Minimal Ramsey graphs, orthogonal Latin squares, and hyperplane coverings

Dissertation zur Erlangung des Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich Mathematik und Informatik
der Freien Universität Berlin

vorgelegt von

Simona Boyadzhiyska

Berlin
2022

Betreuer und Erstgutachter:
Prof. Tibor Szabó, PhD

Zweitgutachter:
Prof. Mathias Schacht, PhD

Tag der Disputation: 05.12.2022

Institut für Mathematik
Freie Universität Berlin

Summary

This thesis consists of three independent parts.

The first part of the thesis is concerned with Ramsey theory. Given an integer $q \geq 2$, a graph G is said to be *q-Ramsey* for another graph H if in any q-edge-coloring of G there exists a monochromatic copy of H. The central line of research in this area investigates the smallest number of vertices in a q-Ramsey graph for a given H. In this thesis, we explore two different directions. First, we will be interested in the smallest possible minimum degree of a minimal (with respect to subgraph inclusion) q-Ramsey graph for a given H. This line of research was initiated by Burr, Erdős, and Lovász in the 1970s. We study the minimum degree of a minimal Ramsey graph for a random graph and investigate how many vertices of small degree a minimal Ramsey graph for a given H can contain. We also consider the minimum degree problem in a more general asymmetric setting. Second, it is interesting to ask how small modifications to the graph H affect the corresponding collection of q-Ramsey graphs. Building upon the work of Fox, Grinshpun, Liebenau, Person, and Szabó and Rödl and Siggers, we prove that adding even a single pendent edge to the complete graph K_t changes the collection of 2-Ramsey graphs significantly.

The second part of the thesis deals with orthogonal Latin squares. A *Latin square of order n* is an $n \times n$ array with entries in $[n]$ such that each integer appears exactly once in every row and every column. Two Latin squares L and L' are said to be *orthogonal* if, for all $x, y \in [n]$, there is a unique pair $(i, j) \in [n]^2$ such that $L(i, j) = x$ and $L'(i, j) = y$; a system of k *mutually orthogonal Latin squares*, or a *k-MOLS*, is a set of k pairwise orthogonal Latin squares. Motivated by a well-known result determining the number of different Latin squares of order n log-asymptotically, we study the number of k-MOLS of order n. Earlier results on this problem were obtained by Donovan and Grannell and Keevash and Luria. We establish new upper bounds for a wide range of values of $k = k(n)$. We also prove a new, log-asymptotically tight, bound on the maximum number of other squares a single Latin square can be orthogonal to.

The third part of the thesis is concerned with grid coverings with multiplicities. In particular, we study the minimum number of hyperplanes necessary to cover all points but one of a given finite grid at least k times, while covering the remaining point fewer times. We study this problem for the grid $\mathbb{F}_2^n$, determining the number exactly when one of the parameters n and k is much larger than the other and asymptotically in all other cases. This generalizes a classic result of Jamison for $k = 1$. Additionally, motivated by the recent work of Clifton and Huang and Sauermann and Wigderson for the hypercube $\{0, 1\}^n \subseteq \mathbb{R}^n$, we study hyperplane coverings for different grids over $\mathbb{R}$, under the stricter condition that the remaining point is omitted completely. We focus on two-dimensional real grids, showing a variety of results and demonstrating that already this setting offers a range of possible behaviors.

Acknowledgements

First and foremost, I would like to thank my supervisor, Tibor Szabó, for his advice and guidance during the past few years and for all the wonderful opportunities he has provided me with during my PhD. I am also incredibly thankful to Shagnik Das and Anurag Bishnoi for all of their support and encouragement; thank you for sharing your knowledge with me and for always being there to answer my questions or give me feedback. Thank you also to all of the current and former members of AG Szabó for providing me with a stimulating (and fun!) work environment!

During my PhD, I have had the pleasure of collaborating with a number of people, and I would like to express my gratitude to all of them. In particular, I thank my co-authors, Benjamin Aram Berendsohn, Anurag Bishnoi, Dennis Clemens, Shagnik Das, Yvonne den Bakker, Pranshu Gupta, Thomas Lesgourgues, Anita Liebenau, László Kozma, Tamás Mészáros, Jonathan Rollin, and Tibor Szabó: I have learned a tremendous amount from you and have greatly enjoyed working with each and every one of you!

My work on minimum degrees of minimal Ramsey graphs and my collaboration with Dennis and Pranshu began during a research visit at the Hamburg University of Technology; I am grateful to Anusch Taraz for inviting me and to his entire research group for their hospitality during my stay. I also thank Dennis for initiating several of the subsequent projects related to Ramsey graphs that I was involved in. Also, thank you to Anurag for inviting me to visit him in Delft, where some of the work related to Chapter 9 of this thesis was done. I am thankful to László and Benjamin for my internship experience, through which I was introduced to several nice problems. I am also grateful to Anita for all of her mathematical and non-mathematical advice. Thank you also to Pranshu for being a great conference buddy and roommate.

I am grateful to my reviewers, Tibor Szabó and Mathias Schacht, for the time they will invest in reading this thesis and providing me with feedback. I also thank Anurag, Jonathan, Shagnik, Thomas, and Tibor for reading parts of this thesis and giving me helpful comments and Olaf Parczyk for helping me with the German abstract (and many other tasks requiring German skills!).

I also thank all the teachers I have had over the years, who have been a great source of inspiration. In particular, a big thank you to Ann Trenk for introducing me to combinatorics research.

I thank the Deutsche Forschungsgemeinschaft (DFG), in particular the Research Training Group "Facets of Complexity" and the Berlin Mathematical School/MATH+, for supporting my doctorate financially and for providing me with funding to attend various conferences and workshops during this time.

In addition, I thank the BMS One-Stop Office for all of their help with bureaucratic matters throughout my time in Berlin. I also thank the BMS for giving me the opportunity to be part of such a wonderful community. All those hours we have spent chatting, playing board games or beach volleyball, or dancing the Macarena have been a fantastic way to unwind after a long day of work.

Thank you to András for all the times he has been there for me and has offered advice; also, thank you for all the postcards! I am grateful to Vlad, whose help (and German skills!) made my move to Berlin much easier; thank you also for all the fun memories, both in Berlin and in Varna (by the way, now you are allowed to ask how my thesis is going!). A huge thank you to Riccardo for the immense amount of support he has given me, for helping me stay sane during the pandemic, and for all the tasty food providing me with energy to keep going! I am also grateful to all my friends and family back in Bulgaria for the warm welcome every time I go back. Last but not least, I owe my parents a huge thank you for everything they have taught me and for all the support they have given me.

Thank you everyone for being part of this journey and for encouraging me to move forward!

Contents

General terminology and notation

We give a brief overview of our notation and general definitions. Much of our terminology is standard; in general, we follow the notation used in [16, 19, 27, 28, 30]. Many of the concepts below are also introduced in the text, together with the additional terminology and notation that will be needed in each specific part of the thesis.

Basics We write $\mathbb{R}$ for the field of real numbers, $\mathbb{F}_q$ for the unique finite field with q elements, and $\mathbb{F}$ for a general field. Further, we write $\mathbb{Z}$ for the set of integers and $\mathbb{Z}_{\geq a}$ for the set $\{z \in \mathbb{Z} : z \geq a\}$, where $a \in \mathbb{R}$. The cyclic group of order d for $d \in \mathbb{Z}_{\geq 1}$ is denoted by $\mathbb{Z}_d$. Given a positive integer n, we write $[n]$ for the set $\{1, 2, \ldots, n\}$. Given integers $n \geq k \geq 0$, we write $\binom{n}{k}$ for the number of ways to choose a k-element subset from an n-element set. Throughout the thesis log stands for the natural logarithm and $\log_2$ for the binary logarithm. Further, $\mathrm{polylog}(n)$ denotes any function that is polynomial in $\log(n)$. We write $z = x \pm y$ to mean that $z \in [x - y, x + y]$. Throughout the thesis, we often omit floors and ceilings when they are not strictly necessary.

General graph theory Most of our graph-theoretic notation follows [151], and all terms not defined here can be found in that textbook. For a graph G, we denote its vertex and edge set by $V(G)$ and $E(G)$, respectively, and their cardinalities by $v(G)$ and $e(G)$, respectively. We often identify a graph with its edge set. We say that a vertex v is incident to an edge e if v is an endpoint of e.

For a vertex $v \in V(G)$, we write $N_G(v) = \{u \in V(G) : uv \in E(G)\}$ for the *neighborhood* of v in G and $d_G(v) = |N_G(v)|$ for the *degree* of v in G. When the graph G is clear from the context, we sometimes omit the subscript and write simply $N(v)$ and $d(v)$. We write $\delta(G) = \min\{d_G(v) : v \in V(G)\}$ for the *minimum degree* of G and $\Delta(G) = \max\{d_G(v) : v \in V(G)\}$ for the *maximum degree* of G. The *average degree* of G is the quantity $\frac{\sum_v d_G(v)}{v(G)}$, which is also equal to $\frac{2e(G)}{v(G)}$ (see e.g. [151, Proposition 1.3.3]).

For two subsets $U, W \subseteq V(G)$, we write $E_G(U, W)$ for the edges with one endpoint in U and one endpoint in W, and we denote the size of $E_G(U, W)$ by $e_G(U, V)$. If $U = W$, we write $E_G(U) = E_G(U, U)$ and $e_G(U) = e_G(U, U)$. Again, we omit subscripts when they are not strictly necessary.

We write $\omega(G)$ and $\alpha(G)$ for the *clique* and *independence number* of G, respectively. The *girth* of G is the smallest length of a cycle in G; if G contains no cycle, then its girth is infinite.

Given two subsets $A, B \subseteq V(G)$, we define the *distance* between A and B to be the length of a shortest path with one endpoint in A and one endpoint in B, where the length of a path is the number of edges it contains.

Let G and H be two graphs. We write $G \cong H$ if the two graphs are isomorphic; in this case we also say that H is a *copy* of G. We say that G *contains* H if G contains a subgraph isomorphic to H; otherwise, we say that G is H*-free*. Given a subset $U \subseteq V(G)$, we denote by $G[U]$ the subgraph *induced* by U. For a vertex $v \in V(G)$, we write $G - v$ for the subgraph induced by $V(G) - v$; similarly for a subset $W \subseteq V(G)$, we write $G - W$ for the subgraph induced by $V(G) \setminus W$. For an edge e, we let $G - e$ denote the graph with vertex set $V(G)$ and edge set $E(G) - e$; the graph $G - F$ for a subset $F \subseteq E(G)$ is defined analogously. If G is a graph and v is a new vertex, not in $V(G)$, we write $G + v$ for the graph obtained by connecting v to all vertices of G.

The *union* of two graphs G and H is the graph with vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H)$.

The *connectivity* of a graph G is the minimum size of a set $S \subseteq V(G)$ such that $G - S$ is disconnected or consists of a single vertex; G is k*-connected* if its connectivity is at least k. Two paths between vertices x and y in a graph G are *internally vertex-disjoint* if they share no vertices other than x and y. It is well known that, if G is k-connected and $x, y \in V(G)$, then there are k pairwise internally vertex-disjoint paths between x and y in G (see for example [151, Theorem 4.2.17]).

We write K_t, P_t, and C_t for the complete graph, path, and cycle on t vertices, respectively, and $K_{s,t}$ for the complete bipartite graph with partite sets of size s and t. A star is a graph of the form $K_{1,t}$.

Ramsey graphs, edge-colorings, and color patterns Unless otherwise specified, we use the term *coloring* to refer to an edge-coloring of a graph. A q*-coloring* is a coloring using at most q different colors; unless otherwise specified, the color palette is assumed to be the set $[q]$. When the number of colors q is clear from the context, we will often simply say coloring, without specifying q. If $\varphi : E(G) \to [q]$ is a q-coloring of G, then we write $\varphi^{-1}(i)$ to denote *ith color class*, that is, the graph on $V(G)$ formed by all edges of color i. A subgraph H of G is *monochromatic (under* φ*)* if φ assigns the same color to all edges of H. If F is a subgraph of G, we write $\varphi_{|F}$ for the coloring *induced* by φ on F. If ψ is a coloring of F and φ is a coloring of G, then we say that φ *extends* ψ (to G) if $\varphi(e) = \psi(e)$ for all $e \in F$. Given colorings φ_1 and φ_2 of G_1 and G_2, respectively, such that $\varphi_1(e) = \varphi_2(e)$ for all $e \in E(G_1) \cap E(G_2)$, we define the coloring $\varphi_1 \cup \varphi_2$ on $G_1 \cup G_2$ by setting

$$\varphi(e) = \begin{cases} \varphi_1(e) & \text{if } e \in E(G_1), \\ \varphi_2(e) & \text{if } e \in E(G_2). \end{cases}$$

Given a tuple of graphs $(H_1, \dots, H_q)$, we say that a q-coloring is $(H_1, \dots, H_q)$*-free* if $\varphi^{-1}(i)$ is H_i-free for all $i \in [q]$.

Given an integer $q \geq 2$ and a vertex set V, a *q-color pattern* on V is a collection of edge-disjoint graphs $G_1, \dots, G_q$ on V. Given a graph G, a *q-color pattern for G* is a collection of edge-disjoint graphs $G_1, \dots, G_q$ on $V(G)$ such that $E(G) = E(G_1) \cup \dots \cup E(G_q)$. For a tuple of graphs $(H_1, \dots, H_q)$, a color pattern $G_1, \dots, G_q$ is $(H_1, \dots, H_q)$*-free* if every graph G_i is H_i-free. If $U \subseteq V$, the color pattern *induced* on U is the color pattern $G_1[U], \dots, G_q[U]$. The color pattern *induced* by a coloring φ of a graph G is the color pattern $\varphi^{-1}(1), \dots, \varphi^{-1}(q)$.

If G has no $(H_1, \dots, H_q)$-free q-coloring, then G is *q-Ramsey* for $(H_1, \dots, H_q)$ and we write $G \to_q (H_1, \dots, H_q)$. We denote the set of all q-Ramsey graphs for $(H_1, \dots, H_q)$ by $\mathcal{R}_q(H_1, \dots, H_q)$ and the set of all minimal (with respect to subgraph inclusion) q-Ramsey graphs for $(H_1, \dots, H_q)$ by $\mathcal{M}_q(H_1, \dots, H_q)$. If $H_i \cong H$ for all $i \in [q]$, then we simplify notation and just write H instead of $(H, \dots, H)$. We denote the Ramsey number of $(H_1, \dots, H_q)$ by $r_q(H_1, \dots, H_q)$ and the smallest minimum degree of a minimal q-Ramsey graph for $(H_1, \dots, H_q)$ by $s_q(H_1, \dots, H_q)$.

We usually refer to the graph being colored as the *host graph* and the graph(s) that we want to find inside the host graph as the *target graph(s)*.

Hypergraphs For a hypergraph $\mathcal{H}$, we denote its vertex set and edge set by $V(\mathcal{H})$ and $E(\mathcal{H})$, respectively; $v(\mathcal{H})$ and $e(\mathcal{H})$ are defined in the natural way. A hypergraph is *t-uniform* if every hyperedge contains exactly t vertices. A *cycle of length s* in a hypergraph $\mathcal{H}$ is a sequence $e_1, v_1, e_2, v_2 \dots, e_s, v_s$ of distinct hyperedges and vertices of $\mathcal{H}$ such that $v_i \in e_i \cap e_{i+1}$ for all $1 \leq i < s$ and $v_s \in e_s \cap e_1$. Note in particular that two edges intersecting in more than one vertex form a cycle of length two in $\mathcal{H}$. The *girth* of a hypergraph $\mathcal{H}$ is the length of the shortest cycle in $\mathcal{H}$ (if no cycle exists, then by convention we say that the girth of $\mathcal{H}$ is infinity).

Asymptotic notation Let $f, g : \mathbb{Z}_{\geq 0} \to \mathbb{R}$ be two functions. We write $f = O(g)$ if there exist constants $n_0 \in \mathbb{Z}_{\geq 0}$ and $C > 0$ such that $|f(n)| \leq C|g(n)|$ for all $n \geq n_0$. When $f = O(g)$, we sometimes write $g = \Omega(f)$ instead. If $f = O(g)$ and $g = O(f)$, we write $f = \Theta(g)$.

In addition, we write $f = o(g)$ to mean that $\lim_{n\to\infty} f(n)/g(n) = 0$ and $f = \omega(g)$ to mean that $\lim_{n\to\infty} f(n)/g(n) = \infty$. When $f = o(g)$, we sometimes write $f \ll g$ or $g \gg f$ instead.

Probability We write $\mathbb{P}[A]$ for the probability of an event A and $\mathbb{E}[Y]$ for the expectation of the random variable Y. We write $\mathrm{Bin}(n, p)$ for the binomial distribution with n trials and probability of success p. For an integer $n \geq 1$ and a probability $p \in [0, 1]$, the *binomial random graph $G(n, p)$* is a graph on $[n]$ in which each potential edge is present with probability p, independently of all other edges. A sequence of events $(A_n)_{n=1}^{\infty}$ occurs *with high probability (w.h.p)* if $\lim_{n\to\infty} \mathbb{P}[A_n] = 1$.

Orthogonal Latin squares Our terminology and notation mostly follow [147]. For an integer $n \geq 1$, a *Latin square of order n* is an $n \times n$ matrix with entries in $[n]$ such that each $x \in [n]$ appears exactly once in every row and every column. Given a Latin square L, we write $L(i, j)$ for the entry in the ith row and jth column of L. We denote the number of Latin squares of order n by $L(n)$. A *transversal* in a Latin square L is a set of n cells from L such that no two share a row, column, or symbol.

Two Latin squares L and L' are *orthogonal* if for any pair $(x, y) \in [n]^2$, there exists a unique pair $(i, j) \in [n]^2$ such that $L(i, j) = x$ and $L'(i, j) = y$. A k-tuple of Latin squares $(L_1, \ldots, L_k)$ is a *k-MOLS* if L_i and L_j are orthogonal for all distinct $i, j \in [k]$. We write $L^{(k)}(n)$ for the number of k-MOLS of order n.

Hyperplane coverings Let $\mathbb{F}$ be any field and $n \geq 1$ be an integer. For a vector $\vec{v} \in \mathbb{F}^n$, we often write v_i for the ith entry of $\vec{v}$. The *Hamming weight* of $\vec{v}$ is the number of nonzero entries $\vec{v}$ contains.

Unless otherwise specified, we use the term subspace for an *affine* subspace. The *codimension* of a subspace of dimension d is given by $n - d$. A *hyperplane* in $\mathbb{F}^n$ is a subspace of codimension 1. A hyperplane H can be written as $H = \{\vec{x} \in \mathbb{F}^n : \vec{a} \cdot \vec{x} = c\}$, where $\cdot$ denotes the usual dot product of two vectors, $\vec{a} = (a_1, \ldots, a_n) \in \mathbb{F}^n \setminus \{\vec{0}\}$ is the *normal vector* of H, and $c \in \mathbb{F}$; in this case we also write that H is given by $\vec{a} \cdot \vec{x} = c$ or $\sum_{i=1}^n a_i x_i = c$ for short. Given a normal vector $\vec{a} = (a_1, \ldots, a_n) \in \mathbb{F}^n \setminus \{\vec{0}\}$, we sometimes write $H_{\vec{a}}$ for the hyperplane given by $\vec{a} \cdot \vec{x} = 1$. In two-dimensional case, we will often use x and y, as opposed to x_1 and x_2, to denote the two coordinates.

Given integers $k \geq 1$ and $n \geq d \geq 1$, a field $\mathbb{F}$, and finite subsets $S_1, \ldots, S_n$ with $0 \in \bigcap_{i=1}^n S_i$, we sometimes denote the grid $S_1 \times \cdots \times S_n$ by $\Gamma = \Gamma(S_1, \ldots, S_n)$. If the sets $S_1, \ldots, S_n$ are clear from the context, we sometimes omit them from the notation. Unless otherwise specified, we always assume that a grid Γ contains the origin $\vec{0}$ and often write $\Gamma^- = \Gamma \setminus \{\vec{0}\}$.

For a grid $\Gamma \subseteq \mathbb{F}^n$, a multiset $\mathcal{H}$ of $(n - d)$-dimensional affine subspaces in $\mathbb{F}^n$ is a *(k, d)-cover* if every nonzero point of Γ is covered at least k times, while $\vec{0}$ is covered at most $k - 1$ times. A $(k, d; s)$-cover is a (k, d)-cover in which $\vec{0}$ is covered exactly s times. We often call a $(k, d; 0)$-cover a *strict (k, d)-cover.*

When $d = 1$, we often suppress it from our notation and write for example k-cover instead of $(k, 1)$-cover.

For any integers $k \geq 1$, $n \geq d \geq 1$, and $s \geq 0$ the extremal function $f(n, k, d)$ is defined to be the minimum possible size of a (k, d)-cover of $\mathbb{F}_2^n$ and $g(n, k, d; s)$ is the minimum size of a $(k, d; s)$-cover of $\mathbb{F}_2^n$.

For a grid Γ, we write $h(\Gamma, k)$ for the minimum size of a strict k-cover of Γ.

Chapter 1

Introduction

Extremal combinatorics is a branch of discrete mathematics, investigating questions of the following type: how large or how small can a structure of a given type be, provided that it satisfies certain conditions? As a concrete example, one of the classical questions in this area asks: given n vertices, what is the largest number of edges we can add between them without creating a triangle? This thesis consists of three parts, studying questions with an extremal flavor related to three different topics: Ramsey theory, orthogonal Latin squares, and grid coverings.

Part I of this thesis focuses on Ramsey theory. Informally speaking, Ramsey theory is concerned with finding orderly substructures within large structures. Specifically, we will work in the realm of graph Ramsey theory. A typical question in this field asks: what properties does a graph need to satisfy in order to ensure that, no matter how we partition its edges into a fixed number of classes, at least one of the classes will always contain a particular type of subgraph? This idea is best illustrated with an example. Consider the following problem: what is the smallest integer n such that the edges of the complete graph K_n cannot be colored red and blue without creating a triangle K_3 with only red edges or only blue edges? This is an exercise many students see in their first combinatorics course, and one easily arrives at the correct answer: six. However, this innocuous-looking question becomes notoriously difficult if we replace the triangle by a slightly larger graph such as K_5 or add extra colors to the palette. Problems of this type have given rise to a very active area of research in combinatorics and have driven the development of many powerful tools and methods.

In this thesis, we consider two different directions. We say that a graph G is q-Ramsey for another graph H if, no matter how we color the edges of G using q colors, we can find a copy of H whose edges all have the same color. In this language, the motivating example above asks for the minimum number of *vertices* in a graph that is 2-Ramsey for K_3, and more generally, for the minimum number of vertices in a graph that is q-Ramsey for H. Here we study how small the vertex degrees of a minimal (with respect to subgraph inclusion) q-Ramsey graph for a given H

can be, continuing a line of research initiated by Burr, Erdős, and Lovász [40]. We are concerned with determining the smallest minimum degree of a minimal q-Ramsey graph for a given H and finding out how many vertices of this degree can occur. Among other results, we answer both questions for moderately sparse random graphs and the latter question for complete graphs.

In the second part, we explore how small changes to the graph H affect the corresponding collection of Ramsey graphs. Considering the cases where H is a complete graph or an odd cycle, we show that even adding a single pendent edge changes the collection significantly.

Part II is concerned with orthogonal Latin squares. A Latin square is an $n \times n$ array whose cells are filled with n different symbols in such a way that no symbol is repeated in any row or column. The study of Latin squares has a long history in combinatorics, dating back to Euler; Latin squares also appear in other areas of mathematics such as algebra, where they generalize the notion of a group. Outside of mathematics, Latin squares are perhaps best known thanks to the popular Sudoku puzzles and their more advanced relatives, but they also find applications in the design of experiments (for instance in agriculture).

Two Latin squares are said to be orthogonal if superimposing them results in a square containing n^2 different ordered pairs of symbols. Going even further, we can consider larger collections of pairwise orthogonal Latin squares. As in the case of Latin squares, orthogonal Latin squares have attracted the attention of researchers for several centuries. They have proven interesting from both a theoretical and a practical point of view: they are related to other classical structures in design theory and finite geometry and find applications in statistics and coding theory.

After the question of existence, it is natural to ask about the number of different structures of a particular type. For Latin squares the enumeration question was answered log-asymptotically as an application of two well-known results concerning permanents of matrices [31, 60, 66]. We consider several enumeration questions in the context of orthogonal Latin squares. For example, we study the maximum number of different Latin squares a single Latin square can be orthogonal to. We also address the enumeration question for larger collections of pairwise orthogonal Latin squares, asking how many different systems of k pairwise orthogonal Latin squares of size n there are. We establish new upper bounds in both cases.

In Part III, we study coverings of grids with hyperplanes. Consider the following simple question: how many affine hyperplanes does it take to cover the vertices of the n-dimensional Boolean hypercube, $\{0,1\}^n$? This simple question has an equally straightforward answer — we can cover all the vertices with a pair of parallel hyperplanes, and it is easy to see that a single hyperplane can cover at most half the vertices, and so two hyperplanes are indeed necessary. However, the waters are quickly muddied with a minor twist to the problem.

Indeed, if we are instead asked to cover all the vertices except the origin, the parallel hyperplane construction is no longer valid. Given a moment's thought, we might come up with the much larger family of n hyperplanes given by $\{\vec{x} : x_i = 1\}$ for $i \in [n]$. This fulfills the task and, surprisingly, turns out to be optimal, although this is far from obvious. This problem was famously resolved in full generality, that is, for any finite grid, by Alon and Füredi [3] in the early 90s. As it turns out, problems of this type have connections to other areas, such as finite geometry and Ramsey theory, and have played an important role in the development of a powerful method in combinatorics, known as the polynomial method.

We explore a couple of generalizations of this covering problem. More specifically, we seek to determine the number of hyperplanes needed to cover every point but one of a given finite grid *at least* $k \geq 1$ *times* while covering the remaining point fewer times. We study this problem for the grid $\mathbb{F}_2^n$ and for different grids in $\mathbb{R}^2$, proving a variety of asymptotic and exact results.

1.1 Minimal Ramsey graphs

In Part I of this thesis, we will investigate several different problems in Ramsey theory. We begin by introducing the necessary background. Throughout Part I of this thesis, unless otherwise specified, all graphs will be finite, simple, and undirected. Our graph-theoretic notation mostly follows [151].

Given a graph H and an integer $q \geq 1$, we say that a graph G is *q-Ramsey* for another graph H if, for any q-coloring of the edges of G, there exists a *monochromatic* copy of H in G, that is, a copy of H whose edges all have the same color. In this case, we write $G \rightarrow_q H$, and we denote the collection of all q-Ramsey graphs for H by $\mathcal{R}_q(H)$. Note that the collection $\mathcal{R}_1(H)$ is simply all graphs containing H as a subgraph. For $q \geq 2$, the seminal result of Ramsey [125] establishes that $\mathcal{R}_q(H) \neq \emptyset$ for any graph H. While we can easily describe the collection $\mathcal{R}_1(H)$, understanding $\mathcal{R}_q(H)$ for $q \geq 2$ is considerably more involved and is our focus in this thesis. Characterizing all graphs in $\mathcal{R}_q(H)$ is a difficult problem resolved in very few special cases, for example, when H is a star or a matching and $q = 2$ [40]. It is then natural to investigate what properties the graphs belonging to $\mathcal{R}_q(H)$ for given H and q must have. This question has led to the development of a very active area of research, known as *(graph) Ramsey theory* (see [51] for a survey of some recent developments in the field). Notice that, if a graph G is q-Ramsey for a graph H, then any graph containing G is also q-Ramsey for H. Thus, to understand the properties of $\mathcal{R}_q(H)$, it suffices to restrict our attention to those graphs that are minimal with respect to their Ramsey property. We say that a graph $G \in \mathcal{R}_q(H)$ is *minimal q-Ramsey* for H if no proper subgraph of G is q-Ramsey for H; we denote the corresponding collection by $\mathcal{M}_q(H)$.

As a first step towards understanding the collection of q-Ramsey graphs for a given H, we can ask how large the graphs in $\mathcal{R}_q(H)$ need to be. This question leads us to the definition of the most well-studied notion in Ramsey theory, the Ramsey number. The *(q-color) Ramsey number* of a graph H, denoted $r_q(H)$, is the minimum number of vertices in a graph that is q-Ramsey for H. Ramsey numbers have been studied extensively over the past few decades and have turned out to be notoriously difficult to understand in most cases. The most prominent such example is when H is a complete graph, that is, $H \cong K_t$ for some positive integer t. While showing that $r_2(K_3) = 6$ is an easy exercise, already the value of $r_2(K_5)$ is unknown (see [4, 65] for the best known bounds on $r_2(K_5)$ and [124] for a dynamic survey of small Ramsey numbers). The asymptotic growth of $r_2(K_t)$ as $t \to \infty$ is also not very well understood. We know from the early work of Erdős [61] and Erdős and Szekeres [64] that $r_2(K_t)$ grows exponentially with t: more precisely, up to lower order terms, they established that $2^{t/2} \leq r_2(K_t) \leq 2^{2t}$. Despite several decades of effort, the only improvements to date have been in the lower order terms; the current best lower bound is due to Spencer [142], while the best upper bound was recently announced by Sah [132] (see also Conlon's paper [48]). In the multicolor setting, the gap is even larger: for a fixed $q \geq 3$ and $t \to \infty$, again up to lower order terms, the best known bounds are $2^{cqt} \leq r_q(K_t) \leq q^{qt}$. The upper bound can be derived using the ideas of Erdős and Szekeres [64]. The current best value of c is due to Sawin [135], who built on the earier work of Conlon and Ferber [50] and Wigderson [152].

One of the prominent lines of research in Ramsey theory, initiated in the 1970s, is concerned with studying the behavior of different graph parameters in the context of Ramsey graphs. The general question is: given a graph parameter γ, what is the minimum or maximum value γ can take in $\mathcal{R}_q(H)$ (or $\mathcal{M}_q(H)$) for given q and H? One of the first examples of such a result appears in the work of Folkman [70], who, motivated by a question of Erdős and Hajnal (see [1]), proved that, for any $t \geq 3$, there exists a 2-Ramsey graph for K_t containing no copy of K_{t+1}. In other words, he showed that $\min\{\omega(G) : G \in \mathcal{R}_2(K_t)\} = t$. This result was later generalized by Nešetřil and Rödl [120], who showed that $\min\{\omega(G) : G \in \mathcal{R}_q(H)\} = \omega(H)$ for any graph H and any integer $q \geq 2$. Another well-known example is the so-called *(q-color) size Ramsey number* of a graph H, defined as the minimum number of edges in a graph that is q-Ramsey for H; size Ramsey numbers were introduced by Erdős, Faudree, Rousseau, and Schelp [62] and have been studied extensively (see [51] and the references therein).

In Part I of this thesis we will be concerned with another parameter, the smallest possible minimum degree, the study of which was initiated by Burr, Erdős, and Lovász [40] in the 1970s. We will introduce the precise problem and give a brief overview of its history in Section 1.1.1.

A natural question stemming from this line of research is: how do the above parameters change when we alter the target graph H slightly? In some cases, this question might be too difficult to tackle, motivating the even simpler question: does the collection of q-Ramsey graphs change if we alter H slightly? What modifications to H preserve the corresponding collection of Ramsey graphs? These questions lead to the notion of Ramsey equivalence, introduced by Szabó, Zumstein, and Zürcher in [143].

Another direction of research that began in the 1970s is concerned with determining whether the collection $\mathcal{M}_2(H)$ is finite or infinite for a given graph H. In a series of papers, it was shown that $\mathcal{M}_2(H)$ is finite if and only if H is the disjoint union of a star with an odd number of leaves and any number of isolated edges [36, 38, 121, 126, 127]. In [42], Burr, Nešetřil, and Rödl showed the stronger result that, if H is 3-connected or isomorphic to K_3, then in fact $\mathcal{M}_2(H)$ contains at least $2^{\Omega(n \log n)}$ graphs on at most n vertices. For certain classes of graphs, even more is known. In [128], Rödl and Siggers introduced the stronger notion of a highly q-Ramsey-infinite graph (the actual term appears in a later paper of Siggers [141]). A graph H is said to be *highly q-Ramsey-infinite* if, for any sufficiently large integer n, the collection $\mathcal{M}_q(K_t)$ contains at least $2^{\Omega(n^2)}$ graphs on at most n vertices. Rödl and Siggers [128] showed that, for any $t \geq 3$ and any $q \geq 2$, the complete graph K_t is highly q-Ramsey-infinite. Similar results were shown later by Siggers for odd cycles [139, 141] and for non-bipartite 3-connected graphs [140].

In Section 1.1.2 and Chapter 6, we will combine the notions of highly Ramsey-infinite graphs and Ramsey equivalence. In particular, we will prove analogs of the results of Rödl and Siggers [128] for cliques and Siggers [139, 141] for cycles in this setting.

1.1.1 Minimum degrees of minimal Ramsey graphs

As mentioned earlier, we will be concerned with minimum degrees in the context of Ramsey graphs, asking how small they can be. Note that the question of determining the smallest possible minimum degree among *all* q-Ramsey graphs for a given H is not particularly interesting: by adding an isolated vertex to an arbitrary q-Ramsey graph for H, we obtain another q-Ramsey graph with minimum degree zero. The question becomes much less trivial if we restrict our attention to the collection of *minimal* q-Ramsey graphs for H. This problem was first considered by Burr, Erdős, and Lovász [40] for complete graphs in the two-color setting. We define the corresponding parameter below.

Definition 1.1.1. For a given graph H and integer $q \geq 1$, we define $s_q(H)$ to be the smallest minimum degree among all minimal q-Ramsey graphs for H, that is,

$$s_q(H) = \min\{\delta(G) : G \in \mathcal{M}_q(H)\}.$$

As a first step towards understanding this parameter, we can show the following easy bounds, valid for any integer $q \geq 1$ and any graph H:

$$q(\delta(H) - 1) + 1 \leq s_q(H) \leq r_q(H) - 1. \tag{1.1.1}$$

For the upper bound, note that the graph $K_{r_q(H)}$ is q-Ramsey for H and thus contains a *minimal* q-Ramsey graph for H as a subgraph; the minimum degree of the latter graph cannot exceed

$r_q(H) - 1$. The lower bound was observed by Fox and Lin [73] and follows from a simple application of the pigeonhole principle; we provide the argument for completeness. Indeed, the statement is trivial if H contains an isolated vertex, so consider a graph H with $\delta(H) \geq 1$. Suppose that G is a minimal q-Ramsey graph for H containing a vertex v of degree at most $q(\delta(H) - 1)$. By the minimality of G, the graph $G - v$ has an H-free q-coloring. Now, we extend this coloring to the edges incident to v by using each of the q colors at most $\delta(H) - 1$ times. Since $G \rightarrow_q H$, there exists a monochromatic copy of H under this coloring; the coloring of $G - v$ was chosen to be H-free, so this monochromatic copy of H must contain the vertex v. But v is incident to at most $\delta(H) - 1$ edges of any given color and can therefore not be part of a monochromatic copy of H, a contradiction.

As mentioned earlier, the first result concerning minimum degrees of minimal Ramsey graphs appears in the work of Burr, Erdős, and Lovász [40], who showed that $s_2(K_t) = (t-1)^2$. This result is surprising for two reasons. First, while the value of $r_2(K_t)$ is still far from being understood, the smallest minimum degree $s_2(K_t)$ was determined precisely. Second, this result shows that a 2-Ramsey graph for K_t, which by the above discussion must have at least an exponential (in t) number of vertices, can contain a vertex of very small (polynomial in t) degree that is essential for the Ramsey property. This study was subsequently extended to multiple colors. It was shown by Fox, Grinshpun, Liebenau, Person, and Szabó [72] that $s_q(K_t)$ is bounded above by a polynomial in q and t, more specifically, that $s_q(K_t) \leq 8(t-1)^6 q^3$; their bound was later improved to $C(t-1)^5 q^{5/2}$, where $C > 0$ is an absolute constant, by Bamberg, Bishnoi, and Lesgourgues [13]. The asymptotic behavior of $s_q(K_t)$ as a function of each of its parameters is known up to a polylogarithmic factor. More precisely, it was shown by Fox, Grinshpun, Liebenau, Person, and Szabó [72] that, if t is a constant and $q \to \infty$, we have $s_q(K_t) = q^2 \text{polylog}(q)$. For the special case $t = 3$, Fox, Grinshpun, Liebenau, Person, and Szabó [72] and Guo and Warnke [84] determined the correct power of the logarithm, showing that $s_q(K_3) = \Theta(q^2 \log q)$. In the other regime, when $q \geq 2$ is fixed and $t \to \infty$, Hàn, Rödl, and Szabó [86] showed that $s_q(K_t) = t^2 \text{polylog}(t)$.

The parameter s_q has also been studied for other graphs H. For example, the value of s_2 has been determined for a large class of bipartite graphs by Fox and Lin [73], Grinshpun [83], and Szabó, Zumstein, and Zürcher [143]. Further examples can be found in [29, 82, 83].

It turns out that even a small change in the target graph H can significantly alter the value of s_2. In particular, Fox, Grinshpun, Liebenau, Person, and Szabó [71] proved that $s_2(K_t \cdot K_2) = t - 1$, where $K_t \cdot K_2$ denotes the graph obtained from the clique K_t by attaching a pendent edge. Thus, the addition of a pendent edge to the complete graph K_t reduces the value of s_2 from quadratic to linear in t.

In this thesis, we consider several problems related to minimum degrees of minimal Ramsey graphs.

1.1.1.1 Minimal Ramsey graphs for the clique with many vertices of small degree

When considering the known results about the parameter s_q, we observe an interesting pattern: in essentially all studied cases, the easy upper bound from (1.1.1) is far from tight. Many of the constructions proving upper bounds on s_q, however, use certain gadget graphs, which tend to force all vertices but one to have large degrees. It is then natural to ask: *how many* vertices of small degree can a minimal q-Ramsey graph for a given H contain? In particular, can a minimal q-Ramsey graph for H have arbitrarily many vertices of degree $s_q(H)$? This question motivates the following definition, introduced by the author, Clemens, and Gupta in [29].

