heuristic ansatz, we now consider the event $\mathcal{E}_{v,c}$ that no edge $f \in E_i$ with $v \in f$ is colored c. Exploiting that no two distinct edges containing v can receive the same color in the algorithm (since this coloring would not be proper), our pseudo-random ansatz and the degree assumption (3.2) then suggests that

$$\mathbb{P}(\neg\mathcal{E}_{v,c}) = \sum_{f \in E(\mathcal{H}) : v \in f} \mathbb{P}(f \text{ in } E_i \text{ and colored } c) \approx D \cdot p = t.$$

Since for every pair of vertices there are only at most $n^{-\sigma} D$ edges containing both (by the codegree assumption), for $\ell = o(\log n)$ distinct vertices $v_1, \dots, v_\ell$ our pseudo-random ansatz also loosely suggests that

$$\mathbb{P}\Big(\bigcap_{i \in [\ell]} \mathcal{E}_{v_i,c}\Big) \approx \prod_{i \in [\ell]} \mathbb{P}(\mathcal{E}_{v_i,c}) + O\big(\ell^2 \cdot n^{-\sigma} D \cdot p\big) \approx (1-t)^\ell.$$

Recalling (3.46) from Section 3.3, using linearity of expectation we then anticipate $|Q_e(i)| \approx$

$\hat{q}(t)$ based on

$$\mathbb{E}\,|Q_e(i)| = \sum_{c\in[q]} \mathbb{P}\big(c \in Q_e(i)\big) = \sum_{c\in[q]} \mathbb{P}\Big(\bigcap_{v\in e} \mathcal{E}_{v,c}\Big) \approx q\cdot(1-t)^r = \hat{q}(t).$$

Mimicking this reasoning, recalling (3.47) we similarly anticipate $|Y_{v,c}(i)| \approx \hat{y}(t)$ based on

$$\mathbb{E}\,|Y_{v,c}(i)| = \sum_{f\in E(\mathcal{H}):v\in f} \mathbb{P}\big(c \in Q_{f\setminus\{v\}}(i)\big) \approx D\cdot(1-t)^{r-1} = \hat{y}(t).$$

In our second *expected one-step changes heuristic* we assume for simplicity that there are deterministic approximations $|Q_e(i)| \approx f(t)q$ and $|Y_{v,c}(i)| \approx g(t)D$. Using these approximations and $q \approx rm/n$, the calculations leading to (3.55)–(3.56) and (3.58)–(3.59) in Section 3.3.2.1 then suggest that

$$\mathbb{E}\big(|Q_e(i+1)| - |Q_e(i)| \,\big|\, \mathcal{F}_i\big) \approx -\frac{f(t)q\cdot r\cdot g(t)D}{nD/r\cdot f(t)q} \approx -\frac{rg(t)q}{m}, \tag{3.82}$$

$$\mathbb{E}\big(|Y_{v,c}(i+1)| - |Y_{v,c}(i)| \,\big|\, \mathcal{F}_i\big) \approx -\frac{g(t)D\cdot(r-1)\cdot g(t)D}{nD/r\cdot f(t)q} \approx -\frac{(r-1)g^2(t)D}{f(t)m}, \tag{3.83}$$

where $\mathcal{F}_i$ denotes the natural filtration of the algorithm after i steps. Since the left-hand sides of (3.82)–(3.83) are approximately equal to $[f(t+1/m) - f(t)]q \approx f'(t)q/m$ and $g'(t)D/m$, respectively, we anticipate

$$f'(t) = -rg(t) \quad \text{and} \quad g'(t) = -(r-1)g^2(t)/f(t). \tag{3.84}$$

Noting $|Q_e(0)| = q$ and $|Y_{v,c}(0)| \approx D$, we also anticipate $f(0) = g(0) = 1$. The solutions $f(t) = (1-t)^r$ and $g(t) = (1-t)^{r-1}$ then make $|Q_e(i)| \approx f(t)q = \hat{q}(t)$ and $|Y_{v,c}(i)| \approx g(t)D = \hat{y}(t)$ plausible.

CHAPTER 4

ON THE POWER OF RANDOM GREEDY ALGORITHMS

4.1 Background and main results

The probabilistic method is a widely used tool for proving the existence of hard-to-construct mathematical objects with certain desirable properties: it works by showing that a randomly chosen object has the desired properties with non-zero probability. In classical textbook approaches to the probabilistic method, the underlying random objects are typically generated in a *static* way, e.g., by choosing a graph uniformly at random from a prescribed class of graphs, or by independently including each possible edge.

In this chapter we illustrate the power of the algorithmic approach to the probabilistic method, where the random objects are generated step-by-step in a *dynamic* way using a randomized algorithm. To this end we consider two examples from graph theory and additive combinatorics, and show that each time random greedy algorithms allow us to go beyond classical applications of the probabilistic method, i.e., prove existence of mathematical objects with better properties. These algorithmic improvements are key for (i) resolving a problem of Esperet, Kang and Thomassé [36], and (ii) answering a question of Li [77], see Theorems 54 and 55.

For the two combinatorial examples considered in this chapter, previous work used the probabilistic method to show that static random objects can avoid certain forbidden substructures, while maintaining other desired pseudo-random properties. Our technical results show that random greedy algorithms, which by construction avoid these forbidden substructures, create objects with superior pseudo-random properties, see Theorems 56 and 58. With the benefit of hindsight, earlier work of Rödl [97], Kahn [62], Wormald [123], Spencer [108], Kim [67], Bohman [10], and others [9, 14, 46] can be interpreted similarly.

This chapter thus reveals the following emerging algorithmic paradigm: one can often take proofs based on classical probabilistic method arguments, and obtain improvements by using an algorithmic approach to the probabilistic method.

4.1.1 Induced bipartite subgraphs in triangle-free graphs

Our first example is from extremal graph theory, concerning a local refinement of the famous Max Cut problem. Here the history starts in 1988, when Erdős, Faudree, Pach and Spencer [29] introduced the problem of searching for large induced bipartite subgraphs in triangle-free graphs. Around 2018 Esperet, Kang and Thomassé [36] further refined this problem, focusing on induced bipartite subgraphs with large minimum degree. More precisely, for fixed $\eta \in (0,1)$ they asked to determine the behavior of the parameter $f_\eta(n)$, which is defined as the maximum f such that every n-vertex triangle-free graph with minimum degree at least n^η contains an induced bipartite subgraph with minimum degree at least f. Recent results of Kwan, Letzter, Sudakov, and Tran [76] and van Batenburg, de Verclos, Kang, and Pirot [113] show that

$$f_\eta(n) = \Theta\big(\max\big\{\log n, n^{2\eta-1}\big\}\big) \qquad \text{for fixed } \eta \in (0,1) \setminus (1/2, 2/3], \tag{4.1}$$

and also determine $f_\eta(n)$ up to logarithmic factors in the remaining range $\eta \in (1/2, 2/3]$. Illustrating the conceptual punchline of this chapter, we use a 'dynamic' randomized greedy algorithm to improve existing upper bound constructions [76, 113], which were based on classical probabilistic method tools applied to the binomial random graph $G_{n,p}$. This algorithmic improvement allows us to close the logarithmic gap for $\eta \in (1/2, 2/3]$, and determine the order of magnitude of $f_\eta(n)$ for any fixed $\eta \in (0,1)$. The following result in particular resolves [36, Problem 4.1] of Esperet, Kang, and Thomassé up to constant factors.

Theorem 54. *For fixed* $\eta \in (0,1)$*, we have* $f_\eta(n) = \Theta(\max\{\log n, n^{2\eta-1}\})$.

In comparison to the previous upper bounds [76, 113] based on the probabilistic analysis of $G_{n,p}$ via the alteration method or the Lovász Local Lemma, our key improvement stems from the fact that via the so-called semi-random triangle-free process we are able to algorithmically construct pseudo-random triangle-free graphs with higher edge density (see Theorem 56), confirming speculations from [36, Section 4] and [113, Section 3].

4.1.2 Van der Waerden numbers

Our second example is from additive combinatorics, concerning a well-known Ramsey-type parameter for arithmetic progressions. The van der Waerden number $W(r,k)$ is defined as the smallest integer n such that every red and blue coloring of numbers in $[n] := \{1, 2, \ldots, n\}$ contains either a monochromatic red r-term arithmetic progression (r-AP) or a monochromatic blue k-AP. The celebrated van der Waerden's theorem guarantees that $W(r,k)$ is finite for all integers $r, k \geq 2$, making it a natural and interesting problem to determine the asymptotic behavior of $W(r,k)$, see [47, 50]. The 'off-diagonal' case, where $r \geq 3$ is fixed and k tends to infinity, was of particular interest to Graham (note that $W(2,k) = \Theta(k)$ holds trivially). Indeed, in the mid 2000s Graham conjectured that $W(3,k) \leq k^{O(1)}$, and mentioned that numerical evidence suggests $W(3,k) = k^{2+o(1)}$, see [48, 49, 52]. Around 2015 Graham even started offering \$250 reward for his conjecture, see [49, p. 19]. In terms of lower bounds, in 2008 Li and Shu [78] showed that

$$W(r,k) = \Omega\big((k/\log k)^{r-1}\big) \qquad \text{for fixed } r \geq 3,$$

by applying the Lovász Local Lemma to a random subset of the integers $[n]$. Subsequently, Li raised in 2009 the natural question [77] whether this probabilistic lower bound can be improved via a randomized greedy algorithm that 'dynamically' constructs an r-AP free subset of the integers $[n]$. The proof of the following theorem answers Li's question affirmatively, see also Sections 4.1.4.2 and 4.3.

Theorem 55. *For fixed $r \geq 3$, we have $W(r,k) = \Omega\big(k^{r-1}/(\log k)^{r-2}\big)$.*

This result was announced in October 2020, see [53]. While this chapter was under slow preparation during the COVID-19 pandemic, Green [51] made a breakthrough and showed $W(3,k) \geq k^{(\log k)^{1/3-o(1)}}$ using very different techniques, which in view of $W(r,k) \geq W(3,k)$ disproves the earlier belief that $W(r,k) = k^{O(1)}$ for fixed $r \geq 3$. The best known upper bound $W(3,k) \leq \exp(k^{1-\Omega(1)})$ was obtained by Schoen [100] in 2020.

4.1.3 Organization

In Section 4.1.4 we state our main technical results, which will imply Theorem 54 and 55 for induced bipartite subgraphs and van der Waerden numbers, respectively. In Sections 4.2 and 4.3 we then prove these technical results using an algorithmic approach to the probabilistic method, i.e., by analyzing randomized algorithms that construct pseudo-random triangle-free graphs and r-AP free subsets of the integers, respectively.

4.1.4 Main technical results

4.1.4.1 Construction of pseudo-random triangle-free graphs

To prove the upper bound on the parameter $f_\eta(n)$ claimed by Theorem 54 for $\eta \in (1/2, 2/3]$, our strategy is to construct a pseudo-random triangle-free graph G_n with $\Theta(n)$ vertices, where pseudo-randomness will intuitively ensure the desired minimum degree properties (in suitable constructions that are based on G_n). Following the conceptual punchline of this chapter, we shall construct the desired graph G_n using a semi-random variant of the triangle-free process, which is a randomized greedy algorithm that sequentially adds more edges to G_n without creating a triangle, see Section 4.2 for the full details. This algorithmic approach to the probabilistic method is key for obtaining our improved upper bound on $f_\eta(n)$ via the following auxiliary result, since earlier approaches based on the binomial random graph $G_{n,p}$ were only able to prove Theorem 56 with weaker minimum and maximum degree bounds $\delta(G_n), \Delta(G_n) = \Theta(\sqrt{n})$, see [76, Lemma 5.1] and [113, Theorem 3.1].

Theorem 56. *There are constants $c, C, C' > 0$ such that for any $0 < \beta < 1/14$ the following holds for any integer $n \geq n_0 = n_0(\beta)$. There exists a triangle-free graph G_n with $v(G_n) \in [n/3, n]$ vertices,*

$$c\sqrt{\beta n \log n} \leq \delta(G_n) \leq \Delta(G_n) \leq C\sqrt{\beta n \log n}, \tag{4.2}$$

and the property that any induced bipartite subgraph $F \subseteq G_n$ has minimum degree $\delta(F) \leq C' \log n$.

We defer the proof of Theorem 56 to Section 4.2: it is based on a careful refinement of the semi-random triangle-free process analysis of Guo and Warnke [56]. Using Theorem 56 we shall in fact establish improved bounds for the more general parameter $g(n, d)$, which denotes the maximum g such that every n-vertex triangle-free graph with minimum degree at least d contains an induced bipartite subgraph with minimum degree at least g. Extending [76, 113], the following result establishes a phase transition of $g(n, d)$ when the minimum degree d is around $\sqrt{n \log n}$, and it also implies Theorem 54 since $f_\eta(n) = g(n, n^\eta)$.

Theorem 57. *For any fixed $\gamma \in (0, 1)$, we have $g(n, d) = \Theta\big(\max\{\log d, d^2/n\}\big)$ for all $n^\gamma \leq d \leq n/2$.*

Similar to $f_\eta(n) = g(n, n^\eta)$, the cases $n^\gamma \leq d \leq \sqrt{n}$ and $n^{2/3} \leq d \leq n/2$ of Theorem 57 follow from [76]. Furthermore, for $\sqrt{n} \leq d \leq n^{2/3}$ we obtain $g(n, d) = \Omega(\max\{\log d, d^2/n\})$ by combining [76, Theorem 1.3] with the fact that $g(n, d)$ is monotone increasing in d. We now close the gap for $\sqrt{n} \leq d \leq n^{2/3}$ by mimicking the upper bound constructions from [76, 113] using the semi-random triangle-free process based graphs G_n from Theorem 56, which have better degree properties than the $G_{n,p}$ based graphs used in [76, 113].

Proof of Theorem 57 based on Theorem 56. Writing $c, C' > 0$ for the constants of Theorem 56, let $\beta := 10^{-2}$ and $A := c\sqrt{\beta}/3$. As discussed, it suffices to prove $g(n,d) = O(\max\{\log d, d^2/n\})$ for $\sqrt{n} \le d \le n^{2/3}$.

We start with the case $\sqrt{n} \le d \le A\sqrt{n \log n}$, where we set $\alpha := 2/(c^2\beta)$ and $n' := \lceil \alpha d^2/\log n \rceil$. Note that $n^{2/3} \ll n' \le \lceil \alpha A^2 n \rceil \le n/2$ and $n' \ll d^2$. By taking the disjoint union of $\lfloor n/n' \rfloor$ copies of $G_{n'}$, we obtain a triangle-free graph H_n with $v(H_n) = \lfloor n/n' \rfloor \cdot v(G_{n'}) \in [n/6, n]$ vertices and minimum degree

$$\delta(H_n) = \delta(G_{n'}) \ge c\sqrt{\beta n' \log n'} \ge \sqrt{c^2\beta\alpha \cdot d^2 \cdot 2/3} > d.$$

Furthermore, every induced bipartite subgraph $F \subseteq H_n$ is a disjoint union of induced bipartite subgraphs from copies of $G_{n'}$ and thus has minimum degree at most $\delta(F) \le C' \log n' \le 2C' \log d$. By 'blowing up' each vertex of H_n into an independent set of suitable sizes between one and six (i.e., after replacing each vertex of H_n by an independent set, we add a complete bipartite graph between every pair of independent sets that correspond to an edge in H_n), we thus obtain an n-vertex triangle-free graph $G_{n,d}$ with $\delta(G_{n,d}) \ge \delta(H_n) \ge d$, where furthermore every induced bipartite subgraph $F \subseteq G_{n,d}$ has minimum degree at most $\delta(F) \le 6 \cdot 2C' \log d$ (by analogous disjoint reasoning as before), establishing that $g(n,d) = O(\log d)$.