Definition 1.1.2. Given an integer $q \geq 2$, a graph H is said to be s_q*-abundant* if, for every $k \geq 1$, there exists a minimal q-Ramsey graph for H with at least k vertices of degree $s_q(H)$.

It is not difficult to see that, if $\mathcal{M}_q(H)$ is finite, then H cannot be s_q-abundant. Thus, for example, stars with an odd number of leaves are not s_2-abundant. As it turns out, stars with an even number of leaves, which have infinitely many 2-Ramsey graphs, are also not s_2-abundant; this result follows easily from the characterization of all graphs in $\mathcal{M}_2(K_{1,t})$ given in [40]. More generally, it was argued in [29] that stars are not s_q-abundant for any $q \geq 2$.

It is not immediate whether s_q-abundant graphs exist at all. In [40], Burr, Erdős, and Lovász noted (without proof) that their construction can be generalized to show that cliques are s_2-abundant. In Chapter 3, we will show that this result holds for any number of colors. More precisely, we will show the following theorem.

Theorem 1.1.3. *For any integers $q \geq 2$ and $t \geq 3$, the clique K_t is s_q-abundant.*

Recall from our earlier discussion that, while $s_2(K_t)$ was determined precisely, the value of $s_q(K_t)$ is still unknown for $q \geq 3$. Yet our approach allows us to prove Theorem 1.1.3 without knowing the precise value of $s_q(K_t)$. In fact, we will only need to argue that the upper bound in (1.1.1) is not tight for any $q \geq 2$ and $t \geq 3$, that is, $s_q(K_t) < r_q(K_t) - 1$. Given the known results about $r_q(K_t)$ and $s_q(K_t)$, this fact is easily seen to be true asymptotically, but we will give a construction that proves it directly for all choices of q and t. We will also rely on a connection between $s_q(K_t)$ and a certain packing problem, established by Fox, Grinshpun, Liebenau, Person, and Szabó [72], and the existence of certain gadget graphs.

In [29], we provide further examples of graphs that are s_q-abundant. For instance, we show that the graph $K_t \cdot K_2$ and the path with three edges are both s_2-abundant and that the cycle C_ℓ for $\ell \geq 4$ is s_q-abundant for any $q \geq 2$. In fact, in the next section and Chapter 4, we will see that a moderately sparse random graph is almost surely s_q-abundant for any $q \geq 2$, showing that s_q-abundance is a much more widespread phenomenon.

References: The main result (Theorem 1.1.3) discussed in this section and proved in Chapter 3 and the tools described in Section 2.5 were obtained in collaboration with Dennis Clemens and Pranshu Gupta. The results appear in [29] and the corresponding parts of this thesis are partially based on that paper.

1.1.1.2 Ramsey simplicity of random graphs

As already discussed above, a common pattern in the study of the parameter s_q is that the upper bound in (1.1.1) tends to be far from the truth. It is then natural to ask whether the lower bound from (1.1.1) can ever be tight. Perhaps surprisingly, it turns out that the answer is yes. Following Grinshpun [83], we call target graphs for which the easy lower bound is tight q-Ramsey simple.

Definition 1.1.4. A graph H without isolated vertices is said to be *q-Ramsey simple* if

$$s_q(H) = q(\delta(H) - 1) + 1.$$

If H has isolated vertices, then we say that H is q-Ramsey simple if the graph obtained from H by removing all isolated vertices is q-Ramsey simple.

It was shown by Fox and Lin [73] that complete bipartite graphs are 2-Ramsey simple. Subsequently this result was extended to a much broader class of bipartite graphs, including all trees, even cycles, and biregular bipartite graphs, by Szabó, Zumstein, and Zürcher [143] and later to all 3-connected bipartite graphs by Grinshpun [83] (both for $q = 2$). In fact, it was conjectured by Szabó, Zumstein, and Zürcher that the same should hold for any bipartite graph.

Conjecture 1.1.5 (Szabó, Zumstein, and Zürcher [143])**.** *Every bipartite graph is 2-Ramsey simple.*

This phenomenon also extends beyond the bipartite setting. It was shown in [29] that all cycles of length at least four are q-Ramsey simple for all $q \geq 2$. Grinshpun [83, Theorem 2.1.3] provided another class of examples of not necessarily bipartite graphs that are 2-Ramsey simple.

Theorem 1.1.6 (Theorem 2.1.3 in [83])**.** *If H is a 3-connected graph containing a vertex u of degree $\delta(H)$ such that $N_H(u)$ is contained in an independent set of size $2\delta(H) - 1$, then H is 2-Ramsey simple.*

These results led Grinshpun to propose the following even bolder conjecture.

Conjecture 1.1.7 (Conjecture 2.8.2 in [83])**.** *Every triangle-free graph is 2-Ramsey simple.*

Some evidence in favor of this conjecture was given in [82], where it was shown that regular 3-connected triangle-free graphs satisfying an additional technical condition are 2-Ramsey simple.

Given the aforementioned results about Ramsey simplicity, it is natural to wonder how common this property is: are most graphs q-Ramsey simple? To formalize this question, we introduce the random graph $G(n, p)$. For an integer $n \geq 1$ and a probability $p \in [0, 1]$, the *(binomial) random graph* $G(n, p)$ is a graph on $[n]$ in which each possible edge is present with probability p, independently of all other edges. As is standard in the study of random graphs, we say that a sequence of events $(A_n)_{n=1}^{\infty}$ occurs *with high probability (w.h.p.)* if $\lim_{n\to\infty} \mathbb{P}(A_n) = 1$. For more background related to the binomial random graph, see for example [24, 75].

The Ramsey properties of the random graph $G(n, p)$ have been studied extensively, culminating in the seminal result of Rödl and Ruciński [126, 127]: for any number of colors $q \geq 2$ and most graphs H, they determined the range of values of p for which a random graph $G \sim G(n, p)$ is q-Ramsey for H. More generally, the random graph has played an important role in Ramsey theory, but in most cases $G(n, p)$ plays the role of the host graph G, while the target graph H is fixed in advance. Surprisingly, there has been considerably less work in the setting where the target graph H is itself random. Some results concerning the 2-Ramsey number of a random graph $H \sim G(n, p)$ appear in the papers of Fox and Sudakov [74] and Conlon [49], providing some lower and upper bounds on $r_2(H)$ for different ranges of p, and of Conlon, Fox, and Sudakov [52], showing that $\log r_2(H)$ is well-concentrated.

Returning to Ramsey simplicity, in the world of random graphs, we can ask a more precise question: for which pairs $p = p(n)$ and $q = q(n, p)$ is a random graph $H \sim G(n, p)$ almost surely q-Ramsey simple? Let us first consider the special case $q = 2$. If $p \ll n^{-1}$, it is well known that $H \sim G(n, p)$ is a forest with high probability (see for example [75, Theorem 2.1]), so the result of Szabó, Zumstein, and Zürcher [143] concerning forests implies that H is almost surely 2-Ramsey simple. On the other hand, if $p = 1$, then $H \cong K_n$ and $s_2(K_n) = (n-1)^2 > 2(n-2)+1$ by the work of Burr, Erdős, and Lovász [40], so H is not 2-Ramsey simple. It is then natural to ask whether there exists a threshold p^* such that $H \sim G(n, p)$ is almost surely 2-Ramsey simple whenever $p \ll p^*$ and not 2-Ramsey simple when $p \gg p^*$. It is well known that such a threshold exists for any nontrivial monotone graph property, that is, any property that is preserved under adding edges and that holds for the complete graph but does not hold for the empty graph (see for example [75, Theorem 1.7]). In the case of 2-Ramsey simplicity it is not clear whether such a threshold exists, as this property is not monotone. For example, from our earlier discussions we know that any tree on t vertices is 2-Ramsey simple, whereas the clique K_t is not. Similarly, there exist graphs that are 2-Ramsey simple but contain subgraphs that are not. For instance, consider the graph obtained from the complete bipartite graph $K_{t,2t-1}$ by adding all possible edges within the smaller vertex class, creating a copy of K_t, and a new vertex connected to exactly t vertices from the larger vertex class. Using Theorem 1.1.6, we conclude that this graph is 2-Ramsey simple for any $t \geq 3$. However, K_t, and hence also the graph obtained by adding $2t$ isolated vertices to K_t, is not 2-Ramsey simple.

The first result concerning the 2-Ramsey simplicity of a random graph appears in the PhD thesis of Grinshpun [83]. Using Theorem 1.1.6, Grinshpun proved that sparse random graphs are 2-Ramsey simple with high probability.

Theorem 1.1.8 (Corollary 2.1.4 in [83])**.** *Let $p = p(n) \in (0, 1)$ and $H \sim G(n, p)$. If $\frac{\log n}{n} \ll p \ll n^{-2/3}$, then w.h.p. H is 2-Ramsey simple.*

Note in particular that Grinshpun's result implies that there are many non-bipartite and non-triangle-free graphs that are 2-Ramsey simple: indeed, it is well known (see for example [75, Theorem 5.3]) that, if $p \gg n^{-1}$, then with high probability $H \sim G(n, p)$ contains cycles of every fixed length. We are interested in extending Grinshpun's result to other ranges of the parameters p and q.

As discussed above, even for two colors, Ramsey simplicity is not a monotone graph property. We will see in Chapter 4, however, that q-Ramsey simplicity *is* monotone in the number of colors q: we will show in Lemma 4.1.1 that, if a graph H is q-Ramsey simple, then it is necessarily $(q-1)$-Ramsey simple. This motivates us to define the *simplicity threshold* $\widetilde{q}(H)$ for a graph H to be the largest number of colors q for which H is q-Ramsey simple, that is,

$$\tilde{q}(H) := \sup\{q : \ H \text{ is } q\text{-Ramsey simple}\}.$$

Observe that every graph H is 1-Ramsey simple, since H is the only minimal 1-Ramsey graph for itself. If H is q-Ramsey simple for all $q \geq 1$, we have $\widetilde{q}(H) = \infty$.

The main result of Chapter 4 estimates the simplicity threshold $\tilde{q}(H)$ when $H \sim G(n, p)$ for different ranges of p. Unless otherwise specified, we assume that $V(H) = [n]$, so the vertices of H come with a natural ordering.

Theorem 1.1.9. *Let $p = p(n) \in (0, 1)$ and $H \sim G(n, p)$. Let $u \in V(H)$ be the smallest (with respect to the natural vertex ordering) vertex of degree $\delta(H)$, and let $F = H[N(u)]$ be the subgraph of H induced by the neighborhood of u. Denote by $\lambda(F)$ the order of the largest connected component in F. Then w.h.p. the following bounds hold:*

(a) $\tilde{q}(H) = \infty$ *if* $0 < p \ll n^{-1}$.

(b) $\tilde{q}(H) = \infty$ *if* $\frac{\log n}{n} \ll p \ll n^{-\frac{2}{3}}$.

(c) $\tilde{q}(H) \geq (1 + o(1)) \max\left\{\frac{\delta(H)}{\lambda(F)^2}, \frac{\delta(H)}{80 \log n}\right\}$ *if* $n^{-\frac{2}{3}} \ll p \ll n^{-\frac{1}{2}}$.

(d) $\tilde{q}(H) \leq (1 + o(1)) \min\left\{\frac{\delta(H)}{\Delta(F)}, \frac{\delta(H)^2}{2e(F)}\right\}$ *if* $n^{-\frac{2}{3}} \ll p \ll 1$.

(e) $\tilde{q}(H) = 1$ *if* $\left(\frac{\log n}{n}\right)^{1/2} \ll p < 1$.

Note that the choice of u as the *smallest* vertex of minimum degree is a technicality. In fact, if $\frac{\log n}{n} \ll p \ll 1$, and in particular in the ranges where the bounds explicitly depend on the neighborhood graph F, a random graph almost surely has a *unique* vertex of minimum degree (see for example [24, Theorem 3.9]). In words, Theorem 1.1.9 shows that, if $p \ll n^{-2/3}$, a random graph $H \sim G(n,p)$ is almost surely q-Ramsey simple for any $q \geq 2$, except possibly in the range $\Omega(n^{-1}) = p = O\left(\frac{\log n}{n}\right)$. In particular, part (b) extends Grinshpun's result to an arbitrary number of colors, even when the number of colors is allowed to depend on n and p. When p is above $n^{-2/3}$, the simplicity threshold becomes finite, as can be seen from part (d). In part (c), we provide lower bounds for the simplicity threshold in the intermediate range $n^{-2/3} \ll p \ll n^{-1/2}$, which are reasonably close to the corresponding upper bounds given by part (d) (this will become clear in Corollary 1.1.10 below). Finally, part (e) shows that, in the dense range $p \gg \left(\frac{\log n}{n}\right)^{1/2}$, a random graph is almost surely not even 2-Ramsey simple.

As mentioned earlier, Theorem 1.1.8 follows from Theorem 1.1.6; in particular, when $\frac{\log n}{n} \ll p \ll n^{-2/3}$, the neighborhood graph F is almost surely empty. As we will see in Chapter 4, in the range $n^{-2/3} \ll p \ll n^{-1/2}$, the neighborhood graph of the minimum degree vertex is almost surely *not* empty. To prove our results, we relate the problem of determining whether a random graph with $\frac{\log n}{n} \ll p \ll 1$ is q-Ramsey simple to a certain packing problem (different from the one introduced in [72] and discussed in the previous section), showing that the answer depends heavily on the structure of the neighborhood graph F. The rest of the proof in that range involves constructing solutions to the packing problem for each of the different intervals, or showing that none exists. To the best of our knowledge, part (c) above provides the first examples of graphs that are q-Ramsey simple for some $q \geq 2$ but contain no minimum degree vertex whose neighborhood is an independent set.

Our ideas show that in the range $\frac{\log n}{n} \ll p \ll 1$ the simplicity threshold for $H \sim G(n,p)$ is infinite whenever the neighborhood graph F is empty and finite otherwise (this can be seen from the proofs of parts (b) and (d)). When $p = \Theta\left(n^{-\frac{2}{3}}\right)$, it can be shown that F is nonempty with probability bounded away from 0 and 1; as a result, in that case $\tilde{q}(H)$ is finite with probability bounded away from 0 and 1.

When $n^{-\frac{2}{3}} \ll p \ll \left(\frac{\log n}{n}\right)^{\frac{1}{2}}$, standard results concerning the value of $\delta(H)$ for $H \sim G(n,p)$ (see Lemma 4.3.2(c)) imply that $\tilde{q}(H) > 2$ with high probability. By analyzing the structure of random graphs and in particular of the neighborhood graph F, we provide more precise quantitative estimates for $\tilde{q}(H)$ in this intermediate range in Corollary 1.1.10 below.

Corollary 1.1.10. *Let $k \geq 2$ be a fixed integer, and let $f = f(n)$ satisfy $1 \ll f = n^{o(1)}$. Further, let $p = p(n)$ satisfy $n^{-\frac{2}{3}} \ll p \ll \left(\frac{\log n}{n}\right)^{\frac{1}{2}}$ and $H \sim G(n,p)$. Then w.h.p. the following bounds hold:*

(a) $(1+o(1))\frac{np}{k^2} \leq \tilde{q}(H) \leq (1+o(1))\frac{np}{k-1}$ *if* $n^{-\frac{k}{2k-1}} \ll p \ll n^{-\frac{k+1}{2k+1}}$.

(b) $(1+o(1))\frac{np}{(k+1)^2} \le \tilde{q}(H) \le (1+o(1))\frac{np}{k-1}$ *if* $p = \Theta\left(n^{-\frac{k+1}{2k+1}}\right)$.

(c) $(1+o(1))\frac{np}{\log n}\max\left\{\frac{16\log^2 f}{\log n}, \frac{1}{80}\right\} \le \tilde{q}(H) \le (1+o(1))\frac{np}{\log n}2\log(f^2\log n)$ *if* $p = n^{-\frac{1}{2}}f^{-1}$.

(d) $1 \le \tilde{q}(H) \le (8+o(1))\frac{1}{p}$ *if* $n^{-\frac{1}{2}} \ll p \ll \left(\frac{\log n}{n}\right)^{\frac{1}{2}}$.

Corollary 1.1.10 shows that in the range $n^{-\frac{2}{3}} \ll p \ll n^{-\frac{1}{2}}$ the simplicity threshold for $H \sim G(n,p)$ is np up to a constant (when $p \ll n^{-\frac{1}{2}-\varepsilon}$) or polylogarithmic (when $p = n^{-\frac{1}{2}-o(1)}$) factor. Most surprisingly, these bounds reveal that the simplicity threshold $\tilde{q}(H)$ evolves in a complicated fashion. Considering that it drops from ∞ to 1 as p ranges from $\frac{\log n}{n}$ to 1, we might expect that the threshold decreases as p grows in that range. However, from the above bounds we see it must actually *increase* in the ranges $p \in \left(n^{-\frac{k}{2k-1}}, n^{-\frac{k+1}{2k+1}}\right)$ for each fixed k: indeed, plugging $p = n^{-\frac{k}{2k-1}}$ into the upper bound given in part (a) and $p = n^{-\frac{k+1}{2k+1}}$ into the lower bound from the same part, we find that the latter exceeds the former. Figure 1.1 summarizes the results given by Theorem 1.1.9 and Corollary 1.1.10.

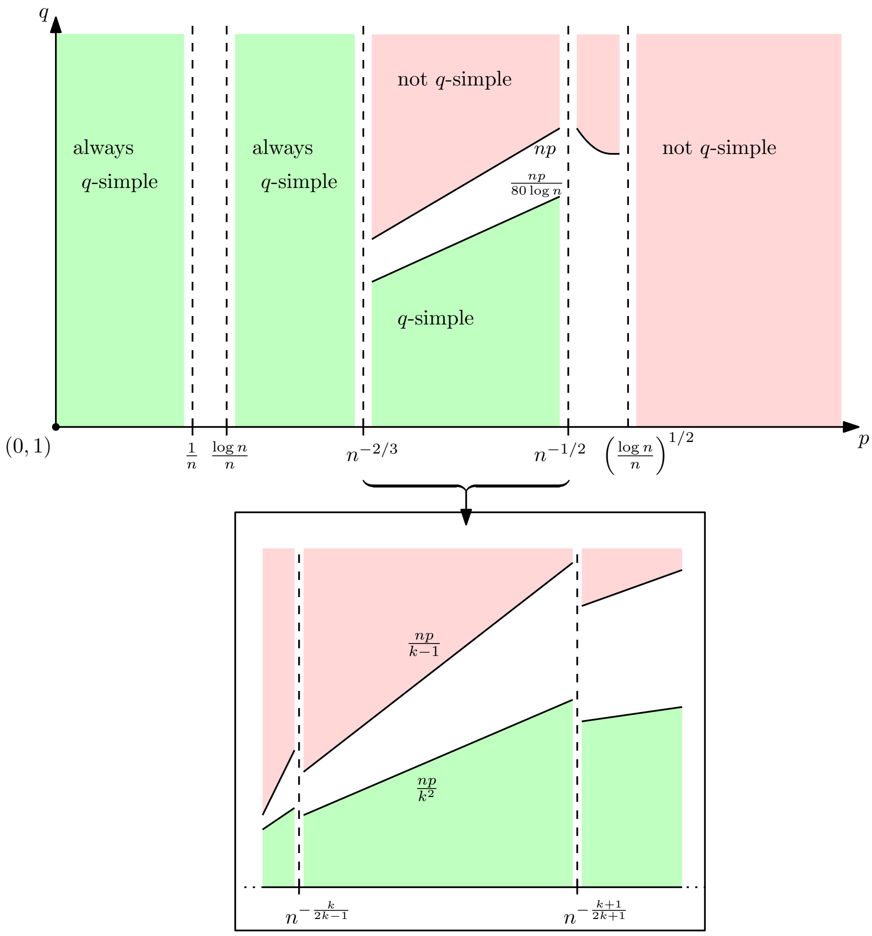

FIGURE 1.1: Bounds on the simplicity threshold $\tilde{q}(G(n,p))$

As discussed above, a key part of the proof of Theorem 1.1.9 involves relating the q-Ramsey simplicity problem to a particular packing problem. In fact, this result applies to any graph H

satisfying particular mild pseudorandom properties (see Definition 4.2.2). This in turn allows us to show Ramsey simplicity for the class of 3-connected triangle-free graphs satisfying those properties, which provides further evidence that Conjecture 1.1.7 might be true.

Our approach also allows us to extend Theorem 1.1.9(b) and (c) and show that, for the same values of q, a random graph $H \sim G(n, p)$ is not just q-Ramsey simple but almost surely s_q-abundant.

Proposition 1.1.11. *Let $p = p(n) \in (0, 1)$ and $H \sim G(n, p)$. Let $u \in V(H)$ be the smallest (with respect to the natural vertex ordering) vertex of degree $\delta(H)$ and let $F = H[N(u)]$ be the subgraph of H induced by the neighborhood of u. Denote by $\lambda(F)$ the order of the largest connected component in F. Then w.h.p. the following is true:*

(a) H is s_q-abundant for all $q \geq 2$ *if* $\frac{\log n}{n} \ll p \ll n^{-\frac{2}{3}}$.

(b) H is s_q-abundant for any $2 \leq q \leq (1 + o(1)) \max\left\{\frac{\delta(H)}{\lambda(F)^2}, \frac{\delta(H)}{80 \log n}\right\}$ *if* $n^{-\frac{2}{3}} \ll p \ll n^{-\frac{1}{2}}$.

Further, we will see in Chapter 4 that our ideas extend easily to the asymmetric setting for pairs of graphs; we introduce the necessary notions in the next section.

References: The results discussed in this section and proved in Chapter 4 (Theorem 1.1.9, Corollary 1.1.10, and Proposition 1.1.11) represent joint work with Dennis Clemens, Shagnik Das, and Pranshu Gupta. The corresponding parts of the thesis were adapted from [27].

1.1.1.3 Minimum degrees of minimal Ramsey graphs in the asymmetric setting

The notion of Ramsey graphs, together with all other related concepts discussed earlier, lends itself to a natural asymmetric generalization. Let $q \geq 2$ and $(H_1, \ldots, H_q)$ be a q-tuple of graphs. We say that a graph G is *q-Ramsey* for $(H_1, \ldots, H_q)$ if, for any q-edge-coloring of G, there exists a color $i \in [q]$ such that G contains a monochromatic copy of H_i in color i. In this case, we write $G \longrightarrow_q (H_1, \ldots, H_q)$. We then define minimal Ramsey graphs, Ramsey numbers, and minimum degrees of minimal Ramsey graphs in the natural way. To simplify notation, as we have been doing until now, in the case $H_1 = \cdots = H_q$, we simply write H instead of $(H, \ldots, H)$. Asymmetric tuples have been considered since the early days of Ramsey theory, already in the work of Erdős and Szekeres [64], and have received considerable attention, especially in the context of Ramsey numbers (see the survey [51] for some examples of such results).

We are again interested in the parameter s_q. Using similar arguments as in the symmetric case, we can easily show the following bounds on $s_q(H_1, \ldots, H_q)$:

$$\sum_{i=1}^{q}(\delta(H_i) - 1) < s_q(H_1, \ldots, H_q) \leq r_q(H_1, \ldots, H_q) - 1. \tag{1.1.2}$$

While in the symmetric case the parameter s_q has been studied for a number of different target graphs, much less has been done in the asymmetric setting. The only published result concerning s_q in the asymmetric setting that we are aware of appears in the original paper of Burr, Erdős, and Lovász [40] and concerns pairs of cliques, showing that $s_2(K_s, K_t) = (s-1)(t-1)$. In particular, we are not aware of any such results for pairs of graphs (H_1, H_2) where H_1 and H_2 have different structures or of any results concerning s_q in the asymmetric setting when $q > 2$. We remark that, to our best knowledge, even the value of s_q for a triple of different cliques is open.

It is then natural to consider pairs of graphs (H, K_t), where H is a very sparse graph such as a tree or a cycle. Pairs of the form (T_ℓ, K_t) and (C_ℓ, K_t), where T_ℓ stands for any tree on ℓ vertices and C_ℓ is the cycle on ℓ vertices, have already been studied in Ramsey theory, in the context of Ramsey numbers. A classic result by Chvátal [45] states that $r_2(T_\ell, K_t) = (\ell-1)(t-1)+1$. In fact, any red/blue-coloring witnessing the inequality $r_2(T_\ell, K_t) > (\ell-1)(t-1)$ is very special; this fact will allow us to easily determine $s_2(T, K_t)$ for any tree T that is not a single vertex and any $t \geq 3$.

Proposition 1.1.12. *For all integers $\ell \geq 2$ and $t \geq 3$ and any tree T_ℓ on ℓ vertices, we have $s_2(T_\ell, K_t) = t-1$.*

The Ramsey number $r_2(C_\ell, K_t)$ has received considerably more attention, as it has turned out to be much more difficult to understand and behaves differently depending on the magnitude of ℓ. After decades of effort by researchers, the study of these Ramsey numbers has resulted in several very recent breakthroughs. The case where $\ell = 3$ defaults to the notoriously difficult asymmetric Ramsey number $r_2(K_3, K_t)$, which is known to be between $(1/4 + o(1))t^2/\log t$ and $(1 + o(1))t^2/\log t$, as shown by Bohman and Keevash [23], Fiz Pontiveros, Griffiths, and Morris [69], and Shearer [138], following earlier results by Ajtai, Komlós, and Szemerédi [2] and Kim [100]. At the other end of the spectrum, Keevash, Long, and Skokan [99] showed that $r_2(C_\ell, K_t) = (\ell-1)(t-1)+1$ for $\ell = \Omega(\log t/\log\log t)$, and that this bound on ℓ is best possible for the equality to hold. For a more detailed discussion of the history of $r_2(C_\ell, K_t)$ and other known results, we refer the reader to [99].

We determine the value of $s_2(C_\ell, K_t)$ precisely for all $t \geq 3$ and $\ell \geq 4$. In particular, we show that, unlike the Ramsey number, our parameter of interest is independent of ℓ. Further, we determine s_2 for pairs of cycles. The study of the Ramsey number in this case was completed already in the 1970s by Rosta [130] and Faudree and Schelp [67], and the answer again depends on the values of ℓ and k. The minimum degree s_2, however, is independent of either cycle length.

Theorem 1.1.13. *For all integers $t \geq 3$ and $k, \ell \geq 4$, we have*

(i) $s_2(C_\ell, C_k) = 3$.

(ii) $s_2(C_\ell, K_t) = 2(t-1)$.

Note that $s_2(K_t, K_s)$ and $s_2(C_\ell)$ have been previously determined by Burr, Erdős, and Lovász [40] and the author, Clemens, and Gupta [29]. Thus, Theorem 1.1.13 completes the study of s_2 for pairs of graphs, each of which is a complete graph or a cycle.

Next, we venture into the multicolor setting. Our goal is to investigate the parameter s_q for tuples consisting of multiple cliques and multiple cycles. Recall that $s_q(K_t)$ was studied by several groups of authors [13, 72, 84, 86], who showed that $s_q(K_t)$ is at most polynomial in both q and t and determined its growth up to a polylogarithmic factor when one of q and t tends to infinity while the other is fixed. The analogous problem for cycles was resolved in [29], where it was established that $s_q(C_\ell) = q + 1$ for all $q \geq 2$ and $\ell \geq 4$.

For given integers $\ell \geq 4$, $t \geq 3$, and $q, q_1, q_2 \geq 0$ satisfying $q = q_1 + q_2$, we define $\mathcal{T} = \mathcal{T}(q_1, q_2, \ell, t)$ to be the q-tuple consisting of q_1 cycles on ℓ vertices and q_2 cliques on t vertices, that is,

$$\mathcal{T}(q_1, q_2, \ell, t) = (\underbrace{C_\ell, \ldots, C_\ell}_{q_1 \text{ times}}, \underbrace{K_t, \ldots, K_t}_{q_2 \text{ times}}), \tag{1.1.3}$$

and let $s_q(\mathcal{T}(q_1, q_2, \ell, t))$ be the smallest minimum degree of a minimal q-Ramsey graph for $\mathcal{T}(q_1, q_2, \ell, t)$. When the parameters are clear from context, we will suppress them from the notation. Our main result in the multicolor setting shows that $s_q(\mathcal{T}(q_1, q_2, \ell, t))$ is independent of ℓ and establishes bounds on $s_q(\mathcal{T}(q_1, q_2, \ell, t))$, relating it to q_1, $s_{q_2}(K_t)$, and $s_q(K_t)$.

Theorem 1.1.14. *For any $t \geq 3$ and any integers $q_1, q_2 \geq 1$, there exists a function $f = f(q_1, q_2, t)$ such that $s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t)) = f(q_1, q_2, t)$ for all $\ell \geq 4$ and*

$$s_{q_2}(K_t) + q_1 \leq f(q_1, q_2, t) \leq s_{q_1+q_2}(K_t). \tag{1.1.4}$$

A key step in the proof of Theorem 1.1.14 is to reformulate the problem of determining $s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t))$ as a packing problem; this step extends the ideas of Fox, Grinshpun, Liebenau, Person, and Szabó [72] and is the content of Lemma 5.2.2. The proof of this equivalence relies on the existence of certain gadget graphs, which were previously known to exist for cliques but not for the tuples we are interested in. Our main technical contribution in the asymmetric setting is the construction of such gadgets, presented in Section 2.6.

Substituting the known bounds for $s_q(K_t)$ into (1.1.4), we can deduce the following quantitative estimates on $s_q(\mathcal{T}(q_1, q_2, \ell, t))$.

Corollary 1.1.15.

(i) For all $t \geq 4$ and $q_1 \geq 1$, there exist constants $c = c(q_1, t) > 0$ and $C = C(q_1, t) > 0$ such that, for all $\ell \geq 4$ and $q_2 \geq 1$, we have

$$c\, q_2^2 \frac{\log q_2}{\log\log q_2} \leq s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t)) \leq C q_2^2 (\log q_2)^{8(t-1)^2}.$$

(ii) For all $q_1 \geq 1$, there exist constants $c = c(q_1) > 0$ and $C = C(q_1) > 0$ such that, for all $\ell \geq 4$ and $q_2 \geq 1$, we have

$$c\, q_2^2 \log q_2 \leq s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, 3)) \leq C q_2^2 \log q_2.$$

(iii) For all $q_1, q_2 \geq 1$, there exists a constant $C = C(q_1, q_2) > 0$ such that, for all $\ell \geq 4$ and $t \geq 3$, we have

$$(t-1)^2 \leq s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t)) \leq C t^2 \log^2 t.$$

Thus, Theorem 1.1.14 is sufficient to determine $s_q(\mathcal{T}(q_1, q_2, \ell, t))$ for arbitrary $\ell \geq 4$ up to polylogarithmic factors when $q_2 \to \infty$ and q_1 and t are fixed, and when $t \to \infty$ and q_2 and q_1 are fixed. Similarly, the bounds in [13, 72] yield bounds on $s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t))$ that are polynomial in all of t, q_1, and q_2.

It remains to consider the case where $q_1 \to \infty$ and q_2 and t are fixed. In this case the lower bound in (1.1.4) is linear in q_1 while the upper bound is essentially quadratic in q_1. Using a different approach, again relying on Lemma 5.2.2, we prove that $s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t))$ is asymptotically equal to q_1.

Theorem 1.1.16. *For all $\ell \geq 4$, $t \geq 3$, $q_2 \geq 1$, and $\varepsilon > 0$, there exists q_0 such that for all $q_1 \geq q_0$, we have*

$$s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t)) \leq (1+\varepsilon) q_1.$$

References: The results discussed in this section and proved in Chapter 5 (Proposition 1.1.12, Theorems 1.1.13, 1.1.14 and 1.1.16, and Corollary 1.1.15), together with the theory of set-senders and set-determiners discussed in Chapter 2 (in particular Section 2.6) represent joint work with Anurag Bishnoi, Dennis Clemens, Pranshu Gupta, Thomas Lesgourgues, and Anita Liebenau. The corresponding parts of the thesis were lightly adapted from [16] (`https://doi.org/10.1137/21M1444953`).

1.1.2 Ramsey non-equivalent graphs

The notion of Ramsey equivalence arose naturally in the work of Szabó, Zumstein, and Zürcher [143] and is motivated by the question: how does the collection of q-Ramsey graphs change when we alter the target graph (slightly)? We can ask more specific questions such as: how does the Ramsey number change or how does the parameter s_q change? In some cases, it can happen that two graphs have the same Ramsey number or the same smallest minimum degree s_q, or simply that these specific parameters are difficult to determine, but we would still like to understand whether these two graphs are "the same" from the point of view of Ramsey graphs.

Definition 1.1.17. Given an integer $q \geq 2$, we say that two graphs H and H' are q-Ramsey equivalent if $\mathcal{R}_q(H) = \mathcal{R}_q(H')$.

After its introduction in [143], the notion of Ramsey equivalence has been studied by several groups of authors, see for example [5, 22, 46, 71, 83, 134]. Szabó also proposed a weaker version of this concept, calling two graphs H_1 and H_2 *q-Ramsey close* if $\mathcal{M}_q(H_1)$ and $\mathcal{M}_q(H_2)$ differ only by a finite number of graphs (see [22] for more about this notion).

One of the questions that has received the most attention in this area asks what graphs are q-Ramsey equivalent to the complete graph K_t. It is not difficult to see that, for any $q \geq 2$ and $t \geq 3$, we can always add an isolated vertex (or even $r_q(K_t) - t$ such vertices) and obtain a graph that is q-Ramsey equivalent to K_t. Further examples of disconnected graphs of the form $K_t \cup H$, where $V(K_t) \cap V(H) = \emptyset$ and $\omega(H) < t$, that are 2-Ramsey equivalent to K_t are given in [22, 143]. Furthermore, from the early work of Nešetřil and Rödl [120], we know that no graph with clique number different from t can be q-Ramsey equivalent to K_t for any $q \geq 2$. In light of these results, it is natural to wonder whether there are any connected graphs that are q-Ramsey equivalent to a complete graph for some $q \geq 2$. Perhaps surprisingly, it turned out that the answer is no. Fox, Grinshpun, Liebenau, Person, and Szabó showed that, for all $t \geq 3$, the graph $K_t \cdot K_2$ obtained by attaching a pendent edge to the clique K_t is not 2-Ramsey equivalent to K_t, thus implying that no connected graph (other than K_t itself) can be 2-Ramsey equivalent to K_t. We remark here that it was shown by Burr, Erdős, Faudree, and Schelp [39] that K_t and $K_t \cdot K_2$ *do* have the same 2-color Ramsey number when $t \geq 4$. The non-equivalence of K_t and $K_t \cdot K_2$ was later established also for arbitrarily many colors by Clemens, Liebenau, and Reding [46]. It is then natural to wonder by *how much* the collections $\mathcal{R}_q(K_t)$ and $\mathcal{R}_q(K_t \cdot K_2)$ differ. In this section, we address this question in the special case $q = 2$, strengthening the result of Fox, Grinshpun, Liebenau, Person, and Szabó.

The precise question we want to study was asked by Tibor Szabó and is inspired by the work of Rödl and Siggers [128], who showed that, for any integers $t \geq 3$ and $q \geq 2$, as $n \to \infty$, the collection $\mathcal{M}_q(K_t)$ contains at least $2^{\Omega(n^2)}$ graphs on at most n vertices. Our goal is to study *how many* graphs there are in $\mathcal{M}_2(K_t) \setminus \mathcal{R}_2(K_t \cdot K_2)$. Our first result in this section provides a common generalization of the results of Fox, Grinshpun, Liebenau, Person, and Szabó and Rödl and Siggers, showing that, for every $t \geq 3$, as $n \to \infty$, there are many graphs in $\mathcal{M}_2(K_t) \setminus \mathcal{R}_2(K_t \cdot K_2)$ on at most n vertices. In particular, this result shows that K_t and $K_t \cdot K_2$ are very far from being 2-Ramsey close.

Theorem 1.1.18. *For every $t \geq 3$, as $n \to \infty$, there are $2^{\Omega(n^2)}$ graphs on at most n vertices that are minimal 2-Ramsey for K_t but are not 2-Ramsey for $K_t \cdot K_2$.*

To prove the theorem, we combine the construction of Rödl and Siggers with an idea of Grinshpun, allowing us to construct gadget graphs with convenient properties.

We also study the analogous problem for cycles. Cycles have a very different structure compared to complete graphs and are much sparser, and it is again natural to wonder whether the collection of Ramsey graphs changes after adding a pendent edge. We show that for $q = 2$ the answer is yes, which to our best knowledge was not previously known. Here $C_\ell \cdot K_2$ denotes the graph obtained by adding a pendent edge to the cycle C_ℓ.

Proposition 1.1.19. *For every $\ell \geq 4$, the graphs C_ℓ and $C_\ell \cdot K_2$ are not 2-Ramsey equivalent.*

In the cycle setting, Siggers [139, 141] showed that, for any odd $\ell \geq 5$ and $q \geq 2$, $\mathcal{M}_q(C_\ell)$ contains at least $2^{\Omega(n^2)}$ graphs on at most n vertices as $n \to \infty$. Note that, as observed in [141], we cannot expect a result of this type for even cycles: even cycles are bipartite and it is well known that their extremal number is subquadratic [106]. Thus, every minimal q-Ramsey graph for an even cycle must contain $o(n^2)$ edges, and thus there are at most $2^{o(n^2)}$ such graphs on at most n vertices. We again provide a common strengthening of the results of Siggers and Proposition 1.1.19 for odd cycles when $q = 2$.

Theorem 1.1.20. *For every odd $\ell \geq 5$, as $n \to \infty$, there are $2^{\Omega(n^2)}$ graphs on at most n vertices that are minimal 2-Ramsey for C_ℓ but are not 2-Ramsey for $C_\ell \cdot K_2$.*

The construction needed to prove Theorem 1.1.20 is considerably simpler than that needed to prove Theorem 1.1.18 and differs slightly from that of Siggers. Once again, an important step of the proof involves showing the existence of certain gadget graphs. Proposition 1.1.19 is a simple corollary of the existence of these gadget graphs.

The main results in this section, namely Theorems 1.1.18 and 1.1.20 and Proposition 1.1.19, are proved in Chapter 6.

1.2 Enumerating orthogonal Latin squares

The goal of Part II of this thesis is to answer questions concerning the enumeration of orthogonal Latin squares. We begin by presenting some of the relevant background.

Given an integer $n \geq 1$, a *Latin square of order n* is an $n \times n$ matrix with entries in $[n]$ such that each $x \in [n]$ appears exactly once in every row and in every column. For a Latin square L, we denote the entry in row i and column j by $L(i, j)$. It is not difficult to see that Latin squares of order n exist for all n; indeed, a rich class of constructions is given by the Cayley tables of groups. We refer the reader to [147] for more definitions, results and proofs related to Latin squares, noting only that the number of Latin squares is log-asymptotically given by

$$L(n) = \left((1 + o(1))\frac{n}{e^2}\right)^{n^2}. \tag{1.2.1}$$

Ryser [131] showed that the lower bound follows from Van der Waerden's conjecture on permanents of matrices, which was famously later proven by Egorychev [60] and Falikman [66]. The upper bound is also closely related to permanents, as it is a consequence of Brègman's Theorem [31] (see [147, Chapter 17] for details).

We will be concerned with orthogonal Latin squares. Informally speaking, two Latin squares are orthogonal if superimposing them yields in an $n \times n$ square containing n^2 different ordered pairs of symbols.

Definition 1.2.1. Two Latin squares L, L' of order n are said to be *orthogonal* if, for all pairs $(x, y) \in [n]^2$, there exist unique $i, j \in [n]$ such that $L(i, j) = x$ and $L'(i, j) = y$. In this case, L' is said to be an *orthogonal mate* of L.

A k-tuple of Latin squares $(L_1, \ldots, L_k)$ forms a system of k *mutually orthogonal Latin squares*, or a *k-MOLS*, if for all $1 \leq i < j \leq k$, the squares L_i and L_j are orthogonal.

The early research related to orthogonal Latin squares concerned the existence of k-MOLS. In particular, there was much interest in the maximum size of a set of mutually orthogonal Latin squares; that is, the function $N(n) = \max\{k : \text{a } k\text{-MOLS of order } n \text{ exists}\}$.

It is well known that $N(n) \leq n - 1$ (see for example [147, Theorem 22.1]). Indeed, suppose $(L_1, \ldots, L_k)$ is a k-MOLS. Observing that orthogonality is preserved under permutations of the symbols within each square, we may assume that the first row of each square is $[1, 2, \ldots, n]$. Considering the entries in position $(2, 1)$, we find that all $L_i(2, 1)$ must be distinct, by orthogonality, and different from 1, since the L_i are Latin squares. Hence $k \leq n - 1$.

It is also well known that equality holds if and only if a projective plane of order n exists (see for example Theorem 22.2 and the corollary following it in [147]). This shows that the precise determination of the function $N(n)$ is likely to be difficult, as the existence of projective planes for orders n that are not prime powers is a longstanding open problem. Still, several polynomial lower bounds on $N(n)$ with ever-improving exponents appear in the literature [44, 129, 153], with the largest one due to Lu [112], who proved $N(n) = \Omega(n^{1/14.3})$.

Given that large sets of mutually orthogonal Latin squares exist, it is natural to extend (1.2.1) and enumerate k-MOLS for $k \geq 2$. Early work in this direction was undertaken by Donovan and Grannell [56], who constructed many k-MOLS, and also sought to bound the number of orthogonal mates a Latin square can have. An important component of their argument is an upper bound on number of transversals in a Latin square, where a *transversal* is a selection of n cells from the square, with no two sharing the same row, column or symbol. Taranenko [145] later proved a sharp upper bound on the number of transversals in a Latin square, which, when used in Donovan and Grannell's proof, shows that a Latin square can have at most

$$\left((1+o(1))\frac{n}{e^{2+1/e}}\right)^{n^2} \tag{1.2.2}$$

orthogonal mates. Coupled with (1.2.1), this can be used to give upper bounds on the number of pairs of orthogonal Latin squares and, more generally, the number of k-MOLS (since in a k-MOLS $(L_1, \ldots, L_k)$, the Latin squares $L_2, \ldots, L_k$ must all be orthogonal mates of L_1).

More recently, tight bounds on the number of k-MOLS follow from the breakthroughs of Luria [113] and Keevash [98]. Through an elegant entropic argument, Luria gives a general upper bound on the number of perfect matchings a regular r-uniform hypergraph can have. Assuming certain pseudorandom conditions, Keevash provides a matching lower bound, coupling randomized constructions with the use of absorbers. When applied to the enumeration of k-MOLS, their theorems imply the following result, where we denote the number of k-MOLS of order n by $L^{(k)}(n)$.

Theorem 1.2.2 (Luria, 2017, and Keevash, 2018)**.** *For every fixed $k \in \mathbb{Z}_{\geq 1}$, the number of k-MOLS of order n is*

$$L^{(k)}(n) = \left((1+o(1))\frac{n^k}{e^{\binom{k+2}{2}-1}}\right)^{n^2}. \tag{1.2.3}$$

1.2.1 Results

The one drawback of Theorem 1.2.2 is that both the lower and upper bounds in (1.2.3) require k to be fixed as n tends to infinity. We seek upper bounds that hold when k grows with n. We combine the approach of Donovan and Grannell [56] with the method of Luria [113], using entropy to bound the number of ways of extending a k-MOLS by adding an additional Latin square.

Before presenting our upper bound, let us discuss a lower bound for this number of extensions. Since every $(k+1)$-MOLS contains a k-MOLS as a prefix, Theorem 1.2.2 implies that, for fixed $k \in \mathbb{Z}_{\geq 0}$, the average number of extensions of a k-MOLS to a $(k+1)$-MOLS is at least

$$\frac{L^{(k+1)}(n)}{L^{(k)}(n)} = \left((1+o(1))\frac{n}{e^{k+2}}\right)^{n^2}. \tag{1.2.4}$$

This clearly gives a lower bound for the maximum number of such extensions. In the following theorem, we provide an upper bound that is valid for all k.

Theorem 1.2.3. *For $0 \leq k \leq n-2$, the logarithm of the number of ways to extend a k-MOLS of order n to a $(k+1)$-MOLS is at most*

$$n^2 \int_0^1 \log(1+(n-1)t^{k+2})\,\mathrm{d}t.$$

We will estimate the value of this integral in Lemma 7.2.2. As a corollary, combining Theorem 1.2.3 with (1.2.4) allows us to determine the number of extensions of a k-MOLS of fixed size. In particular, setting $k = 1$ bounds the number of orthogonal mates a Latin square can have, sharpening the bound in (1.2.2).

Corollary 1.2.4. *For every fixed $k \in \mathbb{Z}_{\geq 1}$, the maximum number of ways to extend a k-MOLS of order n to a $(k+1)$-MOLS is*

$$\left((1+o(1))\frac{n}{e^{k+2}}\right)^{n^2}.$$

As previously stated, our primary goal is to bound the number of k-MOLS when k grows with n. We can do so by building the k-MOLS one Latin square at a time, using Theorem 1.2.3 to bound the number of choices at each step. In this way we can recover the upper bound of Theorem 1.2.2 when k is constant, but the main novelty is the following extension to larger values of k.

Corollary 1.2.5. *As $n \to \infty$,*

(a) $\log L^{(k)}(n) \leq \left(k\log n - \binom{k+2}{2} + 1 + k^2 n^{-1/(k+2)}\right)n^2$ *if $k = o(\log n)$,*

(b) $\log L^{(k)}(n) \leq (c(\beta)+o(1))kn^2\log n$ *if $k = \beta \log n$, for fixed $\beta > 0$,*

(c) $\log L^{(k)}(n) \leq \left(\frac{1}{2}+o(1)\right)(\log k - \log\log n)n^2\log^2 n$ *if $k = \omega(\log n)$,*

where in (b) we define $c(\beta) = 1 - \beta^{-1}\int_0^\beta x(1-e^{-1/x})\,\mathrm{d}x \in [0,1]$.

Note that we trivially have $L^{(k)}(n) \leq L(n)^k$, which, in light of (1.2.1), gives the upper bound $\log L^{(k)}(n) \leq kn^2\log n$. Corollary 1.2.5 provides a significant improvement over this trivial bound. Furthermore, part (a) shows that the upper bound from Theorem 1.2.2 is valid whenever $k = o\left(\frac{\log n}{\log\log n}\right)$.

It is well known that mutually orthogonal Latin squares are equivalent to many other combinatorial structures such as transversal designs, nets and orthogonal arrays, while also being related to certain error correcting codes and affine and projective planes (see [147]), and so our results give upper bounds for the number of structures in each of these classes. In fact, we shall prove Theorem 1.2.3 for the more general class of gerechte designs (see Theorem 7.2.1), which allow us to, for instance, bound the number of sets of mutually orthogonal Sudoku squares. We discuss this particular extension further in our concluding remarks.

References: This section and Chapter 7 are based on joint work with Shagnik Das and Tibor Szabó and are adapted, with minor modifications, from [30] (https://doi.org/10.1007/s10623-020-00771-6).

1.3 Hyperplane coverings with multiplicities

In Part III of this thesis, we explore several questions related to grid coverings. Before introducing our new results, we discuss some motivation and relevant history.

1.3.1 An origin story

Grid covering problems have attracted considerable amount of attention from researchers. Recall the motivating example from earlier, where we were asked to determine the minimum number of hyperplanes needed to cover all nonzero points of the hypercube $\{0,1\}^n$, while leaving $\vec{0}$ uncovered. Note that in this example we did not specify what space the grid $\{0,1\}^n$ resides in, but of course the structure of hyperplanes depends heavily on the ambient vector space, and in particular, on the underlying field.

One way to make our question precise is to ask for the minimum number of hyperplanes needed to cover all points of $\mathbb{F}_2^n$ while omitting the origin. Given this formulation, we can then generalize this problem by replacing $\mathbb{F}_2$ with any finite field $\mathbb{F}_q$ and asking how many hyperplanes we need to cover all nonzero points of $\mathbb{F}_q^n$. This problem turns out to be equivalent to the well-known blocking set problem in finite geometry (see e.g. [90]), and it was in this guise that it was first studied. A *blocking set* in $\mathbb{F}_q^n$ is a set of points that meets every hyperplane, and the objective is to find a blocking set of minimum size. To see why the two problems are equivalent, consider a blocking set S. Note that by translating we may assume that S contains the origin $\vec{0}$, and hence we can write $S = \{\vec{0}, \vec{s_1}, \dots, \vec{s_m}\}$. Now, S is a blocking set if and only if it intersects every hyperplane avoiding the origin. Every such hyperplane has the form $\vec{a} \cdot \vec{x} = 1$ for some normal vector $\vec{a} \in \mathbb{F}_q^n \setminus \{\vec{0}\}$. Thus, S is a blocking set if and only if, for every normal vector $\vec{a} \in \mathbb{F}_q^n \setminus \{\vec{0}\}$, there exists an $i \in [m]$ such that $\vec{a} \cdot \vec{s_i} = 1$, which is equivalent to saying that for every point $\vec{a} \in \mathbb{F}_q^n \setminus \{\vec{0}\}$, there exists an $i \in [m]$ such that $\vec{s}_i \cdot \vec{a} = 1$, that is, $\vec{a}$ is contained in the hyperplane with normal vector $\vec{s}_i$. Hence, minimizing the size of a blocking set in $\mathbb{F}_q^n$ is the same as finding the minimum size of a collection of affine hyperplanes covering every nonzero point of $\mathbb{F}_q^n$. Using the blocking set formulation, it is not difficult to see that our earlier construction with n hyperplanes for $\mathbb{F}_2^n$ is optimal: indeed, if S is a blocking set of size $m+1$ containing the origin, then the vectors in S span a subspace of dimension at most m. If $m < n$, then S is fully contained in some hyperplane containing $\vec{0}$ and is thus disjoint from the affine hyperplane parallel to it.

Going further, one may replace the hyperplanes with affine subspaces of codimension d. In this generality, the problem was studied in the late 1970s by Jamison [90], who proved that the minimum number of affine subspaces of codimension d that cover all nonzero points in $\mathbb{F}_q^n$ while avoiding the origin is $q^d - 1 + (n-d)(q-1)$. In particular, when $q = 2$ and $d = 1$, this expression simplifies to n. A simpler proof of the case $d = 1$ for arbitrary q was given independently by Brouwer and Schrijver [32].

A similar problem for infinite fields was raised in the early 1990s by Komjáth [103]: in order to prove some results in infinite Ramsey theory, Komjáth showed that the number of hyperplanes required to cover all points but one of the real hypercube $\{0,1\}^n \subseteq \mathbb{R}^n$ must go to infinity with n. Shortly afterwards, a celebrated result of Alon and Füredi [3] established a tight bound in the more general setting of covering all but one point of a finite grid. Let $S_1, S_2, \ldots, S_n$ be finite subsets of some arbitrary field $\mathbb{F}$; by translating, we may assume that $0 \in S_1 \cap \cdots \cap S_n$ and that the point to be omitted is $\vec{0}$. Then it is not difficult to check that the collection of all hyperplanes of the form $x_i = s$ for all $i \in [n]$ and $s \in S_i \setminus \{0\}$ covers all points but one of the grid $S_1 \times \cdots \times S_n$ while leaving $\vec{0}$ uncovered. Alon and Füredi [3] showed that this construction is best possible, establishing a matching lower bound of $\sum_{i=1}^{n} (|S_i| - 1)$. Observe that this is a generalization of Jamison's result in the case $d = 1$. Note also that the Alon-Füredi result allows us to give a more general answer to our motivating question: if we take $S_i = \{0,1\} \subseteq \mathbb{F}$ for all $i \in [n]$, where $\mathbb{F}$ is *any* field, applying their theorem shows that we always need n hyperplanes to cover the nonzero points of the hypercube.

1.3.2 The polynomial method

Despite these motivating applications to finite geometry and Ramsey theory, the primary reason this problem has attracted so much attention lies in the proof method used. These hyperplane covers have driven the development of the so-called polynomial method in combinatorics; indeed, in light of Jamison's early results [90], this method is sometimes referred to as the Jamison method in finite geometry [34].

To see how polynomials come into play, suppose we have a set of hyperplanes $\{H_i : i \in [m]\}$ in $\mathbb{F}^n$, with the hyperplane H_i defined by $H_i = \{\vec{x} : \vec{a}_i \cdot \vec{x} = c_i\}$ for some normal vector $\vec{a}_i \in \mathbb{F}^n \setminus \{\vec{0}\}$ and some constant $c_i \in \mathbb{F}$. We can then define the degree-m polynomial $f(\vec{x}) = \prod_{i\in[m]} (\vec{a}_i \cdot \vec{x} - c_i)$, observing that $f(\vec{x}) = 0$ if and only if $\vec{x}$ is covered by one of the hyperplanes H_i. Thus, if $S_1 \times \cdots \times S_n \subseteq \mathbb{F}^n$ is a finite grid and $\vec{p} \in S_1 \times \cdots \times S_n$, then lower bounds on the degrees of polynomials that vanish at all points of $S_1 \times \cdots \times S_n$ except for $\vec{p}$ translate to lower bounds on the number of hyperplanes needed to cover all points of $S_1 \times \cdots \times S_n$ but $\vec{p}$.

This approach has proven very robust, and lends itself to a number of generalizations. For instance, Kós, Mészáros and Rónyai [104] and Bishnoi, Clark, Potukuchi and Schmitt [20] considered variations over rings, while Blokhuis, Brouwer and Szőnyi [21] studied the hyperplane covering problem for different subsets of projective and affine spaces over $\mathbb{F}_q$.

1.3.3 Covering with multiplicity

In this thesis, we study coverings with higher multiplicities. That is, for a given finite grid in $\mathbb{F}^n$, we shall seek the minimum number of hyperplanes in $\mathbb{F}^n$ needed to cover all points of the grid but one at least k times, while the remaining point is covered fewer times. Previous work in this direction has imposed the stricter condition of avoiding the remaining point altogether. Bruen [33] considered this problem for the grid $\mathbb{F}_q^n$ and showed a lower bound, generalizing Jamison's lower bound in the case $k = 1$. Later on, Ball and Serra [11] provided a common generalization of the results of Bruen and Alon and Füredi. More precisely, Theorem 5.3 in [11] implies that, if $\mathbb{F}$ is any field and $S_1, \ldots, S_n \subseteq \mathbb{F}$ are finite subsets, then the number of hyperplanes needed to cover all points but one of the grid $S_1 \times \cdots \times S_n$ at least k times while leaving the remaining point uncovered is at least

$$\sum_{i=1}^{n}(|S_i| - 1) + (k - 1) \max_{i\in[n]}(|S_i| - 1). \tag{1.3.1}$$

While the results of Jamison and Alon and Füredi are known to be tight, in the case of higher multiplicities, the tightness of the Bruen and Ball-Serra lower bounds is understood only in some special cases. Some results addressing this problem appear for example in [8, 10, 107, 154]. Further work related to coverings with multiplicities appears for instance in the paper of Kós and Rónyai [105].

To prove their results, Bruen [33] and Ball and Serra [11] strengthened the polynomial method described above to obtain lower bounds for the covering problem with higher multiplicities. We briefly outline the idea. Recall that in the case of $k = 1$, given a collection $\mathcal{H}$ of m hyperplanes defined by $\vec{a}_i \cdot \vec{x} = c_i$ for all $i \in [m]$, we defined the degree-m polynomial $f(\vec{x}) =$

$\prod_{i\in[m]}(\vec{a}_i\cdot\vec{x}-c_i)$; this polynomial vanishes at a point $\vec{y}$ if and only if $\vec{y}$ is covered by one of the hyperplanes in $\mathcal{H}$. Lower bounds on the degree of a polynomial vanishing at all points but one of a given grid then translate directly into lower bounds on the corresponding hyperplane covering problem. Now, to extend this idea to higher multiplicities, suppose that a point $\vec{y}=(y_1,\dots,y_n)$ is covered by exactly $k\geq 1$ hyperplanes in the collection $\mathcal{H}$, that is, $\vec{a}_i\cdot\vec{y}=c_i$ for exactly k different $i\in[m]$. Then the factors in the product corresponding to the k hyperplanes containing $\vec{y}$ in the polynomial $f(\vec{x}+\vec{y})$ have the form $\vec{a}_i\cdot\vec{x}$, that is, they have no constant term, and the factors corresponding to every other hyperplane do have a constant term. Thus, the smallest degree of a monomial in the expansion of $f(\vec{x}+\vec{y})$ corresponds to the number of times the point $\vec{y}$ is covered by the collection $\mathcal{H}$. We say that a polynomial $P\in\mathbb{F}[x_1,\dots,x_n]$ has a *zero of multiplicity* k at a point $(y_1,\dots,y_n)\in\mathbb{F}^n$ if the smallest degree of a monomial occurring in the expansion of $P(x_1+y_1,\dots,x_n+y_n)$ is k (see for example [11, 33]). Then we can obtain lower bounds on the number of hyperplanes needed to cover all points but one of a given grid at least k times while covering the remaining point (at most) $\ell\geq 0$ times by studying the related polynomial problem, that is, by asking for the smallest possible degree of a polynomial that has a zero of multiplicity at least k at all points of the grid but one and a zero of multiplicity (at most) ℓ at the remaining point. Bruen and Ball-Serra used this framework to prove their bounds on the hyperplane covering problem.

Significant progress in this line of research was made recently for the real hypercube $\{0,1\}^n\subseteq\mathbb{R}^n$. Clifton and Huang [47] studied the number of hyperplanes needed to cover all nonzero points of this hypercube at least k times, while leaving the origin uncovered. Observe that we can remove $k-1$ hyperplanes arbitrarily from such a cover, and the remainder will still cover each nonzero point at least once. Thus, by the theorem of Alon and Füredi, we must be left with at least n planes, giving a lower bound of $n+k-1$ (note that the Ball-Serra bound (1.3.1) yields the same result). By considering the collection of $n+1$ hyperplanes given by $x_i=1$ for all $i\in[n]$ and $\sum_{i=1}^n x_i=1$, it is not hard to verify that this bound is tight for $k=2$. Clifton and Huang used Ball and Serra's Punctured Combinatorial Nullstellensatz [11] to improve the lower bound for larger k, showing in particular that the Ball-Serra bound cannot be tight when $k\geq 3$. More precisely, they proved that for $k=3$ and $n\geq 2$ the correct answer is $n+3$, while for $k\geq 4$ and $n\geq 3$ the lower bound can be further improved to $n+k+1$. The latter result was coupled with an upper bound of $n+\binom{k}{2}$, which they conjectured to be correct when n is large with respect to k. However, they showed that this was far from the case when n is fixed and k is large; in this range, the answer is $(c_n+o(1))k$, where c_n is the nth term in the harmonic series.

A major breakthrough was then made by Sauermann and Wigderson [133], who skipped the geometric motivation and resolved the polynomial problem directly.

Theorem 1.3.1 ([133]). *Let $k \geq 2$ and $n \geq 2k-3$, and let $P \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial that has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{\vec{0}\}$ but does not have a zero of multiplicity at least $k-1$ at $\vec{0}$. Then P must have degree at least $n+2k-3$.*

Furthermore, for every $\ell \in \{0, 1, \ldots, k-2\}$, there exists a polynomial P with degree exactly $n+2k-3$ that has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{\vec{0}\}$ and a zero of multiplicity exactly ℓ at $\vec{0}$.

As explained earlier, any lower bound on the polynomial problem directly translates into a lower bound on the corresponding covering problem; thus, as an immediate corollary, Theorem 1.3.1 improves the lower bound in the Clifton–Huang result from $n+k+1$ to $n+2k-3$. However, Theorem 1.3.1 establishes that $n+2k-3$ is also an upper bound for the polynomial problem, whereas Clifton and Huang conjectured that the answer for the covering problem should be $n+\binom{k}{2}$. Thus, resolving this conjecture will likely require new ideas. An affirmative answer to this conjecture would also demonstrate a separation between the algebraic polynomial problem and the geometric covering problem.

Even though Theorem 1.3.1 is stated for polynomials defined over $\mathbb{R}$, Sauermann and Wigderson note that the proof works over any field of characteristic zero. However, the result need not hold over finite fields. In particular, they showed the existence of a polynomial P_4 over $\mathbb{F}_2$ of degree $n+4$ with zeroes of multiplicity four at all nonzero points in $\mathbb{F}_2^n$ and with $P_4(\vec{0}) \neq 0$. More generally, for every $k \geq 4$, $P_k(\vec{x}) = x_1^{k-4}(x_1-1)^{k-4}P_4(\vec{x})$ is a binary polynomial of degree only $n+2k-4$ with zeroes of multiplicity k at all nonzero points and of multiplicity $k-4$ at the origin. The correct behavior of the critical degree of a polynomial with such properties over finite fields is left as an open problem.

Note also that Theorem 1.3.1 allows the origin to be covered up to $k-2$ times. Sauermann and Wigderson also considered the case where the origin must be covered with multiplicity exactly $k-1$, showing that the degree of such a polynomial has to be at least $n+2k-2$, which is again tight (see [133, Theorem 1.5]). In contrast to Theorem 1.3.1, the proof of this result is valid over all fields.

1.3.4 Results

Before we proceed with describing our results, we introduce some terminology that we will use in Chapters 8 and 9. Note that by translation we may always assume that the origin is contained in the grid we want to cover. Given integers $k \geq 1$ and $n \geq d \geq 1$, a field $\mathbb{F}$, and finite subsets $S_1, \ldots, S_n \subseteq \mathbb{F}$ with $0 \in \bigcap_{i=1}^{n} S_i$, we say a multiset $\mathcal{H}$ of $(n-d)$-dimensional affine subspaces in $\mathbb{F}^n$ is a (k,d)-*cover* if every nonzero point of $S_1 \times \cdots \times S_n$ is covered at least k times, while $\vec{0}$

is covered at most $k-1$ times. A (k,d)-cover $\mathcal{H}$ is a *strict* (k,d)*-cover* if $\vec{0}$ is not covered at all. For the sake of simplicity, when $d=1$, we will often suppress it from our notation and call a (strict) (k,d)-cover simply a (strict) k-cover.

1.3.4.1 Subspace coverings over the binary field

In Chapter 8, we study the problem of covering $\mathbb{F}_2^n$ with multiplicity. We are motivated not only by the body of research described above, but also by the fact that, as we shall show in Proposition 8.2.3, when one forbids the origin from being covered, this problem is equivalent to finding linear binary codes of large minimum distance. This classic problem from coding theory has a long and storied history of its own, and is likely to be very difficult. We focus on the less restrictive setting where we require all nonzero points in $\mathbb{F}_2^n$ to be covered at least k times while the origin can be covered at most $k-1$ times. This is another natural generalization to higher multiplicities of the covering problem addressed earlier by Jamison [90], Brouwer and Schrijver [32], and Alon and Füredi [3].

In light of the previous results, we will abstain from employing the polynomial method, and will instead attack the problem more directly with combinatorial techniques. In particular, we will exploit the connection to coding theory established in Proposition 8.2.3. As an added bonus, our arguments readily generalize to covering points with codimension-d affine subspaces, rather than just hyperplanes, thereby extending Jamison's original results in the case $q=2$.

Given integers $k \geq 1$ and $n \geq d \geq 1$, we define the extremal function $f(n,k,d)$ to be the minimum possible size of a (k,d)-cover of $\mathbb{F}_2^n$. For instance, when we take $k=1$, we obtain the original covering problem, and from the work of Jamison [90] we know $f(n,1,d) = n+2^d-d-1$. At another extreme, if we take $d=n$, then our affine subspaces are simply individual points, each of which must be covered k times, and hence $f(n,k,n) = k\,(2^n-1)$. We study this function for intermediate values of the parameters, determining it precisely when either k is large with respect to n and d, or n is large with respect to k and d, and deriving asymptotic results in every other case.

Theorem 1.3.2. *Let $k \geq 1$ and $n \geq d \geq 1$.*

- *(a) If $n \leq \lfloor \log_2 k \rfloor + d + 1$, then $f(n,k,d) = 2^d k - \left\lfloor \frac{k}{2^{n-d}} \right\rfloor$.*
- *(b) If $k \geq 2$ and $n \geq \lfloor \log_2 k \rfloor + d + 1$, then $n + 2^d k - d - \log_2(2k) \leq f(n,k,d) \leq n + 2^d k - d - 2$.*
- *(c) If $k \geq 2$ and $n > 2^{2^d k - d - k + 1}$, then $f(n,k,d) = n + 2^d k - d - 2$.*

There are a few remarks worth making at this stage. First, observe that, just as in the Clifton–Huang setting, the extremal function $f(n,k,1)$ exhibits different behavior when n is fixed and k is large as compared to when k is fixed and n is large. Second, and perhaps most significantly, Theorem 1.3.2 demonstrates the gap between the hyperplane covering problem and the polynomial degree problem: our result shows that, for any $k \geq 4$ and sufficiently large n, we have $f(n,k,1) = n + 2k - 3$, whereas the answer to the corresponding polynomial problem is at most $n+2k-4$, as explained after Theorem 1.3.1. Our ideas allow us to establish an even stronger separation in the case $k = 4$: while the polynomial P_4 constructed by Sauermann and Wigderson, which has zeroes of multiplicity at least four at all nonzero points of $\mathbb{F}_2^n$ while not vanishing at the origin, has degree only $n + 4$, we shall show in Corollary 8.2.4 that any hyperplane system with the corresponding covering properties must have size at least $n + \log_2\left(\frac{2}{3}n\right)$. Third, we see that in the intermediate range, when both n and k grow moderately, the bounds in (b) determine $f(n,k,d)$ up to an additive error of $\log_2(2k)$, which is a lower-order term. Thus, $f(n,k,d)$ grows asymptotically like $n + 2^d k$. Note that the range considered in (b) is the complement of the range in (a), so these two parts together yield asymptotic results for the full range of values of n and k. Last of all, if one substitutes $k = 2^{n-d-1} - 1$, the lower bound from (b) is larger than the value in (a). This shows that $k \geq 2^{n-d-1}$ is indeed the correct range for which the result in (a) is valid. In contrast, we believe the bound on n in (c) is far from optimal, and discuss this in greater depth in Section 8.3.

References: The results discussed in this section and proved in Chapter 8 were obtained jointly with Anurag Bishnoi, Shagnik Das, and Tamás Mészáros; the corresponding parts of the thesis are adapted, with small modifications, from [18], the arXiv version of [19]. The general introduction in Section 1.3 is adapted from the introduction of [19] (`https://doi.org/10.1017/S0963548323000123`).

1.3.4.2 Hyperplane coverings over the reals

In Chapter 9, we extend the work of Clifton and Huang [47] and Sauermann and Wigderson [133] in a different direction by considering larger grids over the reals. In particular, we initiate the systematic study of coverings of two-dimensional grids over $\mathbb{R}$, which, as we will soon see, offers a variety of interesting problems and results. Some of our results also extend to higher dimensions, but, as this is not our primary focus, we will not discuss higher-dimensional generalizations in detail. Moreover, our focus will be on strict coverings, that is, we will not allow the origin to be covered at all. Once again, our proofs will not rely on algebraic techniques; one of the main tools that will come into play is linear programming and duality.

Before we move on, we introduce some additional notation. Given finite sets $S_1, \ldots, S_n \subseteq \mathbb{R}$ with $0 \in \bigcap_{i=1}^{n} S_i$, we denote the grid $S_1 \times \cdots \times S_n$ by $\Gamma = \Gamma(S_1, \ldots, S_n)$. Note that unless explicitly specified, we do not require that $\vec{0}$ be a corner of the grid. For a grid $\Gamma \subseteq \mathbb{R}^n$ and an integer $k \geq 1$, we define extremal function $h(\Gamma, k)$ to be the minimum size of a *strict* k-cover of Γ. To simplify our notation, when the sets $S_1, \ldots, S_n$ are clear from the context, we will sometimes suppress the tuple $(S_1, \ldots, S_n)$. We will call a nonzero point of a two-dimensional grid $\Gamma(S_1, S_2)$ an *interior point* if it has no zero coordinates; otherwise, we will call it an *axis point*.

When $k = 1$, we always have $h(\Gamma(S_1, \ldots, S_n), 1) = \sum_{i=1}^{n}(|S_i| - 1)$ by the result of Alon and Füredi [3]. Recall that Ball and Serra proved a general lower bound for the covering problem of arbitrary finite grids when $k \geq 2$. As remarked previously, the tightness of this bound is not well understood. In the special case of the binary cube $\{0, 1\}^n$, Clifton and Huang [47] showed that the Ball-Serra bound is not tight for all $k \geq 3$. Our first result shows that the Ball-Serra bound is tight for "long" two-dimensional grids. Note that in the special case of two-dimensional grids $S_1 \times S_2$ with $|S_1| \geq |S_2|$ the Ball-Serra bound simplifies to

$$h(\Gamma(S_1, S_2), k) \geq k(|S_1| - 1) + (|S_2| - 1). \tag{1.3.2}$$

Theorem 1.3.3. *Let $n, m \geq 1$ and $k \geq 2$ be integers such that $n \geq (k-1)(m-1)+1$. Then for any $S_1, S_2 \subseteq \mathbb{R}$ satisfying $|S_1| = n$, $|S_2| = m$, and $0 \in S_1 \cap S_2$, we have $h(\Gamma(S_1, S_2), k) = k(n-1) + (m-1)$, that is, the Ball-Serra bound is tight.*

We complement this result by showing that for certain types of grids the Ball-Serra bound is not tight when $n < (k-1)(m-1)+1$. In fact, Theorem 1.3.4 below shows more: it determines the precise value of $h(\Gamma, k)$ even when the Ball-Serra is not tight for some choices of Γ and k.

Theorem 1.3.4. *Let $n, m \geq 1$ be integers satisfying $n \geq m$ and $S_1, S_2 \subseteq \mathbb{R}$ be sets with $|S_1| = n$, $|S_2| = m$, and $0 \in S_1 \cap S_2$. Assume further that every line that is not parallel to the x- or y-axis contains at most two points of $\Gamma(S_1, S_2)$. Then, for all $k \geq 2$, we have*

$$(n-1)k + \frac{k}{n+m-2}(m-1)^2 \leq h(\Gamma(S_1, S_2), k)$$

and

$$h(\Gamma(S_1, S_2), k) \leq \left\lceil \frac{(n-1)k}{n+m-2} \right\rceil (n-1) + \left\lceil \frac{(m-1)k}{n+m-2} \right\rceil (m-1) + \left\lceil \frac{(n-1)k}{n+m-2} \right\rceil (m-1).$$

Moreover, if $a = \frac{n-1}{\gcd(n-1, m-1)}$ and $b = \frac{m-1}{\gcd(n-1, m-1)}$ and k is divisible by $a + b$, then

$$h(\Gamma(S_1, S_2), k) = (n-1)k + \frac{k}{n+m-2}(m-1)^2.$$

Observe that $\frac{k}{n+m-2}(m-1)^2$ is strictly larger than $m-1$ precisely when $n \leq (k-1)(m-1)$. Thus, for a grid Γ satisfying the conditions of Theorem 1.3.4 the Ball-Serra lower bound is tight if and only if $n \geq (k-1)(m-1)+1$. Second, observe that the seemingly very specific condition required for Theorem 1.3.4 to apply is actually satisfied for *most* grids: indeed, if for example we sample the $n-1$ nonzero points of S_1 and the $m-1$ nonzero points of S_2 uniformly at random from $[-1,1]$, the probability that any line that is not parallel to an axis contains at least three points of Γ tends to 0. Finally, there are examples of grids $n \times m$ grids for which the Ball-Serra bound is tight already when $n = (k-1)(m-1)$, so the bound given in Theorem 1.3.3 is not always best possible; we will see such an example in Chapter 9.

In the remainder of this section, we focus on square grids in which $\vec{0}$ is a corner. More precisely, we consider grids of the form $S_1 \times S_2$, where $|S_1| = n = |S_2|$ and $\min S_1 = 0 = \min S_2$. Our first result provides general lower and upper bounds for square grids of the given form.

Proposition 1.3.5. *Let $n, k \geq 2$ be integers and $S_1, S_2 \subseteq \mathbb{R}$ be sets satisfying $|S_1| = n = |S_2|$, $\min S_1 = 0 = \min S_2$. Then:*

(a) $h(\Gamma(S_1,S_2),k) \leq \left\lceil \frac{3}{2}k \right\rceil (n-1)$.
(b) $h(\Gamma(S_1,S_2),k) \geq (4 - 2\sqrt{2} + o(1))k(n-1)$ *as* $n \to \infty$.

Note that $4 - 2\sqrt{2} \approx 1.1716$, so part (b) shows that the Ball-Serra bound (1.3.1), which in this case is $(k+1)(n-1)$, is not tight when n and k are large.

We next examine several special types of grids. We call a grid Γ Δ-*bounded* if any line incident to $(x_i, 0)$ and $(0, y_j)$ passes through at most Δ other points of Γ. Note that any $n \times n$ grid is $(n-2)$-bounded. Our next result shows that, if Δ is not too large, we can prove a considerably better lower bound on $h(\Gamma, k)$.

Proposition 1.3.6. *Let $n, k \geq 2$ be integers and $S_1, S_2 \subseteq \mathbb{R}$ be sets satisfying $|S_1| = n = |S_2|$, $\min S_1 = 0 = \min S_2$. If $\Gamma(S_1, S_2)$ is Δ-bounded, then*

$$h(\Gamma(S_1,S_2),k) \geq \left[2 - \frac{n-1}{2(n-1)-\Delta}\right] k(n-1).$$

Note that actually *most* grids are 0-bounded: indeed, if we sample the nonzero points of S_1 and S_2 uniformly and independently at random from $[0,1]$, we can show that with high probability the resulting grid is 0-bounded. Then Proposition 1.3.6 yields a lower bound of $\frac{3}{2}k(n-1)$ for most $n \times n$ grids as $n \to \infty$, which asymptotically matches the upper bound from Proposition 1.3.5(a) when also $k \to \infty$. In fact, whenever $\Delta = o(n)$, we obtain a lower bound of the form $\left(\frac{3}{2} - o(1)\right) k(n-1)$ as $n \to \infty$. One example of such a grid is the exponential grid, given by $S_1 = S_2 = \{0, 1, 2, 4, 8, \ldots, 2^{n-2}\}$, which is 1-bounded. Indeed, suppose we have a line of negative slope passing through $(0, 2^j)$. If this line contains an interior point, let $(2^i, 2^{j'})$ be the interior point with the smallest value of i. We then must have $j' \leq j-1$, and so $2^{j'} \leq \frac{1}{2} \cdot 2^j$. Thus, by the time we reach the next x-coordinate of our grid, $2^{i+1} = 2 \cdot 2^i$, the line will already have met the x-axis, and thus cannot contain another interior point.

Perhaps the most natural grid to consider is the *standard* $n \times n$ *grid*, given by $S_1 = S_2 = \{0, 1, 2, \ldots, n-1\}$. It is not difficult to see that in this case $\Gamma(S_1, S_2)$ is Δ-bounded with $\Delta = n-2$, and thus Proposition 1.3.6 does not yield a useful lower bound. It turns out that using a different idea we can still improve on the general bounds from Proposition 1.3.5 asymptotically.

Theorem 1.3.7. *Let $n, k \geq 2$ be integers and $S_1 = S_2 = \{0, 1, 2, \ldots, n-1\}$. Then, as $n, k \to \infty$, we have*

$$(2 - e^{-1/2} + o(1))k(n-1) \leq h(\Gamma(S_1, S_2), k) \leq (\sqrt{2} + o(1))k(n-1). \tag{1.3.3}$$

Note that $2 - e^{-1/2} \approx 1.3935$, while $\sqrt{2} \approx 1.4142$, so there is still a gap between the best lower and upper bounds we obtain for standard grids. As we will see in the proof of the upper bound in Theorem 1.3.7, our construction uses only three types of lines: horizontal, vertical, and lines of slope -1. In fact, we used a computer (more specifically, SageMath [146] and Gurobi [85]) to find the optimal solutions to our problem for some small values of n and k, and we noticed that we were always able to find an optimal solution using only these three types of lines. That led to a natural question: what is the minimum size of a strict k-cover of $\{0, 1, 2, \ldots, n-1\}^2$ using only lines of slope $0, \infty$, and -1? In contrast to the more general problem, for this variant we provide a matching lower bound.