Finally, in the remaining case $A\sqrt{n \log n} \le d \le n^{2/3}$ we set $\alpha := c^2\beta/18$ and $n' := \lfloor \alpha (n/d)^2 \log n \rfloor$. Note that $n^{2/3} \ll n' \le \alpha n/A^2 \le n/2$. By 'blowing up' each vertex of $G_{n'}$ into an independent set of size $\lfloor n/n' \rfloor$, we obtain a triangle-free graph H_n with $v(H_n) = \lfloor n/n' \rfloor \cdot v(G_{n'}) \in [n/6, n]$ vertices and minimum degree

$$\begin{aligned}\delta(H_n) = \lfloor n/n' \rfloor \cdot \delta(G_{n'}) &\ge \frac{n}{2n'} \cdot c\sqrt{\beta n' \log n'} \\ &\ge \sqrt{\frac{c^2\beta n^2 \log(n^{2/3})}{4n'}} \ge \sqrt{\frac{c^2\beta \cdot d^2 \cdot 2/3}{4\alpha}} > d.\end{aligned}$$

Furthermore, every induced bipartite subgraph $F \subseteq H_n$ has minimum degree at most

$\delta(F) \leq \lfloor n/n' \rfloor \cdot C' \log n' \leq 2\alpha^{-1}C' \cdot d^2/n$. By blowing up each vertex of H_n into an independent set of suitable sizes between one and six, we then obtain an n-vertex triangle-free graph $G_{n,d}$ that establishes $g(n,d) = O(d^2/n)$. □

4.1.4.2 Construction of pseudo-random r-AP free sets of integers

To prove the lower bound on the van der Waerden number $W(r,k)$ claimed by Theorem 55, our strategy is to construct a large subset $I \subseteq [n]$ of the integers that is r-AP free and pseudo-random, where pseudo-randomness will intuitively ensure that $[n] \setminus I$ is k-AP free for fairly large $k = k(n)$. For technical reasons, it will be convenient to work with the field $\mathbb{Z}/N\mathbb{Z}$ for a prime number N, where a set of numbers $\{a_1, \ldots, a_r\} \subseteq \mathbb{Z}/N\mathbb{Z}$ is formally called an r-term arithmetic progression (r-AP) in $\mathbb{Z}/N\mathbb{Z}$ if $|\{a_1, \ldots, a_r\}| = r$ and $a_i \equiv_N a_1 + (i-1)d$ for some $d \not\equiv_N 0$. Following the conceptual punchline of this chapter, we shall construct the desired pseudo-random r-AP free subset $I \subseteq \mathbb{Z}/N\mathbb{Z}$ using the so-called random greedy r-AP free process, which is a randomized greedy algorithm that step-by-step adds more random numbers to I without creating an r-AP, see Section 4.3 for the full details. This algorithmic approach to the probabilistic method is key for obtaining our improved lower bound on $W(r,k)$ via the following result, since earlier approaches based on random subsets the integers were only able to prove Theorem 58 with the weaker parameter choice $k = \Theta(N^{1/(r-1)} \log N)$, see [78].

Theorem 58. *For any fixed $r \geq 3$, there are constants $C, N_0 > 0$ such that the following holds for any prime number $N \geq N_0$. There exists a set $I \subseteq \mathbb{Z}/N\mathbb{Z}$ which (i) is r-AP free in $\mathbb{Z}/N\mathbb{Z}$ and (ii) satisfies $|I \cap K| \geq 1$ for all k-APs K in $\mathbb{Z}/N\mathbb{Z}$ of size $k := \lceil C(N/\log N)^{1/(r-1)} \log N \rceil$.*

Proof of Theorem 55 based on Theorem 58. For any integer $n \geq \max\{2, N_0\}$, by Betrand's postulate we may fix a prime number satisfying $n < N < 2n$. For $I \subseteq \mathbb{Z}/N\mathbb{Z}$ as given by Theorem 58, we color $I \cap [n]$ red and $[n] \setminus I$ blue. Properties (i)–(ii) of Theorem 58 ensure that there are no red r-APs or blue k-APs in $[n]$, since any AP in $[n]$ corresponds to an AP in $\mathbb{Z}/N\mathbb{Z}$. It follows that $W(r,k) > n > N/2 = \Theta(k^{r-1}/(\log k)^{r-2})$. $\square$

We defer the proof of Theorem 58 to Section 4.3: it is based on the differential equation method and results of Bohman and Bennett [9] for the so-called random greedy independent set algorithm. Noteworthily, in our analysis we need to ensure that all of the polynomially many k-APs are 'hit' by the set I produced by the r-AP process. This is in great contrast to the analysis of the H-free process arising in graph Ramsey theory, where one typically needs to ensure that an exponential number of substructures are hit [10, 12, 37, 13, 115, 116, 92].

4.2 Semi-random triangle-free process

In this section we prove Theorem 56 by showing that a semi-random variant of the so-called triangle-free process typically finds a triangle-free graph $G_n \subseteq K_n$ with the desired properties. Intuitively, this process starts with an empty graph, and then iteratively adds a large number of carefully chosen edges (instead of just adding a single edge as the original triangle-free process) such that the resulting graph stays triangle-free.

4.2.1 More details and heuristics

The formal details of the semi-random triangle-free process are rather involved, so here we shall only introduce the aspects that are important for the upcoming arguments of this chapter, deferring the full details to [56, Section 2]. The semi-random process starts with

$$E_0 = T_0 := \varnothing \quad \text{and} \quad O_0 := E(K_n), \tag{4.3}$$

and the rough plan is to step-by-step build up a 'random' set of edges $E_i \subseteq E(K_n)$, a triangle-free edge subset $T_i \subseteq E_i$, and a set of 'open' edges $O_i \subseteq E(K_n) \setminus E_i$ that can still be added to E_i without creating triangles. More precisely, in step $i+1 \geq 1$ of the semi-random triangle-free process we sample a random edge subset $\Gamma_{i+1} \subseteq O_i$, where each edge $e \in O_i$ is included independently with probability

$$p := \sigma/\sqrt{n} \quad \text{with} \quad \sigma := (\log n)^{-2}, \tag{4.4}$$

and update the random set of edges by setting

$$E_{i+1} := E_i \cup \Gamma_{i+1}. \tag{4.5}$$

To determine the new triangle-free edge subset $T_{i+1} \subseteq T_i \cup \Gamma_{i+1}$, the idea is to delete a suitable set $D_{i+1} \subseteq \Gamma_{i+1}$ of edges from Γ_{i+1} with $|\Gamma_{i+1} \setminus D_{i+1}| \approx |\Gamma_{i+1}|$, such that

$$T_{i+1} := T_i \cup \big(\Gamma_{i+1} \setminus D_{i+1}\big) \tag{4.6}$$

remains triangle-free, see [56, (13)–(14) in Section 2.1] for the precise definition of D_{i+1} (this construction intuitively works since only few new triangles are created in $E_i \cup \Gamma_{i+1}$ due to the fact that Γ_{i+1} is fairly small). To determine the new open edge set $O_{i+1} \subseteq O_i \setminus \Gamma_{i+1}$, we certainly have to remove the set C'_{i+1} of 'newly closed' edges, which simply contains all edges $e \in O_i$ that form a triangle with some two edges of $E_{i+1} = E_i \cup \Gamma_{i+1}$. For technical reason we also remove an extra random edge subset $S_{i+1} \subseteq O_i$ and set

$$O_{i+1} := O_i \setminus \big(\Gamma_{i+1} \cup C'_{i+1} \cup S_{i+1}\big), \tag{4.7}$$

see [56, (15)–(20) in Section 2.1] for the precise definition of $C'_{i+1} \cup S_{i+1}$ (the removal of extra edges is a technical twist that intuitively makes it easier to prove certain concentration statements).

Stopping this iterative construction after $I \approx n^{\beta}$ steps, the pseudo-random intuition from [56, Section 2] suggests that, with respect to various edge statistics, the resulting n-

vertex triangle-free graph

$$H := \big([n], T_I\big) \quad \text{and} \quad I := \lceil n^{\beta} \rceil \tag{4.8}$$

heuristically resembles a binomial random graph $G(n, \rho)$ with edge probability

$$\rho := \sqrt{\beta(\log n)/n}, \tag{4.9}$$

but with the notable exception that it by construction contains no triangles (such a random graph would typically contain many triangles). This heuristic makes it plausible that $G_n = H$ satisfies the degree properties claimed by Theorem 56, since routine arguments show that the random graph $G(n, \rho)$ typically has these degree properties. To keep the modifications of [56] minimal, we shall in fact find an induced subgraph $G_n \subseteq H$ with the desired degree properties (this extra step is convenient but not necessary, see Remark 59).

4.2.2 Setup and proof of Theorem 56

We now turn to the technical details of our proof of Theorem 56, which extends [56, Sections 2–3]. Here our setup is guided by the pseudo-random heuristic discussed in [56, Section 2.2], which loosely suggests that

$$\mathbb{P}(e \in E_i) \approx \pi_i/\sqrt{n} \quad \text{and} \quad \mathbb{P}(e \in O_i) \approx q_i, \tag{4.10}$$

where the parameters q_i and π_i satisfy the technical properties

$$\pi_i := \sigma + \sum_{0 \le j < i} \sigma q_j, \qquad q_i \in (0, 1] \quad \text{and} \quad \pi_I/\sqrt{n} = (1 + o(1))\rho, \tag{4.11}$$

see [56, Section 2.3 and Lemma 17] for the full details. In particular, to get a handle on the number of edges between large sets of vertices, consistent with (4.10)–(4.11) we introduce the pseudo-random events

$$\mathcal{T}_I^* := \Big\{|T_I(A,B)| \ge (1-\delta)|A||B|\rho \text{ for all disjoint } A,B \subseteq [n] \text{ with } |A| = |B| = s\Big\}, \tag{4.12}$$

$$\mathcal{T}_I^+ := \Big\{|T_I(A,B)| \le (1+\delta)2s|B|\rho \text{ for all disjoint } A,B \subseteq [n] \text{ with } |A| = |B| \le 2s\Big\}, \tag{4.13}$$

where we write $S(A,B) := \{ab \in S : a \in A, b \in B\}$ for the set of edges from S that go between A and B, and tacitly use the carefully chosen (see [56, Section 2.3 and Theorem 9]) size parameter

$$s := \lceil D(\log n)/\rho \rceil \quad \text{with} \quad D := 108/\delta^2 \quad \text{and} \quad \delta := 1/10. \tag{4.14}$$

To eventually get a handle on the maximum degree, we also introduce the auxiliary event

$$\mathcal{N}_{\le I} := \Big\{|N_{\Gamma_i}(v)| \le 2\sigma q_{i-1}\sqrt{n} \text{ for all } v \in [n] \text{ and } 0 \le i \le I\Big\}, \tag{4.15}$$

writing $N_S(v) := \{w \in [n] : vw \in S\}$ for the set of neighbors of v in a given edge set S.

Results of Guo and Warnke, see [56, Theorem 9], imply that

$$\mathbb{P}\big(\mathcal{T}_I^* \cap \mathcal{N}_{\le I}\big) \ge 1 - n^{-\omega(1)}. \tag{4.16}$$

As we shall show next, Theorem 56 then follows from the claim

$$\mathbb{P}\big(\mathcal{T}_I^+\big) \ge 1 - o(1), \tag{4.17}$$

whose stochastic domination based proof we defer to Section 4.2.3.

Proof of Theorem 56 assuming inequality (4.17). Combining (4.16)–(4.17), for all sufficiently large n we infer that $H = ([n], T_I)$ satisfies the event $\mathcal{T}_I^* \cap \mathcal{N}_{\le I} \cap \mathcal{T}_I^+$. We construct the induced triangle-free subgraph $G_n \subseteq H$ by iteratively deleting vertices of degree at most $\delta/4 \cdot n\rho$, and now verify that it has the claimed properties, starting with the degree bound (4.2). Noting $e_H(A,B) = |T_I(A,B)|$, the event $\mathcal{T}_I^*$ implies, via a double-counting argument analogous to the proof of [56, Theorem 5], that the number of edges of H is at least

$$e(H) = \frac{\sum_{A\subseteq[n]:|A|=s}\sum_{B\subseteq[n]\setminus A:|B|=s}|T_I(A,B)|}{2\binom{n-2}{s-1}\binom{n-s-1}{s-1}} \ge \frac{\binom{n}{s}\binom{n-s}{s}\cdot(1-\delta)s^2\rho}{2\binom{n-2}{s-1}\binom{n-s-1}{s-1}} = (1-\delta)\binom{n}{2}\rho. \tag{4.18}$$

Furthermore, by the recursive definition (4.6) of the edge set $T_I \subseteq \bigcup_{0\le i<I}\Gamma_{i+1}$, using the properties (4.11) of π_i we infer, for all sufficiently large $n \ge n_0(\delta)$, that the event $\mathcal{N}_{\le I}$ implies the maximum degree bound

$$\Delta(H) = \max_{v\in[n]}|N_{T_I}(v)| \le \max_{v\in[n]}\sum_{0\le i<I}|N_{\Gamma_{i+1}}(v)| \le \sum_{0\le i<I}2\sigma q_i\sqrt{n} \le 2\pi_I\sqrt{n} \le (2+\delta)\rho n.$$

By construction of $G_n \subseteq H$, using $\delta = 1/10$ we thus infer, for all sufficiently large $n \ge n_0(\delta)$, that

$$v(G_n) \ge \frac{2e(G_n)}{\Delta(G_n)} \ge \frac{2\big[e(H) - n\cdot\delta/4\cdot n\rho\big]}{\Delta(H)} \ge \frac{2(1-2\delta)\binom{n}{2}\rho}{(2+\delta)n\rho} > \frac{n}{3},$$

and so the claimed degree bound (4.2) follows with $c := \delta/4$ and $C := 2+\delta$.

Next, suppose that $F \subseteq G_n$ is an induced bipartite subgraph with two parts A and B, where we may assume that $|A| \ge |B|$. Since $F \subseteq G_n$ and $G_n \subseteq H$ are both induced subgraphs, we have

$$e_F(A,B) = e_{G_n}(A,B) = e_H(A,B) = |T_I(A,B)|.$$

Furthermore, since A and B are both independent sets in F, the event $\mathcal{T}_I^*$ implies that $|B| \le |A| \le \alpha(F) \le \alpha(H) \le 2 \cdot s$. Using a double counting argument similar to (4.18), the event $\mathcal{T}_I^+$ then implies that

$$|T_I(A,B)| = \frac{\sum_{A' \subseteq A: |A'|=|B|} |T_I(A',B)|}{\binom{|A|-1}{|B|-1}} \le \frac{\binom{|A|}{|B|} \cdot (1+\delta)2s|B|\rho}{\binom{|A|-1}{|B|-1}} = (1+\delta)2s|A|\rho.$$

The definitions (4.14) of $s \approx D(\log n)/\rho$ and $\delta = 1/10$ give $(1+\delta)2s\rho \le 3D \log n$ for sufficiently large n. By averaging it follows that $\delta(F) \le e_F(A,B)/|A| \le 3D \log n$, completing the proof with $C' := 3D$. □

Remark 59. *One can in fact show that the minimum and maximum degree of the n-vertex graph $H = ([n], T_I)$ satisfy $(1-\delta)n\rho \le \delta(H) \le \Delta(H) \le (1+\delta)n\rho$ with high probability (by adapting [56, Sections 3.1–3.5]), which would allow us to directly use $G_n = H$ in the above proof of Theorem 56. However, the coarser bounds used above suffice for our purposes, and require less technical modifications of [56].*

4.2.3 Pseudo-randomness: deferred proof of inequality (4.17)

This subsection is devoted to the deferred proof of inequality (4.17), i.e., $\mathbb{P}(\mathcal{T}_I^+) \ge 1 - o(1)$. To this end we shall adapt the strategy from [56, Sections 3.4–3.5] to our setting, i.e., use estimates on the number of open edges $|O_i(A,B)|$ to eventually get a handle on the total number of added edges $|T_I(A,B)|$.