Theorem 1.3.8. *As $n, k \to \infty$, the minimum number of lines of slope $0, \infty$, or -1 needed to cover every nonzero point of Γ_n at least k times, while leaving the origin uncovered, is $\left(\sqrt{2} + o(1)\right)k(n-1)$.*

The upper bounds in Proposition 1.3.5 and Theorem 1.3.7 are shown using explicit constructions. We use linear programming and duality to derive a general recipe for showing lower bounds; we then use these ideas to prove the lower bounds in Propositions 1.3.5 and 1.3.6 and Theorems 1.3.7 and 1.3.8.

References: The results discussed in this section and proved in Chapter 9 were obtained jointly with Anurag Bishnoi, Shagnik Das, and Yvonne den Bakker; following the original submission of this thesis, the corresponding results (with some improvements) appeared in [17]. The initial results leading to this project appeared in the bachelor's thesis of Yvonne den Bakker under the supervision of Anurag Bishnoi [55].

Part I

Minimal Ramsey graphs

Chapter 2

Preliminaries and tools

This chapter collects a number of well-known results about gadget graphs used in Ramsey theory and contains some theory and results appearing in [16, 27–29].

2.1 Safeness

In the constructions that we will present in the following chapters, we will use certain gadget graphs that will allow us to force colors on certain edges. The first gadgets of this type, called signal senders, were introduced by Burr, Erdős, and Lovász [40]. Before we introduce the different gadgets we will use, we define a useful notion in this section.

In this chapter we consider several different kinds of gadget graphs associated with a q-tuple of graphs $(H_1, \ldots, H_q)$. Informally speaking, each of these gadgets G comes with a distinguished subgraph F and has a property of the following form: in every q-coloring of G, either there exists a copy of H_i all of whose edges have color i inside G, or the coloring induced on F is "nice"; moreover, G has a coloring that is $(H_1, \ldots, H_q)$-free. The utility of these gadgets comes from the fact that they allow us to force specific "nice" colorings on particular sets of edges, which simplifies our task of constructing Ramsey graphs. In our constructions, we will usually build a base graph Γ and add gadget graphs by identifying their distinguished subgraphs with subgraphs of Γ so that, in any $(H_1, \ldots, H_q)$-free coloring of the resulting graph, the edges of Γ have a particular type of coloring. While proving that a given graph is q-Ramsey for $(H_1, \ldots, H_q)$ typically means that we have to consider *all* possible q-colorings of the host graph, using our gadget graphs allows us to restrict our attention to a "nice" subset of colorings of a "nice" subgraph. We will see many examples of such constructions in this thesis.

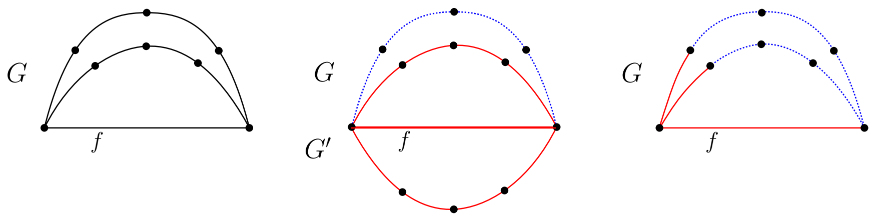

FIGURE 2.1: A graph G (left) with two colorings (in the middle and right figure solid lines represent red edges, while dotted lines represent blue edges). The coloring in the middle is not C_8-safe at $\{f\}$: if G' is a copy of C_5 sharing the edge f with G and G' is monochromatic in red, then neither G nor G' contains a monochromatic copy of C_8, but $G \cup G'$ does. The coloring on the right is C_8-safe.

In order for the gadgets to be useful, we need to be able to control the new copies of $H_1, \ldots, H_q$ that might be created in the identification process. In particular, since we usually use the color-forcing gadgets as black boxes, we would like to be able to obtain an $(H_1, \ldots, H_q)$-free coloring of the entire graph by simply giving each of the building blocks an $(H_1, \ldots, H_q)$-free coloring. This motivates the definition of a safe coloring, given by Siggers in [140]. The definition is illustrated in Figure 2.1.

Definition 2.1.1 (Safe coloring)**.** Let $q \geq 2$ and $(H_1, \ldots, H_q)$ be a q-tuple of graphs. Further, let G be a graph, F be a subgraph of G, and φ be an $(H_1, \ldots, H_q)$-free q-coloring of G. We say that φ is $(H_1, \ldots, H_q)$-*safe at* F if, for any graph G' with $V(G) \cap V(G') = V(F)$ and $E(G) \cap E(G') = E(F)$, a q-coloring ψ of $G \cup G'$ with $\psi_{|G} = \varphi$ is $(H_1, \ldots, H_q)$-free if and only if $\psi_{|G'}$ is $(H_1, \ldots, H_q)$-free.

For simplicity, in some cases we will specify just the set of edges of F, in which case our convention will be that the vertex set of F is the set of vertices incident to the specified edges. When the subgraph F or the tuple $(H_1, \ldots, H_q)$ is clear from the context, as it will be for instance in the context of a specific type of gadget graph, we will sometimes suppress them from the notation.

2.2 Signal senders

We begin with signal senders. These gadgets were first defined by Burr, Erdős, and Lovász [40] for pairs of cliques. We state the more general definition for q-tuples of graphs.

Definition 2.2.1 (Signal sender)**.** Given an integer $q \geq 2$, a q-tuple of graphs $(H_1, \ldots, H_q)$, and two edges e and f, a *negative (resp., positive) signal sender* $S^-((H_1, \ldots, H_q), q, e, f)$ (resp., $S^+((H_1, \ldots, H_q), q, e, f)$) for $(H_1, \ldots, H_q)$ is a graph S that contains e and f and satisfies:

(S1) $S \not\to_q (H_1, \ldots, H_q)$.

(S2) In any $(H_1, \ldots, H_q)$-free q-coloring of S, the edges e and f have different colors (resp., the same color).

(S3) For any colors $c_1, c_2 \in [q]$ with $c_1 \neq c_2$ (resp., $c_1 = c_2$), there exists an $(H_1, \ldots, H_q)$-free coloring φ of S with $\varphi(e) = c_1$ and $\varphi(f) = c_2$.

The edges e and f are called the *signal edges* of S. An *interior* vertex of S is a vertex that is not incident to either of the signal edges. The *interior* of S is the set of all interior vertices.

We call a signal sender *safe* if the $(H_1, \ldots, H_q)$-free colorings guaranteed by (S1) and (S3) can be chosen to be $(H_1, \ldots, H_q)$-safe at $\{e, f\}$.

Burr, Erdős, and Lovász [40] and Burr, Faudree, and Schelp [41] showed that positive and negative signal senders exist for pairs of complete graphs. Subsequently, it was proven that they exist for other graphs as well as for more colors; in particular, Rödl and Siggers [128] and Siggers [141] established their existence for any number of colors when $H_i \cong H$ for all $i \in [q]$ and H is either 3-connected or a cycle. Moreover, these signal senders can be chosen to satisfy additional properties, guaranteeing their safeness. We state the precise results below. Throughout the thesis, we will say that we *connect* or *join* two edges h_1 and h_2 of a graph G by a signal sender S on a disjoint set of vertices to mean that we identify the signal edges of S with the edges h_1 and h_2 (in an arbitrary fashion).

Lemma 2.2.2 ([128, Lemma 2.2])**.** *For any integer $q \geq 2$, any graph H that is either 3-connected or isomorphic to K_3, and any integer $d \geq 1$, there exist positive and negative signal senders for H in which the signal edges are at distance at least d.*

The next corollary is a simple consequence of Lemma 2.2.2 and shows that, if the signal edges of a signal sender as given by Lemma 2.2.2 are far apart, then attaching the signal sender to an arbitrary graph does not create any new copies of the target graph H. We provide the proof for completeness.

Corollary 2.2.3. *Let $q \geq 2$, let H be 3-connected or isomorphic to K_3, let $d > v(H)$, and let $S = S^-(H, q, e, f)$ or $S = S^+(H, q, e, f)$ be a signal sender for H in which the distance between e and f is at least d. Let G be any graph with $V(G) \cap V(S) = V(e) \cup V(f)$ and $E(G) \cap E(S) = \{e, f\}$, and write $G' = G \cup S$. Then any copy H_0 of H in G' is fully contained either in G or in S. In particular, any H-free coloring of S is H-safe at $\{e, f\}$ and S is safe.*

Proof. We prove the first part. The second follows easily from Definition 2.1.1.

Suppose for a contradiction that there exists a copy H_0 of H in G' that is fully contained neither in G nor in S. First note that H_0 must contain at least one vertex from a signal edge of S. Now observe that H_0 must also contain an interior vertex of S. Indeed, if H_0 is not fully contained in G, then it must contain at least one edge from S that is not one of the signal edges. By our assumption on the distance between the signal edges of S, at least one vertex of this edge must be an interior vertex z of S. Without loss of generality, we may assume that z is adjacent to a vertex $x \in V(H_0)$ that is contained in one of the signal edges of S, say e.

Now, we also know that H_0 must contain an edge of G. This edge must contain a vertex from $V(G) \setminus V(S)$ or connect two vertices from different signal edges of S, in which case it contains a vertex of f; let $y \in V(H_0) \cap (V(G) \setminus e)$. Now, since H is 3-connected, there are three internally vertex-disjoint paths from z to y in H_0. For each of these paths, consider the first vertex from G that appears on the path as we traverse it from z to y. These three vertices must all be distinct or equal to y (if $y \in f$), and since $V(G) \cap V(S) = V(e) \cup V(f)$, they must be vertices of the signal edges. In particular, one of those vertices must be in f. But then there is a path from x to a vertex of f in H_0 that is fully contained in S, since z is connected to $x \in e$ by an edge; this path has length at most $v(H_0) = v(H)$, contradicting the fact that the signal edges of S are at distance at least $v(H) + 1$. □

We now turn our attention to cycles.

Lemma 2.2.4 ([141, Lemma 2.2]). *For any integers $q \geq 2$ and $\ell \geq 4$, there exist positive and negative signal senders for the cycle C_ℓ that have girth ℓ and distance at least $\ell + 1$ between their signal edges.*

Again, we show that attaching a signal sender as given by Lemma 2.2.4 to an arbitrary graph cannot create new copies of the cycle C_ℓ. This is again easy to deduce from the above lemma (a proof is given for instance in [141, Proposition 2.4]); we sketch the proof for completeness.

Corollary 2.2.5. *Let $q \geq 2$ and $\ell \geq 4$ be integers, and let $S = S^-(C_\ell, q, e, f)$ or $S = S^+(C_\ell, q, e, f)$ be a signal sender for C_ℓ that has girth ℓ and distance at least $\ell + 1$ between e and f. Let G be any graph with $V(G) \cap V(S) = V(e) \cup V(f)$ and $E(G) \cap E(S) = \{e, f\}$, and write $G' = G \cup S$. Then any copy C of C_ℓ in G' is fully contained either in G or in S. In particular, any C_ℓ-free coloring of S is C-safe at $\{e, f\}$ and S is safe.*

Proof. The argument is similar to that in the proof of Corollary 2.2.3, so we will be brief. Suppose for a contradiction that there exists a copy C of C_ℓ in G' that is fully contained neither in G nor in S. As before, we may assume that C contains vertices $x, z \in S$ such that $x \in e$ and z is an interior vertex of S that is adjacent to x, and a vertex $y \in V(G) \setminus e$. There are two internally vertex-disjoint paths from z to y in C. For each of these paths, consider the first vertex from G that appears on the path as we traverse it from z to y. These two vertices must be distinct, unless $y \in f$ and they are both equal to y, and since $V(G) \cap V(S) = V(e) \cup V(f)$, they must be vertices of the signal edges. If one of those vertices is in f, then as before there is a path from x to a vertex of f in C that is fully contained in S and has length at most ℓ, which is a contradiction. Otherwise, the two paths between x and y in H_0 contain the two vertices of e; the segments of these paths connecting x to the vertices of e, which do not contain y, form a cycle of length less than ℓ in S together with the edge e, which contradicts the fact that S has girth ℓ. □

In a later paper, Siggers [140] showed the existence of safe signal senders for some pairs of the form (C_ℓ, H). We will return to these pairs in the next section.

In some cases, in particular in Chapter 5, we will need a gadget that gives us finer control over what colors certain edges can take. This motivates us to define a more general gadget, which we call a set-signal sender, or a set-sender for short.

Definition 2.2.6 (Set-sender)**.** Let $q \geq 2$ be an integer, $X \subseteq [q]$ be any subset of colors, $(H_1, \ldots, H_q)$ be a q-tuple of graphs, and e and f be two edges. A *negative* (resp., *positive*) *X-sender* for $(H_1, \ldots, H_q)$ is a graph S containing the edges e and f and satisfying the following properties:

(S1) $S \not\to_q (H_1, \ldots, H_q)$.

(S2) For any $(H_1, \ldots, H_q)$-free q-coloring φ of S, there exist colors $c_1, c_2 \in X$ with $c_1 \neq c_2$ (resp., $c_1 = c_2$) such that $\varphi(e) = c_1$ and $\varphi(f) = c_2$.

(S3) For any colors $c_1, c_2 \in X$ with $c_1 \neq c_2$ (resp., $c_1 = c_2$), there exists an $(H_1, \ldots, H_q)$-free coloring φ of S with $\varphi(e) = c_1$ and $\varphi(f) = c_2$.

The signal edges, interior vertices, and interior of a set-sender and the notion of a safe set-sender are defined in the same way as in Definition 2.2.1. Observe that a signal sender as given by Definition 2.2.1 is an X-sender with $X = [q]$.

In Chapter 5, we investigate the parameter s_q in the case of multiple cliques and multiple cycles. Recall the definition of the q-tuple $\mathcal{T} = \mathcal{T}(q_1, q_2, \ell, t)$ consisting of cycles and cliques from (1.1.3). Our main technical results for q-tuples of this form concern the existence of the necessary gadget graphs. The next theorem establishes the existence of set-senders for the color palettes $\{1, \ldots, q_1\}$ and $\{q_1 + 1, \ldots q_1 + q_2\}$; the proof is deferred to Section 2.6.

Theorem 2.2.7. *Let $\ell \geq 4$, $t \geq 3$, and $q_1, q_2 \geq 1$ be integers. If $q_1 > 1$, then there exist safe positive and negative $\{1, \ldots, q_1\}$-senders for $\mathcal{T}(q_1, q_2, \ell, t)$. If $q_2 > 1$, then there exist safe positive and negative $\{q_1 + 1, \ldots q_1 + q_2\}$-senders for $\mathcal{T}(q_1, q_2, \ell, t)$.*

2.3 Determiners

In the asymmetric setting we can build another useful gadget graph, known as a determiner. Determiners were introduced by Burr, Faudree, and Schelp in [41] for two colors. In Chapter 5, we work both in the two-color setting and in the multicolor setting, so we need to generalize their definition to q colors and define what we call set-determiners.

Definition 2.3.1 (Set-determiner)**.** Let $q \geq 2$ be an integer, $X \subseteq [q]$ be any subset of colors, $(H_1, \ldots, H_q)$ be a q-tuple of graphs, and d be an edge. An *X-determiner* for $(H_1, \ldots, H_q)$ is a graph D containing d and satisfying the following properties:

(D1) $D \not\to_q (H_1, \ldots, H_q)$.
(D2) For any $(H_1, \ldots, H_q)$-free q-coloring φ of D, we have $\varphi(d) \in X$.
(D3) For any color $c \in X$, there exists an $(H_1, \ldots, H_q)$-free coloring φ of D such that $\varphi(d) = c$.

The edge d is referred to as the *signal edge* of D.

We call a set-determiner *safe* if the $(H_1, \ldots, H_q)$-free colorings guaranteed by (D1) and (D3) can be chosen to be $(H_1, \ldots, H_q)$-safe at $\{d\}$.

We will say that we *attach* a set-determiner D to an edge e of a graph G on a disjoint set of vertices to mean that we identify the edge e with the signal edge of D. It is not difficult to see that a $\{c\}$-determiner can only exist for a q-tuple $(H_1, \ldots, H_q)$ if $H_c \not\cong H_i$ for all $i \in [q] \setminus \{c\}$. Similarly, an X-determiner for a q-tuple $(H_1, \ldots, H_q)$ can only exist if, for all $c_1 \in X$ and $c_2 \in [q] \setminus X$, we have $H_{c_1} \not\cong H_{c_2}$. When $X = \{c\}$, we sometimes simplify the notation and write c-determiner instead of $\{c\}$-determiner. In particular, when $q = 2$, we sometimes call the corresponding gadgets red-determiner and blue-determiner.

The determiners originally defined by Burr, Faudree, and Schelp in [41] for $q = 2$ are simply c-determiners for $c \in \{1, 2\}$. Such determiners are known to exist for all pairs (H_1, H_2) such that $H_1 \not\cong H_2$ and H_1 and H_2 are 3-connected or isomorphic to K_3 (see [41, 42]). These determiners can be shown to be safe using an argument similar to Corollary 2.2.3, giving the following result.

Lemma 2.3.2 ([41, 42])**.** *Let H_1 and H_2 be non-isomorphic graphs, each of which is 3-connected or isomorphic to K_3. Then there exist safe red- and blue-determiners for (H_1, H_2).*

The only other result in this direction that we are aware of is due to Siggers [140], who used the ideas of Bollobás, Donadelli, Kohayakawa, and Schelp [25] to prove the existence of safe determiners for many pairs of the form (C_ℓ, H), where H is a 2-connected graph satisfying some additional properties. The special cases that will be needed in Chapter 5 are given in the following lemma. While part (ii) of Lemma 2.3.3 also follows from our more general Theorem 2.3.4, we briefly sketch Siggers' proof for both cases below, combining a few arguments from his paper.

Lemma 2.3.3 ([140])**.**

(i) Let $k, \ell \geq 4$ be integers with $\ell > k$. Then there exist safe red-determiners and safe blue-determiners for (C_ℓ, C_k).
(ii) Let $\ell \geq 4$ and $t \geq 3$. Then there exist safe red-determiners and safe blue-determiners for (C_ℓ, K_t).

Proof. We know that C_k and K_t are 2-connected and, since $\ell > k$, these graphs contain no cycle of length at least $\ell + 1$ as an induced subgraph. Therefore, by [140, Corollary 3.12], there exist safe blue-determiners for (C_ℓ, C_k) and (C_ℓ, K_t).

Let C be a copy of C_k, and let e be any edge of C. Attach a copy of the safe blue-determiner for (C_ℓ, C_k) from the previous paragraph to each edge of C except e, and let D be the resulting graph. Clearly, in any (C_ℓ, C_k)-free coloring of D, the edge e is red: indeed, a (C_ℓ, C_k)-free coloring of D induces a (C_ℓ, C_k)-free coloring on each copy of the blue-determiner; by property (D2), every edge of C except for e must be blue, which in turn forces e to be red. Furthermore, giving each copy of the blue-determiner a (C_ℓ, C_k)-safe coloring, as guaranteed by property (D3) and the safeness of the blue-determiner, and coloring e red results in a (C_ℓ, C_k)-free coloring of D. The safeness of the blue-determiner further ensures that this coloring is (C_k, C_ℓ)-safe at $\{e\}$. Therefore D is a safe red-determiner for (C_ℓ, C_k) with signal edge e. A similar argument yields a safe red-determiner for (C_ℓ, K_t). □

As explained earlier, in Chapter 5 we investigate the parameter s_q in the case of multiple cliques and multiple cycles. Along with Theorem 2.2.7, the second main technical result that allows us to study $s_q(\mathcal{T})$ (again, recall the definition of the tuple $\mathcal{T}$ from (1.1.3)) concerns the existence of safe determiners for this tuple.

Theorem 2.3.4. *Let $\ell \geq 4$, $t \geq 3$, and $q_1, q_2 \geq 1$ be integers. Then there exist safe $\{1, \ldots, q_1\}$-determiners and safe $\{q_1 + 1, \ldots, q_1 + q_2\}$-determiners for $\mathcal{T}(q_1, q_2, \ell, t)$.*

Most of Section 2.6 is devoted to the proof of Theorem 2.3.4. We will then deduce Theorem 2.2.7 from Theorem 2.3.4 in combination with Corollary 2.2.5 and Corollary 2.2.3.

2.4 Indicators

We now turn our attention to a different type of gadget, known as an indicator, which will play a role in Chapter 6. Indicators were originally introduced by Burr, Faudree, and Schelp in [41]; we are going to use a slightly modified definition, as given in [29]. Indicators will be used only in the symmetric setting, so for the sake of simplicity we restrict ourselves to the symmetric case in this section. In fact, in most of this section, we will focus on the case $q = 2$.

Definition 2.4.1 (Indicator)**.** Given an integer $q \geq 2$, graphs H and F such that $H \not\subseteq F$, and an edge e disjoint from F, an *indicator* $I = I(H, F, q, e)$ for H is a graph I containing F as an induced subgraph and e that satisfies:

(I1) There exists an H-free q-coloring of I in which F is monochromatic.

(I2) For every H-free q-coloring of I in which F is monochromatic, the edge e has the same color as F.

(I3) For any non-constant coloring $\varphi_F : E(F) \to [q]$ and every $k \in [q]$, there exists an H-free coloring $\widetilde{\varphi} : E(I) \to [q]$ such that $\widetilde{\varphi}_{|F} = \varphi_F$ and $\widetilde{\varphi}(e) = k$.

We call F the *indicator subgraph* and e the *indicator edge* of I. An *interior vertex* of I is a vertex that is not contained in F or e, and the *interior* of I is the set of all interior vertices.

We call an indicator I *safe* if the colorings guaranteed by (I1) and (I3) can be taken to be H-safe at $F \cup \{e\}$.

We will say that we *join* or *connect* a subgraph J and an edge h in a graph G by an indicator I to mean that we identify the indicator subgraph with J and the indicator edge with h.

In the original definition of an indicator, property (I3) is replaced by the following property: for any edge $f \in F$ and for any pair of distinct colors $j, k \in [q]$, there is an H-free coloring of $I - f$ in which $F - f$ is monochromatic in color j and e receives color k. The existence of indicators of this kind when $q = 2$ and H is a complete graph was shown by Burr, Faudree, and Schelp [41]. This result was generalized to multiple colors and all graphs H that are either 3-connected or isomorphic to K_3 by Clemens, Liebenau, and Reding [46]. The slightly modified definition given above was introduced by the author, Clemens, and Gupta [29]; in the arXiv version of [29], we provide an argument showing that such indicators exist when H is 3-connected or isomorphic to a cycle for any $q \geq 2$ (see [28]). Moreover, these indicators were also shown to be safe[1]. For our application in Chapter 6 for $q = 2$, we will need to be able to keep track of the number of vertices that an indicator contains. For indicators to be useful in our construction, we will need the number of vertices of I to grow linearly with $e(F)$. This was not done explicitly in the existence proofs mentioned above, so we will sketch the argument for $q = 2$ in the next lemma.

Lemma 2.4.2. *Let H be a connected graph for which safe positive and negative signal senders exist, and let F be a graph with $e(F) \geq 2$ not containing H. Then there exists a safe indicator $I = I(H, F, e, 2)$ such that $v(I) \leq c_H(e(F) - 1)$, where $c_H > 0$ is a constant depending only on H.*

Proof. Let S^+ and S^- be a safe positive and a safe negative signal sender for H, respectively. Set $c_H = e(H)s_H$, where $s_H = \max\{v(S^+), v(S^-)\}$.

We recall the construction given in [28, 41]. We proceed by induction on the number of edges of F.

[1]In [28, 29], we establish a slightly different property, known as robustness. The fact that the gadgets are safe will be apparent from the construction.

For the base case, assume that $e(F) = 2$. Let $E(F) = \{f_1, f_2\}$. We construct the graph I in the following way. Let H_0 be a copy of H that is vertex-disjoint from F, let $e', e'' \in E(H_0)$ be arbitrary, and let e be an edge disjoint from H_0 and F. For each edge $h \in E(H_0) \setminus \{e', e''\}$, we connect f_1 and h by a copy of S^-. In addition, we connect f_2 and e'' by a copy of S^- and e' and e by a copy of S^+.

It is not difficult to check that I is an indicator for H with indicator subgraph F and indicator edge e (see [28] for the details). Now we briefly argue why I is safe. Consider the coloring φ in which f_1, f_2, and e are colored red and each signal sender is given an H-safe coloring (this coloring is easily seen to satisfy (I1)). Let G be a graph with $V(G) \cap V(I) = V(F) \cup V(e)$ and $E(G) \cap E(I) = F \cup \{e\}$ and suppose G is given an H-free coloring that agrees with φ on $F \cup \{e\}$. Suppose there is a monochromatic copy H' of H in $G \cup I$. Since each signal sender was given an H-safe coloring, we may assume H' is fully contained in the graph $H_0 \cup G$. There are no edges between $V(H_0)$ and $V(G)$, and since H_0 is not monochromatic, it follows that H' is contained in G. The argument showing that the colorings given by (I3) can be taken to be H-safe is similar.

We count the number of vertices in I. Note that each vertex of I is contained in at least one of the copies of S^- or in the copy of S^+. There are $e(H) - 1$ copies of S^-, and hence there are at most $e(H)s_H = c_H(e(F) - 1)$ vertices in I.

For the induction step, assume that, if $2 \leq e(F) \leq k$, then there exists an appropriate safe indicator with at most $c_H(e(F) - 1)$ vertices.

Assume $e(F) = k + 1$. Define the graph I in the following way. Let $f \in E(F)$ be fixed, and let e and e' be edges disjoint from each other and from F. Let $I' = I(H, F - f, 2, e')$ be a safe indicator with $v(I') \leq c_H(e(F - f) - 1) = c_H(e(F) - 2)$ and $I'' = I(H, \{e, f'\}, 2, e)$ be a safe indicator with $v(I'') \leq c_H(e(\{e', f\}) - 1) = c_H$, both of which exist by induction. We now connect $F - f$ and e' by a copy of I' and $\{e, f'\}$ and e by a copy of I''.

As shown in [28], the graph I is an indicator for H with indicator subgraph F and indicator edge e. We can find the required H-safe colorings of I by combining H-safe colorings of I' and I'', which exist by induction. Finally, every vertex of I is in at least one of I' and I'' and hence $v(I) \leq v(I') + v(I'') \leq c_H(e(F) - 2) + c_H = c_H(e(F) - 1)$. □

2.5 Pattern gadgets

In this section, we introduce the last type of gadget graph that we will need, called a pattern gadget, which generalizes all previous gadget graphs and allows for more flexibility. This gadget again comes with a subgraph F on which fixed color patterns are forced in any H-free q-coloring. Pattern gadgets will be useful in the proof of the main result of Chapter 3. Again we restrict ourselves to the symmetric setting. We introduce some additional notation that we will need here and in Chapter 3. Given a graph G, a *q-color pattern for G* is a collection of edge-disjoint graphs $G_1, \dots, G_q$ on $V(G)$ such that $E(G) = E(G_1) \cup \dots \cup E(G_q)$. Given another graph H, a color pattern $G_1, \dots, G_q$ is *H-free* if every graph G_i is H-free. If G' is isomorphic to G and $G'_1, \dots, G'_q$ is a color pattern for G', then we say that $G_1, \dots, G_q$ is *isomorphic* to $G'_1, \dots, G'_q$ if there exists a permutation π of $[q]$ such that $G_i \cong G'_{\pi(i)}$ for every $i \in [q]$.

Definition 2.5.1 (Pattern gadget)**.** Let $q \geq 2$ be an integer and H and F be graphs such that $F \not\to_q H$. Also let $\mathscr{F}$ be a family of H-free q-color patterns for F that is closed under permutations of the graphs within a single pattern. A *pattern gadget* $P = P(H, F, \mathscr{F}, q)$ is a graph containing F as an induced subgraph and satisfying:

(P1) $P \not\to_q H$.

(P2) If $\varphi : E(P) \to [q]$ is an H-free coloring of P, then the color pattern $\varphi^{-1}_{|F}(1), \dots, \varphi^{-1}_{|F}(q)$ is in $\mathscr{F}$.

(P3) For every pattern $F_1, \dots, F_q$ in $\mathscr{F}$, there exists an H-free coloring $\varphi : E(P) \to [q]$ such that $\varphi^{-1}_{|F}(i) \cong F_i$ for all $i \in [q]$.

A pattern gadget P is *safe* if the colorings guaranteed by (P1) and (P3) can be taken to be H-safe at F.

In other words, a pattern gadget allows us to pick any family of H-free color patterns for F so that each of these patterns, and no other, can be extended to an H-free coloring of the whole graph. Pattern gadgets were implicitly introduced by Siggers [141], who showed the existence of a specific kind of pattern gadgets when H is a cycle. Siggers' proof idea applies more generally as well; an alternative, more constructive, proof in the case where H is 3-connected or isomorphic to a cycle was given in [29, Theorem 2.4]. In Chapter 3, we will require the existence of pattern gadgets when H is a complete graph.

Lemma 2.5.2. *Let $q \geq 2$ and $t \geq 3$ be integers. Let F be a graph with $F \not\to_q K_t$ and $\mathscr{F}$ be a family of K_t-free q-color patterns for F. Then there exists a safe pattern gadget $P = P(K_t, F, \mathscr{F}, q)$.*

Once again, instead of safeness, in [29] we show a slightly stronger property, which we refer to as robustness, which guarantees that if G is a graph with $V(G) \cap V(P) \subseteq V(F)$, then any copy of H is fully contained in G or fully contained in P.

2.6 Existence of set-determiners and set-senders

This section is taken from [16] with minor modifications.

In this section we construct set-determiners and set-senders for tuples of the form $(C_\ell, \dots, C_\ell, K_t, \dots, K_t)$, that is, we prove Theorems 2.2.7 and 2.3.4. Our set-senders will be constructed in several stages. Before diving into the proofs, we give a brief overview.

Throughout the rest of the section, assume that $\ell \geq 4$, $t \geq 3$, and $q, q_1, q_2 \geq 1$ are fixed integers such that $q_1 + q_2 = q$ and recall that $\mathcal{T} = \mathcal{T}(q_1, q_2, \ell, t)$ denotes the q-tuple of cycles and cliques as defined in (1.1.3). For convenience, we will sometimes write $\mathcal{S}(C_\ell)$ for the color palette $\{1, \dots, q_1\}$ and refer to it as the *cycle-colors*; similarly, $\mathcal{S}(K_t)$ will denote the color palette $\{q_1 + 1, \dots, q\}$, referred to as the *clique-colors*.

First, we construct a graph Γ that is q-Ramsey for the tuple $\mathcal{T}$ and has certain special properties; for this, we generalize the ideas of Bollobás, Donadelli, Kohayakawa, and Schelp [25] used to construct 2-Ramsey graphs for certain pairs of graphs, including (C_ℓ, K_t), to multiple colors. This graph Γ is built by sampling a random hypergraph, applying alterations to remove all short cycles from it, and then replacing every hyperedge by a large (depending only on t) clique. In order to prove the claimed properties of Γ, we use a number of results, all of which are fairly standard by now. Second, we modify Γ slightly and construct set-determiners for each of the color palettes $\mathcal{S}(C_\ell)$ and $\mathcal{S}(K_t)$. This is a generalization of a construction given by Siggers in [140], valid for certain pairs of the form (C_ℓ, H). Finally, since we need finer control over the color patterns we force on given set of edges when $q_1 > 1$ or $q_2 > 1$, we build set-senders from our set-determiners. This final step is the main novelty in this section.

2.6.1 Preliminary results

We begin by collecting the different results that will be needed for the construction and proof of the claimed properties of the graph Γ.

Hypergraphs with few short cycles. First, we need to construct a uniform hypergraph with no short cycles that is nevertheless not too sparse. This is done using a standard construction due to Erdős and Hajnal [63], starting from a random hypergraph. A *cycle of length s* in a hypergraph $\mathcal{H}$ is a sequence $e_1, v_1, e_2, v_2 \dots, e_s, v_s$ of distinct hyperedges and vertices of $\mathcal{H}$ such that $v_i \in e_i \cap e_{i+1}$ for all $1 \leq i < s$ and $v_s \in e_s \cap e_1$. Note in particular that two edges intersecting in more than one vertex form a cycle of length two in $\mathcal{H}$. The *girth* of a hypergraph $\mathcal{H}$ is the length of the shortest cycle in $\mathcal{H}$ (if no cycle exists, then by convention we say that the girth of $\mathcal{H}$ is infinity). We state the necessary results about random hypergraphs without proof, as these are by now standard applications of the probabilistic method.

Lemma 2.6.1. *Let $\ell, h \geq 2$ be fixed integers, and $p_h = An^{-(h-1)+1/(\ell-1)}$, where A is a constant. For an integer $n \geq 1$, let $\mathcal{H}_{n,p_h}$ be a random h-uniform hypergraph on $[n]$ in which each h-subset* *of $[n]$ is added as an edge with probability p_h, independently of all other h-subsets. Then, as $n \to \infty$, the following hold with high probability:*

(i) $e(\mathcal{H}_{n,p_h}) = (1+o(1))\binom{n}{h}p_h$.

(ii) The number of cycles in $\mathcal{H}_{n,p_h}$ of length less than ℓ is $o(e(\mathcal{H}_{n,p_h}))$.

Part (i) follows from an application of the Chernoff bound (see Lemma 4.3.1), while part (ii) is shown using a first-moment argument.

Quantitative version of Ramsey's theorem. The following lemma is a simple consequence of Ramsey's theorem and is obtained by a straightforward averaging argument. Informally, it says that, for any r-tuple of graphs $(H_1, \ldots, H_r)$, if we r-color a sufficiently large complete graph, then we can find not just one monochromatic H_i in the correct color, but many of them. The proof is a simple generalization of the one given, for example, in [119, Theorem 2].

Lemma 2.6.2 (Quantitative version of Ramsey's theorem)**.** *Let $r \geq 1$ and $H_1, \ldots, H_r$ be graphs. Then there exist a real number $c = c(H_1, \ldots, H_r) > 0$ and an integer $k_0 = k_0(H_1, \ldots, H_r) \geq 1$ such that, if $k \geq k_0$ and the edges of K_k are colored with r colors, then there exists an $i \in [r]$ such that there are at least $ck^{v(H_i)}$ monochromatic copies of H_i in color i.*

Colorful sparse regularity lemma. One of the tools required to show that Γ is q-Ramsey for the tuple $\mathcal{T}$ is a version of Szemerédi's celebrated regularity lemma [144]. More specifically, we will need the colorful sparse version of the lemma, as given for example in [108] (see also [87, Lemma 3.1]). Before giving the precise statement in Lemma 2.6.4 below, we again need several definitions.

Definition 2.6.3. Let G be a graph on n vertices, $D \geq 1$, and $0 < \varepsilon, p, \eta \leq 1$. Also let U and W be disjoint subsets of $V(G)$. The *p-density* of the pair (U, W) is defined to be

$$d_{G,p}(U, W) = \frac{e_G(U, W)}{p|U||W|}.$$

The pair (U, W) is said to be *(ε, p)-regular* if, for all $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$, we have

$$|d_{G,p}(U', W') - d_{G,p}(U, W)| \leq \varepsilon.$$

If (U, W) is (ε, p)-regular with $p = \frac{e_G(U,W)}{|U||W|}$, then we say that (U, W) is *(ε)-regular* for short.

A partition $(V_1, \ldots, V_k)$ of $V(G)$ is an *equipartition* if $|V_i| \in \left\{\left\lfloor \frac{v(G)}{k} \right\rfloor, \left\lceil \frac{v(G)}{k} \right\rceil\right\}$ for all $i \in [k]$. An equipartition is said to be an *(ε, p)-regular partition* if all but at most $\varepsilon\binom{k}{2}$ pairs (V_i, V_j) are (ε, p)-regular.

A graph G is said to be (η, D, p)-*upper uniform* if, for all disjoint $U, W \subseteq V(G)$ with $|U|, |W| \geq \eta v(G)$, we have $d_{G,p}(U, W) \leq D$.

We are now ready to state the version of the regularity lemma that we are going to use.

Lemma 2.6.4 (Colorful sparse regularity lemma)**.** *Let $\varepsilon > 0$ and $D > 1$ be fixed reals and $k_0 \geq 1$ and $r \geq 1$ be integers. Then there exist constants $\eta = \eta(\varepsilon, k_0, D, r)$ and $K_0 = K_0(\varepsilon, k_0, D, r)$ for which the following holds: If $0 < p \leq 1$ and $G_1, \ldots, G_r$ are (η, D, p)-upper uniform graphs on vertex set $[n]$, then there is an equipartition $(V_1, \ldots, V_k)$ of $[n]$ for some $k_0 \leq k \leq K_0$ such that all but at most $\varepsilon\binom{k}{2}$ of the pairs (V_i, V_j) are (ε, p)-regular in G_s for all $s \in [r]$.*

We will also need the following additional technical lemma, which can be found for example in [78, Lemma 4.3].