Turning to the details, let $\mathcal{S}$ denote the set of all pairs of vertex disjoint $A, B \subseteq [n]$ with $|A| = |B| \le 2s$. To keep the changes to [56] minimal, for each pair $(A,B) \in \mathcal{S}$ we enlarge A to A^+ by adding the lexicographic first $2s - |A|$ vertices from $[n] \setminus (A \cup B)$. Note that the vertex set A^+ is determined by A. Consistent with the heuristic approximations (4.10), we then introduce the 'open edges' related pseudo-random events

$$\tilde{\mathcal{Q}}_i^+ := \big\{\big|O_i(A^+,B)\big| \le q_i|A^+||B| \text{ for all } (A,B)\in\mathcal{S}\big\} \quad \text{and} \quad \tilde{\mathcal{Q}}_{\le I}^+ := \bigcap_{0\le i\le I} \tilde{\mathcal{Q}}_i^+. \tag{4.19}$$

Note that $|B| \le |A^+| = 2s$ for all pairs $(A,B)\in\mathcal{S}$. Furthermore, there are at most n^{2j} pairs $(A,B)\in\mathcal{S}$ with $|B|=j$. With these two key properties in mind, the proof of [56, Lemma 24] carries over to the pairs (A^+,B) virtually unchanged (that proof merely exploits that $|A|$ is large, and only uses $|A|=|B|$ to control the final union bound estimate over all pairs (A,B) of vertex subsets), giving

$$\max_{0\le i<I} \mathbb{P}\big(\neg\tilde{\mathcal{Q}}_{i+1}^+ \mid \mathfrak{X}_{\le i}\cap\tilde{\mathcal{Q}}_{\le i}^+\big) \le \sum_{(A,B)\in\mathcal{S}} n^{-\omega(|B|)} \le n^{-\omega(1)},$$

where $\mathfrak{X}_{\le i}$ is a 'good' event determined by $(O_j,E_j,T_j,\Gamma_j,S_j)_{0\le j\le i}$ that is formally defined in [56, Section 2.4]; here we shall only use that the event $\mathfrak{X}_{\le i+1}$ implies $\mathfrak{X}_{\le i}$, and that $\mathbb{P}(\neg\mathfrak{X}_{\le I}) \le n^{-\omega(1)}$ by [56, Theorem 9]. Since the event $\mathfrak{X}_{\le i+1}\cap\tilde{\mathcal{Q}}_{\le i+1}^+$ implies $\mathfrak{X}_{\le i}\cap\tilde{\mathcal{Q}}_{\le i}^+$, using $I\approx n^\beta$ it then follows that

$$\mathbb{P}\big(\neg\tilde{\mathcal{Q}}_{\le I}^+\big) \le \mathbb{P}\big(\neg\mathfrak{X}_{\le I}\big) + \sum_{0\le i<I} \mathbb{P}\big(\neg\tilde{\mathcal{Q}}_{i+1}^+ \mid \mathfrak{X}_{\le i}\cap\tilde{\mathcal{Q}}_{\le i}^+\big) \le (I+1)\cdot n^{-\omega(1)} \le n^{-\omega(1)}. \tag{4.20}$$

Turning to the total number of added edges $|T_I(A,B)|$ for $(A,B)\in\mathcal{S}$, using $A\subseteq A^+$ and $T_I\subseteq E_I$ together with the recursive definition (4.5) of the edge set $E_I=\bigcup_{0\le i<I}\Gamma_{i+1}$, it follows that

$$|T_I(A,B)| \le \big|E_I(A^+,B)\big| = \sum_{0\le i<I} \big|O_i(A^+,B)\cap\Gamma_{i+1}\big|. \tag{4.21}$$

Recall that the event $\tilde{\mathcal{Q}}_i^+$ implies $|O_i(A^+,B)| \le q_i|A^+||B|$, and that $\Gamma_{i+1}\subseteq O_i$ is the random subset where each edge $e\in O_i$ is included independently with probability p. Combining these properties, by mimicking the stochastic domination arguments from the proof of [56, Claim 30] it then follows that

$$\mathbb{P}\big(|E_I(A^+, B)| \ge t \text{ and } \tilde{\mathcal{Q}}^+_{\le I}\big) \le \mathbb{P}\big(Z^+ \ge t\big) \quad \text{with} \quad Z^+ \overset{\mathrm{d}}{=} \mathrm{Bin}\Big(\sum_{0\le i<I} \big\lfloor q_i|A^+||B|\big\rfloor,\ p\Big).$$

Using $p = \sigma/\sqrt{n}$, the properties (4.11) of π_i and $|A^+| = 2s$, similar to [56, Section 3.5] we have

$$\mu^+ := \mathbb{E}\, Z^+ \sim \sum_{0\le i<I} \sigma q_i/\sqrt{n} \cdot |A^+||B| = (\pi_I - \sigma)/\sqrt{n} \cdot |A^+||B| \sim 2s|B|\rho.$$

Using the definitions (4.14) of the parameter $s \approx D(\log n)/\rho$ and the constant $D = 108/\delta^2$, for sufficiently large n it follows that $(1+\delta)2s|B|\rho \ge (1+\delta/2)\mu^+$ and $\delta^2\mu^+/12 > 12|B|\log n$, say. Similar to [56, (97)–(98)], standard Chernoff bounds such as [60, Theorem 2.1] thus routinely give

$$\begin{aligned}\mathbb{P}\big(|E_I(A^+, B)| \ge (1+\delta)2s|B|\rho \text{ and } \tilde{\mathcal{Q}}^+_{\le I}\big) &\le \mathbb{P}\big(Z^+ \ge (1+\delta/2)\mu^+\big)\\ &\le \exp\big(-\delta^2\mu^+/12\big) \le n^{-12|B|}.\end{aligned}$$

Recalling that the vertex set A^+ is determined by A, and that there are at most most n^{2j} pairs $(A, B) \in \mathcal{S}$ with $|B| = j$, in view of inequality (4.21) it then follows via a standard union bound argument that

$$\mathbb{P}\big(\neg\mathcal{T}^+_I \cap \tilde{\mathcal{Q}}^+_{\le I}\big) \le \sum_{(A,B)\in\mathcal{S}} n^{-12|B|} = o\big(n^{-9}\big),$$

which together with (4.20) implies $\mathbb{P}(\mathcal{T}^+_I) \ge 1 - o(1)$. This completes the proof of inequality (4.17) and thus Theorem 56, as discussed. $\square$

4.3 Random greedy r-AP free process

In this section we prove Theorem 58 by showing that the random greedy r-AP free process typically finds an r-AP free subset $I \subseteq \mathbb{Z}/N\mathbb{Z}$ with the desired properties. Intuitively, this process starts with an empty set $I = \varnothing$, and then iteratively adds new random numbers from $\mathbb{Z}/N\mathbb{Z}$ such that the resulting set I stays r-AP free. More formally, fixing $r \geq 3$, the random greedy r-AP free process starts with

$$I(0) := \varnothing \quad \text{and} \quad S(0) := \mathbb{Z}/N\mathbb{Z}. \tag{4.22}$$

Here $I(i)$ denotes the growing r-AP free set found after i steps, and $S(i)$ denotes the set of 'available' numbers in $\mathbb{Z}/N\mathbb{Z} \setminus I(i)$, i.e., that can be added to $I(i)$ without creating an r-AP. In step $i+1 \geq 1$ of the random greedy r-AP free process, we then choose $x_{i+1} \in S_i$ uniformly at random and update the r-AP free set and available set via

$$I(i+1) := I(i) \cup \{x_{i+1}\}, \tag{4.23}$$

$$S(i+1) := S(i) \setminus \big(\{x_{i+1}\} \cup Y_{x_{i+1}}(i)\big), \tag{4.24}$$

tacitly writing $Y_{x_{i+1}}(i)$ for the set of numbers that become 'unavailable' when x_{i+1} is added, i.e.,

$$Y_x(i) := \Big\{y \in S(i) \setminus \{x\} \ : \ \text{there is } A \in \mathcal{A}_{N,r} \text{ such that } x, y \in A \text{ and } A \setminus \{x,y\} \subseteq I(i)\Big\}, \tag{4.25}$$

where $\mathcal{A}_{N,\ell}$ is a shorthand for the collection of all ℓ-APs in $\mathbb{Z}/N\mathbb{Z}$.

4.3.1 Proof strategy

In this subsection we discuss our proof strategy for Theorem 58. To this end, let us first record the basic observation that each number $x \in \mathbb{Z}/N\mathbb{Z}$ is contained in exactly

$$D := r|\mathcal{A}_{N,r}|/N = \Theta(N) \tag{4.26}$$

many r-APs $A \in \mathcal{A}_{N,r}$. Our strategy is then to analyze the random greedy r-AP free process for

$$m := \xi \cdot ND^{-\frac{1}{r-1}}(\log N)^{\frac{1}{r-1}} \tag{4.27}$$

steps, and show that the r-AP free set $I := I(m) \subseteq \mathbb{Z}/N\mathbb{Z}$ typically satisfies $I \cap K \neq \varnothing$ for all k-APs $K \in \mathcal{A}_{N,k}$ of size

$$k := 9\xi^{-1} \cdot (D/\log N)^{1/(r-1)} \log N = \Theta\big((N/\log N)^{1/(r-1)} \log N\big), \tag{4.28}$$

deferring the choices of the sufficiently small constants $0 < \xi, \delta < 1/(2r)$. As usual, we are henceforth treating both m and k as integers (since rounding has an asymptotically negligible effect on our arguments).

The outlined proof strategy is consistent with the pseudo-random heuristic that $I = I(m)$ resembles a random m-element subset of $\mathbb{Z}/N\mathbb{Z}$. Indeed, noting $km = 9N \log N$, this heuristic suggests that

$$\mathbb{P}\big(I \cap K = \varnothing\big) \approx \frac{\binom{N-k}{m}}{\binom{N}{m}} = \prod_{0 \le j < k} \Big(1 - \frac{m}{N-j}\Big) \le \exp\Big(-\frac{km}{N}\Big) \ll N^{-2},$$

which is small enough to employ a union bound argument over the at most N^2 many k-APs $K \in \mathcal{A}_{N,k}$. In (4.37) below and Section 4.3.3 we will essentially make this heuristic reasoning rigorous, albeit in a slightly roundabout way (via several pseudo-random events and the differential equation method).

4.3.2 Setup and proof of Theorem 58

We now turn to the technical details of our proof of Theorem 58, which require some setup. In order to relate the discrete steps of the process to continuous trajectories, we introduce the convenient scaling

$$t_i := i/M \quad \text{with} \quad M := ND^{-\frac{1}{r-1}}. \tag{4.29}$$

To get a handle on all k-APs $K \in \mathcal{A}_{N,k}$, we denote the number of available numbers in K by

$$S_K(i) := S(i) \cap K. \tag{4.30}$$

We then define $\mathcal{K}_{\leq j}$ as the pseudo-random event that for all $0 \leq i \leq j$ we have

$$|S_K(i)| = \big(1 \pm e(t_i)\big)kq(t_i) \qquad \text{for all } K \in \mathcal{A}_{N,k}, \tag{4.31}$$

and similarly define $\mathcal{S}_{\leq j}$ as the pseudo-random event that for all $0 \leq i \leq j$ we have

$$|S(i)| = \big(1 \pm D^{-\delta}\big)Nq(t_i) \quad \text{and} \quad \max_{x \in S(i)} \big||Y_x(i)| - s_2(t_i)\big| \leq D^{\frac{1}{r-1}-\delta}, \tag{4.32}$$

tacitly using the deterministic functions

$$q(t) := e^{-t^{r-1}}, \quad s_2(t) := (r-1)D^{\frac{1}{r-1}}t^{r-2}q(t) \quad \text{and} \quad e(t) := e^{5(t+t^{r-1})} \cdot D^{-\delta}. \tag{4.33}$$

Note that, by choosing $\xi = \xi(r,\delta) > 0$ small enough compared to $r, \delta > 0$, we may assume that for all steps $0 \leq i \leq m$ we have $0 \leq t_i \leq t_m = m/M = \xi(\log N)^{\frac{1}{r-1}}$ as well as

$$0 < D^{-\delta} \leq e(t) = o(1) \quad \text{and} \quad 0 \leq t \leq D^{o(1)} \quad \text{for} \quad 0 \leq t \leq t_m. \tag{4.34}$$

Results of Bohman and Bennett, see [9, Section 4], imply that for sufficiently[1] small $\xi, \delta > 0$ we have

$$\mathbb{P}(\neg \mathcal{S}_{\le m}) \le \exp\bigl(-N^{\Omega(1)}\bigr). \tag{4.35}$$

As we shall show next, Theorem 58 then follows easily from the claim

$$\mathbb{P}(\neg \mathcal{G}_{\le m}) = o(1) \quad \text{for} \quad \mathcal{G}_{\le i} := \mathcal{S}_{\le i} \cap \mathcal{K}_{\le i}, \tag{4.36}$$

whose differential equation method based proof we defer to Section 4.3.3.

Proof of Theorem 58 assuming inequality (4.36). For any k-AP $K \in \mathcal{A}_{N,k}$ in $\mathbb{Z}/N\mathbb{Z}$, whenever the event $\mathcal{G}_{\le i}$ holds, by combining the concentration bounds (4.31)–(4.32) with the error estimate (4.34) we infer that

$$\mathbb{P}(x_{i+1} \notin S(i) \cap K \mid \mathcal{F}_i) = 1 - \frac{|S_K(i)|}{|S(i)|} \le 1 - \frac{\frac{1}{2}kq(t_i)}{2Nq(t_i)} = 1 - \frac{k}{4N},$$

where $\mathcal{F}_i$ denotes the natural filtration associated with the algorithm after i steps (which intuitively keeps track of the 'history' of the algorithm, i.e., all the information available up to and including step i). Since the event $\mathcal{G}_{\le m}$ implies the event $\mathcal{G}_{\le i}$ for all $0 \le i \le m$, using $km = 9N \log N$ it routinely follows that

$$\mathbb{P}\bigl(I(m) \cap K = \varnothing \text{ and } \mathcal{G}_{\le m}\bigr) \le \prod_{0 \le i \le m-1} \Bigl(1 - \frac{k}{4N}\Bigr) \le \exp\Bigl(-\frac{km}{4N}\Bigr) \ll N^{-2}. \tag{4.37}$$

Taking a union bound over the at most N^2 many k-APs $K \in \mathcal{A}_{N,k}$ in $\mathbb{Z}/N\mathbb{Z}$ then completes the proof of Theorem 58 with $I := I(m)$, since $\mathbb{P}(\neg \mathcal{G}_{\le m}) = o(1)$ by the assumed inequality (4.36). $\square$

[1] For example, the explicit choices $\delta = 1/(40r^2)$, $\xi = \delta/500$ satisfy all constraints of this chapter and [9].

4.3.3 Dynamic concentration: deferred proof of inequality (4.36)

This subsection is devoted to the deferred proof of inequality (4.36), i.e., $\mathbb{P}(\neg\mathcal{G}_{\le m}) = o(1)$, which in view of the probability estimate (4.35) and the definition of the event $\mathcal{K}_{\le i}$ requires us to establish the dynamic concentration estimate (4.31) for $|S_K(i)|$. To this end, following the differential equation method approach to dynamic concentration [123, 10, 119], we introduce the auxiliary random variables

$$X_K^{\pm}(i) := \pm[|S_K(i)| - kq(t_i)] - kq(t_i)e(t_i). \tag{4.38}$$

The point is that the desired estimate (4.31) follows when both inequalities $X_K^{\pm}(i) \le 0$ hold. In the following we shall use supermartingale arguments to establish these inequalities, by analyzing the (expected and worst-case) one-step changes of $X_K^{\pm}(i)$ and $|S_K(i)|$.

4.3.3.1 Expected one-step changes

We start by estimating the expected one-step changes $\Delta S_K(i) := |S_K(i+1)| - |S_K(i)|$ of the number of available numbers in any k-AP $K \in \mathcal{A}_{N,k}$, assuming that $0 \le i < m$ and $\mathcal{G}_{\le i}$ hold. Note that $|S_K(i)|$ is monotone decreasing. Furthermore, a number $x \in S_K(i)$ is removed from the set of available numbers if the algorithm chooses a number x_{i+1} from $Y_x(i) \cup \{x\}$. Since $x_{i+1} \in S(i)$ is chosen uniformly at random, using the estimates (4.31)–(4.32) implied by $\mathcal{G}_{\le i}$ it follows that

$$\mathbb{E}(\Delta S_K(i) \mid \mathcal{F}_i) = -\sum_{x \in S_K(i)} \frac{|Y_x(i)| \pm 1}{|S(i)|} = \frac{-[1 \pm e(t_i)]kq(t_i) \cdot [s_2(t_i) \pm 2D^{\frac{1}{r-1}-\delta}]}{[1 \pm D^{-\delta}]Nq(t_i)}. \tag{4.39}$$

Recalling that $0 < D^{-\delta} \le e(t_i) = o(1)$ by (4.34), using $s_2(t_i)/D^{\frac{1}{r-1}} = (r-1)t_i^{r-2}q(t_i) = -q'(t_i)$ and $D^{\frac{1}{r-1}}/N = 1/M$ it follows that

$$\begin{aligned}\mathbb{E}(\Delta S_K(i) \mid \mathcal{F}_i) &= -\Big(1 \pm 4e(t_i)\Big)\Big(s_2(t_i) \pm 2D^{\frac{1}{r-1}-\delta}\Big)\frac{k}{N} \\ &= \frac{kq'(t_i)}{M} \pm \Big(4(r-1)t_i^{r-2} \cdot q(t_i)e(t_i) + 4D^{-\delta}\Big)\frac{k}{M}.\end{aligned} \tag{4.40}$$

In preparation for the upcoming supermartingale arguments, we now show that the expected one-step changes of the associated auxiliary variables $\Delta X_K^{\pm}(i)$ are negative, again assuming that $0 \le i < m$ and $\mathcal{G}_{\le i}$ hold. Set $f(t) := q(t)e(t)$. Recalling the shorthand $t_i = i/M$ and the definition (4.38) of $\Delta X_K^{\pm}(i)$, by applying Taylor's theorem with remainder it follows that

$$\begin{aligned}\mathbb{E}(\Delta X_K^{\pm}(i) \mid \mathcal{F}_i) = \pm &\left[\mathbb{E}(\Delta S_K(i) \mid \mathcal{F}_i) - \frac{kq'(t_i)}{M}\right] - \frac{kf'(t_i)}{M} \\ &+ O\left(\max_{0\le t\le t_m} \frac{k\big(|q''(t)| + |f''(t)|\big)}{M^2}\right).\end{aligned} \tag{4.41}$$

Using (4.40) we see that in (4.41) the main $kq'(t_i)/M$ term cancels up to second order terms. In the following we shall show that the main error term $-kf'(t_i)/M$ is large enough to make the expected change (4.41) negative. Indeed, noting $f(t) = e(t)q(t) \ge D^{-\delta}$, we have

$$f'(t_i) = \big(5 + 4(r-1)t_i^{r-2}\big)e(t_i)q(t_i) \ge 5D^{-\delta} + 4(r-1)t_i^{r-2}e(t_i)q(t_i).$$

Furthermore, $D = \Theta(N)$ and $\delta \le 1/(2r)$ imply $M = N/D^{\frac{1}{r-1}} \gg D^{2\delta}$. Recalling that $q(t) \le 1$, $f(t) \le e(t) \ll 1$ and $0 \le t \le D^{o(1)}$, see (4.34), it then routinely follows that

$$\frac{|q''(t)| + |f''(t)|}{M} \le \frac{O\big(\sum_{0\le j\le 2r} t^j\big) \cdot \big[q(t) + f(t)\big]}{M} \le \frac{D^{o(1)}}{D^{2\delta}} \ll D^{-\delta}.$$

Inserting these estimates and (4.40) into the expected one-step changes (4.41) of $\Delta X_K^{\pm}(i)$, it follows that

$$\mathbb{E}(\Delta X_K^{\pm}(i) \mid \mathcal{F}_i) \le -\big(1-o(1)\big)kD^{-\delta}/M < 0.$$

4.3.3.2 Bounds on the one-step changes

We next bound the expected one-step changes $|\Delta S_K(i)| = \big||S_K(i+1)| - |S_K(i)|\big|$, tacitly assuming that $0 \le i < m$ and $\mathcal{G}_{\le i}$ hold. Since $|S_K(i)|$ is step-wise decreasing, by combining $s_2(t_i) \le rD^{\frac{1}{r-1}}t_i^{r-2}$ with the first estimate of the expected one-step changes (4.40), using $e(t_i) = o(1)$ and $0 \le t_i \le D^{o(1)}$ it follows that

$$\mathbb{E}(|\Delta S_K(i)| \mid \mathcal{F}_i) = -\,\mathbb{E}(\Delta S_K(i) \mid \mathcal{F}_i) \le O\Big(D^{\frac{1}{r-1}}t_i^{r-2} + D^{\frac{1}{r-1}-\delta}\Big)\cdot\frac{k}{N} \ll kD^{\frac{1}{r-1}+\delta/2}/N. \tag{4.42}$$

Turning to the worst-case one-step changes of $|S_K(i)|$, we introduce the auxiliary event

$$\mathcal{N}_{\le j} := \Big\{\max_{x\in S(i)} |Y_x(i)\cap K| \le D^{\frac{1}{r-1}-3\delta} \text{ for all } K\in\mathcal{A}_{N,k} \text{ and } 0\le i\le j\Big\}. \tag{4.43}$$

Recalling the reasoning leading to (4.39), the crux is that when $\mathcal{N}_{\le i}$ holds, then we have

$$|\Delta S_K(i)| \le 1 + \max_{x\in S(i)} |Y_x(i)\cap K| \le 2D^{\frac{1}{r-1}-3\delta}. \tag{4.44}$$

We now claim that the auxiliary event $\mathcal{N}_{\le m}$ typically holds, i.e., more precisely that

$$\mathbb{P}\big(\neg\mathcal{N}_{\le m} \text{ and } \mathcal{S}_{\le m}\big) \le \exp\big(-N^{\Omega(1)}\big). \tag{4.45}$$

Turning to the proof details, with an eye on $|Y_x(i)\cap K|$ let

$$\mathcal{I} = \mathcal{I}(K,x) := \big\{W \,:\, |W| = r-2,\ W\cup\{x,y\}\in\mathcal{A}_{N,r} \text{ for some } y\in K\big\}. \tag{4.46}$$

Note that $|\mathcal{I}| \le kr^2$, as there are at most r^2 many r-APs containing two distinct numbers $\{x, y\}$. Let

$$N_{K,x} := \sum_{W \in \mathcal{I}} Y_W \quad \text{with} \quad Y_W := \mathbb{1}_{\{W \subseteq I(m) \text{ and } \mathcal{S}_{\le m}\}}. \tag{4.47}$$

Since $\{x\} \cup W$ contains $r-1 \ge 2$ elements, by similar reasoning as for $|\mathcal{I}|$ it follows that

$$\max_{0 \le i \le m} |Y_x(i) \cap K| \cdot \mathbb{1}_{\{\mathcal{S}_{\le m}\}} \le N_{K,x} \cdot r^2. \tag{4.48}$$

We shall bound $N_{K,x}$ via the following Chernoff-type upper tail estimate for combinatorial random variables with 'controlled dependencies', which is a convenient corollary of [120, Theorem 7 and Remarks 9–10].

Lemma 60. *Let $(Y_\alpha)_{\alpha \in \mathcal{I}}$ be a finite family of variables with $Y_\alpha \in [0,1]$ and $\sum_{\alpha \in \mathcal{I}} \lambda_\alpha \le \mu$, where $(\lambda_\alpha)_{\alpha \in \mathcal{I}}$ satisfies $\mathbb{E}(\prod_{i \in [s]} Y_{\alpha_i}) \le \prod_{i \in [s]} \lambda_{\alpha_i}$ for all $(\alpha_1, \dots, \alpha_s) \in \mathcal{I}^s$ with $\alpha_i \cap \alpha_j = \varnothing$ for $i \ne j$. Set $Y := \sum_{\alpha \in \mathcal{I}} Y_\alpha$. If $\max_{\alpha \in \mathcal{I}} |\{\beta \in \mathcal{I} : \alpha \cap \beta \ne \varnothing\}| \le C$, then $\mathbb{P}(Y \ge z) \le (e\mu/z)^{z/C}$ for all $z > \mu$.*

With an eye on $\mathbb{E}\, N_{K,x}$, we first record the basic observation that when $\mathcal{S}_{\le m}$ holds, then in every step $i \le m$ there are at least $|S(i)| \ge |S(m)| \gg ND^{-\delta/4}$ available numbers, say. For any set U of numbers from $\mathbb{Z}/N\mathbb{Z}$ a straightforward adaptation of the proof of [12, Lemma 4.1] (which proceeds by taking a union bound over all possible steps where the numbers of U could appear) then ensures that

$$\mathbb{P}\big(U \subseteq I(m) \text{ and } \mathcal{S}_{\le m}\big) \le m^{|U|} \cdot \Big(\frac{1}{ND^{-\delta/4}}\Big)^{|U|} \le \pi^{|U|} \quad \text{with} \quad \pi := D^{-\frac{1}{r-1}+\delta/2}, \tag{4.49}$$

where we used that $m/N = D^{-\frac{1}{r-1}}(\log N)^{O(1)} \ll D^{-\frac{1}{r-1}+\delta/4}$, say. In particular, for any sequence of sets $(W_1, \dots, W_s) \in \mathcal{I}^s$ satisfying $W_i \cap W_j = \varnothing$ for $i \ne j$, using the definition (4.47) of Y_W it follows that

$$\mathbb{E}\Big(\prod_{i\in[s]} Y_{W_i}\Big) \le \pi^{s(r-2)} = \prod_{i\in[s]} \lambda_{W_i} \quad \text{with} \quad \lambda_W := \pi^{r-2}.$$

Furthermore, combining $|\mathcal{I}| \le kr^2$ with (4.49) and the definition (4.28) of k, it also follows that

$$\sum_{W\in\mathcal{I}} \lambda_W \le kr^2 \cdot \pi^{r-2} = 9\xi^{-1} r^2 (\log n)^{1-\frac{1}{r-1}} D^{\frac{3-r}{r-1}} D^{(r-2)\delta/2} \ll D^{\frac{1}{r-1}} D^{-4\delta} =: \mu.$$

To estimate the associated C-parameter of Lemma 60, note that any set $W \in \mathcal{I}$ satisfies

$$\big|\{W' \in \mathcal{I} \,:\, W \cap W' \neq \varnothing\}\big| \le \sum_{w\in W} \sum_{A\in\mathcal{A}_{N,r}:\{x,w\}\subseteq A} 2^{|A|} \le r \cdot r^2 \cdot 2^r =: C.$$

Using inequality (4.48), by invoking Lemma 60 with $z := \mu D^{\delta}/r^2 \ge D^{\Omega(1)}$ it follows that

$$\mathbb{P}\Big(\max_{0\le i\le m} |Y(i,x) \cap K| \ge D^{\frac{1}{r-1}} D^{-3\delta} \text{ and } \mathcal{S}_{\le m}\Big) \le \mathbb{P}(N_{K,x} \ge z) \le (e\mu/z)^{z/C} \le \exp\big(-N^{\Omega(1)}\big).$$

Taking a union bound over all of the at most $N \cdot N^2 = N^{O(1)}$ possible pairs (x, K) then establishes the claimed inequality (4.45).

4.3.3.3 Supermartingale arguments

We are now ready to prove $\mathbb{P}(\neg\mathcal{G}_{\le m}) \le \exp(-N^{\Omega(1)})$, by showing that $X_K^{\pm}(i) \ge 0$ is extremely unlikely. Here our main probabilistic tool is the following supermartingale inequality [54, Lemma 19], which allows us to exploit that $X_K^{\pm}(i)$ is defined (4.38) as the sum of a random variable and a deterministic function.

Lemma 61. *Let $(S_i)_{i\ge0}$ be a supermartingale adapted to the filtration $(\mathcal{F}_i)_{i\ge0}$. Assume that $S_i = X_i + D_i$, where X_i is $\mathcal{F}_i$-measurable and D_i is $\mathcal{F}_{\max\{i-1,0\}}$-measurable. Writing $\Delta X_i := X_{i+1} - X_i$, assume that $\max_{i\ge0} |\Delta X_i| \le C$ and $\sum_{i\ge0} \mathbb{E}(|\Delta X_i| \mid \mathcal{F}_i) \le V$. Then,* *for all* $z > 0$*,*

$$\mathbb{P}\big(S_i \ge S_0 + z \textit{ for some } i \ge 0\big) \;\le\; \exp\left(-\frac{z^2}{2C(V+z)}\right). \tag{4.50}$$

Turning to the details, we define the stopping time T as the minimum of m and the first step $i \ge 0$ where the 'good' event $\mathcal{G}_{\le i} \cap \mathcal{N}_{\le i}$ fails. For brevity, set $i \wedge T := \min\{i, T\}$. Recalling the definition (4.36) of $\mathcal{G}_{\le m}$, by the discussion below (4.38) it follows that

$$\mathbb{P}(\neg\mathcal{G}_{\le m}) \;\le\; \mathbb{P}(\neg\mathcal{S}_{\le m} \text{ or } \neg\mathcal{N}_{\le m}) + \sum_{\sigma\in\{+,-\}}\sum_{K\in\mathcal{A}_{N,k}} \mathbb{P}\left(X_K^{\sigma}(i\wedge T) \ge 0 \text{ for some } i \ge 0\right). \tag{4.51}$$

For any $K \in \mathcal{A}_{N,k}$, we initially have $S_K(0) = |K| = k$. By definition (4.38) of $X_K^{\pm}(i)$ we thus have

$$X_K^{\pm}(0\wedge T) = X_K^{\pm}(0) = \pm[|S_K(0)| - k] - ke(0) = -kD^{-\delta}.$$

Note that the estimates in Sections 4.3.3.1–4.3.3.2 apply for $0 \le i \le T-1$ (since then $0 \le i \le m-1$ and $\mathcal{G}_{\le i} \cap \mathcal{N}_{\le i}$ hold). The stopped sequence $S_i := X_K^{\sigma}(i \wedge T)$ thus is a supermartingale with $S_0 = -kD^{-\delta}$, to which Lemma 61 can be applied with $X_i = \sigma|S_K(i\wedge T)|$, $C = O(D^{\frac{1}{r-1}-3\delta})$ and $V = m \cdot kD^{\frac{1}{r-1}+\delta/2}/N = O(kD^{2\delta/3})$. Invoking inequality (4.50) with $z = kD^{-\delta}$, using the definition (4.28) of k and $D = \Theta(N)$ it follows that

$$\mathbb{P}\left(X_K^{\sigma}(i\wedge T) \ge 0 \text{ for some } i \ge 0\right) \le \exp\big(-\Omega\big(kD^{\delta/3}/D^{\frac{1}{r-1}}\big)\big) \le \exp\big(-N^{\Omega(1)}\big). \tag{4.52}$$

Inserting (4.52) and $|\mathcal{A}_{N,k}| \le N^2$ into (4.51), then $\mathbb{P}(\neg\mathcal{G}_{\le m}) \le \exp(-N^{\Omega(1)})$ follows from (4.35) and (4.45), which completes the proof of inequality (4.36) and thus Theorem 58, as discussed. □

CHAPTER 5

BOUNDS ON RAMSEY GAMES VIA ALTERATIONS

5.1 Background and main results

The probabilistic method is a widely-used tool in discrete mathematics. Many of its powerful approaches have been developed in the pursuit of understanding the graph Ramsey number $R(H,k)$, which is defined as the the minimum number n so that any n-vertex graph contains either a copy of H or an independent set of size k. For example, in 1947 Erdős pioneered the *random coloring approach* to obtain the lower bound $R(K_k,k) = \Omega(k2^{k/2})$, and in 1961 he developed the *alteration method* in order to obtain $R(K_3,k) = \Omega(k^2/(\log k)^2)$, see [27]. In 1975 and 1977 Spencer [106, 107] reproved these results via the *Lovász Local Lemma*, and also extended them to lower bounds on $R(H,k)$ for $H \in \{K_s, C_\ell\}$. In 1994 Krivelevich [71] further extended this to general graphs H via a new *(large-deviation based) alteration approach*, obtaining the lower bound

$$R(H,k) = \Omega\big((k/\log k)^{m_2(H)}\big) \quad \text{with} \quad m_2(H) := \max_{F \subseteq H}\Big(\mathbb{1}_{\{v_F \geq 3\}}\tfrac{e_F-1}{v_F-2} + \mathbb{1}_{\{F=K_2\}}\tfrac{1}{2}\Big), \tag{5.1}$$

where the implicit constants may depend on H (writing $v_F := |V(F)|$ and $e_F := |E(F)|$, as usual). By analyzing *(semi-random) H-free processes*, in 1995 Kim [67] and in 2010 Bohman–Keevash [12] have further improved the logarithmic factors in (5.1) for some graphs H such as triangles K_3, cliques K_s, and cycles C_ℓ. However, despite considerable effort, for $H \neq K_3$ the best known lower and upper bounds are still polynomial factors apart, see [12, 13, 37]. Unsurprisingly, to further advance the proof methods, the field has thus stretched in several directions. One such widely-studied direction investigates online graph Ramsey games, with the goal of understanding what happens to various Ramsey numbers when decisions need to be made online.

In this chapter, we present a *refinement of* the above-mentioned *widely-used alteration approaches of Erdős and Krivelevich* (see e.g., [29, 67, 72, 73, 111, 42, 11, 56, 20, 76, 39]) that enables analysis of online graph Ramsey games. As two concrete applications we consider Ramsey, Paper, Scissors games and online Ramsey numbers, each time extending recent bounds of Fox–He–Wigderson [39] and Conlon–Fox–Grinshpun–He [20].

5.1.1 Applications: Online Ramsey games

Our first application concerns the widely-studied *online Ramsey game* (see, e.g., [8, 75, 66, 19, 20]) that was introduced independently by Beck [8] and Kurek–Ruciński [75]. This is a game between two players, Builder and Painter, that starts with an infinite set $V = \{1, 2, \dots\}$ of isolated vertices. In each turn, Builder places an edge between two non-adjacent vertices from V, and Painter immediately colors it either red or blue. The *online Ramsey number* $\tilde{r}(H, k)$ is defined as the smallest number of turns N that Builder needs to guarantee the existence of either a red copy of H or a blue copy of K_k (regardless of Painter's strategy).

Our refined alteration approach enables us to prove a lower bound on $\tilde{r}(H, k)$ that, up to logarithmic factors, is about k times the best-known general lower bound for the usual Ramsey number $R(H, k)$, cf. (5.1).