Lemma 2.6.5. *Given $0 < \varepsilon < 1/6$, there exists a constant $\beta > 0$ such that the following holds. For any bipartite graph F with vertex classes V_1 and V_2 such that the pair (V_1, V_2) is (ε)-regular in F, and for all M satisfying $\beta v(F) \leq M \leq e(F)$, there exists a subgraph $F' \subseteq F$ with $V(F') = V(F)$ and $e(F') = M$ and such that (V_1, V_2) is (2ε)-regular in F'.*

Enumeration lemma for C_ℓ-free graphs. Let $m, M \geq 1$ and $\ell \geq 4$ be integers, and let $\varepsilon > 0$ be a real number. Let $V_1, \ldots, V_\ell$ be disjoint sets, each of size m. Let $\mathcal{G}(\ell, m, (V_i)_{i=1}^{\ell}, M, \varepsilon)$ denote the collection of graphs G such that

- $V(G) = V_1 \cup \cdots \cup V_\ell$, where $|V_i| = m$ for each $i \in [\ell]$;
- each V_i is an independent set in G;
- the pair (V_i, V_{i+1}) is $\left(\varepsilon, \frac{M}{m^2}\right)$-regular in G with $e_G(V_i, V_{i+1}) = M$ for all $i \in [\ell]$[2];
- there are no edges between any other pair (V_i, V_j).

In other words, the graphs in $\mathcal{G}(\ell, m, (V_i)_{i=1}^{\ell}, M, \varepsilon)$ are blow-ups of the cycle C_ℓ in which each vertex v_i of C_ℓ is blown-up to an independent set V_i of size m and such that each edge $v_i v_{i+1}$ of C_ℓ corresponds to an $\left(\varepsilon, \frac{M}{m^2}\right)$-regular pair (V_i, V_{i+1}). Let $\mathcal{F}(\ell, m, (V_i)_{i=1}^{\ell}, M, \varepsilon)$ denote the set of graphs in $\mathcal{G}(\ell, m, (V_i)_{i=1}^{\ell}, M, \varepsilon)$ that do *not* contain C_ℓ as a subgraph.

The following enumeration lemma was shown by Gerke, Kohayakawa, Rödl, and Steger [77, Theorem 5.2]; it is a special case of a well-known conjecture by Kohayakawa, Łuczak, and Rödl [102] (the so-called KŁR conjecture), which was famously resolved in the general case using the container method [12, 136].

Lemma 2.6.6 (Counting Lemma)**.** *For any real number $\alpha > 0$ and integer $\ell \geq 4$, there are constants $\varepsilon_0 = \varepsilon_0(\ell, \alpha) > 0, C_0 = C_0(\ell, \alpha) > 0$, and $m_0 = m_0(\ell, \alpha) \geq 1$ such that, for all*

[2]For convenience, we define $V_{\ell+1} = V_1$.

$m \ge m_0$, $0 < \varepsilon \le \varepsilon_0$, and $M \ge C_0 m^{1+1/(\ell-1)}$, we have

$$|\mathcal{F}(\ell, m, (V_i)_{i=1}^{\ell}, M, \varepsilon)| \le \alpha^M \binom{m^2}{M}^{\ell}.$$

2.6.2 Construction of a special graph Γ

For the rest of the section, assume that n is a sufficiently large integer with respect to ℓ, t, q, q_1, and q_2; in all asymptotic estimates in this section, we assume that n tends to infinity. We begin by fixing some constants. Let $h = r_{q_2}(K_t)$; it is not difficult to check that K_h is minimal q_2-Ramsey for K_t. Let

$$k_0 = k_0(\underbrace{C_\ell, \dots, C_\ell}_{q_1 \text{ times}}, K_h, K_2), \quad c = c(\underbrace{C_\ell, \dots, C_\ell}_{q_1 \text{ times}}, K_h, K_2)$$

be the constants given by Lemma 2.6.2. We next set

$$\rho = \frac{c}{2q_1}, \quad \alpha = \frac{\rho^\ell}{e^{\ell+1}}, \quad D = 3h^2.$$

Let $\varepsilon_0 = \varepsilon_0(\ell, \alpha), m_0 = m_0(\ell, \alpha)$, and $C_0 = C_0(\ell, \alpha)$ be the constants given by Lemma 2.6.6, and set

$$\varepsilon = \min\{\rho\varepsilon_0/2, \rho/10\}, \quad C = \max\{C_0, 1\}.$$

Further, let

$$\eta = \eta(\varepsilon, k_0, D, q_1), \quad K_0 = K_0(\varepsilon, k_0, D, q_1), \quad \beta = \beta(\varepsilon/\rho)$$

be the constants from Lemmas 2.6.4 and 2.6.5. Finally, define

$$A = \max\{(h+1)e^{-h}, \rho^{-1} K_0^{1-1/(\ell-1)} C\}, \quad p_h = A n^{-(h-1)+1/(\ell-1)}, \quad p_e = A n^{-1+1/(\ell-1)}.$$

Let $\mathcal{H}$ be a hypergraph on $[n]$ sampled from $\mathcal{H}_{n,p_h}$ as in Lemma 2.6.1. Let $\mathcal{G}$ be the hypergraph obtained from $\mathcal{H}$ after the removal of one hyperedge from each cycle of length less than ℓ. Then $\mathcal{G}$ contains no cycles of length less than ℓ; by Lemma 2.6.1(i) and (ii), we also know that $e(\mathcal{G}) = (1 + o(1))\binom{n}{h} p_h$.

Let Γ be the graph on $[n]$ obtained by embedding a copy of K_h into every hyperedge of $\mathcal{G}$, i.e., Γ is the graph on $[n]$ in which two vertices are adjacent if and only if they are contained in a common hyperedge of the hypergraph $\mathcal{G}$. The main difference between this construction and the one given in [25] is that, in order to deal with multiple colors, instead of placing just a copy of our target graph K_t in each hyperedge of $\mathcal{G}$, we place a Ramsey graph for it. For a given graph F and a subgraph $\Gamma' \subseteq \Gamma$, we call a copy F' of F in Γ' a *hyperedge copy* if the vertex set of F' is contained within a single hyperedge of $\mathcal{G}$. All remaining copies of F in Γ' are referred to as *non-hyperedge copies*. In addition, we call a subgraph $\Gamma' \subseteq \Gamma$ *transversal* if there exists a bijection $f : E(\Gamma') \to E(\mathcal{G})$ such that $e \subseteq f(e)$ for all $e \in E(\Gamma')$; that is, Γ' is transversal if it contains exactly one edge from each hyperedge copy of K_h in Γ.

Before showing that with high probability $\Gamma \to_q \mathcal{T}(q_1, q_2, \ell, t)$ in Theorem 2.1, we discuss some properties of the graph Γ in Lemma 2.6.7 below. The proofs of parts (a), (b), (d) are essentially the same as those given in [25]. The proof of (c) is similar to the proof of Proposition 9 in [25] and is by now also standard in light of the recently resolved KŁR conjecture; as we believe that our version (using more modern results) can be generalized more easily to other tuples of graphs, we include the details.

Lemma 2.6.7. *The graph Γ satisfies each of the following properties with high probability:*

(a) If F is a 2-connected graph with no induced cycles of length ℓ or more, then every copy of F in Γ is a hyperedge copy; in particular, every copy of K_h, K_t, and $C_{\ell'}$ for any $\ell' < \ell$ in Γ is a hyperedge copy.

(b) Γ is (η, D, p_e)-upper uniform.

(c) Let m be an integer satisfying $\frac{n}{K_0} \leq m \leq \frac{n}{k_0}$, let $(V_1, \ldots, V_\ell)$ be any ℓ-tuple of disjoint subsets of $V(\Gamma)$ such that $|V_i| = m$ for all $i \in [\ell]$, and let $\Gamma' \subseteq \Gamma$ be transversal. If the pairs (V_i, V_{i+1}) are (ε, p_e)-regular in Γ' with p_e-density at least ρ for all $i \in [\ell]$, then $\Gamma'[V_1 \cup \cdots \cup V_\ell]$ contains a copy of C_ℓ.

(d) Let m be an integer satisfying $\frac{n}{\log n} \leq m \leq \frac{n}{h}$ and $(W_1, \ldots, W_h)$ be an h-tuple of pairwise disjoint subsets of $V(\Gamma)$ with $|W_i| = m$ for all $i \in [h]$. Then there are at least $\frac{1}{4} m^h p_h$ distinct copies of K_h contained in the multipartite subgraph of Γ spanned by $W_1 \cup \cdots \cup W_h$.

Proof of Lemma 2.6.7(c). Let m satisfy $\frac{n}{K_0} \leq m \leq \frac{n}{k_0}$; we can write $p_e = Bm^{-1+1/(\ell-1)}$, where $B = A\left(\frac{n}{m}\right)^{-1+1/(\ell-1)}$. Notice that B satisfies $AK_0^{-1+1/(\ell-1)} \leq B \leq Ak_0^{-1+1/(\ell-1)}$.

Let $(V_1, \ldots, V_\ell)$ and Γ' be as given. Suppose that the pairs (V_i, V_{i+1}) for $i \in [\ell]$ are (ε, p_e)-regular with p_e-density at least ρ in Γ'. Then we have $e_{\Gamma'}(V_i, V_{i+1}) \geq \rho p_e m^2$ for all $i \in [\ell]$. Let M be an integer satisfying

$$\rho p_e m^2 \leq M \leq \min_{i \in [\ell]} e_{\Gamma'}(V_i, V_{i+1}).$$

Notice that this integer M satisfies

$$\begin{aligned} M &\geq \rho p_e m^2 = \rho B m^{1+1/(\ell-1)} \geq \rho A K_0^{-1+1/(\ell-1)} m^{1+1/(\ell-1)} \\ &\geq C m^{1+1/(\ell-1)} \geq 2\beta m = \beta |V_i \cup V_{i+1}|, \end{aligned}$$

since $A \geq K_0^{1-1/(\ell-1)} C/\rho$ and n, and hence m, is taken to be sufficiently large.

Consider the pair (V_1, V_2) and let $d = \frac{e_{\Gamma'}(V_1,V_2)}{m^2}$; then we have $d \geq \rho p_e$, and thus $p_e \leq \frac{d}{\rho}$. By definition, it then follows that the pair (V_1, V_2) is $\left(\frac{\varepsilon}{\rho}, d\right)$-regular, or simply $\left(\frac{\varepsilon}{\rho}\right)$-regular. By Lemma 2.6.5, there is a subset $E_{1,2} \subseteq E_{\Gamma'}(V_1, V_2)$ such that $|E_{1,2}| = M$ and the pair (V_1, V_2) is $\left(\frac{2\varepsilon}{\rho}\right)$-regular in $(V_1 \cup V_2, E_{1,2})$. Repeating this argument for all pairs of the form (V_i, V_{i+1}), we find that $\Gamma'[V_1 \cup \cdots \cup V_\ell]$ contains at least one graph in $\mathcal{G}\left(\ell, m, (V_i)_{i=1}^{\ell}, M, \frac{2\varepsilon}{\rho}\right)$.

Our goal now is to show that, with high probability, there is no collection of subsets $(V_i)_{i=1}^{\ell}$ and subgraph $\Gamma' \subseteq \Gamma$ as given in the statement such that $\Gamma'[V_1 \cup \cdots \cup V_\ell]$ contains a subgraph belonging to $\mathcal{F}\left(\ell, m, (V_i)_{i=1}^{\ell}, M, \frac{2\varepsilon}{\rho}\right)$. Again, let the ℓ-tuple $(V_1, \ldots, V_\ell)$ be fixed. If $F \in \mathcal{F}\left(\ell, m, (V_i)_{i=1}^{\ell}, M, \frac{2\varepsilon}{\rho}\right)$ has edges $e_1, \ldots, e_{M\ell}$ and there exists a transversal Γ' such that $F \subseteq \Gamma'[V_1 \cup \cdots \cup V_\ell]$, there must exist distinct hyperedges $\mathcal{E}_1, \ldots, \mathcal{E}_{M\ell} \in E(\mathcal{H}_{n,p_h})$ such that $e_i \subseteq \mathcal{E}_i$ for all $i \in [M\ell]$. Therefore

$$\begin{aligned}
\mathbb{P}[\exists \text{transversal } \Gamma' : \ F \subseteq \Gamma'[V_1 \cup \cdots \cup V_\ell]] &\leq \left(\binom{n-2}{h-2} p_h\right)^{M\ell} \\
&\leq \left((n-2)^{h-2} A n^{-(h-1)+1/(\ell-1)}\right)^{M\ell} \\
&\leq \left(A n^{-1+1/(\ell-1)}\right)^{M\ell} = p_e^{M\ell}. \qquad (2.6.1)
\end{aligned}$$

Note that, when n is sufficiently large, we have $m \geq m_0$. By the choice of $\varepsilon \leq \rho\varepsilon_0/2$ and the fact that $M \geq C m^{1+1/(\ell-1)} \geq C_0 m^{1+1/(\ell-1)}$, applying Lemma 2.6.6 and the union bound, we obtain

$$\begin{aligned}
&\mathbb{P}\left[\exists \text{ transversal } \Gamma', F \in \mathcal{F}\left(\ell, m, (V_i)_{i=1}^{\ell}, M, \frac{2\varepsilon}{\rho}\right) : F \subseteq \Gamma'[V_1 \cup \cdots \cup V_\ell]\right] \\
&\leq \alpha^M \binom{m^2}{M}^{\ell} p_e^{M\ell} \leq \alpha^M \left(\frac{m^2 e}{M}\right)^{M\ell} p_e^{M\ell} \leq \alpha^M \left(\frac{e}{\rho}\right)^{M\ell} = e^{-M},
\end{aligned}$$

where the last inequality follows from the fact that $M \geq \rho p_e m^2$ and the final step follows by the choice of α.

This implies that, for any fixed integers m and M and any collection of disjoint subsets $V_1, \ldots, V_\ell$ of $[n]$, each of size m, the probability that there exists a transversal Γ' such that $\Gamma'[V_1 \cup \cdots \cup V_\ell]$ contains some graph in $\mathcal{F}\left(\ell, m, (V_i)_{i=1}^{\ell}, M, \frac{2\varepsilon}{\rho}\right)$ is at most e^{-M}.

Now, for any choice of $\frac{n}{K_0} \leq m \leq \frac{n}{k_0}$ and $C m^{1+1/(\ell-1)} \leq M \leq m^2 \leq n^2$, there are at most $n^{m\ell}$ choices for the sets $V_1, \ldots, V_\ell$. Summing over the possible choices for the sets $V_1, \ldots, V_\ell$ and the possible choices for m and M, we find that the probability that (c) fails is bounded from above by the probability that there exist m, M, $(V_i)_{i=1}^{\ell}$ and Γ' such that $\Gamma'[V_1 \cup \cdots \cup V_\ell]$ contains a member of $\mathcal{F}\left(\ell, m, (V_i)_{i=1}^{\ell}, M, \frac{2\varepsilon}{\rho}\right)$, which is at most

$$\begin{aligned}
\sum_m \sum_M n^{m\ell} e^{-M} &\le \sum_m \sum_M \exp(-Cm^{1+1/(\ell-1)} + m\ell \log n) \\
&\le \sum_m \sum_M \exp\left(-C\left(\frac{n}{K_0}\right)^{1+1/(\ell-1)} + \frac{n}{k_0}\ell \log n\right) \\
&\le n^3 \exp\left(-C\left(\frac{n}{K_0}\right)^{1+1/(\ell-1)} + \frac{n}{k_0}\ell \log n\right) = o(1).
\end{aligned}$$

□

We are now ready to show the main result of this section.

Theorem 2.1. *With high probability, we have $\Gamma \to_q \mathcal{T}$.*

Proof. We condition on Γ having all of the properties given in Lemma 2.6.7. For convenience, we may assume also that $\frac{n}{k}$ is an integer for all $k_0 \le k \le K_0$. Consider an arbitrary q-coloring φ of the graph Γ. If any copy of K_h receives only colors in $\mathcal{S}(K_t)$, then we are done since $h = r_{q_2}(K_t)$. So suppose that each such copy has at least one edge whose color comes from $\mathcal{S}(C_\ell)$. Let Γ' be a graph on $V(\Gamma) = [n]$ obtained by taking exactly one edge that has a cycle-color from each hyperedge copy of K_h in Γ; since there are only hyperedge copies of K_h in Γ, we know that Γ' is a transversal subgraph. We claim that Γ' contains a copy of C_ℓ in some cycle-color.

For each $s \in \mathcal{S}(C_\ell)$, let G_s be the subgraph of Γ' on vertex set $[n]$ consisting of all edges that have color s under φ. By Lemma 2.6.7(b), we know that Γ is (η, D, p_e)-upper uniform, and hence G_s is (η, D, p_e)-upper uniform for all $s \in \mathcal{S}(C_\ell)$. So by Lemma 2.6.4, there exists an equipartition $(V_1, \ldots, V_k)$ of $[n]$ in which all but at most $\varepsilon\binom{k}{2}$ pairs (V_i, V_j) are (ε, p_e)-regular in every G_s for $s \in \mathcal{S}(C_\ell)$. Let $m = \frac{n}{k}$; by our choice of k_0, K_0, and n, we know that m is an integer and $\frac{n}{K_0} \le \frac{n}{k} = m \le \frac{n}{k_0}$.

Let K_k be the complete graph on vertex set $\{V_1, \ldots, V_k\}$. Consider the following (q_1+2)-coloring of the edges of K_k with the color palette $\{c_1, \ldots, c_{q_1+2}\}$. If the pair (V_i, V_j) is (ε, p_e)-regular in all G_s for $s \in \mathcal{S}(C_\ell)$ and has p_e-density at least ρ in some G_s, give the edge between V_i and V_j in K_k color c_s (breaking ties arbitrarily). If the pair (V_i, V_j) is (ε, p_e)-regular in G_s for all $s \in \mathcal{S}(C_\ell)$, but its p_e-density is less than ρ in every such G_s, then color the edge between V_i and V_j in K_k with color c_{q_1+1}. Finally, if (V_i, V_j) is not (ε, p_e)-regular in G_s for some $s \in \mathcal{S}(C_\ell)$, let the edge between V_i and V_j in K_k have color c_{q_1+2}.

By the fact that $k \ge k_0$ and our choice of k_0 (from Lemma 2.6.2), we know that at least one of the following must occur:

(a) For some $s \in [q_1]$, there are at least ck^ℓ copies of C_ℓ in color c_s.

(b) There are at least ck^h copies of K_h that are monochromatic in color c_{q_1+1}.

(c) There are at least ck^2 edges of color c_{q_1+2}.

If (a) occurs for some color $c_s \in [q_1]$, the fact that $ck^\ell \geq ck_0^\ell > 0$, together with property (c) in Lemma 2.6.7, implies that there is a copy of C_ℓ in Γ' in color s. It remains to show that neither of the other cases can occur.

First consider option (c). We know that there are at most $\varepsilon\binom{k}{2}$ pairs (V_i, V_j) that are not (ε, p_e)-regular in G_s for some $s \in \mathcal{S}(C_\ell)$, and we have

$$\varepsilon\binom{k}{2} \leq \frac{1}{10}\rho\binom{k}{2} \leq \frac{1}{10}c\binom{k}{2} < ck^2,$$

where the first two inequalities follow by the definitions of ε and ρ. Hence, option (c) is indeed impossible.

We now prove that option (b) cannot occur. Suppose it does. We estimate the number of edges of Γ' corresponding to pairs of color c_{q_1+1} in two different ways. First note that if there is an edge of color c_{q_1+1} between vertices V_i and V_j, then the (ε, p_e)-regular pair (V_i, V_j) has p_e-density at most ρ in G_s for each $s \in \mathcal{S}(C_\ell)$. Hence, in total, the pair (V_i, V_j) has p_e-density at most $q_1\rho$ in Γ'. Hence, the number of edges in Γ' between pairs (V_i, V_j) corresponding to color c_{q_1+1} is at most

$$\binom{k}{2}q_1\rho p_e m^2 = \binom{k}{2}q_1\rho p_e\left(\frac{n}{k}\right)^2 < \frac{1}{2}q_1\rho A n^{1+1/(\ell-1)} = \frac{c}{4}An^{1+1/(\ell-1)}. \tag{2.6.2}$$

Now, since option (b) occurs, we have at least ck^h copies of K_h that are monochromatic in color c_{q_1+1} in K_k. Denote these by $K_h^1, K_h^2, \ldots, K_h^x$, where $x = \lceil ck^h \rceil$. The vertex set $V(K_h^i)$ of each such copy gives an h-partite subgraph $J_i \subseteq \Gamma$ induced by the sets V_j corresponding to the vertices of K_h^i. As each partite set of J_i has size $m \geq \frac{n}{K_0} \geq \frac{n}{\log n}$, Lemma 2.6.7(d) guarantees that J_i contains a family $\mathcal{H}_i$ of at least $\frac{1}{4}m^h p_h$ distinct hyperedge copies of K_h, for every $i \in [x]$. As each hyperedge copy in $\mathcal{H}_i$ intersects each partite set of K_h^i, it is immediate that $\mathcal{H}_i \cap \mathcal{H}_j \neq \emptyset$ for $i \neq j$. Hence, there exist $\left|\bigcup_{i\in[x]}\mathcal{H}_i\right| \geq \frac{1}{4}ck^h m^h p_h$ copies of K_h in Γ. Since every copy of K_h in Γ is a hyperedge copy and no two hyperedge copies share an edge, we find that Γ' has at least

$$\frac{1}{4}ck^h m^h p_h \geq ck^h\frac{1}{4}\left(\frac{n}{k}\right)^h An^{-h+1+1/(\ell-1)} = \frac{c}{4}An^{1+1/(\ell-1)} \tag{2.6.3}$$

edges corresponding to pairs (V_i, V_j) in color c_{q_1+1}, contradicting (2.6.2). $\square$

2.6.3 Construction of set-determiners

This section uses ideas from [140] to prove Theorem 2.3.4. Recall that $\mathcal{S}(C_\ell)$ and $\mathcal{S}(K_t)$ denote the cycle-colors $\{1, \ldots, q_1\}$ and clique-colors $\{q_1+1, \ldots, q\}$, respectively. By construction and by Lemma 2.6.7, we know that Γ satisfies the following properties:

(i) Every copy of K_t in Γ is a hyperedge copy.
(ii) Every copy of $C_{\ell'}$ for $\ell' < \ell$ is a hyperedge copy.
(iii) Each edge of Γ belongs to a unique copy of K_h.

Now, let $G \subseteq \Gamma$ be a minimal q-Ramsey graph for the q-tuple $\mathcal{T}(q_1, q_2, \ell, t)$; it is not difficult to see that G satisfies properties (i) and (ii) given above. In fact, we have a good understanding of what G needs to look like, as given in the following lemma. Naturally, the lemma also establishes that G satisfies property (iii) above.

Lemma 2.6.8. *The graph G is the union of hyperedge copies of K_h, that is, every edge of G belongs to a hyperedge copy of K_h in G.*

Proof. Suppose there is an edge e that does not belong to a copy of K_h in G. We know that e *does* belong to a copy of K_h in $\Gamma \supseteq G$; let H denote this copy of K_h in Γ and let F denote the set of edges on $V(H)$ that are in Γ but not in G. Notice that $\emptyset \subsetneq F \subsetneq E(H)$ by our assumption.

By the minimality of G, we know that $G - H$ has a $\mathcal{T}$-free q-coloring φ. Additionally, since K_h is minimal q_2-Ramsey for K_t, the graph $H - F$ has a K_t-free q_2-coloring $\varphi' : E(H-F) \to \mathcal{S}(K_t)$. We now define a q-coloring $\widetilde{\varphi}$ of G by setting $\widetilde{\varphi} = \varphi \cup \varphi'$.

We claim that $\widetilde{\varphi}$ is a $\mathcal{T}$-free q-coloring of G. Indeed, since φ is a $\mathcal{T}$-free coloring of $G - H$, there are no monochromatic cycles in any cycle-color, and since in the coloring of $H - F$ we add no further edges in these colors, we know that there are no monochromatic copies of C_ℓ in any cycle-color in all of G. Furthermore, since there are no non-hyperedge copies of K_t in G and neither φ nor φ' contains a monochromatic copy of K_t in any color in $\mathcal{S}(K_t)$, we know that there are also no monochromatic copies of K_t in any clique-color in all of G. Hence $\widetilde{\varphi}$ is a $\mathcal{T}$-free q-coloring of G, contradicting the fact that $G \to_q \mathcal{T}$. □

Now, let e be a fixed edge of G and let H be the copy of K_h in G containing e. Let D be the graph obtained from G by removing all edges of H except for e, that is, $D = G - (H - e)$. We now claim that D is a $\mathcal{S}(K_t)$-determiner for the tuple $\mathcal{T}$. This construction generalizes the one presented by Siggers [140].

Lemma 2.6.9. *The graph D is a safe $\mathcal{S}(K_t)$-determiner for the tuple $\mathcal{T}$ with signal edge e.*

Proof. We first show property (D2). For a contradiction, suppose ψ is a $\mathcal{T}$-free coloring of D in which $\psi(e) \in \mathcal{S}(C_\ell)$. Then, by an argument similar to the one used in Lemma 2.6.8, putting together this $\mathcal{T}$-free coloring of D and a K_t-free q_2-coloring of $H - e$ (with colors in $\mathcal{S}(K_t)$), we obtain a $\mathcal{T}$-free coloring of G, which is a contradiction to the fact that $G \to_q \mathcal{T}$.

To see properties (D1) and (D3), note that D is a proper subgraph of G, so D has a $\mathcal{T}$-free q-coloring φ. Further, by permuting the clique-colors in φ appropriately, we can obtain a $\mathcal{T}$-free coloring of D in which the edge e has any color in $\mathcal{S}(K_t)$.

It remains to show that φ is safe at $\{e\}$. Let F be any graph such that $V(D) \cap V(F) \subseteq e$. Let φ' be a $\mathcal{T}$-free q-coloring of F that agrees with φ on the edge e. We claim that the coloring $\widetilde{\varphi}$, given by $\widetilde{\varphi} = \varphi \cup \varphi'$, is a $\mathcal{T}$-free q-coloring of $D \cup F$. We know that the restrictions of $\widetilde{\varphi}$ to both D and F are $\mathcal{T}$-free; it remains to show that there are no monochromatic cliques or cycles in the appropriate colors intersecting both $V(D) - e$ and $V(F) - e$.

First, it is not difficult to see that there can be no such copy of K_t. For $t = 3$, this is clear. If $t \geq 4$ and there is a t-clique K intersecting both $D - e$ and $F - e$, then we can disconnect K by removing the vertices of e, which is impossible. Suppose there is such a copy C of C_ℓ. Note first that C must contain both vertices of e because C_ℓ is 2-connected. Now, let v be a vertex of C contained in $V(D) - e$, and let w be a vertex of C contained in $V(F) - e$. There are no non-hyperedge cycles of length less than ℓ in D, so every cycle containing e in D has length at least ℓ. Hence, the vertices v and w cannot be contained in a cycle of length ℓ with both endpoints of e, and therefore C cannot exist. Thus the coloring $\widetilde{\varphi}$ is $\mathcal{T}$-free, implying that φ is safe. This completes the verification of the safeness property. □

Now we construct a safe $\mathcal{S}(C_\ell)$-determiner D' by taking a copy H of K_h, fixing one edge f, and attaching copies of the $\mathcal{S}(K_t)$-determiner D constructed above to all remaining edges of H. This again generalizes a construction of Siggers [140].

Lemma 2.6.10. *The graph D' is a safe $\mathcal{S}(C_\ell)$-determiner for the tuple $\mathcal{T}$ with signal edge f.*

Proof. We again begin with property (D2). Take an arbitrary $\mathcal{T}$-free coloring of D'. This coloring induces a $\mathcal{T}$-free coloring on each copy of D, so, by property (D2) of D, all edges of $H - f$ have colors in $\mathcal{S}(K_t)$. If f has one of these colors too, then H is fully colored with colors in $\mathcal{S}(K_t)$. Since H is q_2-Ramsey for K_t, there exists a monochromatic copy of K_t in H, contradicting the fact that the coloring φ is $\mathcal{T}$-free. So the color of f must be in the set $\mathcal{S}(C_\ell)$.

We show properties (D1) and (D3) next. By minimality, we know that $H - f$ is not q_2-Ramsey for K_t, and hence it has a K_t-free coloring ψ from the palette $\mathcal{S}(K_t)$. Let φ be a q-coloring extending ψ in which each copy of the determiner D has a $\mathcal{T}$-safe coloring and the edge f has an arbitrary color from $\mathcal{S}(C_\ell)$; this coloring $\widetilde{\varphi}$ exists by property (D3) of D. Since the coloring of each copy of D is safe and since H has a $\mathcal{T}$-free q-coloring, the coloring φ of D' is also $\mathcal{T}$-free.

Finally, to see the safeness of φ, let F be a graph such that $V(D') \cap V(F) \subseteq f$. If F is given a $\mathcal{T}$-free q-coloring φ' that agrees with φ on f, then the coloring $\widetilde{\varphi} = \varphi \cup \varphi'$ is a $\mathcal{T}$-free q-coloring of $D' \cup F$. Indeed, since each copy of D is safe and the only edge of H that has color in $\mathcal{S}(C_\ell)$ is f, we know that there can be no monochromatic copy of C_ℓ in a cycle-color inside $D' \cup F$. Similarly, since we cannot disconnect K_t by removing at most two vertices, we know that there can be no copy of K_t intersecting both $V(D') - f$ and $V(F) - f$, and hence there can be no monochromatic copy of K_t in a clique-color in $D' \cup F$. Hence, $\widetilde{\varphi}$ is a $\mathcal{T}$-free q-coloring and thus φ is a safe coloring of D'. □

2.6.4 Construction of set-senders

So far we have constructed an $\mathcal{S}(K_t)$-determiner D and an $\mathcal{S}(C_\ell)$-determiner D', generalizing ideas from [25] and [140]. We now take the constructions a step further and use our set-determiners to build set-senders for these sets of colors when $q_1 > 1$ or $q_2 > 1$, proving Theorem 2.2.7.

If $q_1 > 1$, let S be a safe negative (resp., positive) signal sender for C_ℓ with q_1 colors, as guaranteed by Corollary 2.2.5; let e and f denote its signal edges. Let R be the graph obtained from S by attaching a copy of D' to every edge of S.

Lemma 2.6.11. *If S is a negative (resp., positive) signal sender for C_ℓ with signal edges e and f as above, then R is a safe negative (resp., positive) $\mathcal{S}(C_\ell)$-sender for $\mathcal{T}$ with signal edges e and f.*

Proof. Assume S is a negative signal sender for C_ℓ with q_1 colors; the other case is similar. We first show properties (S1) and (S3). Let $c_1, c_2 \in \mathcal{S}(C_\ell)$ be distinct. We know that $S \not\to_{q_1} C_\ell$, so S has a safe C_ℓ-free coloring from the set $\mathcal{S}(C_\ell)$, and by property (S3) of S, we can ensure that e and f receive colors c_1 and c_2, respectively. Now, since the signal edge of each copy of D' has color in $\mathcal{S}(C_\ell)$, by property (D3) of D', this coloring of S can be extended to each copy of D' so that each copy of D' has a $\mathcal{T}$-safe q-coloring. The coloring of each copy of D' is safe, so the q-coloring defined on R is $\mathcal{T}$-free. To see the safeness of this coloring, notice that the coloring of each copy D' is safe at its signal edge and the coloring of S, containing only colors from $\mathcal{S}(C_\ell)$, is safe at $\{e, f\}$. □

Finally, if $q_2 > 1$, we build $\mathcal{S}(K_t)$-senders for $\mathcal{T}$. Let S' be a safe negative (resp., positive) signal sender for K_t with q_2 colors taken as $\mathcal{S}(K_t)$, as guaranteed by Corollary 2.2.3; let e and f denote its signal edges. Let R' be a graph obtained from S' by attaching a copy of D to every edge of S'. We omit the proof that R' is a set-sender for K_t, as it is essentially the same as that of Lemma 2.6.11.

Lemma 2.6.12. *If S' is a negative (resp., positive) signal sender for K_t with signal edges e and f, then R' is a safe negative (resp., positive) $\mathcal{S}(K_t)$-sender R' for $\mathcal{T}$ with signal edges e and f.*

Chapter 3

Cliques are s_q-abundant

The goal of this chapter is to prove Theorem 1.1.3.

Theorem 1.1.3. *For any integers $q \geq 2$ and $t \geq 3$, the clique K_t is s_q-abundant.*

The results proved in this chapter were obtained jointly with Dennis Clemens and Pranshu Gupta and appear in [29]. We note that in [29] Theorem 1.1.3 is derived as a consequence of a more general construction for 3-connected graphs (see Theorem 3.1 in [29]; a special case is given in Theorem 4.6.2). Here we provide a more direct proof using similar ideas (parts of the text are adapted from [29]).

We begin by introducing a few preliminaries. A *q-color pattern* on vertex set V is a collection of edge-disjoint graphs $\Gamma_1, \dots, \Gamma_q$ on the same vertex set V, and a color pattern is said to be *H-free* if every graph in it is H-free. Fox, Grinshpun, Liebenau, Person, and Szabó [72] showed a connection between $s_q(K_t)$ and a particular packing problem (as the authors of [72] note, this idea can already be found implicitly in [40]).

Definition 3.0.1 ([72])**.** Given integers $q \geq 2$ and $t \geq 2$, the *q-color t-clique packing parameter* $P_q(t)$ is the smallest integer n for which there exists a K_{t+1}-free color pattern $\Gamma_1, \dots, \Gamma_q$ on a vertex set V of size n satisfying the following property:

(P) For every partition $V = \cup_{j \in [q]} V_j$, there exists a copy H of K_t and an integer $i \in [q]$ such that $H \subseteq \Gamma_i[V_i]$.

Fox, Grinshpun, Liebenau, Person, and Szabó showed that $s_q(K_t) = P_q(t-1)$ for all $q \geq 2$ and $t \geq 3$ (see Theorem 1.5 in [72]).

Recall that, for any graph H and $q \geq 2$, we have $s_q(H) \leq r_q(H) - 1$. In our construction, we will need the slightly stronger claim that $s_q(K_t) \leq r_q(K_t) - 2$ for any $q \geq 2$ and $t \geq 3$. This is of course known to be true asymptotically, since $s_q(K_t)$ is polynomial in both q and t, whereas $r_q(K_t)$ is at least exponential in both (see [135]). However, the bounds on $s_q(K_t)$ and $r_q(K_t)$ mentioned in the introduction do not allow us to deduce the desired inequality directly; we use a different simple argument instead, generalizing the ideas of Burr, Erdős, and Lovász [40] establishing the upper bound $s_2(K_t) \le (t-1)^2$ (this idea also appears in [54] in the context of a related problem in Ramsey theory).

Lemma 3.0.2. *For all $q \ge 2$ and $t \ge 3$, we have $s_q(K_t) \le r_q(K_t) - 2$.*

Proof. Let $t \ge 3$ be fixed. For all $q \ge 2$, define $N_q = (t-1)^q$. To show the claim it suffices to prove that K_{N_q} satisfies the following properties:

(i) There is a K_t-free q-coloring φ_q of K_{N_q} that cannot be extended to a K_t-free coloring of K_{N_q+1}.
(ii) There exists a K_t-free coloring ψ_q of K_{N_q+1}.

Note that property (ii) implies that $N_q + 1 < r_q(K_t)$. We now claim that property (i) implies that $P_q(t-1) \le N_q$. The argument is similar to the proof of Lemma 2.1 in [72]. Let $G_i = \varphi_q^{-1}(i)$ for all $i \in [q]$. This color pattern is K_t-free. Now consider an arbitrary partition $V(K_{N_q}) = \bigcup_{j=1}^{q} V_j$. Let w be a new vertex connected to all vertices of K_{N_q}. Extend the edge-coloring φ_q to all of $K_{N_q} + w \cong K_{N_q+1}$ by giving color j to every edge of the form wu for $u \in V_j$. By property (i), this coloring contains a monochromatic copy of K_t in some color i and this copy must contain the vertex w. By the definition of the extended coloring, the remaining vertices, forming a copy of K_{t-1} in G_i, lie in V_i. Thus, the color pattern $G_1, \ldots, G_q$ satisfies property (P), implying the required inequality. The two inequalities implied by properties (i) and (ii), together with the fact that $s_q(K_t) = P_q(t-1)$, prove the claim.

We now show properties (i) and (ii); for that, we proceed by induction on q. First consider the base case $q = 2$. We can use the idea of Burr, Erdős, and Lovász [40]. Partition the vertices of the graph K_{N_2} into $t-1$ sets $Q_1, \ldots, Q_{t-1}$, each of size $t-1$, and let φ_2 be the coloring in which all edges lying within a single Q_i have color 1 and all edges with endpoints in different sets Q_i have color 2. It is not difficult to check that this coloring is K_t-free but there is no way to extend it to K_{N_2+1} without creating a monochromatic K_t, establishing property (i). On the other hand, we construct a K_t-free 2-coloring ψ_2 of K_{N_2+1} as follows. Let $Q_1, \ldots, Q_{t-1}$ be as before; fix an arbitrary vertex $v_i \in Q_i$ for every $i \in [t-1]$. Color all edges of K_{N_2} as before, except for the edge v_1v_2, to which we now assign color 1. Let w be a new vertex connected to all vertices of K_{N_2}. Assign color 2 to wv_i for all $i \in [t-1]$ and color 1 to all other edges incident to w. It is not difficult to check that this coloring is K_t-free.

Assume that (i) and (ii) hold for some $q \ge 2$. Consider the graph $K_{N_{q+1}}$. Partition its vertex set into $t-1$ equally-sized sets $Q_1, \ldots, Q_{t-1}$, each inducing a copy of K_{N_q}. Now, let w be a vertex connected to all vertices of $K_{N_{q+1}}$. Let φ_{q+1} be the coloring of $K_{N_{q+1}}$ in which the edges inside each Q_i are colored according to the K_t-free coloring φ_q and the edges between two different Q_i are given color $q+1$. Again, it is easily seen that this coloring is K_t-free. Consider any coloring of $K_{N_{q+1}} + w \cong K_{N_{q+1}+1}$ extending φ_{q+1}. If all edges from w to some Q_i have colors in $[q]$, then by induction the graph induced by $Q_i \cup \{w\}$ contains a monochromatic copy of K_t. So we may assume that, for all $i \in [t-1]$, there is a vertex $v_i \in Q_i$ such that the edge wv_i has color $q+1$. But then the vertices $v_1, \ldots, v_{t-1}, w$ induce a monochromatic copy of K_t in color $q+1$. Hence property (i) is satisfied. For property (ii), notice that coloring the graph induced by $Q_i \cup \{w\}$ according to ψ_q for all $i \in [t-1]$ and giving all edges with endpoints in different Q_i color $q+1$ gives the required K_t-free coloring ψ_{q+1} of $K_{N_{q+1}+1}$. □

We are now ready to prove the theorem. For the proof we will use the pattern gadgets introduced in Section 2.5.