Theorem 62 (Online Ramsey Game)**.** *If H is a graph with $e_H \geq 1$, then $\tilde{r}(H, k) = \Omega\big(k \cdot (k/\log k)^{m_2(H)}\big)$ as $k \to \infty$, where the implicit constant may depend on H.*

For general graphs H, Theorem 62 gives the best known lower bounds for online Ramsey numbers. For s-vertex cliques we obtain $\tilde{r}(K_s, k) = \Omega\big(k^{(s+3)/2}/(\log k)^{(s+1)/2}\big)$, which generalizes a recent bound of Conlon–Fox–Grinshpun–He [20, Theorem 1.4] for triangles, and also improves [20, Corollary 1.3] for small cliques. The best-known upper bounds $\tilde{r}(K_s, k) = O\big(k^s/(\log k)^{\lfloor s/2 \rfloor + 1}\big)$ differ by a polynomial factor for $s \geq 4$, (see [20, Theorem 5]), analogous to the known gaps for $R(K_s, k)$. It would be interesting to investigate whether the lower bound of Theorem 62 can be improved if one replaces our alteration approach by an H-free process [12] based approach or semi-random variants thereof [67, 56]; see also [20, Section 6].

Our second application concerns the fairly new *Ramsey, Paper, Scissors game* that was introduced by Fox–He–Wigderson [39]. For a graph H, this is a game between two players, Proposer and Decider, that starts with a finite set $V = \{1, 2, \ldots, n\}$ of n isolated vertices. In each turn, Proposer proposes a pair of non-adjacent vertices from V, and Decider simultaneously decides whether or not to add it as an edge to the current graph (without knowing which pair is proposed). Proposer cannot propose vertex-pairs that would form a copy of H together the current graph, nor vertex-pairs that have been proposed before. The ***RPS number*** $\mathrm{RPS}(H, n)$ is defined[1] as the largest number k for which Proposer can guarantee that, with probability at least $1/2$ (regardless of Decider's strategy), the final graph has an independent set of size k.

Our refined alteration approach enables us to prove an upper bound on $\mathrm{RPS}(H, n)$ for all *strictly* 2-*balanced graphs* H, i.e., which satisfy $m_2(H) > m_2(F)$ for all $F \subsetneq H$. This well-known class contains many graphs of interest, including cliques K_s, cycles C_ℓ, complete multipartite graphs $K_{t_1,\ldots,t_r}$, and hypercubes Q_d.

Theorem 63 (Ramsey, Paper, Scissors Game)**.** *If H is a strictly 2-balanced graph, then $\mathrm{RPS}(H, n) = O(n^{1/m_2(H)} \log n)$ as $n \to \infty$, where the implicit constant may depend on H.*

For all strictly 2-balanced graphs H, Theorem 63 gives the best known upper bounds for RPS numbers. For s-vertex cliques we obtain $\mathrm{RPS}(K_s, n) = O\big(n^{2/(s+1)} \log n\big)$, which generalizes the upper bound part of the very recent $\mathrm{RPS}(K_3, n) = \Theta(\sqrt{n} \log n)$ result of Fox–He–Wigderson [39]. It would be interesting to obtain good (and perhaps again

[1] For imperfect-information games such as Ramsey, Paper, Scissors (both players make simultaneous moves) one usually considers randomized strategies, see [89, pp. 14, 169], motivating why the definition of $\mathrm{RPS}(H, n)$ includes probability of winning.

matching) lower bonds on $\mathrm{RPS}(H, n)$ for other strictly 2-balanced graphs H.

5.1.2 Main tool: Refined alteration approach

To motivate our refined alteration approach, we shall review related arguments for the Ramsey bound (5.1). Here Erdős [27] and Krivelevich [71] use a binomial random graph $G_{n,p}$ with $n = \Theta((k/\log k)^{m_2(H)})$ vertices and edge-probability $p = \Theta((\log k)/k)$ to construct an n-vertex graph $G \subseteq G_{n,p}$ that (i) is H-free and (ii) contains at least one edge in each k-vertex subset K, which implies $R(H, k) > n$. Standard Chernoff bounds suggest that the number X_K of edges of $G_{n,p}$ inside K is around $\binom{k}{2}p$, so for property (ii) it intuitively suffices to show that the alteration from $G_{n,p}$ to G does not remove 'too many' edges from each k-vertex subset K.

To illustrate that this is a non-trivial task, let us consider the natural upper bound $e_H \cdot |\mathcal{H}_K|$ on the number of removed edges from K, where $\mathcal{H}_K$ denotes the collection of all H-copies that have at least one edge inside K. For any $\delta > 0$ it turns out that $\mathbb{P}(|\mathcal{H}_K| \geq \delta\binom{k}{2}p) \geq e^{-o(k)}$ due to 'infamous' upper tail [61, 101] behavior (see Appendix for the details). This lower bound not only rules out simple union bound arguments, but also suggests that one has to more carefully handle edges that are contained in multiple H-copies.

For triangles $H = K_3$, Erdős [27] overcame these difficulties in 1961 by a clever ad-hoc greedy alteration argument, showing that whp[2] the following works: If one sequentially traverses the edges of $G_{n,p}$ in any order, only accepting edges that do not create a triangle together with previously accepted edges, then the resulting 'accepted' subgraph $G \subseteq G_{n,p}$ satisfies (ii), and trivially (i). The fact that any edge-order works was exploited by Conlon et.al [20] and Fox et.al [39] in their analysis of triangle-based online Ramsey games.

To handle general graphs H, Krivelevich [71] developed in 1994 an elegant alteration argument, showing that whp the following works: If one constructs $G \subseteq G_{n,p}$ by

[2] In this chapter *whp* (with high probability) always means with probability tending to 1 as $k \to \infty$.

deleting all edges that are in some maximal (under inclusion) collection $\mathcal{C}$ of edge-disjoint H-copies in $G_{n,p}$, then this (a) removes less than $X_K \approx \binom{k}{2}p$ edges from each k-vertex subset K, and (b) yields an H-free graph by maximality of $\mathcal{C}$, establishing both (ii) and (i). Unfortunately, this slick maximality argument is hard to adapt to online Ramsey games, where players cannot foresee whether in future turns a given edge will be contained in an H-copy or not.

Our refined alteration approach overcomes the above-discussed difficulties, by showing that whp the desired properties (i) and (ii) remain valid even if one deletes *all* edges from $G_{n,p}$ that are in some H-copy (and not just some carefully chosen *subset* of these edges, as in the influential alteration approaches of Erdős and Krivelevich, cf. [27, 29, 71, 67, 72, 73, 111, 42, 11, 56, 20, 76, 39]). To state our main technical result, let Y_K denote the number of edges in $E(G_{n,p}[K])$ that are in some H-copy of $G_{n,p}$. Recall that $X_K = |E(G_{n,p}[K])|$.

Theorem 64 (Main technical result)**.** *Let H be a strictly 2-balanced graph. Then, for any $\delta > 0$, the following holds for all $C \geq C_0(\delta, H)$ and $0 < c \leq c_0(C, \delta, H)$. Setting $n := \lfloor c(k/\log k)^{m_2(H)} \rfloor$ and $p := C(\log k)/k$, whp $G_{n,p}$ satisfies $Y_K \leq \delta\binom{k}{2}p$ for all k-vertex sets K.*

Remark 65. *For any $\delta \in (0, 1]$, the following holds for all $C \geq C_0(\delta, H)$ and $c > 0$. Setting n and p as in Theorem 64, whp $G_{n,p}$ satisfies $X_K \geq (1 - \delta)\binom{k}{2}p$ for all k-vertex sets K.*

As discussed, our basic alteration idea is to construct $G \subseteq G_{n,p}$ by deleting all edges that are in some H-copy of $G_{n,p}$, so (i) holds trivially, and for suitable n, p then Theorem 64 and Remark 65 suggest that whp $|E\big(G[K]\big)| = X_K - Y_K \geq (1 - 2\delta)\binom{k}{2}p > 0$ for all k-vertex subsets K, establishing (ii). It is noteworthy that the largest independent sets of G (which have size less than k) are not much larger than those of $G_{n,p}$, which are well-known to be of order $\log(np)/p = \Theta(k)$ for $p \gg n^{-1}$ and thus $m_2(H) > 1$, see [60, Section 7.1].

As we shall see in Section 5.2, variants of the above-discussed alteration argument carry over to certain online Ramsey games (where it will be useful that we can allow for arbitrary deletion of edges in H-copies), and we believe that the bound on Y_K from Theorem 64 might also be useful in other contexts. We remark that the restriction to strictly 2-balanced graphs in Theorem 64 is often immaterial, since for (5.1) and related Ramsey bounds one can usually obtain the desired general bound by simply forbidding a strictly 2-balanced subgraph $H_0 \subseteq H$ with $m_2(H_0) = m_2(H)$, cf. Section 5.2.2. Finally, in Section 5.4 we also discuss some further extensions of our alteration approach, including variants which forbid multiple hypergraphs.

5.1.3 Organization

In Section 5.2 we prove the discussed online Ramsey game results (Theorems 62–63) using the main technical result of our refined alteration approach (Theorem 64), which we subsequently prove in Section 5.3. Finally, in Section 5.4 we discuss some extensions of our alteration approach, including hypergraph variants.

5.2 Online Ramsey games

5.2.1 Ramsey, Paper, Scissors: Proof of Theorem 63

The following argument is based on a Decider strategy that randomly accepts edges (this strategy is completely oblivious, i.e., does not require knowledge of any proposed or accepted edges).

Proof of Theorem 63. For $\delta := 1/4$ we choose $C > 0$ large enough and then $c > 0$ small enough so that Remark 65 and Theorem 64 both apply to $G_{n,p}$ with $n := \lfloor c(k/\log k)^{m_2(H)} \rfloor$ and $p := C(\log k)/k$. We shall analyze the following strategy: in each turn Decider accepts the (unknown) proposed vertex-pair as an edge independently with probability p. Let G denote the resulting final graph at the end of the game, i.e., which contains all accepted edges. Since all edges that do not create H-copies are eventually proposed, there is a natural coupling between $G_{n,p}$ and G which satisfies the following two properties:

(a) that $E(G) \subseteq E(G_{n,p})$, and (b) that every edge in $E(G_{n,p}) \setminus E(G)$ is contained in an H-copy of $G_{n,p}$. Invoking Theorem 64 and Remark 65, it follows that this coupling satisfies the following whp: for any k-vertex set K of G we have

$$\left|E\big(G[K]\big)\right| \;\geq\; X_K - Y_K \;\geq\; (1-2\delta)\tbinom{k}{2}p = \tfrac{1}{2}\tbinom{k}{2}p > 0,$$

which implies that the final graph G has whp no independent set of size k. It follows that $\mathrm{RPS}(H, n) < k = O(n^{1/m_2(H)} \log n)$ as $n \to \infty$ (where the implicit constant depends on H). $\square$

5.2.2 Online Ramsey numbers: Proof of Theorem 62

The following argument is based on a Painter strategy that attempts to randomly color edges between high-degree vertices. The analysis is complicated by the fact that the game is played on an infinite set $V = \{1, 2, \ldots\}$ of vertices, which requires some care in the coupling and union bound arguments below.

Proof of Theorem 62. For convenience we first suppose that H is strictly 2-balanced. For $\delta := 1/8$ we choose $C \geq 64e_H$ large enough and then $c > 0$ small enough so that Theorem 64 applies to $G_{n,p}$ with $n := \lfloor c(k/\log k)^{m_2(H)} \rfloor$ and $p := C(\log k)/k$. Set $L := \lfloor (k-1)/4 \rfloor$. At any moment of the game, we define $U \subseteq V$ as the set of all vertices that, in the current graph, are adjacent to at least L edges placed by builder (to clarify: the growing vertex set U is updated at the end of each turn).

We shall analyze the following strategy: Painter's default color is blue, but if an edge $e = \{x, y\}$ is placed inside U, then Painter does the following independently with probability p ($\star$): it colors the edge e red, unless this would create a red H-copy (†), in which case the edge e is still colored blue. By construction there are no red H-copies, and blue cliques K_k can only appear inside U (since all vertices in copy of K_k must be adjacent to at least $k-1 > L$ vertices). To prove $\tilde{r}(H, k) > N := \lfloor L \cdot n/2 \rfloor = \Omega(k \cdot (k/\log k)^{m_2(H)})$

as $k \to \infty$ (with implicit constants depending on H), by the usual reasoning it remains to show that after N steps there are whp no blue cliques K_k inside U. Let $\mathcal{K}$ denote the collection of all k-vertex sets $K \subseteq U$ after N steps. Intuitively, the plan is to show that, inside each vertex set $K \in \mathcal{K}$ that can become a blue clique K_k, there are more red-coloring attempts ($\star$) than 'discarded' red-coloring attempts (†), which enforces a red edge inside K.

Turning to details, note that $|U| \le 2N/L \le n$ during the first N steps. Using the order in which vertices enter U (breaking ties using lexicographic order), at any moment during the first N steps we thus obtain an injection $\Phi : U \mapsto \{1, \ldots, n\} = V(G_{n,p})$. After N steps, we abbreviate this injection by Φ_N, and write $\Phi_N(K) := \{\Phi_N(v) : v \in K\}$. Define $\mathcal{B}_K$ as the event that, during the first N steps, the number of 'discarded' red-coloring attempts (†) inside K is at most $\frac{1}{8}\binom{k}{2}p$. There is a natural turn-by-turn inductive coupling between $G_{n,p}$ and Painter's strategy, where the red-coloring attempt ($\star$) occurs if $\Phi(e) := \{\Phi(x), \Phi(y)\}$ is an edge of $G_{n,p}$. A moments thought reveals that, during the first N steps, under this coupling the total number of 'ignored' red-colorings (†) inside $K \in \mathcal{K}$ is at most $Y_{\Phi_N(K)}$ defined with respect to $G_{n,p}$ (since (†) can only happen when a red-coloring of $e \subseteq K$ creates a red H-copy, which under the coupling implies that $\Phi(e) \subseteq \Phi(K)$ is contained in an H-copy of $G_{n,p}$). Applying Theorem 64 with $\delta = 1/8$ to $G_{n,p}$, using the described coupling and $|\Phi_N(K)| = |K| = k$ it then follows that, whp, the event $\mathcal{B}_K$ occurs for all $K \in \mathcal{K}$.

Intuitively, we shall next show that, for all k-vertex sets $K \in \mathcal{K}$ that contain $\binom{k}{2}$ edges (a prerequisite for having a blue clique K_k inside K), the number of red-coloring attempts ($\star$) inside K is at least $\frac{1}{4}\binom{k}{2}p$. To make this precise, define $\mathcal{T}_K$ as the event that builder places less than $\binom{k}{2}$ edges inside K during the first N steps. Let $X_K^\star$ denote the number of red-coloring attempts ($\star$) inside K during the first N steps, and define $\mathcal{A}_K$ as the event that $X_K^\star \ge \frac{1}{4}\binom{k}{2}p$. Let $\mathcal{K}'$ denote the collection of all k-vertex sets $K' \subseteq V(G_{n,p})$. Since Φ_N defines an injection from $\mathcal{K}$ to $\mathcal{K}'$, writing $\Phi_N^{-1}(K') := \{v \in V : \Phi_N(v) \in K'\}$ it follows that

$$\mathbb{P}(\neg\mathcal{A}_K \cap \neg\mathcal{T}_K \text{ for some } K \in \mathcal{K}) \;\le\; \sum\nolimits_{K' \in \mathcal{K}'} \mathbb{P}\Big(X^\star_{\Phi^{-1}_N(K')} \le \tfrac{1}{4}\tbinom{k}{2}p \;\text{ and }\; \neg\mathcal{T}_{\Phi^{-1}_N(K')}\Big). \tag{5.2}$$

Fix $K' \in \mathcal{K}'$, and set $K := \Phi^{-1}_N(K')$. Note that, by checking in each turn for red-coloring attempts $(\star)$ inside $\Phi^{-1}(K') := \{v \in V : \Phi(v) \in K'\}$, we can determine $X^\star_K$ without knowing Φ^{-1}_N in advance. Furthermore, since every vertex is adjacent to at most L vertices before entering U, the event $\neg\mathcal{T}_K$ implies that during the first N steps at least $\binom{k}{2} - |K| \cdot L \ge \frac{1}{2}\binom{k}{2}$ red-coloring attempts $(\star)$ happen inside K, each of which is (conditional on the history) successful with probability p. It follows that $X^\star_K$ stochastically dominates a binomial random variable $Z \sim \mathrm{Bin}\big(\lceil \frac{1}{2}\binom{k}{2}\rceil, p\big)$, unless the event $\mathcal{T}_K$ occurs. Noting $kp = C\log k \ge 64e_H \log k$ and $n \ll k^{e_H}$, by invoking standard Chernoff bounds (see, e.g., [60, Theorem 2.1]) it then follows that

$$\begin{aligned}\mathbb{P}\Big(X^\star_{\Phi^{-1}_N(K')} \le \tfrac{1}{4}\tbinom{k}{2}p \;\text{ and }\; \neg\mathcal{T}_{\Phi^{-1}_N(K')}\Big) \;&\le\; \mathbb{P}\Big(Z \le \tfrac{1}{4}\tbinom{k}{2}p\Big)\\ &\le\; \exp\Big(-\tbinom{k}{2}p/16\Big) \ll k^{-e_H k} \ll n^{-k}.\end{aligned} \tag{5.3}$$

Combining (5.2)–(5.3) with $|\mathcal{K}'| \le n^k$, we readily infer that, whp, the event $\mathcal{A}_K \cup \mathcal{T}_K$ occurs for all $K \in \mathcal{K}$.