Proof of Theorem 1.1.3. Let $k \geq 1$ be arbitrary. We will construct a graph G satisfying the following properties:

(i) $G \to_q K_t$.
(ii) G contains k vertices $v_1, \ldots, v_k$, each of degree $s_q(K_t)$.
(iii) For all $i \in [k]$, we have $G - v_i \not\to_q K_t$.

To see why this implies the desired result, consider a minimal q-Ramsey graph $G' \subseteq G$ and note that from (iii) we can conclude that G' must contain all of $v_1, \ldots, v_k$. Moreover, we have $d_{G'}(v_i) \leq d_G(v_i) = s_q(K_t)$ for all $i \in [k]$, and the definition of $s_q(K_t)$ implies that each v_i must have degree $s_q(K_t)$ in G', as required. Since k is arbitrary, it then follows that K_t is s_q-abundant.

Let $\Gamma_1, \ldots, \Gamma_q$ be a K_t-free color pattern on a set V of size $P_q(t-1)$ as in Definition 3.0.1, and let $T = \bigcup_{i=1}^q \Gamma_i$. Let J be the graph obtained from T by adding a new vertex v and connecting it to all vertices of T. By Lemma 3.0.2, we know that $v(J) = v(T) + 1 \leq r_q(K_t) - 1$, so J has a K_t-free coloring ψ, which induces another K_t-free color pattern $\Gamma'_1, \ldots, \Gamma'_q$ on V.

Before we proceed with the construction, we state and prove a key property of the color pattern $\Gamma_1, \ldots, \Gamma_q$. The proof follows easily from Definition 3.0.1 (see the proof of Theorem 2.3 in [72]); we include the short argument here for completeness.

Claim 3.0.3. *Suppose φ is a q-coloring of T such that, for all $i \in [q]$, the edges in Γ_i all have color i. Then φ cannot be extended to a K_t-free q-coloring of J.*

Proof. Let φ' be a coloring extending φ to J. Consider the edges incident to v, and define a partition $V = \bigcup_{i=1}^q V_i$, where V_i contains all vertices $w \in V$ such that $\varphi'(vw) = i$ for all $i \in [q]$. By property (P) of the color pattern $\Gamma_1, \ldots, \Gamma_q$, we know that there exists a copy H of K_{t-1} and a color $i \in [q]$ such that $V(H) \subseteq \Gamma_i[V_i]$. By the definition of the partition $\bigcup_{i=1}^q V_i$, all edges from v to $V(H)$ have color i, yielding a monochromatic copy of K_t in color i. $\square$

Let $T^1, \dots, T^k$ be k vertex-disjoint copies of T and set $F = \bigcup_{i=1}^k T^i$. Let $\mathscr{F}$ be the collection of all color patterns on F satisfying the following: there exists $j \in [k]$ such that the pattern induced on $V(T^j)$ is isomorphic to $\Gamma_1, \dots, \Gamma_q$ and the pattern induced on $V(T^\ell)$ for all $\ell \in [k] \setminus \{j\}$ is isomorphic to $\Gamma'_1, \dots, \Gamma'_q$. By our choice of the color patterns $\Gamma_1, \dots, \Gamma_q$ and $\Gamma'_1, \dots, \Gamma'_q$, every color pattern in $\mathscr{F}$ is K_t-free, so by Theorem 2.5.2, there exists a safe pattern gadget $P = P(K_t, F, \mathscr{F}, q)$.

Now, add k new vertices $v_1, \dots, v_k$ and, for all $i \in [k]$, connect v_i to all vertices of T^i so that they form a copy of J. Call the resulting graph G. We now claim that G satisfies properties (i)–(iii) above. Since $v(T) = P_q(t-1) = s_q(K_t)$, property (ii) is satisfied.

We next argue that G satisfies property (i), that is, $G \to_q K_t$. Suppose there exists a K_t-free q-coloring φ of G. This coloring induces a K_t-free q-coloring on P; property (P2) of the pattern gadget P then implies that the color pattern $\varphi^{-1}_{|F}(1), \dots, \varphi^{-1}_{|F}(q)$ is in $\mathscr{F}$. Hence, there exists a $j \in [k]$ such that the color pattern $\varphi^{-1}_{|T^j}(1), \dots, \varphi^{-1}_{|T^j}(q)$ is isomorphic to $\Gamma_1, \dots, \Gamma_q$. Without loss of generality, we may assume that $\varphi^{-1}_{|T^j}(i) \cong \Gamma_i$ for all $i \in [q]$. The desired result then follows from Claim 3.0.3, since $T^j + v_j \cong J$.

Finally, we verify property (iii). Let $i \in [k]$, and consider the following coloring of the graph $G' = G - v_i$. For each $j \in [k] \setminus \{i\}$, color the copy of J induced by $V(T^j) \cup \{v_j\}$ according to the K_t-free coloring ψ; note that this coloring induces a color pattern isomorphic to $\Gamma'_1, \dots, \Gamma'_q$ on T^j. Color T^i so that the graph formed by the edges of color ℓ is isomorphic to Γ_ℓ for all $\ell \in [q]$. Note that this partial coloring is K_t-free. By property (P3), we can extend this coloring to the pattern gadget P in such a way that P has a K_t-safe coloring. By the definition of safeness, it follows that the coloring of the entire graph $G - v_i$ is K_t-free, completing the proof. $\square$

It would be interesting to continue this line of research and understand better which graphs are s_q-abundant. As discussed in the introduction, in [29] we provide further examples of abundant graphs. A natural next step would be to study the abundance question for the class of 3-connected graphs. In [29, Theorem 3.1], we provide a sufficient condition for such a graph to be s_q-abundant, and we tend to believe that this condition should be satisfied by all 3-connected graphs for any $q \geq 2$.

As a second possible direction, it would be interesting to find further examples of graphs that are not s_q-abundant for some $q \geq 2$. So far the only such examples we are aware of are stars (see [29] for the details).

Chapter 4

Ramsey simplicity of random graphs

The goal of this chapter is to prove Theorem 1.1.9, Corollary 1.1.10, and Proposition 1.1.11. We first recall the statement of our main theorem.

Theorem 1.1.9. *Let $p = p(n) \in (0, 1)$ and $H \sim G(n, p)$. Let $u \in V(H)$ be the smallest (with respect to the natural vertex ordering) vertex of degree $\delta(H)$, and let $F = H[N(u)]$ be the subgraph of H induced by the neighborhood of u. Denote by $\lambda(F)$ the order of the largest connected component in F. Then w.h.p. the following bounds hold:*

(a) $\tilde{q}(H) = \infty$ *if* $0 < p \ll n^{-1}$.

(b) $\tilde{q}(H) = \infty$ *if* $\frac{\log n}{n} \ll p \ll n^{-\frac{2}{3}}$.

(c) $\tilde{q}(H) \geq (1 + o(1)) \max\left\{\frac{\delta(H)}{\lambda(F)^2}, \frac{\delta(H)}{80 \log n}\right\}$ *if* $n^{-\frac{2}{3}} \ll p \ll n^{-\frac{1}{2}}$.

(d) $\tilde{q}(H) \leq (1 + o(1)) \min\left\{\frac{\delta(H)}{\Delta(F)}, \frac{\delta(H)^2}{2e(F)}\right\}$ *if* $n^{-\frac{2}{3}} \ll p \ll 1$.

(e) $\tilde{q}(H) = 1$ *if* $\left(\frac{\log n}{n}\right)^{1/2} \ll p < 1$.

Before turning to studying the random graph, we will first show several more general results concerning Ramsey simplicity. In Section 4.1 we prove that q-Ramsey simplicity is monotone in the number of colors, justifying the definition of the simplicity threshold. In Section 4.2, we discuss a necessary and a sufficient condition for a graph H satisfying certain conditions to be q-Ramsey simple, relating Ramsey simplicity to a particular packing problem. Before we prove our main results concerning random graphs in Sections 4.4 and 4.5, we discuss some useful properties of the random graph $G(n, p)$ in Section 4.3. Finally, we present some further results, including our abundance result and a result for asymmetric pairs, in Section 4.6. We close with a discussion of open problems and future directions. This chapter represents joint work with Dennis Clemens, Shagnik Das, and Pranshu Gupta and is based on [27].

4.1 Monotonicity in q

In this section, we will prove that the property of being q-Ramsey simple is monotone decreasing in the number of colors, justifying the definition of the simplicity threshold; that is, we will show that if a graph is not q-Ramsey simple for some q, then it cannot be q'-Ramsey simple for any $q' \geq q$. Recall that q-Ramsey simplicity is not a monotone property with respect to the graph H, that is, it is not preserved under adding or removing edges.

Lemma 4.1.1. *If H is not q-Ramsey simple, then H is not $(q+1)$-Ramsey simple.*

Proof. We may assume H has no isolated vertices. Suppose for a contradiction that H is not q-Ramsey simple but it is $(q+1)$-Ramsey simple. Let G be a minimal $(q+1)$-Ramsey graph for H containing a vertex w of degree $(q+1)(\delta(H)-1)+1$. Let e be an arbitrary edge incident to w. By the minimality of G, the graph $G-e$ has an H-free coloring φ.

If at most $\delta(H)-2$ of the edges incident to w have color $q+1$ under φ, then we can extend φ to all of G by giving color $q+1$ to e. The resulting $(q+1)$-coloring of G is H-free: $G-e$ has an H-free coloring and the edge e cannot be part of a monochromatic copy of H, since w is incident to at most $\delta(H)-1$ edges of color $q+1$ in G. This contradicts the fact that $G \to_{q+1} H$.

Thus we may assume that there are least $\delta(H)-1$ edges of color $q+1$ under φ that are incident to w. Let G' be the graph obtained from G by removing all edges that have color $q+1$ under φ, that is, $G' = G - \varphi^{-1}(q+1)$. Then $G' \to_q H$, since otherwise we can take an H-free q-coloring of G' and extend it to an H-free coloring of G by assigning color $q+1$ to the edges in $E(G) \setminus E(G')$.

Consider a minimal q-Ramsey subgraph G'' of G'. If the vertex w is part of G'', we have $d_{G''}(w) \leq d_{G'}(w) \leq d_G(w) - (\delta(H)-1) = q(\delta(H)-1)+1 < s_q(H)$, where the last inequality follows by our assumption that H is not q-Ramsey simple, a contradiction. Hence w cannot be in G'', that is, $G'' \subseteq G' - w = (G - \varphi^{-1}(q+1)) - w$. But then φ induces an H-free q-coloring on G'', contradicting the fact that $G'' \to_q H$. Hence H cannot be $(q+1)$-Ramsey simple. $\square$

4.2 Conditions for Ramsey (non-)simplicity

Throughout the section, H will be a connected graph containing a unique vertex of minimum degree u. We will show that in this case determining whether or not H is q-Ramsey simple is related to a certain packing problem. In particular, we will see that the structure of the neighborhood of u plays a key role in this problem.

Proposition 4.2.1. *Let $q \geq 2$ and H be a connected graph containing a unique vertex u of minimum degree. Let $F = H[N(u)]$ be the subgraph induced by the neighborhood of u. If*

H is q-Ramsey simple, then there exists a color pattern $\Gamma_1, \dots, \Gamma_q$ on a vertex set V of size $q(\delta(H)-1)+1$ such that, for every subset $U \subseteq V$ of size $\delta(H)$ and for every color $i \in [q]$, there exists a copy $F_{U,i}$ of F in $\Gamma_i[U]$.

Proof. Assume H is q-Ramsey simple, and let G be a minimal q-Ramsey graph for H with a vertex w of degree $q(\delta(H)-1)+1$. By the minimality of G, there exists an H-free q-coloring ψ of $G - w$. Let $V = N(w)$ and Γ_i consist of all edges in $G[N(w)]$ that have color i under ψ. We claim that the color pattern $\Gamma_1, \dots, \Gamma_q$ satisfies the required property.

Let $U \subseteq V$ be a subset of size $\delta(H)$ and $i \in [q]$. We extend ψ to all of G by coloring the edges from w to U with color i and coloring the remaining edges incident to w in such a way that every other color is used exactly $\delta(H)-1$ times. Since $G \to_q H$, we know that this new coloring contains a monochromatic copy H' of H, and since the coloring ψ of $G - w$ is H-free, H' must contain w. Furthermore, H has minimum degree $\delta(H)$ and w is incident to $\delta(H)-1$ edges of every color other than i, so H' must have color i. Finally, we know that u is the unique vertex of degree $\delta(H)$ in H, so w must play the role of u in H', and therefore we must find a monochromatic copy of F in color i inside $G[N(w)]$. Hence, $\Gamma_i[U]$ contains a copy of F, as required. $\square$

It turns out that, under some additional conditions, the above simple necessary condition also becomes sufficient. Before we state the precise result, we introduce the notion of a well-behaved graph. Roughly speaking, a well-behaved graph has a unique vertex of minimum degree and satisfies some mild pseudorandom properties on its connectivity, expansion, and co-degrees. As we will see later in this chapter, in the ranges of p we will focus on, a random graph $H \sim G(n,p)$ is very likely to be well-behaved.

Definition 4.2.2 (Well-behaved)**.** We say that an n-vertex graph H is *well-behaved* if it satisfies the following properties:

(W1) H has a unique vertex u of minimum degree $\delta(H)$.
(W2) Every vertex $v \in V(H) \setminus \{u\}$ has at most $\frac{1}{2}\delta(H)$ common neighbors with u.
(W3) H is 3-connected.
(W4) Removing at most $\delta(H)$ vertices from H cannot create a component of size $k \in \left[\frac{1}{2}\delta(H), \frac{1}{2}n\right]$.

We now show that, if H is well-behaved, then the existence of a color pattern as in Proposition 4.2.1 in which the maximum degree of each Γ_i not too large is sufficient to guarantee that H is q-Ramsey simple. We remark here that the extra conditions needed to prove Proposition 4.2.3, that is, the well-behavedness of H and the bound on the maximum degree of each Γ_i, might not be necessary, but they allow us to keep track of the different copies of H when constructing a desired minimal q-Ramsey graph.

Proposition 4.2.3. *Let $q \geq 2$ and H be a well-behaved graph on n vertices. Let u be the unique vertex of minimum degree in H and $F = H[N(u)]$ be the subgraph induced by the neighborhood of u. Suppose there exists a color pattern $\Gamma_1, \ldots, \Gamma_q$ on a set V of $q(\delta(H) - 1) + 1$ vertices such that:*

(i) For every subset $U \subseteq V$ of size $\delta(H)$ and for every color $i \in [q]$, there exists a copy $F_{U,i}$ of F in $\Gamma_i[U]$.

(ii) For each $i \in [q]$, we have $\Delta(\Gamma_i) \leq \delta(H) - 1$.

Then H is q-Ramsey simple.

Proof. We will construct a graph G satisfying the following properties:

(1) $G \to_q H$.

(2) G contains a vertex w of degree $q(\delta(H) - 1) + 1$.

(3) $G - w \not\to_q H$.

To see how this implies that H is q-Ramsey simple, consider a minimal q-Ramsey graph $G' \subseteq G$ and notice that w must be a vertex of G' by (3). In addition, we must have $q(\delta(H) - 1) + 1 \leq s_q(H) \leq \delta(G') \leq d_{G'}(w) \leq d_G(w) = q(\delta(H) - 1) + 1$.

We now explain how to construct the graph G. Let $R = H - (N(u) \cup \{u\})$. We start with the graph $\Gamma = \bigcup_{i=1}^q \Gamma_i$; let φ be a q-coloring of Γ assigning color i to the edges in Γ_i for all $i \in [q]$. Then, for each subset $U \subseteq V(\Gamma)$ of size $\delta(H)$ and for each color $i \in [q]$, there exists a copy $F_{U,i}$ of the graph F in $\Gamma[U]$ all of whose edges have color i under φ. We now create a copy $R_{U,i}$ of R on a disjoint set of vertices and add edges between $V(F_{U,i})$ and $V(R_{U,i})$ in such a way that $F_{U,i} \cup R_{U,i} \cup E(V(F_{U,i}), V(R_{U,i})) \cong H - u$. We denote the resulting graph by Γ^+ and extend the q-coloring φ to a q-coloring φ^+ of Γ^+ as follows: for every subset $U \subseteq V(\Gamma)$ of size $\delta(H)$ and for each color $i \in [q]$, we assign color i to all edges incident to a vertex of $R_{U,i}$, that is, to all edges in $R_{U,i}$ and to all edges between $V(F_{U,i})$ and $V(R_{U,i})$.

We now construct the final graph G from Γ^+. For this, let S^+ and S^- be a positive and a negative signal sender for H in which the distance between the signal edges is at least $n + 1$. These signal senders exist by Lemma 2.2.2 and are both safe by Corollary 2.2.3, since H is 3-connected. Let $e_1, \ldots, e_q$ be a matching disjoint from Γ^+. For all $i, j \in [q]$ with $i < j$, we connect e_i and e_j by a copy of the negative signal sender S^-. In addition, for each edge f of Γ^+, we connect f and $e_{\varphi^+(f)}$ by a copy of the positive signal sender S^+. Finally, we add a new vertex w and connect it to all vertices of Γ to obtain the final graph G. The construction is illustrated in Figure 4.1. We now show that G satisfies the claimed properties. Notice that property (2) holds, since $d_G(w) = |V(\Gamma)| = q(\delta(H) - 1) + 1$.

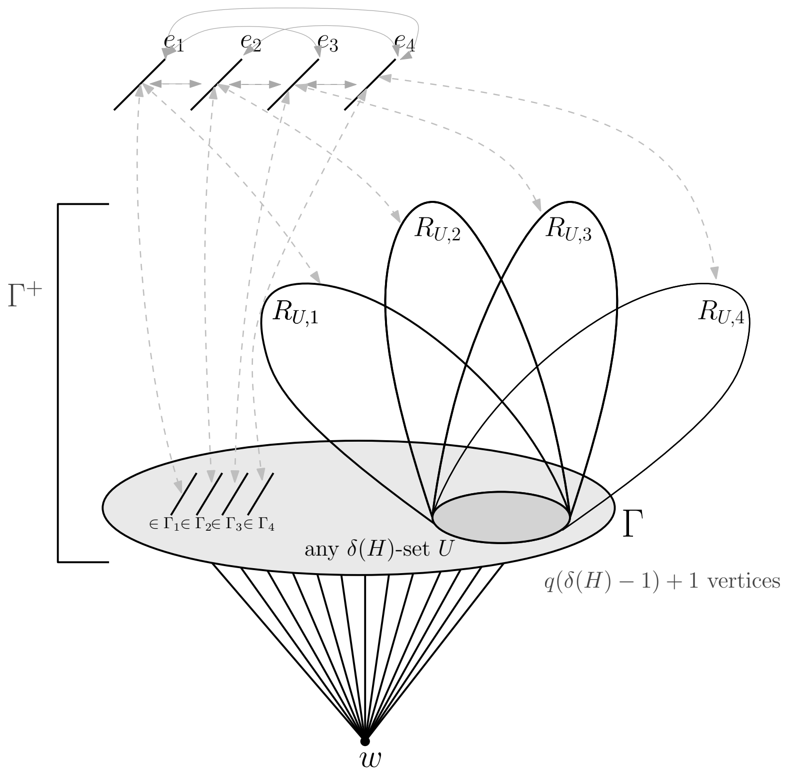

FIGURE 4.1: The graph G in the proof of Proposition 4.2.3 for $q = 4$ (light thin lines represent signal senders: solid lines represent negative signal senders, while dashed lines represent positive signal senders).

We now turn our attention to property (1). Suppose for a contradiction that there exists an H-free q-coloring of G. This coloring induces an H-free coloring on each copy of S^+ and S^-. Property (S2) of the signal senders S^+ and S^- then implies that the edges $e_1, \ldots, e_q$ all have different colors and that each edge f of Γ^+ has the same color as $e_{\varphi^+(f)}$. Without loss of generality, we may assume that e_i has color i for all $i \in [q]$ and hence that each edge f of Γ^+ has color $\varphi^+(f)$.

Now consider the vertex w. By the pigeonhole principle, since $d_G(w) = q(\delta(H) - 1) + 1$, there exists a color i that appears at least $\delta(H)$ times on the edges incident to w. Let $U \subseteq V(\Gamma)$ be a set of $\delta(H)$ vertices such that all edges from w to U have color i. Since φ^+ extends φ, we know that there exists a copy $F_{U,i}$ of F contained in $\Gamma[U] \subseteq G[U]$ whose edges all have color i. Recall that $F_{U,i}$ forms a copy of $H - u$ together with $R_{U,i}$ and the edges between $V(F_{U,i})$ and $V(R_{U,i})$. Further, all edges of this copy of $H - u$ have color i under φ^+; adding w results in a copy of H all of whose edges have color i, proving the claim.

It remains to prove property (3). For this, we provide an H-free q-coloring of $G - w$. We start by coloring Γ^+ according to the coloring φ^+ and giving color i to the edge e_i for all $i \in [q]$. Observe that the signal edges of each copy of S^- are colored differently and that the signal edges of each copy of S^+ have the same color. Thus, by property (S3) of S^+ and S^-, we can extend the partial coloring to the copies of S^+ and S^- in such a way that each signal sender has an H-safe

q-coloring. We now claim that this is an H-free coloring of $G - w$. First notice that, since the colorings of all copies of S^+ and S^- are chosen to be safe, it suffices to show that the graph $\Gamma^+ \cup \{e_1, \dots, e_q\}$ contains no monochromatic copy of H. Further, since H is 3-connected and each edge e_i for $i \in [q]$ is isolated in this graph, we only need to consider the graph Γ^+.

Suppose that there exists a monochromatic copy H' of H that has color $i \in [q]$ in Γ^+. First assume that H' contains vertices from two different copies of R. Since each $R_{U,j}$ is monochromatic in color j under φ^+, the two copies of R intersecting H' must be attached at two different subsets $U_1, U_2 \subseteq V(\Gamma)$. Since the vertex sets of $R_{U_1,i}$ and $R_{U_2,i}$ are disjoint, we may without loss of generality assume that $|V(R_{U_1,i}) \cap V(H')| \le n/2$. We now claim that $|V(R_{U_1,i}) \cap V(H')| \ge \delta(H)/2$. To see why this is true, consider an arbitrary vertex $z \in V(R_{U_1,i}) \cap V(H')$. By definition, we know that z has at least $\delta(H') = \delta(H)$ neighbors in H'. But recall that z is also part of a copy of H formed by $F_{U_1,i}$, $R_{U_1,i}$, w, and the edges between them in G; in this copy of H the vertex w plays the role of u and $F_{U_1,i}$ plays the role of the neighborhood of u. Since H is well-behaved, by property (W2), we know that z has at most $\delta(H)/2$ neighbors in U_1 in the graph Γ^+, that is, $|N_{\Gamma^+}(z) \cap U_1| \le \delta(H)/2$; thus, we must have $|N_{H'}(z) \cap V(R_{U_1,i})| \ge \delta(H)/2$. But then the component of z in $H' - U_1$ has size at least $\delta(H)/2$ and at most $n/2$, which contradicts property (W4) of H. Therefore, we may assume that $V(H')$ intersects at most one copy of R, call it $R_{U,i}$.

Since $|V(R_{U,i})| < n$, we know that H' must contain at least one vertex of Γ. By property (ii), we have $\Delta(\Gamma_i) \le \delta(H) - 1$, so each vertex $y \in V(\Gamma)$ is incident to at most $\delta(H) - 1$ edges of color i inside Γ. Thus, each vertex $y \in V(\Gamma) \cap V(H')$ must have a neighbor in $V(R_{U,i})$. But then $V(\Gamma) \cap V(H') \subseteq U$, as the only edges leaving $V(R_{U,i})$ go to U. Hence $V(H') \subseteq U \cup V(R_{U,i})$, but the latter set contains only $n - 1$ vertices, a contradiction. □

In the next sections, we will use Proposition 4.2.3 and Proposition 4.2.1 to show the claimed results concerning the Ramsey simplicity of $G(n, p)$, namely Theorem 1.1.9. We conclude this section by discussing a simple application. Recall that it was conjectured that all triangle-free graphs are 2-Ramsey simple (see Conjecture 1.1.7). This conjecture is known to be true for a class of regular 3-connected triangle-free graphs [82]; we prove that it holds for all well-behaved triangle-free graphs for any number of colors.

Proposition 4.2.4. *Let H be a well-behaved triangle-free graph. Then H is q-Ramsey simple for all $q \ge 2$.*

Proof. Since H is well-behaved, property (W1) implies that there is a unique vertex u of minimum degree and (W3) implies that $\delta(H) \ge 3$. Since H is triangle-free, we know that the neighborhood of u in H is the empty graph. We take $\Gamma_1, \dots, \Gamma_q$ to be empty graphs on $q(\delta(H) - 1) + 1$ vertices; it is not difficult to check that the (i) and (ii) are both satisfied, and thus applying Proposition 4.2.3 yields the desired result. □

4.3 Properties of $G(n, p)$

Recall that $G(n, p)$ denotes a random graph on vertex set $[n]$ in which each possible edge appears with probability p, independently of all other edges. In this section, we collect some properties of the random graph $G(n, p)$ which will be useful in the proofs of our main results. Many of these results are well-known, so we will often be brief in our proofs. More background on random graphs can be found for example in [24, 75]. Throughout the rest of the chapter, unless otherwise specified, we assume that the number of vertices n tends to infinity.

4.3.1 Probabilistic preliminaries

In the proofs of some of the following lemmas we will use the following well-known concentration (Chernoff) bounds. Part (a) is given for example in [116, Theorem 2.3] and [75, Corollary 21.7]. Part (b) follows easily from [75, Theorem 21.6].

Lemma 4.3.1. *Let $X \sim Bin(n, p)$ and $\mu = \mathbb{E}[X]$.*

(a) If $0 < \varepsilon \leq 1$, then $\mathbb{P}(X \geq (1+\varepsilon)\mu) \leq \exp\left(-\frac{\mu\varepsilon^2}{3}\right)$ and $\mathbb{P}(X \leq (1-\varepsilon)\mu) \leq \exp\left(-\frac{\mu\varepsilon^2}{2}\right)$.

(b) For all $t \geq 7\mu$, we have $\mathbb{P}(X \geq t) \leq \exp(-t)$.

4.3.2 Facts about $G(n, p)$

We begin by proving several facts about the degrees in $G(n, p)$.

Lemma 4.3.2 (Degrees in $G(n, p)$)**.** *Let $p = p(n) \in (0, 1)$, and let $H \sim G(n, p)$. Then w.h.p. the following bounds on the maximum degree hold:*

(a) For any fixed integer $k \geq 2$, if $p \gg n^{-\frac{k}{k-1}}$, then $\Delta(H) \geq k - 1$.

(b) For any $f = f(n)$ satisfying $1 \ll f = n^{o(1)}$, if $p = \frac{1}{nf}$, then $\Delta(H) \geq \frac{\log n}{\log(f \log n)}$.

Moreover, if $p \gg \frac{\log n}{n}$, then with probability at least $1 - n^{-2}$ we have

(c) $d_H(v) = (1 \pm o(1))np$ for every $v \in V(H)$.

Proof. For part (a), it suffices to show that, if $p \gg n^{-\frac{k}{k-1}}$, then w.h.p. H contains the star $K_{1,k-1}$ with $k - 1$ edges. This statement is well known and can be shown using a straightforward application of the second moment method; see for example [75, Theorem 5.3] for more details.

Part (b) again can be shown using the second moment method. To simplify the presentation, we apply directly Theorem 3.1(ii) from [24], implying that, if $n^{-3/2} \ll p \ll 1$ and the expected number of vertices of degree $d = d(n)$ in $H \sim G(n,p)$ tends to infinity, then w.h.p. H contains a vertex of degree d. Now let $p = \frac{1}{nf}$ and $d = \frac{\log n}{\log(f \log n)}$ be as given in (b). We will show that the condition of Theorem 3.1(ii) in [24] is satisfied; we have

$$\begin{aligned}
\mathbb{E}[\text{number of vertices of degree } d \text{ in } H] &= n\binom{n-1}{d}p^d(1-p)^{n-1-d} \\
&\geq n\left(\frac{n-1}{d}\right)^d p^d(1-np) \geq \frac{n}{2}\left(\frac{1}{2df}\right)^d \\
&= \frac{1}{2}\exp(\log n - d\log(2df)) \\
&= \frac{1}{2}\exp\left(\log n - \frac{\log n}{\log(f\log n)}\log\left(\frac{2f\log n}{\log(f\log n)}\right)\right) \\
&= \frac{1}{2}\exp\left(\frac{\log n}{\log(f\log n)}\log(1/2\log(f\log n))\right) \to \infty,
\end{aligned}$$

where the first inequality follows from the standard estimates $\binom{n-1}{d} \geq (\frac{n-1}{d})^d$ and $(1-p)^{n-1-d} \geq (1-p)^n \geq 1-np$ and the second inequality comes from the fact that $np = \frac{1}{f} \ll 1$ and $\frac{n-1}{n} \geq \frac{1}{2}$. Therefore, we conclude that w.h.p. H must contain a vertex of degree d, implying in turn that $\Delta(H) \geq d$, as required.

Finally, part (c) follows from a straightforward application of Lemma 4.3.1(a). Let $\varepsilon > 0$. For each vertex $v \in V(H)$, we have $\mathbb{E}(d_H(v)) = (n-1)p$ and thus the probability that the degree of v is not in $(1 \pm \varepsilon)(n-1)p$ is at most $2\exp\left(-\frac{(n-1)p\varepsilon^2}{3}\right) \leq 2\exp(-4\log n)$, since $p \gg \frac{\log n}{n}$. Taking a union bound over all n vertices, we obtain that the probability that the degree of some vertex is not in $(1 \pm \varepsilon)(n-1)p$ is at most n^{-2}, as required. □

We next consider the number of edges in $G(n,p)$. We will need to understand both the total number of egdes and the number of edges in every large enough induced subgraph.

Lemma 4.3.3 (Edge counts in $G(n,p)$)**.** *Let $p = p(n) \in (0,1)$ with $p \gg n^{-2}$, and let $H \sim G(n,p)$. Then w.h.p. the following statements hold:*

(a) $e(H) = (1 \pm o(1))\frac{n^2 p}{2}$.

(b) If $p \gg \frac{\log n}{n}$, then with probability at least $1-n^{-2}$, every set $S \subseteq V(H)$ of size $s \geq \frac{20\log n}{p}$ satisfies $e_H(S) \geq \frac{1}{4}s^2 p$.

Proof. We again use Lemma 4.3.1(a). Let $S \subseteq V(H)$ with $|S| = s$. Then the expected number of edges in $H[S]$ is $\binom{s}{2}p$.

Now, to show part (a), consider $S = V(H)$ and any $\varepsilon > 0$; applying Lemma 4.3.1(a), we conclude that the probability that the number of edges in H is not in $(1 \pm \varepsilon)\binom{n}{2}p$ is at most $2\exp\left(-\frac{\binom{n}{2}p\varepsilon^2}{3}\right) \to 0$, since $n^2 p \to \infty$.

For part (b), note that, since $s^2 p \gg 1$, we have $\frac{1}{2}s^2 p = (1+o(1))\binom{s}{2}p$. Then using Lemma 4.3.1(a) with $\varepsilon = \frac{1}{2} - o(1)$ and taking a union bound over all sets S with $\frac{20\log n}{p} \leq |S| \leq n$, we obtain that the probability that S contains fewer than $\frac{1}{4}s^2 p$ edges is at most

$$\sum_{s=\frac{20\log n}{p}}^{n} \binom{n}{s} \exp\left(-(1-o(1))\frac{\varepsilon^2}{4}s^2 p\right) \leq \sum_{s=\frac{20\log n}{p}}^{n} \exp\left(s\log n - (1-o(1))\frac{20}{16}s\log n\right)$$
$$\leq \sum_{s=\frac{20\log n}{p}}^{n} \exp(-0.1 s\log n) \leq \exp(\log n - 0.1(\log n)^2) < n^{-2}.$$

□

We will also need the following facts concerning the structure of a sparse random graph and the size of its components.

Lemma 4.3.4. *Let $p = p(n) \in (0, 1)$ with $p \ll n^{-1}$, and let $H \sim G(n, p)$. Then w.h.p. H is a forest and the order $\lambda(H)$ of its largest component satisfies the following:*

(a) $\lambda(H) \leq \log n$.
(b) If $p \ll n^{-\frac{k+1}{k}}$ for some constant $k \in \mathbb{N}$, then $\lambda(H) \leq k$.
(c) If $p = \frac{1}{nf}$ for some $f = f(n)$ satisfying $1 \ll f = n^{o(1)}$, then $\lambda(H) \leq (1 + o(1))\frac{\log n}{\log f}$.

Proof. It is well known that if $p \ll n^{-1}$, then w.h.p. H is a forest (see for example Theorem 2.1 in [75]); the proof is a simple application of the first moment method, taking a union bound over all possible cycles.

Parts (a) and (b) can be deduced easily for instance from Lemma 2.12(ii) and Theorem 5.3 in [75]. In fact, all three can be shown using the same simple argument, similar for example to the proof of Theorem 2.5 in [75], which we now present. Since $p \ll n^{-1}$, we write $p = \frac{1}{nf}$ for some $f = f(n)$ satisfying $1 \ll f \leq n$.

A well-known result due to Cayley states that, for any given t, there are t^{t-2} labeled trees on t vertices (see for example [151, Theorem 2.2.3]). For any fixed tree T on t vertices, the probability that T appears as a subgraph of H is $\binom{n}{t}p^{t-1}$. Taking a union bound over all possible trees, we find that the probability that H contains a tree on t vertices is at most

$$t^{t-2}\binom{n}{t}p^{t-1} \leq \left(\frac{ne}{t}\right)^t \frac{(pt)^t}{t^2 p} \leq \frac{1}{p}(nep)^t = nf\left(\frac{e}{f}\right)^t = \exp(\log(nf) + t - t\log f), \qquad (4.3.1)$$

where the first step follows from the well-known inequality $\binom{n}{t} \leq \left(\frac{ne}{t}\right)^t$.

Now, for part (a), substitute $t = \log n$ into (4.3.1) to obtain that the probability that H contains a tree on $\log n$ vertices is at most

$$\exp(\log(nf) + t - t\log f) \le \exp(3\log n - \log n \log f) \to 0,$$

since $\log f \to \infty$.

For part (b), we have $p \ll n^{-\frac{k+1}{k}}$ for some constant k and therefore $f = f' n^{1/k}$, where $f' = f'(n)$ satisfies $f' \gg 1$. Substituting $t = k+1$ into (4.3.1), we obtain that the corresponding probability is bounded above by

$$\exp(\log n + \log f + t - t\log f) \le \exp(\log n + k + 1 - k\log(f' n^{1/k})) = \exp(k + 1 - k\log f') \to 0,$$

since $\log f' \to \infty$.

Finally, for part (c), let $\varepsilon > 0$. Then $f = n^{o(1)} \le n^{\varepsilon/2}$, and we substitute $t = (1+\varepsilon)\frac{\log n}{\log f}$ into (4.3.1) to obtain an upper bound of

$$\begin{aligned}\exp(\log(nf) + t - t\log f) &\le \exp\left(\left(1 + \frac{\varepsilon}{2}\right)\log n + (1+\varepsilon)\frac{\log n}{\log f} - (1+\varepsilon)\frac{\log n}{\log f}\log f\right) \\ &= \exp\left(-\frac{\varepsilon}{2}\log n + (1+\varepsilon)\frac{\log n}{\log f}\right) \to 0,\end{aligned}$$

again since $\log f \to \infty$.

Thus, in all three cases the probability that H contains a tree on t vertices is $o(1)$, implying that $\lambda(H) \le t - 1$ with high probability. □

We now switch to a much denser range, and show that in that case not only is $G(n,p)$ likely contain triangles, but in fact every edge is likely to be contained in a triangle.

Lemma 4.3.5. *Let $p = p(n) \in (0,1)$ be such that $p \gg \sqrt{\frac{\log n}{n}}$, and let $H \sim G(n,p)$. Then w.h.p. every edge of H is contained in a triangle.*

Proof. For a given edge $xy \in E(H)$, if xy is not contained in a triangle, then for every vertex $z \in V(H) \setminus \{x,y\}$ at least one of the edges xz and yz must be missing. Thus the probability that xy is not contained in a triangle is equal to $(1-p^2)^{n-2}$. Taking a union bound over all (potential) edges, we obtain

$$\begin{aligned}\mathbb{P}[\exists xy \in E(H) : xy \text{ is not in a triangle}] &\le \binom{n}{2}(1-p^2)^{n-2} \le n^2\exp(-p^2(n-2)) \\ &= \exp\left(2\log n - p^2(n-2)\right) \to 0,\end{aligned}$$

where for the second step we use the inequality $1 - x \leq e^{-x}$ and the final step follows by the assumption that $p^2 n \gg \log n$. $\square$

In order to be able to apply Proposition 4.2.3, we need to ensure that our target graph is well-behaved. As expected, this is likely to be true for a random graph when the probability p is neither too small nor too large, which is the content of the next lemma.

Lemma 4.3.6. *If* $\frac{\log n}{n} \ll p \ll 1$, *then w.h.p.* $H \sim G(n, p)$ *is well-behaved.*

Proof. It is well known that, if $p \gg \frac{\log n}{n}$, then w.h.p. H is 3-connected (see for example Theorem 4.3 in [75]), showing property (W3). Theorem 3.9(i) in [24] states that, if $\frac{\log n}{n} \ll p \leq \frac{1}{2}$, then w.h.p. $H \sim G(n, p)$ has a unique vertex of minimum degree, verifying property (W1). Further, by Lemma 4.3.2(c), we know that w.h.p. $\delta(H) = (1 \pm o(1))np$. We condition on the latter two properties from now on. It remains to show properties (W2) and (W4).

For property (W2), observe that, for any pair of distinct vertices x and y, the number of common neighbors of x and y, which we denote by $d_H(x, y)$, is distributed according to $\mathrm{Bin}(n-2, p^2)$ and has expected value $(n-2)p^2$. We consider two cases. First assume that $np^2 \geq 10 \log n$. Then, for every pair x, y of distinct vertices of H, by Lemma 4.3.1(a), we have

$$\mathbb{P}[d_H(x, y) \geq 2(n-2)p^2] \leq \exp\left(-\frac{(n-2)p^2}{3}\right) \leq \exp\left(-\frac{10(n-2)\log n}{3n}\right) \leq \exp(-3\log n).$$

Taking a union bound over all choices of x, y, we obtain that w.h.p. $d_H(x, y) \leq 2(n-2)p^2 \leq \frac{1}{2}\delta(H)$ for all pairs x, y, since $p \ll 1$. Assume next that $np^2 \leq 10 \log n$. In this case we apply Lemma 4.3.1(b) to conclude that, for every pair of distinct vertices x, y, we have $\mathbb{P}[d_H(x, y) \geq 100 \log n] \leq \exp(-100 \log n)$. Again taking a union bound over all pairs of vertices, we find that w.h.p. $d_H(x, y) \leq 100 \log n \leq \frac{1}{2}\delta(H)$ for all pairs x, y, since $p \gg \frac{\log n}{n}$.

Finally, we show (W4). Fix integers $k \in \left[\frac{\delta(H)}{2}, \frac{n}{2}\right]$ and $1 \leq s \leq \delta(H)$, a subset $U \subseteq V(H)$ of size s, and a subset $K \subseteq V(H) \setminus U$ of size k. We first determine the probability that K is a component of $H - U$. For K to be a component of $H - U$, there must be no edges between K and $V(H) \setminus (K \cup U)$, which happens with probability $(1-p)^{k(n-k-s)}$. Now taking a union bound over all possible choices of k, s, U, and K, we obtain that the probability that property (W4) fails is at most

$$\begin{aligned}\sum_{s=1}^{\delta(H)}\sum_{k=\delta(H)/2}^{n/2}\binom{n}{s}\binom{n-s}{k}(1-p)^{k(n-k-s)} &\le \sum_{s=1}^{\delta(H)}\sum_{k=\delta(H)/2}^{n/2}\binom{n}{\delta(H)}\binom{n}{k}(1-p)^{k(n-k-\delta(H))}\\ &\le n\sum_{k=\delta(H)/2}^{n/2}\binom{n}{\delta(H)}\binom{n}{k}\exp(-pk(n-k-\delta(H)))\\ &\le n\sum_{k=\delta(H)/2}^{n/2}\exp\left(\delta(H)\log n + k\log n - \frac{1}{4}npk\right)\\ &\le n\sum_{k=\delta(H)/2}^{n/2}\exp\left(-\frac{5}{32}npk\right) \le n\sum_{k=\delta(H)/2}^{n/2}\exp\left(-\frac{5}{2}\log n\right)\\ &\le \exp(2\log n - 2.5\log n) \to 0.\end{aligned}$$

For the first inequality we use the fact that $s \le \delta(H) \le 2np \le \frac{n}{2}$, the third inequality follows from the fact that $k \le \frac{n}{2}$ and $\delta(H) \le (1+o(1))np \le \frac{1}{4}n$, for the fourth and fifth inequality we use the fact that $npk \ge 32n\frac{\log n}{n}k \ge 32k\log n \ge 32\frac{\delta(H)}{2}\log n = 16\delta(H)\log n$, since $p \gg \frac{\log n}{n}$. □

4.3.3 Transference lemma

As is evident from the statement of Theorem 1.1.9, our bounds on the simplicity threshold of a random graph depend on the graph induced by the neighborhood of the unique vertex of minimum degree. Our goal is to understand this subgraph. The next lemma shows that this subgraph behaves essentially like a smaller random graph, allowing us to use our knowledge about a random graph sampled from $G(cnp, p)$ to understand the neighborhood subgraph. As usual, a graph property $\mathcal{P}_s$ is a subset of all labeled graphs on vertex set $[s]$.

Lemma 4.3.7. *Let $p = p(n) \in (0,1)$ be such that $\frac{\log n}{n} \ll p \ll 1$. For every $s \in [0.5np, 2np]$, let $\mathcal{P}_s$ be a graph property, and assume that a random graph $G_s \sim G(s,p)$ satisfies*

$$\mathbb{P}[G_s \in \mathcal{P}_s] = 1 - o(1).$$

Then $H \sim G(n,p)$ w.h.p. has a unique minimum degree vertex u and $H[N_H(u)] \in \mathcal{P}_{d_H(u)}$.

Proof. We will follow an approach similar to that used in the proof of Corollary 2.1.4 in [83]. Before we proceed with the proof, we introduce some notation and facts that we will need later on.

We begin by fixing some $\beta_n = o(1)$ such that

$$\mathbb{P}\left[G_s \notin \mathcal{P}_s\right] = o(\beta_n) \tag{4.3.2}$$

for every $s \in [0.5np, 2np]$. Moreover, let X_δ denote the event that H has a unique vertex of minimum degree and $0.5np \le \delta(H) \le 2np$. By Lemma 4.3.6, we know that H is well-behaved with high probability; more specifically, by property (W1) and Lemma 4.3.2 we know that X_δ holds with high probability. In particular, we can find $\delta_n = o(1)$ such that

$$\mathbb{P}[0.5np \le \delta(H) \le 2np] \ge \mathbb{P}[X_\delta] = 1 - \delta_n.$$

In addition, we will need the following two facts:

(1) There exists $\gamma_n = o(1)$ such that, for any $d \ge 0$, we have $\mathbb{P}[\delta(H) = d] \le \gamma_n$.
(2) For any $d \ge 0$, if $H' \sim G(n-1,p)$, we have $\mathbb{P}[\delta(H') \ge d-1] \ge \mathbb{P}[\delta(H) \ge d]$.

To see fact (2), note that we can sample from $G(n,p)$ in the following way: first sample a graph $H' \sim G(n-1,p)$; then add a new vertex v and, for each $u \in V(H')$, add the edge uv with probability p independently of all other edges. It is not difficult to see that the resulting graph, which we denote by H, indeed has the distribution $G(n,p)$ and that, if $\delta(H) \ge d$, then $\delta(H') \ge d-1$, implying the claim.

Fact (1) follows from the proof of Theorem 3.9(i) in [24]; we briefly sketch how to deduce the statement. The proof of Theorem 3.9(i) shows that, if $\frac{\log n}{n} \ll p \le \frac{1}{2}$ and $H \sim G(n,p)$, there exists an integer $m = m(n,p)$ satisfying $m < np$ and $n\binom{n-1}{m}p^m(1-p)^{n-1-m} = o(1)$, such that w.h.p. H has a vertex of degree at most m (note that the proof given in [24] is written in terms of maximum degrees). Now let $d \ge 0$ be an integer. If $d > m$, then $\mathbb{P}[\delta(H) = d] = o(1)$, since with high probability H contains a vertex of smaller degree. Assume next that $d \le m$. We know that the probability that any fixed vertex has degree d is $\binom{n-1}{d}p^d(1-p)^{n-1-d}$; it is not difficult to check that, for $d \le m-1 \le np-1$, we have $\binom{n-1}{d+1}p^{d+1}(1-p)^{n-1-(d+1)} \ge \binom{n-1}{d}p^d(1-p)^{n-1-d}$. Therefore, taking a union bound over all n vertices, we have $\mathbb{P}[\text{there exists a vertex of degree } d] \le n\binom{n-1}{m}p^m(1-p)^{n-1-m} = o(1)$. Thus, we again obtain that $\mathbb{P}[\delta(H) = d] = o(1)$.

Next, let $\varepsilon_n = o(1)$ be chosen such that $\varepsilon_n = \omega(\max\{\beta_n, \gamma_n, \delta_n\})$. We further let t_n be the smallest integer such that $\mathbb{P}[\delta(H) \le t_n] \ge 1 - \varepsilon_n$. Note that, by the minimality of t_n, we have $\mathbb{P}[\delta(H) \le t_n - 1] < 1 - \varepsilon_n$. Using fact (1) for $d = t_n$, we conclude

$$1 - \varepsilon_n \le \mathbb{P}[\delta(H) \le t_n] = \mathbb{P}[\delta(H) \le t_n - 1] + \mathbb{P}[\delta(H) = t_n] \le 1 - \varepsilon_n + \gamma_n. \tag{4.3.3}$$

Moreover, since $\varepsilon_n > \gamma_n + \delta_n$, we obtain $\mathbb{P}[\delta(H) \le t_n] < 1 - \delta_n \le \mathbb{P}[\delta(H) \le 2np]$ and thus $t_n \le 2np$.

Since $H \sim G(n,p)$, the subgraph $H - v$, for any fixed vertex v, has the distribution $G(n-1,p)$. In the following, we will condition on the event X_δ, and whenever we do so, we will always let u denote the unique minimum degree vertex in H and write $d = d_H(u)$. We will be interested in the subgraph $H' = H - u$, and first need to determine how conditioning on X_δ affects its distribution.

Suppose $S \subseteq V(H')$ is the neighborhood of u. As u is the only vertex of degree at most d in H, we must have $d_{H'}(v) \geq d + 1$ for all $v \in V(H') \setminus S$, and $d_{H'}(v) \geq d$ for all $v \in S$; let C_S be the event that these lower bounds on the degrees in H' hold. Aside from C_S, however, X_δ yields no further information about the graph H', as the edges in $G(n, p)$ are independent. Thus, we have

$$\mathbb{P}_{G(n,p)}[H[S] \in \mathcal{P}_d | X_\delta \wedge \{N_H(u) = S\}] = \mathbb{P}_{G(n-1,p)}[H'[S] \in \mathcal{P}_d | C_S]. \tag{4.3.4}$$

Now, by the Law of Total Probability,

$$\begin{aligned}
&\mathbb{P}_{G(n,p)}[H[N_H(u)] \in \mathcal{P}_{d_H(u)} | X_\delta] \\
&= \sum_{0 \leq d \leq n-1} \sum_{\substack{S \subseteq V(H') \\ |S|=d}} \mathbb{P}_{G(n,p)}[H[S] \in \mathcal{P}_d | X_\delta \wedge \{N_H(u) = S\}] \mathbb{P}_{G(n,p)}[N_H(u) = S | X_\delta] \\
&\geq \sum_{0.5np \leq d \leq t_n} \sum_{\substack{S \subseteq V(H') \\ |S|=d}} \mathbb{P}_{G(n-1,p)}[H'[S] \in \mathcal{P}_d | C_S] \mathbb{P}_{G(n,p)}[N_H(u) = S | X_\delta].
\end{aligned} \tag{4.3.5}$$

To estimate the first factor, we observe that

$$\begin{aligned}
\mathbb{P}_{G(n-1,p)}[C_S] &\geq \mathbb{P}[\delta(H') \geq d + 1] \geq \mathbb{P}[\delta(H) \geq d + 2] \\
&\geq \mathbb{P}[\delta(H) \geq t_n + 2] \geq \mathbb{P}[\delta(H) \geq t_n + 1] - \gamma_n \geq \varepsilon_n/2,
\end{aligned} \tag{4.3.6}$$

where the second inequality follows from fact (2), for the third inequality we use $d \leq t_n$, the fourth inequality follows from fact (1), and the last inequality comes from (4.3.3) and since $\varepsilon_n = \omega(\gamma_n)$. Hence we have

$$\begin{aligned}
\mathbb{P}_{G(n-1,p)}[H'[S] \in \mathcal{P}_d | C_S] &= 1 - \mathbb{P}_{G(n-1,p)}[H'[S] \notin \mathcal{P}_d | C_S] \\
&= 1 - \frac{\mathbb{P}_{G(n-1,p)}[\{H'[S] \notin \mathcal{P}_d\} \wedge C_S]}{\mathbb{P}_{G(n-1,p)}[C_S]} \\
&\geq 1 - \frac{\mathbb{P}_{G(n-1,p)}[H'[S] \notin \mathcal{P}_d]}{\mathbb{P}_{G(n-1,p)}[C_S]} \\
&\geq 1 - \frac{\mathbb{P}_{G(d,p)}[G_d \notin \mathcal{P}_d]}{\varepsilon_n/2} = 1 - o(1),
\end{aligned}$$

where for the second inequality we use (4.3.6) and that $H'[S] \sim G(d,p)$ and the final estimate uses (4.3.2) and $\beta_n = o(\varepsilon_n)$. Putting this into (4.3.5), we conclude that

$$\begin{aligned}
\mathbb{P}_{G(n,p)}[H[N_H(u)] \in \mathcal{P}_{d_H(u)} | X_\delta] &\geq \sum_{0.5np \leq d \leq t_n} \sum_{\substack{S \subseteq V(H') \\ |S|=d}} (1-o(1))\mathbb{P}_{G(n,p)}[N_H(u) = S | X_\delta] \\
&= (1-o(1))\mathbb{P}_{G(n,p)}\left[0.5np \leq \delta(H) \leq t_n | X_\delta\right] \\
&= (1-o(1))\frac{\mathbb{P}_{G(n,p)}[\{\delta(H) \leq t_n\} \wedge X_\delta)]}{\mathbb{P}[X_\delta]} \\
&\geq (1-o(1))\frac{1-\mathbb{P}_{G(n,p)}[\delta(H) > t_n] - \mathbb{P}_{G(n,p)}[\overline{X_\delta}]}{\mathbb{P}[X_\delta]} \\
&\geq (1-o(1))\frac{1-\varepsilon_n-\delta_n}{1-\delta_n} = 1-o(1).
\end{aligned}$$

This proves the lemma. □

4.3.4 The smallest neighborhood and quantitative simplicity

Throughout this section we will always assume that $\frac{\log n}{n} \ll p \ll 1$, in which case by Lemma 4.3.6 and (W1) we know that a random graph $H \sim G(n,p)$ almost surely has a unique vertex of minimum degree. We now use Lemma 4.3.7 to show some properties of the subgraph F induced by the neighborhood of the unique vertex of minimum degree for several different ranges of p. The results in this section will also allow us to deduce Corollary 1.1.10 from Theorem 1.1.9. We begin by considering the sparsest range, when $p \ll n^{-2/3}$, and show that in this case the graph F is almost surely empty.

Corollary 4.3.8. *Let $p = p(n) \in (0,1)$ be such that $\frac{\log n}{n} \ll p \ll n^{-\frac{2}{3}}$, and let $H \sim G(n,p)$. Then w.h.p. H has a unique minimum degree vertex u and $e(N_H(u)) = 0$.*

Proof. Let $s \in [0.5np, 2np]$ and $G_s \sim G(s,p)$. Then, taking a union bound over all edges, we know that the probability that G_s contains an edge is at most $\binom{s}{2}p \leq 4n^2p^3 \to 0$ since $p \ll n^{-2/3}$. By Lemma 4.3.7, it follows that with high probability $e(N_H(u)) = 0$. □

For larger values of p, we start to see edges in the graph F. One of the bounds given in Theorem 1.1.9(d) depends on the number of edges in the graph F, so to obtain quantitative bounds on the simplicity threshold we need to study this number. This is the content of our next corollary. The upper bound will also be useful in our simplicity proofs.

Corollary 4.3.9. *Let $p = p(n) \in (0,1)$ satisfy $\Omega(n^{-\frac{2}{3}}) = p \ll 1$, and let $H \sim G(n,p)$. Then w.h.p. H has a unique minimum degree vertex u, and the number of edges in the graph $F = H[N(u)]$ satisfies:*

(a) $e(F) = o(np)$ *if* $p \ll n^{-\frac{1}{2}}$.

(b) $\frac{1}{16}n^2p^3 \le e(F) \le 4n^2p^3$ *if* $n^{-\frac{2}{3}} \ll p$.

Proof. Let $s \in [0.5np, 2np]$ and $G_s \sim G(s,p)$. By Lemma 4.3.7, it suffices to show that the statements are true with high probability for G_s. The expected number of edges in G_s is $\binom{s}{2}p \le s^2p$. If $p \ll n^{-\frac{1}{2}}$, then by Markov's inequality (see for example [75, Lemma 20.1]), for any $\varepsilon > 0$, we have $\mathbb{P}[e(G_s) \ge \varepsilon np] \le \frac{s^2p}{\varepsilon np} \le \frac{4n^2p^3}{\varepsilon np} \to 0$, showing part (a). If $p \gg n^{-\frac{2}{3}}$, we have $p \gg s^{-2}$, so by Lemma 4.3.3(a) we know that w.h.p. $e(G_s) = (1 \pm o(1))\frac{s^2p}{2}$ and thus $\frac{1}{16}n^2p^3 \le e(G_s) \le 4n^2p^3$, proving (b). □

In order to construct suitable packings to prove Theorem 1.1.9(c), we will need the fact that F is likely to be a very sparse forest. In addition, to obtain quantitative estimates from the bounds given by Theorem 1.1.9(c) and (d) we will need to control the maximum degree and the largest component of the graph F. These tasks are accomplished by the following corollary.

Corollary 4.3.10. *Let* $p = p(n) \in (0,1)$ *be such that* $\Omega(n^{-\frac{2}{3}}) = p \ll n^{-\frac{1}{2}}$, *and let* $H \sim G(n,p)$. *Then w.h.p.* H *has a unique minimum degree vertex* u, *the graph* $F = H[N(u)]$ *induces a forest, and the order* $\lambda(F)$ *of the largest component in* F *satisfies the following:*

(a) $\lambda(F) \le \frac{1}{2}\log n$,

(b) *If* $p \ll n^{-\frac{k+1}{2k+1}}$ *for some fixed integer* $k \ge 2$, *then* $\lambda(F) \le k$.

(c) *If* $p = n^{-\frac{1}{2}}f^{-1}$ *for some* $f = f(n)$ *satisfying* $1 \ll f = n^{o(1)}$, *then* $\lambda(F) \le \left(\frac{1}{4} + o(1)\right)\frac{\log n}{\log f}$.

Moreover, the maximum degree $\Delta(F)$ *of* F *w.h.p. satisfies the following:*

(d) *If* $p \gg n^{-\frac{k}{2k-1}}$ *for some fixed integer* $k \ge 2$, *then* $\Delta(F) \ge k-1$.

(e) *If* $p = n^{-1/2}f^{-1}$ *for some* $1 \ll f = f(n) = n^{o(1)}$, *then* $\Delta(F) \ge \left(\frac{1}{2} - o(1)\right)\frac{\log n}{\log(f^2\log n)}$.

Proof. Let $s \in [0.5np, 2np]$ and $G_s \sim G(s,p)$. Again, by Lemma 4.3.7, it suffices to show that each of the statements is true with high probability for G_s. Note that $p \ll n^{-\frac{1}{2}}$ implies that $p \ll s^{-1}$ and therefore by Lemma 4.3.4, we know that G_s is a forest with high probability.

We start with the three statements concerning the size of the maximum component of G_s. For part (a), we apply Lemma 4.3.4(a) to deduce that with high probability $\lambda(G_s) \le \log s \le \frac{1}{2}\log n$, since $s \le 2np \ll n^{1/2}$. For (b), we have $p \ll n^{-\frac{k+1}{2k+1}}$, implying that $p \ll s^{-\frac{k+1}{k}}$ and thus that $\lambda(G_s) \le k$ with high probability by Lemma 4.3.4(b). For (c), we have $p = n^{-\frac{1}{2}}f^{-1}$, so $\frac{1}{2sf^2} \le p \le \frac{2}{sf^2}$ and $s = n^{1/2-o(1)}$. Thus, by Lemma 4.3.4(c), with high probability we have $\lambda(G_s) \le (1+o(1))\frac{\log s}{\log(f^2/2)} = \left(\frac{1}{4} + o(1)\right)\frac{\log n}{\log f}$.

Now we turn to the statements concerning the maximum degree of G_s. For (d), note that $p \gg n^{-\frac{k}{2k-1}}$ implies $p \gg s^{-\frac{k}{k-1}}$, so by Lemma 4.3.2(a), with high probability we have $\Delta(G_s) \ge k-1$.

Finally, for part (e), as before we have $\frac{1}{2sf^2} \le p \le \frac{2}{sf^2}$, and by Lemma 4.3.2(b), we have $\Delta(H) \ge \frac{\log s}{\log(2f^2 \log n)} \ge \left(\frac{1}{2} - o(1)\right)\frac{\log n}{\log(f^2 \log n)}$ with high probability. □

We are now ready to prove Corollary 1.1.10. We recall the statement below.

Corollary 1.1.10. *Let $k \ge 2$ be a fixed integer, and let $f = f(n)$ satisfy $1 \ll f = n^{o(1)}$. Further, let $p = p(n)$ satisfy $n^{-\frac{2}{3}} \ll p \ll \left(\frac{\log n}{n}\right)^{\frac{1}{2}}$ and $H \sim G(n,p)$. Then w.h.p. the following bounds hold:*

(a) $(1+o(1))\frac{np}{k^2} \le \tilde{q}(H) \le (1+o(1))\frac{np}{k-1}$ *if* $n^{-\frac{k}{2k-1}} \ll p \ll n^{-\frac{k+1}{2k+1}}$.

(b) $(1+o(1))\frac{np}{(k+1)^2} \le \tilde{q}(H) \le (1+o(1))\frac{np}{k-1}$ *if* $p = \Theta\left(n^{-\frac{k+1}{2k+1}}\right)$.

(c) $(1+o(1))\frac{np}{\log n}\max\left\{\frac{16\log^2 f}{\log n}, \frac{1}{80}\right\} \le \tilde{q}(H) \le (1+o(1))\frac{np}{\log n}2\log(f^2\log n)$ *if* $p = n^{-\frac{1}{2}}f^{-1}$.

(d) $1 \le \tilde{q}(H) \le (8+o(1))\frac{1}{p}$ *if* $n^{-\frac{1}{2}} \ll p \ll \left(\frac{\log n}{n}\right)^{\frac{1}{2}}$.

Proof of Corollary 1.1.10. Note that by Lemma 4.3.2(c), w.h.p. $\delta(H) = (1 \pm o(1))np$. By parts (c) and (d) of Theorem 1.1.9, we know that w.h.p. $(1+o(1))\max\left\{\frac{\delta(H)}{\lambda(F)^2}, \frac{\delta(H)}{80\log n}\right\} \le \tilde{q}(H) \le \min\left\{\frac{\delta(H)}{\Delta(F)}, \frac{\delta(H)^2}{2e(F)}\right\}$.

If $n^{-\frac{k}{2k-1}} \ll p \ll n^{-\frac{k+1}{2k+1}}$, the required bounds follow directly by applying Corollary 4.3.10(b) and (d). If $p = \Theta\left(n^{-\frac{k+1}{2k+1}}\right)$, then noting that $n^{-\frac{k}{2k-1}} \ll p \ll n^{-\frac{(k+1)+1}{2(k+1)+1}}$ and again using Corollary 4.3.10(b) and (d) implies the claimed bounds. If $p = n^{-\frac{1}{2}}f^{-1}$, then a direct application of Corollary 4.3.10(c) and (e) yields the result. Finally, if $n^{-\frac{1}{2}} \ll p \ll \left(\frac{\log n}{n}\right)^{\frac{1}{2}}$, then we can use the lower bound in Corollary 4.3.9(b) to obtain the required upper bound on the simplicity threshold. □

4.4 Simplicity for $G(n,p)$

In this section we prove the lower bounds on $\tilde{q}(G(n,p))$ from Theorem 1.1.9. These are the positive results, showing that with high probability $H \sim G(n,p)$ is q-Ramsey simple for the appropriate values of q.

4.4.1 The case $p \ll n^{-1}$

We first consider (a). Note that there is nothing new to prove here: by Lemma 4.3.4 we know H is a forest with high probability when $p \ll n^{-1}$; Szabó, Zumstein, and Zürcher [143] proved that all forests are 2-Ramsey simple, and their proof extends directly to show q-Ramsey simplicity for all $q \ge 2$ as well. The result of Szabó, Zumstein, and Zürcher follows from a more general argument for a large class of bipartite graphs; for completeness, we provide a simpler version of the argument for the special case of forests below (the construction is similar to that in [143]).

Proposition 4.4.1. *For every forest F without isolated vertices and every integer $q \geq 2$, we have $s_q(F) = 1$. In particular, every forest is q-Ramsey simple for all $q \geq 2$.*

Proof. Let F be a forest with no isolated vertices. We know that every forest is bipartite; among all partitions $V(F) = A \cup B$ such that $E(F) = E(A, B)$, fix one such that $|A| \leq |B|$ and $|A|$ is as small as possible. Write $a = |A|$ and $b = |B|$. Further, let $B_1 \subseteq B$ be the set containing all leaves in B, that is, $B_1 = \{v \in B : d_F(v) = 1\}$, and $B_{\geq 2} = B \setminus B_1$.

We first claim that $|B_{\geq 2}| \leq a - 1$. For that, let F' be an arbitrary component of F, and let $v \in V(F') \cap A$ be an arbitrary vertex, which exists since F has no isolated vertices. Consider F' as a rooted tree with root v. Then each vertex in $V(F') \cap B_{\geq 2}$ must have a child in $V(F') \cap A$. Further, since F' is a tree, different vertices in $B_{\geq 2}$ have different children from A and the root v is not a child of any vertex, implying that $|V(F') \cap B_{\geq 2}| \leq |V(F') \cap A| - 1$. Summing over all components of F, we obtain the desired result. Note that this claim also implies B_1 is not empty.

Now, set $x = q(a-1)$, $y = q^{x+1}v(F)$, and $z = ybq$, and construct a graph G as follows. Let X and Y be a sets consisting of x and y vertices respectively; for each $w \in Y$, let Z_w be a set of bq vertices and $Z = \bigcup_{w \in Y} Z_w$. All of these sets are taken to be disjoint. We then add a complete bipartite graph between X and Y and, for each $w \in Y$, we connect w to all vertices in Z_w.

Observe that each vertex in Z has degree 1. We now claim that $G \to_q F$ but $G - Z \not\to_q F$. These two claims then imply the proposition, since any minimal q-Ramsey graph $G' \subseteq G$ must contain a vertex of degree one from Z.

We first argue that $G - Z \not\to_q F$. For this, we partition X arbitrarily into q sets $X_1, \ldots, X_q$, each of size $a - 1$, and assign color i to the edges between X_i and Y for all $i \in [q]$. We claim that this coloring is F-free. Indeed, each color class is a complete bipartite graph with one vertex class of size $a - 1$. From the choice of a we then know that there can be no monochromatic copy of F.

We now want to show that $G \to_q F$. Consider an arbitrary coloring φ of G. Write $X = \{t_1, \ldots, t_x\}$, and, for each $w \in Y$, let $\vec{c}_w = (\varphi(t_1), \ldots, \varphi(t_x))$ be the *color profile* of w. Observe that there are q^x possible color profiles, and so the pigeonhole principle implies that there is a subset $Y' \subseteq Y$ of size at least $y/q^x = qv(F)$ such that all elements in Y' have the same color profile $\vec{c} = (c_1, \ldots, c_x)$. If some color $j \in [q]$ appears at least a times in $\vec{c}$, then the vertices in Y' together with a vertices from the set $\{t_i : c_i = j\} \subseteq X$ form a monochromatic copy of the complete bipartite graph $K_{a,qv(F)}$ and in which we can easily embed F.

Thus, we may assume that $\vec{c}$ contains each color exactly $a - 1$ times. Now, for each $w \in Y'$, the set Z_w has size bq, so again using the pigeonhole principle we can find a subset $Z'_w \subseteq Z_w$ of size

b such that the star between w and Z'_w is monochromatic. Another application of the pigeonhole principle implies that there exists a subset $Y'' \subseteq Y'$ of size at least $|Y'|/q = v(F)$ and a color $k \in [q]$ such that, for all $w \in Y''$, the star between w and Z'_w has color k. Then we can find a monochromatic copy of F in color k in the following way. First, embed the vertices of A into Y'' arbitrarily. Then, for each $s \in A$, embed the vertices in $N_F(s) \cap B_1$ into the corresponding set Z'_w. Finally, since there is a complete bipartite graph between Y'' and $\{t_i : c_i = k\}$, all of whose edges have color k, and $|B_{\geq 2}| \leq a - 1$, we can embed the vertices in $B_{\geq 2}$ into $\{t_i : c_i = k\}$ arbitrarily. □

4.4.2 Constructing colored neighborhoods when $\mathbf{p \gg \log n/n}$

The goal of this section is to prove Theorem 1.1.9(b) and (c). For that we will utilize Proposition 4.2.3, so our task will be to construct appropriate packings satisfying the required conditions. We start with the sparse range, where $p \ll n^{-2/3}$.

Proof of Theorem 1.1.9(b). Let $q \geq 2$ be an integer, p satisfy $\frac{\log n}{n} \ll p \ll n^{-2/3}$, and $H \sim G(n, p)$. By Lemma 4.3.6 and Corollary 4.3.8, we know that w.h.p. H is well-behaved and the graph F induced by the neighborhood of the unique vertex of minimum degree u is empty. Then we can take $\Gamma_1, \ldots, \Gamma_q$ to all be the empty graph on $q(\delta(H) - 1) + 1$ vertices. This color pattern satisfies conditions (i) and (ii) of Proposition 4.2.3, so an application of Proposition 4.2.3 completes the proof. □

When $p \gg n^{-2/3}$, edges start to appear in the neighborhood graph F, so the packing problem becomes nontrivial. However, in the range $n^{-2/3} \ll p \ll n^{-1/2}$, the graph F is still simple in structure (by Corollaries 4.3.9 and 4.3.10, F is likely to be a very sparse forest). This fact will allow us to construct packings satisfying the required conditions and giving reasonably sharp bounds, as seen in Corollary 1.1.10. We give two constructions, proving the two lower bounds in Theorem 1.1.9(c); our first construction is geometric, while the second is random. The former construction yields a better bound than the latter when the components of F are all small, that is, $p \leq n^{-1/2-\varepsilon}$ for some $\varepsilon > 0$, while the latter construction gives a better bound when p gets closer to $n^{-1/2}$ and the components of F become larger.

Proof of Theorem 1.1.9(c), first bound. Let p satisfy $n^{-2/3} \ll p \ll n^{-1/2}$ and $H \sim G(n, p)$. Let $\varepsilon > 0$. By Lemma 4.3.6 and Corollaries 4.3.9 and 4.3.10, we know that w.h.p. H is well-behaved and that the graph F induced by the neighborhood of the unique vertex u of minimum degree is a forest with $o(np)$ edges. Let $T_1, \ldots, T_t$ be the components of F containing at least one edge. By Lemma 4.3.2(c), w.h.p. we have $\delta(H) = (1 \pm o(1))np$, implying that $\sum_{j=1}^{t} v(T_j) \leq 2 \sum_{j=1}^{t} e(T_j) \leq \varepsilon\delta(H)$. We condition on these properties from now on.

Let $2 \le q \le (1-5\varepsilon)\frac{\delta(H)}{\lambda(F)^2}$ be an integer. We will show that w.h.p. H is q-Ramsey simple. Let s be the largest prime number not exceeding $(1-\varepsilon)\frac{\delta(H)}{\lambda(F)}$. By the result of Baker, Harman, and Pintz [7], we know that that $s \ge (1-2\varepsilon)\frac{\delta(H)}{\lambda(F)}$. Consider the affine plane $\mathbb{F}_s^2$, which has point set $\mathbb{F}_s^2$ and line set consisting of all sets of the form $\{(x,y) \in \mathbb{F}_s^2 : mx + y = b\}$ for $m, b \in \mathbb{F}_s$ and all sets of the form $\{(x,y) \in \mathbb{F}_s^2 : x = c\}$ for $c \in \mathbb{F}_s$. It is not difficult to see that this affine plane satisfies the following:

- There are s^2 points.
- There are $s^2 + s$ lines.
- Each line contains exactly s points.
- Any two points are contained in a unique line.
- The set of lines can be partitioned into $s+1$ parallel classes, i.e., sets of pairwise disjoint lines, $C_1, \dots, C_{s+1}$, each containing s lines, where $C_j = \left\{\{(x,y) \in \mathbb{F}_s^2 : jx + y = b\} : b \in \mathbb{F}_s\right\}$ for all $j \in \mathbb{F}_s$ and $C_{s+1} = \left\{\{(x,y) \in \mathbb{F}_s^2 : x = c\} : c \in \mathbb{F}_s\right\}$.

We are ready to construct the color pattern $\Gamma_1, \dots, \Gamma_q$. Note that s satisfies $q \le s \le \delta(H)$ and $q(\delta(H)-1)+1 \le q\delta(H) \le (1-5\varepsilon)\frac{\delta(H)}{\lambda(F)^2}\delta(H) \le (1-2\varepsilon)^2\frac{\delta(H)^2}{\lambda(F)^2} \le s^2$. Let $V \subseteq \mathbb{F}_s^2$ be an arbitrary subset of size $q(\delta(H)-1)+1$, and set $V(\Gamma_i) = V$ for all $i \in [q]$. For each $i \in [q]$ and any distinct vertices $x, y \in V$, add an edge between x and y in Γ_i if and only if the unique line containing both x and y is contained in C_i.

It remains to verify that Γ satisfies conditions (i) and (ii) of Proposition 4.2.3. Note that, for each $i \in [q]$, the graph Γ_i consists of at most s disjoint cliques, each corresponding to (part of) a line in the affine plane $\mathbb{F}_s^2$ and thus containing at most s vertices. Thus, each component of Γ_i has maximum degree at most $s - 1 \le \delta(H) - 1$, which shows that condition (ii) is satisfied.

We now turn our attention to condition (i). Let $U \subseteq V$ be a subset of size $\delta(H)$ and $i \in [q]$ be any color. We will show that there exists a copy of F in $\Gamma_i[U]$. For this, we embed the nontrivial components $T_1, \dots, T_t$ of F into $\Gamma_i[U]$ one by one. Let $\ell \ge 1$ and suppose we have already embedded $T_1, \dots, T_{\ell-1}$; we now want to embed T_ℓ. Since $\sum_{j=1}^t v(T_j) \le \varepsilon\delta(H)$, we have used up at most $\varepsilon\delta(H)$ vertices so far. Let $U' \subseteq U$ be the set of unused vertices; we have $|U'| \ge (1-\varepsilon)|U| = (1-\varepsilon)\delta(H)$. We know that Γ_i consists of at most s disjoint cliques. Thus, by the pigeonhole principle, there exists a clique that intersects U' in a set T of at least $\frac{|U'|}{s} \ge \frac{(1-\varepsilon)\delta(H)\lambda(F)}{(1-\varepsilon)\delta(H)} = \lambda(F)$ vertices. Since $v(T_\ell) \le \lambda(F)$, we can embed T_ℓ into $\Gamma_i[T]$. Repeating the process until all of $T_1, \dots, T_t$ are embedded and then embedding the remaining vertices of F arbitrarily shows that $F \subseteq \Gamma_i[U]$, completing the verification of condition (i). Applying Proposition 4.2.3 then completes the proof. □

Proof of Theorem 1.1.9(c), second bound. Let p satisfy $n^{-2/3} \ll p \ll n^{-1/2}$ and $H \sim G(n,p)$. Let $\varepsilon > 0$. By Lemma 4.3.6 and Corollaries 4.3.9 and 4.3.10, we know that w.h.p. H is well-behaved and that the graph F induced by the neighborhood of the unique vertex u of minimum degree is a forest with $o(np)$ edges. Let $T_1, \dots, T_t$ be the components of F containing at least one edge. By Lemma 4.3.2(c), w.h.p. we have $\delta(H) = (1 \pm o(1))np$, implying that $\sum_{j=1}^{t} v(T_j) \leq \varepsilon\delta(H)$. By Corollary 4.3.10(a), we know that $\lambda(F) \leq \log n$. We condition on these properties from now on.

Let $q \leq \frac{\delta(H)}{80 \log n}$. We will show that w.h.p. H is q-Ramsey simple. Let $N = q(\delta(H) - 1) + 1$ and $V = [N]$. We now construct the required color pattern $\Gamma_1, \dots, \Gamma_q$. Begin by sampling a graph $\Gamma \sim G(N, \frac{1}{2})$. Then partition the edges of Γ randomly to create the graphs $\Gamma_1, \dots, \Gamma_q$, that is, for each edge $e \in \Gamma$, choose uniformly at random an index $i \in [q]$ and add e to Γ_i (all choices are made independently). Note that each Γ_i has the distribution $G(N, \frac{1}{2q})$. We claim that the color pattern $\Gamma_1, \dots, \Gamma_q$ satisfies the conditions of Proposition 4.2.3 with high probability. As before we begin with (ii). By Lemma 4.3.2(c), we know that, for every $i \in [q]$, the probability that $\Delta(\Gamma_i) > \frac{3}{2}\frac{N}{2q}$ is at most N^{-2}. Taking a union bound over all $i \in [q]$ shows that with high probability we have $\Delta(\Gamma_i) < \frac{3}{2}\frac{N}{2q} < \delta(H)$ for all $i \in [q]$.