To sum up, the following holds whp after N steps: every k-vertex set $K \subseteq U$ contains either (a) at least $\frac{1}{4}\binom{k}{2}p - \frac{1}{8}\binom{k}{2}p = \frac{1}{4}\binom{k}{2}p > 0$ red edges, or (b) less than $\binom{k}{2}$ edges in total. Both possibilities prevent a blue clique K_k inside K, and so the desired lower bound $\tilde{r}(H,k) > N$ follows (as discussed above).

Finally, in the remaining case where H is not strictly 2-balanced, we pick a minimal subgraph $H_0 \subsetneq H$ with $m_2(H_0) = m_2(H)$. It is straightforward to check that, by construction, H_0 is strictly 2-balanced. Furthermore, since any H_0-free graph is also H-free, we also have $\tilde{r}(H,k) \ge \tilde{r}(H_0,k)$. Repeating the above proof with H replaced by H_0 then gives the claimed lower bound on $\tilde{r}(H,k)$. $\square$

5.3 Refined alteration approach

5.3.1 Bounding Y_K: Proof of Theorem 64

For Theorem 64 the core strategy is to approximate Y_K by more tractable auxiliary random variables, inspired by ideas from [61, 118, 120, 103]. In particular, we expect that the main contribution to Y_K should come from H-copies that share exactly two vertices and one edge with K; in the below proof we denote the collection of such 'good' H-copies by $\mathcal{H}_K^*$. Note that when multiple good H-copies from $\mathcal{H}_K^*$ contain some common edge f inside K, they together only contribute one edge to Y_K. It follows that, by arbitrarily selecting one 'representative' copy $H_f \in \mathcal{H}_K^*$ for each relevant edge f, we should obtain a sub-collection $\mathcal{H} \subseteq \mathcal{H}_K^*$ of good H-copies with $|\mathcal{H}| \approx Y_K$. The H-copies in $\mathcal{H}$ share no edges inside K by construction, and it turns out that all other types of edge-overlaps are 'rare', i.e., make a negligible contribution to Y_K. We thus expect that there is an edge-disjoint sub-collection $\mathcal{H}' \subseteq \mathcal{H} \subseteq \mathcal{H}_K^*$ of good H-copies with $|\mathcal{H}'| \approx |\mathcal{H}| \approx Y_K$, and here the crux is that the upper tail of $|\mathcal{H}'|$ is much easier to estimate than the upper tail of Y_K (see Claim 67 below). The following proof implements a rigorous variant of the above-discussed heuristic ideas.

Proof of Theorem 64. Noting that the claimed bounds are trivial when $m_2(H) \leq 1$ (since then there are no k-vertex sets K in $G_{n,p}$ due to $n \ll k$), we may henceforth assume $m_2(H) > 1$.

Fix a k-vertex set K. Let $\mathcal{H}_K$ denote the collection of all H-copies in $G_{n,p}$ that have at least one edge inside K, and let $\mathcal{H}_K^* \subseteq \mathcal{H}_K$ denote the sub-collection of H-copies that moreover share exactly two vertices with K. Let $\mathcal{I}_K$ denote a size-maximal collection of edge-disjoint $H \in \mathcal{H}_K^*$. Clearly $|\mathcal{I}_K| \leq Y_K$, and Claim 66 below establishes a related upper bound. Let $\mathcal{T}_K$ denote a size-maximal collection of edge-disjoint $H \in \mathcal{H}_K \setminus \mathcal{H}_K^*$. Let $\mathcal{P}_K$ denote a size-maximal collection of edge-disjoint $H_1 \cup H_2$ with distinct $H_1, H_2 \in \mathcal{H}_K^*$ that satisfy $|E(H_1) \cap E(H_2)| \geq 1$ and $V(H_1) \cap K \neq V(H_2) \cap K$. Let $\Delta_{H,f}$ denote the number of H-copies in $G_{n,p}$ that contain the edge f, and define Δ_H as the maximum of $\Delta_{H,f}$ over all $f \in E(K_n)$.

Claim 66. *We have $Y_K \leq |\mathcal{I}_K| + 2e_H^2(|\mathcal{T}_K| + |\mathcal{P}_K|)\Delta_H$.*

Proof of Claim 66. We divide the H-copies in $\mathcal{H}_K$ into two disjoint groups: those which share at least one edge with some $H \in \mathcal{T}_K$ or $H_1 \cup H_2 \in \mathcal{P}_K$, and those which do not; we denote these two groups by $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. For $j \in \{1, 2\}$, let $\mathcal{E}_j$ denote the collection of edges from K that are contained in at least one H-copy from $\mathcal{H}_j$. Note that $Y_K \leq |\mathcal{E}_1| + |\mathcal{E}_2|$ and $|\mathcal{E}_1| \leq e_H|\mathcal{H}_1| \leq e_H \cdot (e_H|\mathcal{T}_K| + 2e_H|\mathcal{P}_K|)\Delta_H$. Turning to $\mathcal{E}_2$, by maximality of $\mathcal{T}_K$ and $\mathcal{P}_K$ we infer the following two properties of $\mathcal{H}_2$: (a) all H-copies intersect with K in exactly two vertices, so $\mathcal{H}_2 \subseteq \mathcal{H}_K^*$, and (b) any two distinct H-copies are edge-disjoint, unless they both intersect K in the same two vertices. For each $f \in \mathcal{E}_2 \subseteq \binom{K}{2}$ we now arbitrarily select one H-copy from $\mathcal{H}_2$ that contains f. By properties (a)–(b) of $\mathcal{H}_2$ and size-maximality of $\mathcal{I}_K$, this yields a sub-collection $\mathcal{H}_2' \subseteq \mathcal{H}_2 \subseteq \mathcal{H}_K^*$ of edge-disjoint H-copies satisfying $|\mathcal{E}_2| = |\mathcal{H}_2'| \leq |\mathcal{I}_K|$, and the claim follows. $\square$

The remaining upper tail bounds for $|\mathcal{I}_K|$, $|\mathcal{T}_K|$, $|\mathcal{P}_K|$ and Δ_H hinge on the following four key estimates. First, $m_2(H) > 1$ and strictly 2-balancedness of H imply $m_2(H) = (e_H - 1)/(v_H - 2)$, so that

$$n^{v_H-2}p^{e_H-1} = \left(np^{m_2(H)}\right)^{v_H-2} \leq \left(cC^{m_2(H)}\right)^{v_H-2}. \tag{5.4}$$

Second, $n = k^{m_2(H)-o(1)}$ and $m_2(H) > 1$ imply that there is $\tau = \tau(H) > 0$ such that

$$\tfrac{k}{n} \ll k^{-\tau}/\log k. \tag{5.5}$$

Third, using $p = k^{-1+o(1)}$ and strictly 2-balancedness of H (implying that $(e_J - 1)/(v_J - 2) < m_2(H)$ for all $J \subsetneq H$ with $e_J \geq 2$), it follows that there is $\gamma =$

$\gamma(H) > 0$ such that

$$n^{v_J-2}p^{e_J-1} = \left(np^{(e_J-1)/(v_J-2)}\right)^{v_J-2} \gg k^{\gamma} \qquad \text{for all } J \subsetneq H \text{ with } e_J \geq 2. \tag{5.6}$$

The below-claimed fourth estimate can be traced back to Erdős and Tetali [35]; we include an elementary proof for self-containedness (see [118, Section 2] for related estimates that also allow for overlapping edge-sets).

Claim 67. *Let $\mathcal{S}$ be a collection of edge-subsets from $E(K_n)$. Define Z as the largest number of disjoint edge-sets from $\mathcal{S}$ that are present in $G_{n,p}$. Then $\mathbb{P}(Z \geq x) \leq (e\mu/x)^x$ for all $x > \mu := \sum_{\beta\in\mathcal{S}} \mathbb{P}(\beta \subseteq E(G_{n,p}))$.*

Proof of Claim 67. Set $s := \lceil x \rceil \geq 1$. Exploiting edge-disjointness and $s! \geq (s/e)^s$, it follows that

$$\begin{aligned}\mathbb{P}(Z \geq x) &\leq \sum_{\substack{\{\beta_1,\ldots,\beta_s\}\subseteq\mathcal{S}:\\ \text{all edge-disjoint}}} \underbrace{\mathbb{P}\big(\beta_1 \cup \cdots \cup \beta_s \subseteq E(G_{n,p})\big)}_{=\prod_{1\leq i\leq s}\mathbb{P}(\beta_i\subseteq E(G_{n,p}))} \\ &\leq \frac{1}{s!}\left(\sum_{\beta\in\mathcal{S}} \mathbb{P}\big(\beta \subseteq E(G_{n,p})\big)\right)^s \leq \big(e\mu/s\big)^s,\end{aligned}$$

which completes the proof by noting that the function $s \mapsto (e\mu/s)^s$ is decreasing for positive $s \geq \mu$. $\square$

We are now ready to bound the probability that $|\mathcal{I}_K|$ is large. Since H is strictly 2-balanced, it contains no isolated vertices and thus is uniquely determined by its edge-set. This enables us to apply Claim 67 to $|\mathcal{I}_K| = Z$ (as $\mathcal{I}_K$ is a size-maximal collection of edge-disjoint H-copies from $\mathcal{H}^*_K$). Using estimate (5.4), it is routine to see that, for $c \leq c_0(C, \delta, H)$, the associated parameter μ from Claim 67 satisfies

$$\mu \leq O\big(k^2 n^{v_H-2} \cdot p^{e_H}\big) \leq \tbinom{k}{2}p \cdot \Theta(n^{v_H-2}p^{e_H-1}) \leq \tfrac{\delta}{2e^2}\tbinom{k}{2}p. \tag{5.7}$$

Noting $\delta kp = \delta C \log k$ and $n \ll k^{e_H}$, now Claim 67 (with $Z = |\mathcal{I}_K|$) implies that, for $C \geq C_0(\delta, H)$,

$$\mathbb{P}\Big(|\mathcal{I}_K| \geq \tfrac{\delta}{2}\tbinom{k}{2}p\Big) \leq \left(\frac{e\mu}{\frac{\delta}{2}\binom{k}{2}p}\right)^{\frac{\delta}{2}\binom{k}{2}p} \leq e^{-\frac{\delta}{2}\binom{k}{2}p} \ll k^{-e_H k} \ll n^{-k}. \tag{5.8}$$

Next, we similarly use Claim 67 to bound the probability that $|\mathcal{T}_K|$ is large. For the associated parameter μ we shall proceed similar to (5.7) above: using estimates (5.4)–(5.5), for $c \leq c_0(C, \delta, H)$ we obtain

$$\mu \leq O\big(k^3 n^{v_H-3} \cdot p^{e_H}\big) \leq \tbinom{k}{2}p \cdot \tfrac{k}{n} \cdot \Theta(n^{v_H-2}p^{e_H-1}) \leq k^{-\tau} \cdot \tfrac{\delta}{e}\tbinom{k}{2}p/\log k. \tag{5.9}$$

With similar considerations as for (5.8) above, for $C \geq C_0(\tau, \delta, H)$ Claim 67 (with $Z = |\mathcal{T}_K|$) then yields

$$\mathbb{P}\Big(|\mathcal{T}_K| \geq \delta\tbinom{k}{2}p/\log k\Big) \leq k^{-\tau\delta\binom{k}{2}p/\log k} = e^{-\tau\delta\binom{k}{2}p} \ll k^{-e_H k} \ll n^{-k}. \tag{5.10}$$

We shall analogously use Claim 67 to bound the probability that $|\mathcal{P}_K|$ is large. For the associated parameter μ, the basic idea is to distinguish all possible subgraphs $J \subsetneq H$ in which the relevant $H_1, H_2 \in \mathcal{H}_K^*$ can intersect. Also taking into account the number of vertices which H_1 and H_2 have inside K, i.e., $\big|\big(V(H_1) \cup V(H_2)\big) \cap K\big| \in \{3, 4\}$, by definition of $\mathcal{P}_K$ it now follows via estimates (5.4)–(5.6) that

$$\begin{aligned}\mu &\leq \sum_{J \subsetneq H : e_J \geq 1} O\Big(k^3 n^{2(v_H-2)-(v_J-1)} \cdot p^{2e_H-e_J} + k^4 n^{2(v_H-2)-v_J} \cdot p^{2e_H-e_J}\Big) \\ &\leq \tbinom{k}{2}p \cdot \Big[\tfrac{k}{n} + \big(\tfrac{k}{n}\big)^2\Big] \cdot \sum_{J \subsetneq H : e_J \geq 1} \frac{\Theta\big((n^{v_H-2}p^{e_H-1})^2\big)}{n^{v_J-2}p^{e_J-1}} \leq k^{-\tau} \cdot \tfrac{\delta}{e}\tbinom{k}{2}p/\log k.\end{aligned} \tag{5.11}$$

(To clarify: in (5.11) above we used that (5.6) implies $n^{v_J-2}p^{e_J-1} \geq 1$ for all $J \subsetneq H$ with $e_J \geq 1$.) Similarly to inequalities (5.8) and (5.10), for $C \geq C_0(\tau, \delta, H)$ now Claim 67 (with $Z = |\mathcal{P}_K|$) yields

$$\mathbb{P}\Big(|\mathcal{P}_K| \geq \delta\tbinom{k}{2}p/\log k\Big) \leq k^{-\tau\delta\binom{k}{2}p/\log k} = e^{-\tau\delta\binom{k}{2}p} \ll k^{-e_H k} \ll n^{-k}. \tag{5.12}$$

Finally, combining (5.8), (5.10) and (5.12) with Claim 66, a standard union bound argument gives

$$\mathbb{P}\Big(Y_K \geq \delta\tbinom{k}{2}p \cdot \big(\tfrac{1}{2} + 4e_H^2\Delta_H/\log k\big) \text{ for some } k\text{-vertex set } K\Big) \leq \tbinom{n}{k} \cdot o(n^{-k}) = o(1). \tag{5.13}$$

To complete the proof of (5.13), it thus remains to show that, for $c \leq c_0(C, H)$, we have

$$\mathbb{P}\Big(\Delta_H \geq (\log k)/(8e_H^2)\Big) = o(1). \tag{5.14}$$

Using (5.4), (5.6) and $n \ll k^{e_H}$, this upper tail estimate for $\Delta_H = \max_f \Delta_{H,f}$ follows routinely from standard concentration inequalities such as [120, Theorem 32], but we include an elementary proof for self-containedness (based on ideas from [110, 118]). Turning to the proof of (5.14), let $\Delta_{H,f,g}$ denote the number of H-copies in $G_{n,p}$ that contain the edges $\{f, g\}$, and define $\Delta_H^{(2)}$ as the maximum of $\Delta_{H,f,g}$ over all distinct $f, g \in E(K_n)$. We call an r-tuple $(H_1, \ldots, H_r)$ of H-copies an (r, f, g)-*star* if each H_j contains the edges $\{f, g\}$ and satisfies $H_j \not\subseteq H_1 \cup \cdots \cup H_{j-1}$. Define $Z_{r,f,g}$ as the number of (r, f, g)-stars $(H_1, \ldots, H_r)$ that are present in $G_{n,p}$. Summing over all $(r+1, f, g)$-stars $(H_1, \ldots, H_{r+1})$, by noting that the intersection of H_{r+1} with $F_r := H_1 \cup \cdots \cup H_r$ is isomorphic to some proper subgraph $J \subsetneq H$ containing at least $e_J \geq 2$ edges, using estimates (5.4) and (5.6) it then is routine to see that, for $1 \leq r \leq r_0 := 1 + \lceil (v_H e_H + 4e_H)/\gamma \rceil$, we have

$$\begin{aligned}
\mathbb{E}\, Z_{r+1,f,g} &= \sum_{(H_1,\ldots,H_{r+1})} p^{e_{H_1\cup\cdots\cup H_{r+1}}} = \sum_{(H_1,\ldots,H_r)} p^{e_{F_r}} \sum_{H_{r+1}} p^{e_H - e_{H_{r+1}\cap F_r}} \\
&\leq \sum_{(H_1,\ldots,H_r)} p^{e_{F_r}} \cdot \sum_{J\subsetneq H: e_J \geq 2} O\Big((v_H r)^{v_J} n^{v_H - v_J} \cdot p^{e_H - e_J}\Big) \leq \mathbb{E}\, Z_{r,f,g} \cdot k^{-\gamma}.
\end{aligned}$$

Since trivially $\mathbb{E}\, Z_{1,f,g} = O(n^{v_H})$, using $n \ll k^{e_H}$ we infer $\mathbb{E}\, Z_{r_0,f,g} \le k^{v_H e_H - (r_0-1)\gamma} \le k^{-4e_H} \ll n^{-4}$. Consider a maximal length (r,f,g)-star $(H_1,\ldots,H_r)$ in $G_{n,p}$, and note that in $G_{n,p}$ any H-copy containing the edges $\{f,g\}$ is completely contained in $H_1 \cup \cdots \cup H_r$ (by length maximality), so that $\Delta_{H,f,g} \le (e_H r)^{e_H}$ holds (using that H is uniquely determined by its edge-set). For $D := (e_H r_0)^{e_H}$ it follows that

$$\begin{aligned}\mathbb{P}\Big(\Delta_H^{(2)} \ge D\Big) &\le \sum_{f \neq g} \mathbb{P}\big(\Delta_{H,f,g} \ge D\big) \\ &\le \sum_{f\neq g} \mathbb{P}(Z_{r_0,f,g} \ge 1) \le \sum_{f \neq g} \mathbb{E}\, Z_{r_0,f,g} \le \tbinom{n}{2}^2 \cdot o(n^{-4}) = o(1).\end{aligned} \tag{5.15}$$

With an eye on $\Delta_{H,f}$, let $\mathcal{H}_f$ denote the collection of all H-copies in K_n that contain the edge f. We pick a subset $\mathcal{I} \subseteq \mathcal{H}_f$ of H-copies in $G_{n,p}$ that is size-maximal subject to the restriction that all H-copies are edge-disjoint after removing the common edge f. For any $H' \in \mathcal{H}_f$, note that in $G_{n,p}$ there are a total of at most $e_H \Delta_H^{(2)}$ copies of H that share f and at least one additional edge with H'. Hence $\Delta_{H,f} \ge (\log k)/(8e_H^2)$ and $\Delta_H^{(2)} \le D$ imply $|\mathcal{I}| \ge \lceil (\log k)/A \rceil =: z$ for $A := 8e_H^3 D$ (by maximality of $\mathcal{I}$). As the union of all H-copies in $\mathcal{I}$ contains exactly $1 + (e_H - 1)|\mathcal{I}|$ edges, using $\binom{m}{z} \le (em/z)^z$ and $|\mathcal{H}_f| = O(n^{v_H - 2})$ it follows that

$$\mathbb{P}\Big(\Delta_{H,f} \ge (\log k)/(8e_H^2) \text{ and } \Delta_H^{(2)} \le D\Big) \le \binom{|\mathcal{H}_f|}{z} \cdot p^{1+(e_H-1)z} \le \left(\frac{O(n^{v_H-2}p^{e_H-1})}{z}\right)^z. \tag{5.16}$$

Using estimate (5.4), for $c \le c_0(A, C, H)$ the right-hand side of (5.16) is at most $(\log k)^{-(\log k)/A} \ll k^{-2e_H}$. Recalling $n \ll k^{e_H}$, by taking a union bound over all edges $f \in E(K_n)$ it then follows that

$$\mathbb{P}\Big(\Delta_H \ge (\log k)/(8e_H^2) \text{ and } \Delta_H^{(2)} \le D\Big) \le \tbinom{n}{2} \cdot o(k^{-2e_H}) = o(1), \tag{5.17}$$

which together with (5.15) completes the proof of estimate (5.14) and thus Theorem 64.

□

The above proof of (5.14) can easily be sharpened to $\mathbb{P}\big(\Delta_H \geq B(\log k)/\log\log k\big) = o(1)$ for suitable $B = B(H) > 0$, see (5.16)–(5.17). Together with the proof of (5.13) and $|\mathcal{I}_K| \leq Y_K$, this implies that whp $Y_K = |\mathcal{I}_K| + o\big(\delta\binom{k}{2}p\big)$ for all k-vertex sets K, which intuitively suggests that Y_K is well-approximated by $|\mathcal{I}_K|$.

5.3.2 Bounding X_K: Proof of Remark 65

Remark 65 follows easily from Chernoff bounds; we include the routine details for completeness.

Proof of Remark 65. Noting $\delta^2 kp = \delta^2 C \log k$ and $n \ll k^{e_H}$, by invoking standard Chernoff bounds (see, e.g., [60, Theorem 2.1]) it follows, for $C \geq C_0(\delta, H)$ large enough, that

$$\mathbb{P}\big(X_K \leq (1-\delta)\tbinom{k}{2}p\big) \;\leq\; \exp\Big(-\delta^2\tbinom{k}{2}p/2\Big) \ll k^{-e_H k} \ll n^{-k}. \qquad (5.18)$$

Taking a union bound over all set k-vertex sets K completes the proof of Remark 65. □

5.4 Extensions

In applications of the alteration approach outlined in Section 5.1.2, it often is beneficial to keep track of further properties of the resulting H-free n-vertex graph $G \subseteq G_{n,p}$, including vertex-degrees and the number of edges (see, e.g., [29, Section 3], [11, Section 2], and [76, Section 5.1]). Using the arguments and intermediate results from Section 5.3.1, oftentimes it is routine to show that G resembles a random graph $G_{n,p}$ in many ways. For example, with standard results for $G_{n,p}$ in mind, the following simple lemma intuitively implies that whp the resulting G is approximately np regular, with about $\binom{n}{2}p$ edges. (Note that $k \gg n$ when $m_2(H) \leq 1$.)

Lemma 68. *Let H be a strictly 2-balanced graph with $m_2(H) > 1$. Define Y as the number of H-copies in $G_{n,p}$, and define Y_v as the number of H-copies in $G_{n,p}$ that contain the* *vertex v. For any $\delta > 0$, the following holds for all $C \geq C_0(\delta, H)$ and $0 < c \leq c_0(C, \delta, H)$. Setting n and p as in Theorem 64, whp $G_{n,p}$ satisfies $Y_v \leq \delta np$ for all vertices v, and $Y \leq \delta\binom{n}{2}p$.*

Proof. Since $m_2(H) > 1$ implies $v_H \geq 3$, noting $Y = \sum_{v \in [n]} Y_v/v_H$ it suffices to prove the claimed bounds on the Y_v. Fix a vertex v. Similar to estimate (5.7), using (5.4) it is standard to see that the expected number of H-copies containing v is at most $\mu \leq O(n^{v_H-1}p^{e_H}) \leq \frac{\delta}{e^2}np$ for $c \leq c_0(C, \delta, H)$. Furthermore, if $\Delta_H \leq (\log k)/(8e_H^2)$ holds (see (5.14) in Section 5.3.1), then any H-copy edge-intersects a total of at most $e_H \cdot \Delta_H < \log k$ many H-copies, say. Applying the upper tail inequality [56, Theorem 15] instead of Claim 67, using $\delta np = \delta c C k^{m_2(H)-1-o(1)} \gg (\log k)^2$ it then is, similar to (5.8) and (5.17), routine to see that

$$\mathbb{P}\Big(Y_v \geq \delta np \ \text{ and } \ \Delta_H \leq (\log k)/(8e_H^2)\Big) \leq \left(\frac{e\mu}{\delta np}\right)^{\delta np/\log k} \leq e^{-\delta np/\log k} \ll n^{-1}.$$

Taking a union bound over all vertices v now completes the proof together with estimate (5.14). □

It is straightforward, and useful for many applications (see, e.g., [72, 42, 11]), to extend the alteration approach to r-uniform hypergraphs, where every edge contains $r \geq 2$ vertices. Indeed, to forbid a given r-uniform hypergraph H, similarly to the graph case ($r = 2$) discussed in Section 5.1.2, here the idea is to delete edges from a *binomial r-uniform hypergraph* $G^{(r)}_{n,p}$ (where each of the $\binom{n}{r}$ possible edges appears independently with probability p) to construct an n-vertex r-uniform hypergraph $G \subseteq G^{(r)}_{n,p}$ that is H-free. Defining

$$m_r(H) := \max_{F \subseteq H}\Big(\mathbb{1}_{\{v_F \geq r+1\}}\tfrac{e_F-1}{v_F-r} + \mathbb{1}_{\{v_F=r, e_F=1\}}\tfrac{1}{r}\Big),$$

we say that H is *strictly r-balanced* if $m_r(H) > m_r(F)$ for all $F \subsetneq H$. Noting $G_{n,p} = G^{(2)}_{n,p}$, now the proofs of Theorem 64 and Remark 65 routinely carry over with only obvious notational changes (including the definitions of Y_K and X_K), yielding the following extension of our refined alteration approach to hypergraphs.

Theorem 69. *Given $r \geq 2$, let H be a strictly r-balanced r-uniform hypergraph. Then, for any $\delta \in (0,1]$, the following holds for all $C \geq C_0(\delta, H)$ and $0 < c \leq c_0(C, \delta, H)$. Setting $n := \lfloor c(k^{r-1}/\log k)^{m_r(H)} \rfloor$ and $p := C(\log k)/k^{r-1}$, whp $G_{n,p}^{(r)}$ satisfies $Y_K \leq \delta\binom{k}{r}p$ and $X_K \geq (1-\delta)\binom{k}{r}p$ for all k-vertex sets K.*

Finally, numerous applications of the alteration method require forbidding a collection of hypergraphs $\mathcal{H} = \{H_1, \ldots, H_s\}$ (see, e.g., [72, 73, 42, 11]). The crux is that the bounds on Y_K and X_K from Theorem 69 trivially remain valid for $n \leq \lfloor c(k^{r-1}/\log k)^{m_r(H)} \rfloor$. So, applying this result to all forbidden $H_i \in \mathcal{H}$ (using δ/s instead of δ to sum the different Y_K-bounds), we readily obtain the following corollary.

Corollary 70. *Given $r \geq 2$ and $s \geq 1$, let $\mathcal{H} = \{H_1, \ldots, H_s\}$ be a collection of strictly r-balanced r-uniform hypergraphs. Define $m_r(\mathcal{H}) := \min_{i \in [s]} m_r(H_i)$, and let Y_K' denote the number of edges in $E(G_{n,p}^{(r)}[K])$ that are in at least one H_i-copy of $G_{n,p}^{(r)}$ for some $H_i \in \mathcal{H}$. Then, for any $\delta \in (0,1]$, the following holds for all $C \geq C_0(\delta, \mathcal{H})$ and $0 < c \leq c_0(C, \delta, \mathcal{H})$. Setting $n := \lfloor c(k^{r-1}/\log k)^{m_r(\mathcal{H})} \rfloor$ and $p := C(\log k)/k^{r-1}$, whp $G_{n,p}^{(r)}$ satisfies $Y_K' \leq \delta\binom{k}{r}p$ and $X_K \geq (1-\delta)\binom{k}{r}p$ for all k-vertex sets K.*

Acknowledgements. We would like to thank Jacob Fox for helpful clarifications regarding [20].

5.5 Appendix: Lower bound on the upper tail of $|\mathcal{H}_K|$

Given a fixed graph H with $v_H \geq 3$, let us consider a binomial random graph $G_{n,p}$ with edge-probability $p = \Theta\big((\log k)/k\big)$ as $k \to \infty$. Fix a k-vertex subset K of $G_{n,p}$ (which tacitly requires $k \leq n$), and let $\mathcal{H}_K$ denote the collection of all H-copies that have at least one edge inside K. Given $\delta > 0$, we fix v_H disjoint vertex subsets of K, each of size $t := \left\lceil \left(\delta\binom{k}{2}p\right)^{1/v_H} \right\rceil$. Then $G_{n,p}$ contains with probability $p^{\binom{v_H}{2}t^2}$ a complete v_H-partite subgraph on these v_H sets, which enforces $|\mathcal{H}_K| \geq t^{v_H} \geq \delta\binom{k}{2}p$. It readily follows that

$$\mathbb{P}\Big(|\mathcal{H}_K| \geq \delta\tbinom{k}{2}p\Big) \;\geq\; p^{\binom{v_H}{2}t^2} \;\geq\; e^{-o(k)},$$

as claimed in Section 5.1.2 (since $t^2 \cdot \log(1/p) \leq k^{2/v_H + o(1)} \cdot O(\log k) = o(k)$ as $k \to \infty$).

REFERENCES

[1] T. Ahmed, O. Kullmann, and H. Snevily. On the van der Waerden numbers $w(2;3,t)$. *Discrete Appl. Math.* **174** (2014), 27–51.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A* **29** (1980), 354–360.

[3] M. Ajtai, J. Komlós, and E. Szemerédi. A dense infinite Sidon sequence. *European J. Combin.* **2** (1981), 1–11.

[4] N. Alon. Covering graphs by the minimum number of equivalence relations. *Combinatorica* **6** (1986), 201–206.

[5] N. Alon and R. Alweiss. On the product dimension of clique factors. *European J. Combin.* **86** (2020), 103097, 10 pp.

[6] N. Alon, J.H. Kim, and J. Spencer. Nearly perfect matchings in regular simple hypergraphs. *Israel J. Math.* **100** (1997), 171–187.

[7] J. Beck. On size Ramsey number of paths, trees, and circuits. I. *J. Graph Theory* **7** (1983), 115–129.

[8] J. Beck. Achievement games and the probabilistic method. In *Combinatorics, Paul Erdős is eighty*, Bolyai Soc. Math. Stud **1** (1993), pp. 51–78.

[9] P. Bennett and T. Bohman. A note on the random greedy independent set algorithm. *Rand. Struct. Algor.* **49** (2016), 479–502.

[10] T. Bohman. The triangle-free process. *Adv. Math.* **221** (2009), 1653–1677.

[11] T. Bohman, A. Frieze, and D. Mubayi. Coloring H-free hypergraphs. *Rand. Struct. Algor.* **36** (2010), 11–25.

[12] T. Bohman and P. Keevash. The early evolution of the H-free process. *Invent. Math.* **181** (2010), 291–336.

[13] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. *Rand. Struct. Algor.* **58** (2021), 221–293.

[14] T. Bohman and L. Warnke. Large girth approximate steiner triple systems. *J. Lond. Math. Soc.* **100** (2019), 895–913.

[15] B. Bollobás, P. Erdős, J. Spencer, and D. West. Clique coverings of the edges of a random graph. *Combinatorica* **13** (1993), 1–5.

[16] S. A. Burr, P. Erdős, and L. Lovász. On graphs of Ramsey type. *Ars Combinatoria* **1** (1976), 167–190.

[17] D. de Caen. Extremal clique coverings of complementary graphs. *Combinatorica* **6** (1986), 309–314.

[18] M. Cavers and J. Verstraëte. Clique partitions of complements of forests and bounded degree graphs. *Discrete Math.* **308** (2008), 2011–2017.

[19] D. Conlon. On-line Ramsey numbers. *SIAM J. Discrete Math.* **23** (2009), 1954–1963.

[20] D. Conlon, J. Fox, A. Grinshpun, and X. He. Online Ramsey numbers and the subgraph query problem. In *Building Bridges II*, Bolyai Soc. Math. Stud **28** (2019), pp. 159–194.

[21] D. Conlon, J. Fox, and B. Sudakov. Short proofs of some extremal results. *Combin. Probab. Comput.* **23** (2014), 8–28.

[22] D. Conlon, J. Fox, and B. Sudakov. Recent developments in graph Ramsey theory. In *Surveys in Combinatorics 2015*, pp. 49–118. Cambridge Univ. Press, Cambridge (2015).

[23] M. Cygan, M. Pilipczuk, and M. Pilipczuk. Known algorithms for edge clique cover are probably optimal. *SIAM J. Comput.* **45** (2016), 67–83.

[24] N. Eaton and V. Rödl. Graphs of small dimensions. *Combinatorica* **16** (1996), 59–85.

[25] S. Ehard, S. Glock, and F. Joos. Pseudorandom hypergraph matchings. Preprint (2019). `arXiv:1907.09946`.

[26] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.* **53** (1947), 292–294.

[27] P. Erdős. Graph theory and probability. II. *Canad. J. Math.* **13** (1961), 346–352.

[28] P. Erdős, R. Faudree, and E. Ordman. Clique partitions and clique coverings. *Discrete Math.* **72** (1988), 93–101.

[29] P. Erdös, R. Faudree, J. Pach, and J. Spencer. How to make a graph bipartite. *J. Combin. Theory Ser. B* **45** (1988), 86–98.

[30] P. Erdős, R.J. Faudree, C.C. Rousseau, and R.H. Schelp. The size Ramsey number. *Period. Math. Hungar.* **9** (1978), 145–161.

[31] P. Erdős, A. Goodman, and L. Pósa. The representation of a graph by set intersections. *Canad. J. Math.* **18** (1966), 106–112.

[32] P. Erdős, E. Ordman, and Y. Zalcstein. Clique partitions of chordal graphs. *Combin. Probab. Comput.* **2** (1993), 409–415.

[33] P. Erdős, S. Suen, and P. Winkler. On the size of a random maximal graph. *Rand. Struct. Algor.* **6** (1995), 309–318.

[34] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.* **2** (1935), 463–470.

[35] P. Erdős and P. Tetali. Representations of integers as the sum of k terms. *Rand. Struct. Algor.* **1** (1990), 245–261.

[36] L. Esperet, R. Kang, and S. Thomassé. Separation choosability and dense bipartite induced subgraphs. *Combin. Probab. Comput.* **28** (2019), 720–732.

[37] G. Fiz Pontiveros, S. Griffiths, and R. Morris. The triangle-free process and the Ramsey number $R(3,k)$. *Mem. Amer. Math. Soc.* **263** (2020), no. 1274.

[38] J. Fox, A. Grinshpun, A. Liebenau, Y. Person, and T. Szabó. On the minimum degree of minimal Ramsey graphs for multiple colours. *J. Combin. Theory Ser. B* **120** (2016), 64–82.

[39] J. Fox, X. He, and Y. Wigderson. Ramsey, Paper, Scissors. *Rand. Struct. Algor.* **57** (2020), 1157–1173.

[40] J. Fox and K. Lin. The minimum degree of Ramsey-minimal graphs. *J. Graph Theory* **54** (2007), 167–177.

[41] D. Freedman. On tail probabilities for martingales. *Ann. Probab.* **3** (1975), 100–118.

[42] A. Frieze and D. Mubayi. On the Chromatic Number of Simple Triangle-Free Triple Systems. *Electron. J. Combin.* **15** (2008), Research Paper 121, 27 pp.

[43] A. Frieze and B. Reed. Covering the edges of a random graph by cliques. *Combinatorica* **15** (1995), 489–497.

[44] Z. Füredi. On the Prague dimension of Kneser graphs. In *Numbers, Information and Complexity* (Bielefeld, 1998), pp. 143–150, Kluwer Acad. Publ., Boston (2000).

[45] Z. Füredi and I. Kantor. Kneser ranks of random graphs and minimum difference representations. *SIAM J. Discrete Math.* **32** (2018), 1016–1028.

[46] S. Glock, D. Kühn, A. Lo, and D. Osthus. On a conjecture of Erdős on locally sparse Steiner triple systems. *Combinatorica* **40** (2020), 363–403.

[47] W.T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.* **11** (2001), 465–588.

[48] R. Graham. On the growth of a van der Waerden-like function. *Integers* **6** (2006), A29, 5 pp.

[49] R. Graham and S. Butler. *Rudiments of Ramsey theory*. 2nd ed., Amer. Math. Soc., Providence (2015).

[50] B. Green. Arithmetic progressions in sumsets. *Geom. Funct. Anal.* **12** (2002), 584–597.

[51] B. Green. New lower bounds for van der Waerden numbers. Preprint (2021). `arXiv:2102.01543`.

[52] B. Green and J. Wolf. A note on Elkin's improvement of Behrend's construction. In *Additive number theory*, Springer (2010), pp. 141–144.

[53] H. Guo. On the power of random greedy algorithms. *Talk at AMS Fall Southeastern Sectional Meeting (virtual)*, October 10, 2020. Abstract available at: http://www.ams.org/amsmtgs/2281_abstracts/1161-05-280.pdf

[54] H. Guo, K. Patton, and L. Warnke. Prague dimension of random graphs. Preprint (2020). `arXiv:2011.09459`.

[55] H. Guo and L. Warnke. Bounds on Ramsey Games via Alterations. Preprint (2019). `arXiv:1909.02691`.

[56] H. Guo and L. Warnke. Packing nearly optimal Ramsey $R(3,t)$ graphs. *Combinatorica* **40** (2020), 63–103.

[57] H. Guo and L. Warnke. On the power of random greedy algorithms. Preprint (2021). `arXiv:2104.07854`.

[58] T.E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.* **56** (1960), 13–20.

[59] P. Hell and J. Nešetřil. *Graphs and homomorphisms*. Oxford University Press, Oxford (2004).

[60] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience, New York (2000).

[61] S. Janson and A. Ruciński. The infamous upper tail. *Rand. Struct. Algor.* **20** (2002), 317–342.

[62] J. Kahn. Asymptotically good list-colorings. *J. Combin. Theory Ser. A* **73** (1996), 1–59.

[63] J. Kahn and J. Park. Tuza's conjecture for random graphs. Preprint (2020). `arXiv:2007.04351`.

[64] J. Kahn, A. Steger, and B. Sudakov. Combinatorics (January 1–7, 2017). *Oberwolfach Rep.* **14** (2017), 5–81. Available at `http://www.mfo.de/document/1701/OWR_2017_01.pdf`

[65] I. Kantor. *Graphs, codes, and colorings*. PhD thesis, University of Illinois at Urbana-Champaign (2010). Available at `http://hdl.handle.net/2142/18247`

[66] H.A. Kierstead and G. Konjevod. Coloring number and on-line Ramsey theory for graphs and hypergraphs. *Combinatorica* **29** (2009), 49–64.

[67] J.H. Kim. The Ramsey number $R(3, t)$ has order of magnitude $t^2/\log t$. *Rand. Struct. Algor.* **7** (1995), 173–207.

[68] J. Körner and K. Marton. Relative capacity and dimension of graphs. *Discrete Math.* **235** (2001), 307–315.

[69] J. Körner and A. Orlitsky. Zero-error information theory. *IEEE Trans. Inform. Theory* **44** (1998), 2207–2229.

[70] L. Kou, L. Stockmeyer and C. Wong. Covering edges by cliques with regard to keyword conflicts and intersection graphs. *Comm. ACM* **21** (1978), 135–139.

[71] M. Krivelevich. Bounding Ramsey numbers through large deviation inequalities. *Rand. Struct. Algor.* **7** (1995), 145–155.

[72] M. Krivelevich. Approximate set covering in uniform hypergraphs. *J. Algor.* **25** (1997), 118–143.

[73] M. Krivelevich. On the minimal number of edges in color-critical graphs. *Combinatorica* **17** (1997), 401–426.

[74] V. Kurauskas and K. Rybarczyk. On the chromatic index of random uniform hypergraphs. *SIAM J. Discrete Math.* **29** (2015), 541–558.

[75] A. Kurek and A. Ruciński. Two variants of the size Ramsey number. *Discuss. Math. Graph Theory* **25** (2005), 141–149.

[76] M. Kwan, S. Letzter, B. Sudakov, and T. Tran. Dense induced bipartite subgraphs in triangle-free graphs. *Combinatorica* **40** (2020), 283–305.

[77] Y. Li. A random process on graphs. *Talk at Random graphs and complex networks Conference (Shanghai Jiaotong Univeristy)*, November 14, 2009. Slides available at: http://math.sjtu.edu.cn/institution/C_S/09-11-15/4-LYS.pdf

[78] Y. Li and J. Shu. A lower bound for off-diagonal van der Waerden numbers. *Adv. in Appl. Math.* **44** (2010), 243–247.

[79] A. Liebenau. *Orientation Games and Minimal Ramsey Graphs*. PhD thesis, FU Berlin (2013).

[80] L. Lovász, J. Nešetřil, and A. Pultr. On a product dimension of graphs. *J. Combin. Theory Ser. B* **29** (1980), 47–67.

[81] C. McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics 1989*, pp. 148–188. Cambridge Univ. Press, Cambridge (1989).

[82] C. McDiarmid. Concentration. In *Probabilistic methods for Algorithmic Discrete Mathematics*, pp. 195–248. Springer, Berlin (1998).

[83] M. Molloy and B. Reed. Near-optimal list colorings. *Rand. Struct. Algor.* **17** (2000), 376–402.

[84] J. Nešetřil and A. Pultr. A Dushnik-Miller type dimension of graphs and its complexity. In *Fundamentals of Computation Theory (Proc. Internat. Conf., Poznań-Kórnik, 1977)*, pp. 482–493. Springer, Berlin (1977).

[85] J. Nešetřil and V. Rödl. A simple proof of the Galvin-Ramsey property of the class of all finite graphs and a dimension of a graph. *Discrete Math.*, **23** (1978), 49–55.

[86] J. Nešetřil and V. Rödl. Products of graphs and their applications. In *Graph theory (Łagów, 1981)*, pp. 151–160. Springer, Berlin (1983).

[87] J. Orlin. Contentment in graph theory: covering graphs with cliques. *Indag. Math.* **80** (1977), 406–424.

[88] D. Osthus and A. Taraz. Random maximal H-free graphs. *Rand. Struct. Algor.* **18** (2001), 61–82.

[89] G. Owen. *Game Theory*. 3rd ed., Academic Press, San Diego (1995).

[90] Y. Person. Personal communication (RSA 2013 conference in Poznań), 2013.

[91] M. Picollelli. The diamond-free process. *Rand. Struct. Algor.* **45** (2014), 513–551.

[92] M. Picollelli. The final size of the C_ℓ-free process. *SIAM J. Discrete Math.* **28** (2014), 1276–1305.

[93] N. Pippenger and J. Spencer. Asymptotic behavior of the chromatic index for hypergraphs. *J. Combin. Theory Ser. A* **51** (1989), 24–42.

[94] S. Poljak, V. Rödl and D. Turzík. Complexity of representation of graphs by set systems. *Discrete Appl. Math.* **3** (1981), 301-–312.

[95] F.P. Ramsey. On a Problem of Formal Logic. *Proc. Lond. Math. Soc.* **30** (1930), 264–286.

[96] F. Roberts. Applications of edge coverings by cliques. *Discrete Appl. Math.* **10** (1985), 93–109.

[97] V. Rödl. On a packing and covering problem. *European J. Combin.* **6** (1985), 69–78.

[98] V. Rödl and M. Siggers. On Ramsey minimal graphs. *SIAM J. Discrete Math.* **22** (2008), 467–488.

[99] V. Rödl and E. Szemerédi. On size Ramsey numbers of graphs with bounded degree. *Combinatorica* **20** (2000), 257–262.

[100] T. Schoen. A subexponential upper bound for van der Waerden numbers $W(3,k)$. Preprint (2020). `arXiv:2006.02877`.

[101] M. Šileikis and L. Warnke. A counterexample to the DeMarco-Kahn upper tail conjecture. *Rand. Struct. Algor.* **55** (2019), 775–794.

[102] M. Šileikis and L. Warnke. Counting extensions revisited. Preprint (2019). arXiv:1911.03012.

[103] M. Šileikis and L. Warnke. Upper tail bounds for stars. *Electron. J. Combin.* **27** (2020), Paper no. 1.67, 23 pp.

[104] P. Skums and L. Bunimovich. Graph fractal dimension and structure of fractal networks: a combinatorial perspective. *J. Complex Netw.*, **8** (2020), No. 4, cnaa037.

[105] N. J. A. Sloane. The on-line encyclopedia of integer sequences. http://www.oeis.org.

[106] J. Spencer. Ramsey's theorem – a new lower bound. *J. Combin. Theory Ser. A* **18** (1975), 108–115.

[107] J. Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Math.* **20** (1977), 69–76.

[108] J. Spencer. Asymptotic packing via a branching process. *Rand. Struct. Algor.* **7** (1995), 167–172.

[109] J. Spencer. Maximal triangle-free graphs and Ramsey $R(3,t)$. Unpublished manuscript (1995). http://cs.nyu.edu/spencer/papers/ramsey3k.pdf.

[110] R. Spöhel, A. Steger, and L. Warnke. General deletion lemmas via the Harris inequality. *J. Combin.* **4** (2013), 251–271.

[111] B. Sudakov. Ramsey numbers and the size of graphs. *SIAM J. Discrete Math.* **21** (2007), 980—986.

[112] T. Szabó, P. Zumstein, and S. Zürcher. On the minimum degree of minimal Ramsey graphs. *J. Graph Theory* **64** (2010), 150–164.

[113] W. van Batenburg, R. de Verclos, R. Kang, and F. Pirot. Bipartite induced density in triangle-free graphs. *Electron. J. Combin.* **27** (2020), Paper 2.34, 19 pp.

[114] W. Wallis. Asymptotic values of clique partition numbers. *Combinatorica* **2** (1982), 99–101.

[115] L. Warnke. The C_ℓ-free process. *Rand. Struct. Algor.* **44** (2014), 490–526.

[116] L. Warnke. When does the K_4-free process stop? *Rand. Struct. Algor.* **44** (2014), 355–397.

[117] L. Warnke. On the method of typical bounded differences. *Combin. Probab. Comput.* **25** (2016), 269–299.

[118] L. Warnke. Upper tails for arithmetic progressions in random subsets. *Israel J. Math.* **221** (2017), 317–365.

[119] L. Warnke. On Wormald's differential equation method. *Combin. Probab. Comput.*, to appear. `arXiv:1905.08928`.

[120] L. Warnke. On the missing log in upper tail estimates. *J. Combin. Theory Ser. B* **140** (2020), 98–146.

[121] D. West. *Introduction to graph theory*. Prentice Hall, New Jersey (1996).

[122] G. Wolfovitz. Triangle-free subgraphs in the triangle-free process. *Rand. Struct. Algor.* **39** (2011), 539–543.

[123] N. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.* **5** (1995), 1217–1235.

BIOGRAPHY

HE GUO

He Guo (Chinese name: 郭赫) was born and spent his childhood in Xi'an, an ancient city where the journey of the Silk Road to the west began, at the center of China. He was admitted to Zhejiang University and moved to Hangzhou, a city decorated by the beautiful West Lake on the eastern coast of China. He was honored to be an undergraduate student of Shing-Tung Yau's Mathematics Elite Class, a class established by the Fields Medalist Shing-Tung Yau, of Chu Kochen Honors College in 2011, when he chose the career path to be trained as a mathematician. The first-year analysis course, which introduced the ϵ-δ language to prove many beautiful mathematical statements rigorously, made him fall in love with math. At the beginning of the course, he felt surprised that a reference book by Fichtenholz is the one he ran across in a bookstore couple of years ago when he was a junior high school student. At first glance he was attracted by the proof of the statement that $\sqrt{2}$ is not a rational number using contradiction in the first page of the book and he couldn't wait to purchase that book to read more at home. In 2015 he obtained his Bachelor of Science degree with distinction from Zhejiang University. From December 2014 to May 2015 in his last year as an undergraduate student, he was selected to be a visiting student to study at Harvard University. That was the first time for him to live abroad. He felt happy to find that his previous experience of self-studying mathematics in English made his communication smooth in the math class at Harvard. That half-year studying experience helped him get familiar with the daily life in the United States. An interesting thing for daily English he found almost at the first day in the US is that "How are you?" "I am fine. Thank you!", what he had learned, are likely not the most common greetings in the US. He enjoyed museums in Boston and New York and he enjoyed exploring the cities by the railway system. He likes visiting the historical sites in Boston to learn the history of New England

and the establishing of the United States. From that time, he forms the habit of visiting the museums when he travels to a new city. In August 2015, he started his PhD career in the Algorithms, Combinatorics, and Optimization (ACO) Program at Georgia Institute of Technology and spent six years at Atlanta, Georgia in the United States, a warm city with warm southern US people. His advisor is Lutz Warnke. He is expected to obtain his PhD degree in the summer of 2021. During his PhD career, the academic community all over the world offered him great opportunities to travel to many countries to attend conferences and workshops, which make him have better understanding of different cultures.

He enjoys and is proud of his hobbies. He likes reading literature, history and social science books. He likes art and music. He likes watching movies, soccer, hiking, running, swimming, ping-pong, badminton, billiards, poker, chess, frisbee, computer games (like many Millennials and Gen Z), and many other activities. Those habits together with studying mathematics make him happy and energetic.

VITA

HE GUO

Chinese Name: 郭赫

Education

2021 (expected)	*Georgia Institute of Technology*, Atlanta, USA
	Ph.D. in Algorithms, Combinatorics, and Optimization (ACO)
	Advisor: Lutz Warnke
2015	*Zhejiang University*, Hangzhou, China
	B.S. (with distinction)
	Shing-Tung Yau's Mathematics Elite Class, Chu Kochen Honors College