It remains to verify condition (i). Let $U \subseteq V$ be a subset of size $\delta(H)$ and $i \in [q]$. We will show that, for every subset $U' \subseteq U$ satisfying $|U'| \geq \frac{1}{2}np$ and for any tree T on at most $\log n$ vertices, we can embed T in $\Gamma_i[U']$. This will then imply the claim, since $\sum_{j=1}^{t} v(T_j) \leq \varepsilon\delta(H)$ and thus we can embed the nontrivial components of F one by one and then embed the isolated vertices arbitrarily, as in the proof of the first bound above. Observe that $\frac{1}{2q} \gg \frac{\log N}{N}$ and $\frac{20 \log N}{1/(2q)} \leq 40 \log(q\delta(H)) \frac{np}{80 \log n} \leq \frac{1}{2}np$, since $q\delta(H) \leq (np)^2 \ll n$ by our assumption on p. So, applying Lemma 4.3.3(b) together with a union bound over the q colors, we know that with high probability $\Gamma_i[U']$ contains at least $\frac{1}{4}\left(\frac{1}{2}np\right)^2 \frac{1}{2q} > 2np \log n$ edges, implying that w.h.p. the average degree of $\Gamma_i[U']$ is at least $2 \log n$. Then removing all vertices of small degree from $\Gamma_i[U']$ results in a subgraph of minimum degree at least $\log n$, in which we can greedily embed the tree T. Thus, conditions (i) and (ii) are both satisfied, allowing us to apply Proposition 4.2.3 and conclude that H is q-Ramsey simple. □

4.5 Non-simplicity for $G(n,p)$

In this section we prove the upper bounds on $\tilde{q}(H)$ from Theorem 1.1.9(d) and (e). For one of the bounds in (d) we will need the following result due to Kogan [101].

Theorem 4.5.1 ([101])**.** *Let G be an n-vertex graph of average degree d and let $k \geq 0$ be an integer. Then there is a set U of at least $(k+1)n/(d+k+1)$ vertices such that $\Delta(G[U]) \leq k$.*

We are now ready to prove the bounds in (d). We remark that the proofs of these bounds do not actually use the fact that H is a random graph and are valid for any graph H with a unique vertex of minimum degree, provided that the neighborhood of that vertex is not an independent set.

Proof of Theorem 1.1.9(d). Let p satisfy $n^{-2/3} \ll p \ll 1$ and $H \sim G(n,p)$. By Lemma 4.3.6, we know that w.h.p. H has a unique vertex u of minimum degree. We condition on this property and write $F = H[N_H(u)]$. Let $q \geq 1$ and suppose that H is q-Ramsey simple. We will show that $q \leq (1+o(1))\frac{\delta(H)}{\Delta(F)}$ and $q \leq (1+o(1))\frac{\delta(H)^2}{2e(F)}$, which will then imply the desired result about $\widetilde{q}(H)$. By Proposition 4.2.1, there exists a color pattern $\Gamma_1,\dots,\Gamma_q$ on a vertex set V of size $N = q(\delta(H)-1)+1$ such that, for every $U \subseteq V$ of size $\delta(H)$ and every $i \in [q]$, there exists a copy of F in $\Gamma_i[U]$.

First we show that $q \leq (1+o(1))\frac{\delta(H)}{\Delta(F)}$. We know that $\bigcup_{i=1}^q \Gamma_i$ has at most $\binom{N}{2}$ edges and that the graphs Γ_i are all edge-disjoint, so there is some $i \in [q]$ such that Γ_i has at most $\frac{1}{q}\binom{N}{2}$ edges and hence average degree at most $\frac{2}{qN}\binom{N}{2} = \frac{N-1}{q} = \delta(H)-1$. By Theorem 4.5.1, Γ_i contains a set of $\frac{\Delta(F)N}{\delta(H)+\Delta(F)-1}$ vertices that induces a graph with maximum degree at most $\Delta(F)-1$. But, for every $U \subseteq V$ of size $\delta(H)$, the induced subgraph $\Gamma_i[U]$ contains a copy of F and thus a vertex of degree $\Delta(F)$, so we must have

$$\frac{\Delta(F)N}{\delta(H)+\Delta(F)-1} \leq \delta(H)-1.$$

Thus, using the definition of N and rearranging, we obtain $q \leq \frac{\delta(H)+\Delta(F)-1}{\Delta(F)} - \frac{1}{\delta(H)-1} \leq (1+o(1))\frac{\delta(H)}{\Delta(F)}$.

Next we show that $q \leq (1+o(1))\frac{\delta(H)^2}{2e(F)}$. Let $U \subseteq V(\Gamma)$ have size $\delta(H)$. As before, this subset induces at most $\binom{\delta(H)}{2}$ edges in $\bigcup_{i=1}^q \Gamma_i$, and thus there exists some $i \in [q]$ such that $\Gamma_i[U]$ contains at most $\frac{1}{q}\binom{\delta(H)}{2}$ edges. But we know that F is a subgraph of $\Gamma_i[U]$, so we must have $e(F) \geq \frac{1}{q}\binom{\delta(H)}{2}$, which rearranges to $q \leq \frac{1}{e(F)}\binom{\delta(H)}{2} = (1+o(1))\frac{\delta(H)^2}{2e(F)}$, as required. □

We end this section with a proof of Theorem 1.1.9(e). As p gets close to 1, $G(n,p)$ is no longer expected to have a unique vertex of minimum degree (see for example Theorem 3.9(ii) in [24]), which is why we will abstain from using Lemma 4.3.7 in this proof and will give a more direct argument instead.

Proof of Theorem 1.1.9(e). Let $p \gg \sqrt{\frac{\log n}{n}}$ and $H \sim G(n,p)$. By Lemma 4.1.1, it suffices to show that w.h.p. H is not 2-Ramsey simple.

By Lemma 4.3.5, we know that w.h.p. every edge of H belongs to a triangle. We condition on this property. Suppose now for a contradiction that H is 2-Ramsey simple. Let G be a minimal 2-Ramsey graph for H with a vertex w of degree $2\delta(H)-1$. By the minimality of G, we know that the graph $G-w$ has an H-free 2-coloring. Now let $v \in N_G(w)$ be arbitrary. We know that v has at most $2\delta(H)-2$ edges going to $N_G(w)\setminus\{v\}$, so by the pigeonhole principle, some color, say color 1, appears at most $\delta(H)-1$ times on those edges. Thus there is a subset $W \subseteq N_G(w)\setminus\{v\}$ of size $\delta(H)-1$ containing all vertices v' connected to v by edges of color 1. Now we extend the coloring to the edges incident to w by assigning color 2 to all edges between w and W and color 1 to the remaining edges. Since $G \to_2 H$, we know that there exists a monochromatic copy H' of H. Moreover, since the coloring of $G - w$ is H-free, it follows that H' must contain w. But w is only incident to $\delta(H) - 1$ edges of color 2, so H' must have color 1 and must in particular contain the edge wv. However, v has no edge of color 1 going to any vertex in $N_G(w) \setminus W$, so the edge wv is not contained in a color-1 monochromatic triangle and thus cannot be contained in a monochromatic copy of H, a contradiction. □

4.6 Further results

4.6.1 Abundance

In this section, we prove Proposition 1.1.11.

Proposition 1.1.11. *Let $p = p(n) \in (0,1)$ and $H \sim G(n,p)$. Let $u \in V(H)$ be the smallest (with respect to the natural vertex ordering) vertex of degree $\delta(H)$ and let $F = H[N(u)]$ be the subgraph of H induced by the neighborhood of u. Denote by $\lambda(F)$ the order of the largest connected component in F. Then w.h.p. the following is true:*

(a) H is s_q-abundant for all $q \geq 2$ *if* $\frac{\log n}{n} \ll p \ll n^{-\frac{2}{3}}$.

(b) H is s_q-abundant for any $2 \leq q \leq (1+o(1)) \max\left\{\frac{\delta(H)}{\lambda(F)^2}, \frac{\delta(H)}{80 \log n}\right\}$ *if* $n^{-\frac{2}{3}} \ll p \ll n^{-\frac{1}{2}}$.

This result is an immediate consequence of Proposition 4.6.1 below. Note that Proposition 4.6.1 again applies to any well-behaved graph and does not use the fact that H is random. Thus, in particular, by the proof of Proposition 4.2.4, we actually know that all of those graphs are not just q-Ramsey simple but also s_q-abundant.

Proposition 4.6.1. *Let $q \geq 2$ and let H be a well-behaved n-vertex graph. If there is a q-color pattern graph $\Gamma_1, \ldots, \Gamma_q$ on a vertex set V of size $q(\delta(H) - 1) + 1$ satisfying the conditions of Proposition 4.2.3, and if either $e(\Gamma_i) = 0$ for all $i \in [q]$ or $n > q(\delta(H) - 1) + 2$, then H is s_q-abundant.*

Before we prove Proposition 4.6.1, we explain how to derive Proposition 1.1.11 from it.

Proof of Proposition 1.1.11. Recall that, if $\frac{\log n}{n} \ll p \ll 1$, then by Lemma 4.3.6 $H \sim G(n,p)$ is well-behaved with high probability. To prove parts (b) and (c) of Theorem 1.1.9, we found suitable color patterns satisfying the conditions of Proposition 4.2.3.

First suppose we are in case (a). To prove part (b) of Theorem 1.1.9, we took the color pattern in which each Γ_i is an empty graph, which also satisfies the conditions of Proposition 4.6.1, so the claim follows.

Now consider case (b). By Lemma 4.3.2(c), we know that with high probability $\delta(H) = (1+o(1))np$. By assumption, we have $q \le \delta(H)$, and so $q(\delta(H)-1)+2 \le (1+o(1))(np)^2 \ll n$. So we can again apply Proposition 4.6.1 to complete the argument. $\square$

It remains to prove Proposition 4.6.1. We will do so by using the following result, which is a corollary of Theorem 3.1 from [29] and gives a sufficient condition for the existence of minimal q-Ramsey graphs with several vertices of a given degree.

Theorem 4.6.2. *Let $q \ge 2$ be an integer and H be 3-connected. Suppose there exists a minimal q-Ramsey graph G' for H, together with a vertex $v_0 \in V(G')$ and an edge $e \in E(G')$ such that v_0 end e do not share a copy of H in G'. Then, for any $k \ge 1$, there exists a minimal q-Ramsey graph for H that has at least k vertices of degree $d_{G'}(v_0)$.*

Proof of Proposition 4.6.1. Our goal is to show that the q-Ramsey graph G we built in the proof of Proposition 4.2.3 admits a subgraph $G' \subseteq G$ satisfying the conditions of Theorem 4.6.2.

Consider the graph G constructed in the proof of Proposition 4.2.3. We know that $G - w \not\to_q H$, so any subgraph $G' \subseteq G$ that is minimal q-Ramsey for H will contain the vertex w. Let M be the matching $\{e_1, \dots, e_q\}$; we will show next that $G - M \not\to_q H$, implying that for any minimal q-Ramsey graph for H contained in G must contain an edge from M. Once we have shown that, we can apply Theorem 4.6.2, taking $G' \subseteq G$ to be any minimal q-Ramsey graph, v_0 to be the vertex w, and e to be any edge from M contained in G'. Since the edge e is only connected to Γ^+ by signal senders in G and the distance between the signal edges of each signal sender is at least $n+1$, we know that the distance between e and v_0 is at least $n+1$, implying that they cannot share a copy of H.

It thus remains to prove that $G - M \not\to_q H$. For that, consider the following q-coloring of $G - M$: assign color 1 to all edges that are either contained in Γ or incident to w and color 2 to all edges incident to a vertex of $V(R_{U,i})$ for some $U \subseteq V(\Gamma)$ and $i \in [q]$; then extend the coloring to the (partial) signal senders so that each signal sender has an H-free q-coloring, which is possible since each signal sender is missing at least one of its signal edges.

We now show that there is no monochromatic copy of H under this coloring. By Corollary 2.2.3, it suffices to show that there is no monochromatic copy of H in the graph Γ^+. Suppose there exists a monochromatic copy H' of H in Γ^+. If H' has color 1, then we know that $V(H') \subseteq V(\Gamma) \cup \{w\}$. If $e(\Gamma_i) = 0$ for all $i \in [q]$, then $e(\Gamma) = 0$ and thus the edges of color 1 form a star centered at w, which cannot contain H' (which is 3-connected by (W3)). If instead $v(H) = n > q(\delta(H) - 1) + 2 = v(\Gamma) + 1$, then again H' cannot be contained in the subgraph $\Gamma + w$.

So we may assume that H' has color 2. But arguing in a similar way as in the proof of Proposition 4.2.3, we can show that H' cannot contain vertices from two different subgraphs R_{U_1,i_1} and

R_{U_2,i_2}. Hence $V(H') \subseteq V(R_{U,i}) \cup V(\Gamma) \cup \{w\}$ for some $U \subseteq V(\Gamma)$ and $i \in [q]$; in particular, since H' has color 2, it must contain a vertex in $V(R_{U,i}) \cup U$. But $|V(R_{U,i}) \cup U| = n - 1$, so H' must also use a vertex in $(V(\Gamma) - U) \cup \{w\}$, but all edges between $(V(\Gamma) - U) \cup \{w\}$ and $V(R_{U,i}) \cup U$ have color 1, a contradiction. □

4.6.2 An asymmetric result

In this section, we extend some of the theory developed in this chapter to the asymmetric setting. To the best of our knowledge, determiners for asymmetric tuples of 3-connected graphs are only known to exist in the two-color setting, which is why we restrict our attention to the case $q = 2$ in this section.

Proposition 4.6.3. *Let H_1 and H_2 be well-behaved graphs on n vertices. For $i \in [2]$, let u_i be the unique vertex of minimum degree in H_i and $F_i = H_i[N_{H_i}(u_i)]$ be the subgraph induced by the neighborhood of u_i in H_i. Suppose there exists a color pattern Γ_1, Γ_2 on a set V of $\delta(H_1) + \delta(H_2) - 1$ vertices satisfying the following:*

(i) For every $i \in [2]$ and for every subset $U \subseteq V$ of $\delta(H_i)$ vertices, there exists a copy $F_{U,i}$ of F_i in $\Gamma_i[U]$.
(ii) For each $i \in [2]$, we have $\Delta(\Gamma_i) \leq \delta(H_i) - 1$.

Then $s_2(H_1, H_2) = \delta(H_1) + \delta(H_2) - 1$.

Proof. The lower bound on $s_2(H_1, H_2)$ is given by (1.1.2). The proof of the upper bound is very similar to that of Proposition 4.2.3, so we give only a sketch here, highlighting the differences. As before, we will construct a graph G satisfying the following properties:

(1) $G \rightarrow_2 (H_1, H_2)$.
(2) G contains a vertex w of degree $\delta(H_1) + \delta(H_2) - 1$.
(3) $G - w \not\rightarrow_2 (H_1, H_2)$.

We now explain how to construct the graph G. We may assume $H_1 \not\cong H_2$, as otherwise we can use Proposition 4.2.3. For $i \in [2]$, write $R_i = H_i - (N_{H_i}(u_i) \cup \{u_i\})$. We start with the graph $\Gamma = \Gamma_1 \cup \Gamma_2$; let φ be a 2-coloring of Γ assigning color i to the edges in Γ_i for all $i \in [2]$. Then, for all $i \in [2]$ and for each subset $U \subseteq V(\Gamma)$ of size $\delta(H_i)$, there exists a copy $F_{U,i}$ of the graph F_i in $\Gamma_i[U]$. We then add a copy $R_{U,i}$ of R_i on a disjoint set of vertices and add edges between $V(F_{U,i})$ and $V(R_{U,i})$ in such a way that $F_{U,i} \cup R_{U,i} \cup E(V(F_{U,i}), V(R_{U,i})) \cong H_i - u_i$. We denote the resulting graph by Γ^+. We extend the 2-coloring φ to a 2-coloring of Γ^+ as follows: for every $i \in [2]$ and every subset $U \subseteq V(\Gamma)$ of size $\delta(H_i)$, assign color i to all edges incident to a vertex of $R_{U,i}$. Let φ^+ denote the resulting 2-coloring of Γ^+.

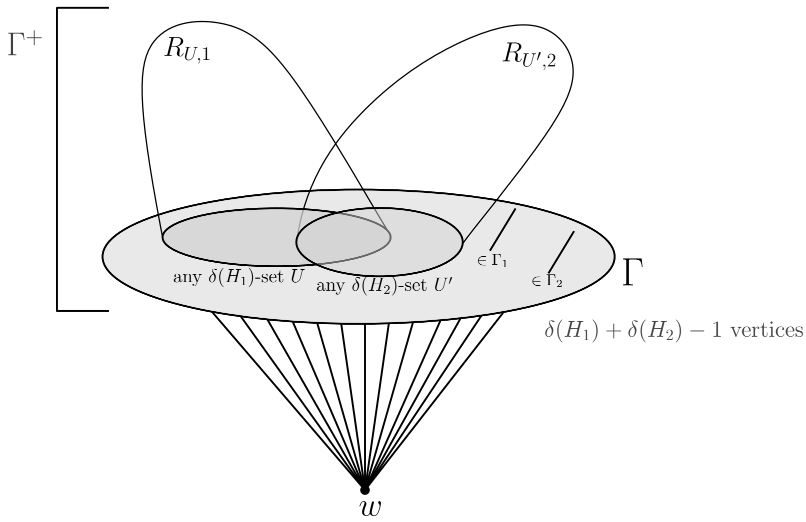

FIGURE 4.2: The graph G in the proof of Proposition 4.6.3 (determiners not shown).

We now construct the final graph G from Γ^+. For this, let D_1 and D_2 be a safe 1-determiner and 2-determiner for (H_1, H_2), respectively, which exist by Lemma 2.3.2, since H_1 and H_2 are both 3-connected and non-isomorphic. For every edge f of Γ^+, attach a copy of $D_{\varphi^+(f)}$ on a new set of vertices to f. Finally, add a new vertex w and connect it to all vertices of Γ to obtain the final graph G. The construction is illustrated in Figure 4.2. It remains to verify that G satisfies the desired properties. As before, property (2) is easy to check.

We now turn our attention to property (1). Suppose for a contradiction that there exists an (H_1, H_2)-free 2-coloring of G. This coloring induces an (H_1, H_2)-free coloring on each copy of D_1 and D_2, so by property (D2) of D_1 and D_2, every edge f of Γ^+ receives color $\varphi^+(f)$. Now, considering the vertex w and noting that there exists an $i \in [2]$ such that color i appears at least $\delta(H_i)$ times on the edges incident to w and arguing as in the proof of Proposition 4.2.3 using property (i) of the color pattern Γ_1, Γ_2 proves the claim.

It remains to prove property (3). For this, we provide an (H_1, H_2)-free 2-coloring of $G - w$. We start by coloring Γ^+ according to the coloring φ^+. Observe that, for each $i \in [2]$, the signal edge of every copy of D_i receives color i. Thus, by property (D3) and the safeness of the copies of D_1 and D_2, we can extend the partial coloring to the determiners so that each determiner receives an (H_1, H_2)-safe 2-coloring. We now claim that this is an (H_1, H_2)-free coloring of $G - w$. Suppose on the contrary that there exists a color $i \in [2]$ such that there is a monochromatic copy H' of H_i. Since D_1 and D_2 were chosen to be safe and each determiner has an (H_1, H_2)-free coloring, we know that H' must be fully contained in Γ^+. Notice that H' cannot contain edges from any copy of R_j for $j \neq i$. Arguing as before, we can show that H' can contain vertices from only one copy $R_{U,i}$ of R_i and that all the remaining vertices need to be contained in U. But $v(R_i) + |U| = v(H_i) - 1$, a contradiction. ◻

We now present an application of Proposition 4.6.3.

Proposition 4.6.4. *Let $p_1 = p_1(n) \in (0,1)$ and $p_2 = p_2(n) \in (0,1)$ be such that $\frac{\log n}{n} \ll p_1 \le p_2 \ll n^{-1/2}$, and let $H_1 \sim G(n,p_1)$ and $H_2 \sim G(n,p_2)$. Then w.h.p. $s_2(H_1,H_2) = \delta(H_1)+\delta(H_2)-1$.*

Proof. By Lemma 4.3.6 and Lemma 4.3.2(c), we know that w.h.p. H_1 and H_2 are both well-behaved and $\delta(H_1) = (1 \pm o(1))np_1$ and $\delta(H_2) = (1 \pm o(1))np_2$. We condition on these properties and let u_i denote the unique vertex of minimum degree in H_i and $F_i = H_i[N_{H_i}(u_i)]$ for each $i \in [2]$. We may further assume that $H_1 \ncong H_2$.

First consider the case where $p_1 \le p_2 \ll n^{-2/3}$. In this case, by Corollary 4.3.8, each of F_1 and F_2 is empty with high probability, and taking Γ_1, Γ_2 to be empty graphs on $\delta(H_1)+\delta(H_2)-1$ vertices satisfies properties (i) and (ii) from Proposition 4.6.3.

Suppose next that $p_1 \ll n^{-2/3}$ and $p_2 = \Omega(n^{-2/3})$. Again by Corollary 4.3.8 and Corollaries 4.3.9 and 4.3.10 we know that w.h.p. F_1 is empty and F_2 is a forest with $o(np_2)$ edges and hence $o(np_2)$ vertices in nontrivial components. We condition on all of these properties. Now, let V be a vertex set of size $\delta(H_1)+\delta(H_2)-1$; take Γ_1 to be the empty graph on V and Γ_2 to be the graph on V consisting of two cliques of size $\left\lfloor \frac{\delta(H_1)+\delta(H_2)-1}{2} \right\rfloor$ and $\left\lceil \frac{\delta(H_1)+\delta(H_2)-1}{2} \right\rceil$. We now show that this choice of Γ_1, Γ_2 satisfies properties (i) and (ii) in Proposition 4.6.3. Property (ii) is not difficult to see, as $\Delta(\Gamma_1) = 0 < \delta(H_1)$ and $\Delta(\Gamma_1) \le \left\lceil \frac{\delta(H_1)+\delta(H_2)-1}{2} \right\rceil \le \frac{1.5\delta(H_2)}{2} < \delta(H_2)$, where in the second to last step we used the fact that $\delta(H_i) = (1 \pm o(1))np_i$ and thus $\delta(H_1) \ll \delta(H_2)$. We now verify (i). The property clearly holds for $i = 1$, so assume $i = 2$ and let $U \subseteq V$ with $|U| = \delta(H_2)$. This set U intersects one of the two connected components of Γ_2 in at least $\frac{\delta(H_2)}{2} \ge \frac{1}{4}np_2$ vertices; these vertices form a clique, in which we can embed the nontrivial components of F_2. We can then embed the isolated vertices of F_2 arbitrarily in $\Gamma_2[U]$. Thus (i) also holds, so Proposition 4.6.3 implies the claim.

Finally, we assume that $p_1, p_2 = \Omega(n^{-2/3})$. By Corollaries 4.3.9 and 4.3.10 we know that with high probability, for each $i \in [2]$, the graph F_i is a forest with at most $o(np_i)$ edges, and hence $o(np_i)$ vertices in nontrivial components, and that the largest component $\lambda(F_i)$ of F_i has size at most $\frac{1}{2}\log n$. We condition on all of these properties.

Now let $s = \left\lfloor \frac{\delta(H_1)}{2\log n} \right\rfloor$. Then $s \le \frac{\delta(H_1)}{2\log n} \le \frac{np_1}{\log n} \le \frac{np_2}{2} \le \delta(H_2)$. Set $t = \min\left\{ \left\lfloor \frac{\delta(H_2)}{2\log n} \right\rfloor, \delta(H_1) \right\}$. Let J be the graph with vertex set $V(J) = \{(i,j) : i \in [s], j \in [t]\}$, in which two distinct vertices (i,j) and (i',j') are adjacent if and only if $i = i'$ or $j = j'$. Let $J_1 \subseteq J$ consist of all edges of the form $\{(i,j),(i,j')\}$ for $j \ne j'$ and $J_2 \subseteq J$ consist of all edges of the form $\{(i,j),(i',j)\}$ for $i \ne i'$. Note that J_1 consists of s cliques, each on t vertices, and J_2 consists of t cliques, each on s vertices. As a result, $\Delta(J_1) = t - 1 \le \delta(H_1) - 1$ and $\Delta(J_2) = s - 1 \le \delta(H_2) - 1$. Now, consider a subset $U \subseteq V(J)$ of size $\delta(H_1)$ and a subset $U' \subseteq U$ of size at least $\frac{\delta(H_1)}{2}$.

By the pigeonhole principle, there exists some $\ell \in [s]$ such that the set U' intersects the set $\{(\ell, j) : j \in [t]\}$, forming a t-clique in J_1, in at least $\frac{\delta(H_1)}{2s} \geq \log n$ vertices. Thus, as in the proof of Theorem 1.1.9(c), we can embed the nontrivial tree components of F_1 one at a time in $J_1[U']$ and then embed the remaining vertices of F_1 arbitrarily to obtain a copy of F_1 in $J_1[U]$. Using a similar argument, we can show that, if $U \subseteq V(J)$ of size $\delta(H_2)$, then there is a copy of F_2 inside $J_2[U]$.

Now, we let V be an arbitrary subset of $V(J)$ of size $\delta(H_1) + \delta(H_2) - 1$ and $\Gamma_1 = J_1[V]$ and $\Gamma_2 = J_2[V]$. This is possible since $st \geq \delta(H_1) + \delta(H_2)$. Indeed, note that $s \geq \frac{\delta(H_1)}{4 \log n} \geq \frac{np_1}{8 \log n}$ and $t = \min\left\{\left\lfloor \frac{\delta(H_2)}{2 \log n} \right\rfloor, \delta(H_1)\right\} \geq \min\left\{\frac{np_2}{8 \log n}, \frac{np_1}{2}\right\} \geq \frac{np_1}{8 \log n}$. Then $st \geq \frac{np_1}{8 \log n} \frac{np_1}{8 \log n} = n \frac{np_1^2}{64 \log^2 n} \geq 4np_2 \geq \delta(H_1) + \delta(H_2)$, since $\frac{np_1^2}{\log^2 n} \gg \frac{n^{-1/3}}{\log^2 n} \gg n^{-1/2} \gg p_2$. Then, by the above argument, the color pattern Γ_1, Γ_2 satisfies properties (i) and (ii) from Proposition 4.6.3. □

4.7 Concluding remarks and open problems

We built upon the work of Grinshpun [83] and studied the q-Ramsey simplicity of $H \sim G(n, p)$ for a wide range of values of p and q. We encountered three different types of behavior: for very sparse ranges, i.e., when $p \ll \frac{1}{n}$ or $\frac{\log n}{n} \ll p \ll n^{-\frac{2}{3}}$, we showed that w.h.p. H is q-Ramsey simple for every possible number of colors q; for much denser ranges, i.e., when $p \gg \left(\frac{\log n}{n}\right)^{\frac{1}{2}}$, w.h.p. we do not have Ramsey simplicity even when $q = 2$; in between these ranges, when $n^{-\frac{2}{3}} \ll p \ll n^{-\frac{1}{2}}$, w.h.p. there exists a finite threshold value $\tilde{q}(H) \geq 2$ on the number of colors q such that H is q-Ramsey simple if and only if $q \leq \tilde{q}(H)$. We determined this threshold up to a constant or, when $p = n^{-\frac{1}{2} - o(1)}$, polylogarithmic factor. Several natural questions remain open.

First, our main result does not provide any information on the Ramsey simplicity of $G(n, p)$ when p is between $\frac{1}{n}$ and $\frac{\log n}{n}$.

Question 1. What is $\tilde{q}(H)$ when $H \sim G(n, p)$ for $\Omega\left(\frac{1}{n}\right) = p = O\left(\frac{\log n}{n}\right)$? In particular, is H likely to be 2-Ramsey simple in this case?

In the range $p \gg \frac{\log n}{n}$ our simplicity proofs rely heavily on the fact that $H \sim G(n, p)$ is almost surely 3-connected, implying the existence of signal senders for H, which in turn allows us to deduce a fairly general recipe for constructing suitable Ramsey graphs in Proposition 4.2.3. When $p \ll \frac{1}{n}$, we know that $H \sim G(n, p)$ is w.h.p. a forest, and simplicity follows from the construction of Szabó, Zumstein, and Zürcher [143], which works for certain bipartite graphs. When $\frac{1}{n} \ll p \ll \frac{\log n}{n}$, however, a random graph $H \sim G(n, p)$ becomes more complex: it is well-known (see for example Theorems 4.1 and 5.3 in [75]) that in that this range H is with high probability not bipartite (because it contains triangles) and not connected. As a result, resolving the above question will likely require new ideas.

Second, in the range $\Omega\left(n^{-\frac{1}{2}}\right) = p = O\left(\left(\frac{\log n}{n}\right)^{\frac{1}{2}}\right)$, we proved that $\tilde{q}(H) = O(p^{-1})$, which shows that the threshold value here is of smaller order than when $p = n^{-\frac{1}{2}-o(1)}$, as demonstrated in Corollary 1.1.10. However, we did not provide any nontrivial lower bounds, and we wonder if that is because the simplicity threshold in that range is 1.

Question 2. Is it true that H is w.h.p. not 2-Ramsey simple when $H \sim G(n,p)$ for $\Omega\left(n^{-\frac{1}{2}}\right) = p = O\left(\left(\frac{\log n}{n}\right)^{\frac{1}{2}}\right)$?

In this case, signal senders for H do exist, but the neighborhood of the minimum degree vertex becomes more complex than just a forest, making it more difficult to construct a color pattern satisfying the conditions of Proposition 4.2.3. On the other hand, the argument used to prove Theorem 1.1.9(e) relies on the fact that the neighborhood graph contains no isolated vertices; it is well-known (see for example Lemma 1.11 in [75]) that if $p \ll \frac{\log s}{s}$ then a random graph $G_s \sim G(s,p)$ almost surely contains an isolated vertex, and Lemma 4.3.7 implies that the same is true for the neighborhood graph. The presence of isolated vertices makes it likely that a more delicate argument than the one used in part (e) would be needed to show non-simplicity for smaller q.

The bounds on $\tilde{q}(H)$ presented in cases (a) and (c) of Corollary 1.1.10 are already quite close, but it would be interesting to close the remaining gaps.

Question 3. Let $H \sim G(n,p)$ with $n^{-2/3} \ll p \ll n^{-1/2}$. What are the asymptotics of $\tilde{q}(H)$?

In this range, as we have seen in Section 4.4, the question about q-Ramsey simplicity is tightly linked to the problem of finding a q-color pattern $\Gamma_1, \dots, \Gamma_q$ on a set V of $q(\delta(H)-1)+1$ vertices such that the following holds: for every set $U \subseteq V$ of $\delta(H)$ vertices and for every color $i \in [q]$, there exists a copy of $F = H[N(u)]$ in $\Gamma_i[U]$. The proofs of our lower bounds in Section 4.4 are obtained by finding such a color pattern (with additional properties as given in Proposition 4.2.3) through explicit constructions or probabilistic arguments. In order to prove that w.h.p. H is not q-Ramsey simple, it would suffice to prove that such a color pattern does not exist, that is, every q-color pattern $\Gamma_1, \dots, \Gamma_q$ on $q(\delta(H) - 1) + 1$ vertices contains *at least* one vertex-subset U of size $\delta(H)$ such that some $\Gamma_i[U]$ does not contain a copy of F. Note that in the proof of our second bound in part (d) of Theorem 1.1.9 we obtain such a result by a simple counting argument which guarantees that we cannot pack q copies of F into any graph on $\delta(H)$ vertices. Related to this argument, it seems challenging to determine how many copies of a given random graph can be packed into a complete graph, leading us to suggest the following question.

Question 4. Let $H \sim G(n,p)$ with $0 < p < 1$. How many copies of H can be packed into K_n?

In the densest range, that is, when $p \gg \left(\frac{\log n}{n}\right)^{\frac{1}{2}}$, we know that $H \sim G(n,p)$ is w.h.p. not q-Ramsey simple for any $q \geq 2$. We wonder, however, what the behavior of $s_q(H)$ in this case is; in particular, it would be interesting to determine whether $s_q(H)$ is still typically close to the easy lower bound $q(\delta(H)-1)+1$. Note that the answer is no if $p=1$ and $q=2$, since $s_2(K_n)=(n-1)^2$, as shown by Burr, Erdős, and Lovász [40]. However, when $\left(\frac{\log n}{n}\right)^{\frac{1}{2}} \ll p \ll 1$, we do not know of any bounds other than the general ones mentioned in the introduction. In particular, we propose the following problem, similar to one posed by Grinshpun, Raina, and Sengupta [82].

Question 5. How large is $s_2(H)$ for $H \sim G(n, \frac{1}{2})$ w.h.p.?

In Section 4.6.2 we showed that, if $\frac{\log n}{n} \ll p_1 \leq p_2 \ll n^{-1/2}$ and $H_1 \sim G(n,p_1)$ and $H_2 \sim G(n,p_2)$, the trivial lower bound from (1.1.2) is almost surely tight for the pair (H_1,H_2). The argument used to prove Proposition 4.4.1 can also be extended to the asymmetric setting, yielding a similar result in the case $p_1, p_2 \ll n^{-1}$; similarly, the argument used to prove Theorem 1.1.9(e) extends to the asymmetric setting. The following questions remain.

Question 6. Let $H_1 \sim G(n,p_1)$ and $H_2 \sim G(n,p_2)$ with $p_1 \ll n^{-1}$ and $\frac{\log n}{n} \ll p_2 \ll n^{-1/2}$. Is the pair (H_1,H_2) w.h.p. 2-Ramsey simple? What happens if one of the graphs comes from the dense range?

Finally, let us emphasize that there has been little study of (minimal) Ramsey graphs for $G(n,p)$. The only results we are aware of concern the Ramsey number of $G(n,p)$, as mentioned in the introduction. Hence, as a more general direction for future research, it would be interesting to explore other aspects of the Ramsey behavior of $G(n,p)$ as the target graph.

Chapter 5

Minimum degrees of minimal Ramsey graphs in the asymmetric setting

The results presented in this chapter are joint work with Anurag Bishnoi, Dennis Clemens, Pranshu Gupta, Thomas Lesgourgues, and Anita Liebenau; the text is adapted, with small modifications, from [16] (https://doi.org/10.1137/21M1444953).

5.1 Two-color cases

Throughout this section the number of colors q is fixed to be two; as usual in this case, for convenience we will refer to color 1 and color 2 as red and blue, respectively. Recall the general bounds from (1.1.2), which in the two-color setting give the following:

$$\delta(H_1) + \delta(H_2) - 2 < s_q(H_1, H_2) \leq r_q(H_1, H_2) - 1.$$

In this section, we determine $s_2(T_\ell, K_t)$, $s_2(C_\ell, C_k)$, and $s_2(C_\ell, K_t)$. We prove that the lower bound in (1.1.2) is tight for $s_2(T_\ell, K_t)$ and $s(C_\ell, C_k)$, but not for $s(C_\ell, K_t)$. In the latter two cases, we make use of the gadgets introduced in Chapter 2. We begin with the case of one clique and one tree. Our proof relies on a result of Burr, Erdős, Faudree, Rousseau, and Schelp [37], stating that there is a unique (T_ℓ, K_t)-free red/blue-coloring of $K_{(\ell-1)(t-1)}$.

Proposition 1.1.12. *For all integers $\ell \geq 2$ and $t \geq 3$ and any tree T_ℓ on ℓ vertices, we have $s_2(T_\ell, K_t) = t - 1$.*

Proof of Proposition 1.1.12. Let $\ell \geq 2$ and $t \geq 3$. First note that the inequality $s_2(T_\ell, K_t) > t-2$ follows directly from the general lower bound in (1.1.2). For the upper bound, we construct a graph G of minimum degree $t-1$ as follows. Let F be a copy of K_t, and let v be an arbitrary vertex of F. For each vertex u of $F-v$, create a copy H_u of $K_{(\ell-1)(t-1)}$ on a new set of vertices and identify u with an arbitrary vertex of H_u. Note that $d_G(v) = t-1$.

We claim that $G \rightarrow_2 (T_\ell, K_t)$ but $G - v \not\rightarrow_2 (T_\ell, K_t)$. For the former, suppose for a contradiction that φ is a (T_ℓ, K_t)-free red/blue-coloring of G. Then φ induces a (T_ℓ, K_t)-free coloring on $H_u \cong K_{(\ell-1)(t-1)}$ for every $u \in V(F-v)$. By [37, Lemma 9], there is a unique (T_ℓ, K_t)-free red/blue-coloring $\widetilde{\varphi}$ of $K_{(\ell-1)(t-1)}$, in which the red subgraph of $K_{(\ell-1)(t-1)}$ is a collection of $(t-1)$ vertex-disjoint cliques, each of size $(\ell-1)$. In particular, in the coloring φ, every vertex $u \in V(F-v)$ is incident to a red copy of $K_{\ell-1}$ in H_u. Since there is no red copy of T_ℓ, every edge of F must be blue, creating a monochromatic blue copy of K_t, a contradiction.

For the second claim, we color the edges of $F-v$ blue and use the (T_ℓ, K_t)-free coloring $\widetilde{\varphi}$ for every H_u. It is easy to see that this red/blue-coloring of $G-v$ is (T_ℓ, K_t)-free. Thus, any subgraph G' of G that is minimal 2-Ramsey for (T_ℓ, K_t) must contain v. This proves that $s_2(T_\ell, K_t) \leq d_G(v) = t-1$. □

We now turn our attention to pairs of graphs involving cycles and prove Theorem 1.1.13.

Theorem 1.1.13. *For all integers $t \geq 3$ and $k, \ell \geq 4$, we have*

(i) $s_2(C_\ell, C_k) = 3$.

(ii) $s_2(C_\ell, K_t) = 2(t-1)$.

We first consider pairs of distinct cycles. Let $k, \ell \geq 4$ be integers such that $k \neq \ell$. It follows from (1.1.2) that $s_2(C_\ell, C_k) > 2$. We show a matching upper bound in Proposition 5.1.1 below. We use the existence of safe determiners given by Lemma 2.3.3(i) to exhibit a minimal 2-Ramsey graph for (C_ℓ, C_k) with minimum degree three when $\ell > k$. Theorem 1.1.13(i) then follows by symmetry, since $s_2(C_k, C_\ell) = s_2(C_\ell, C_k)$.

Proposition 5.1.1. *For any $4 \leq k < \ell$, we have*

$$s_2(C_\ell, C_k) \leq 3.$$

Proof. We construct an appropriate minimal 2-Ramsey graph. We start with an empty graph on three vertices, x, y, and z, and between any pair of these vertices we add two paths, one of length $k-2$ and one of length $\ell-2$, so that all six paths are internally vertex-disjoint. Let D_r and D_b be a safe red- and blue-determiner for (C_ℓ, C_k), respectively, as guaranteed by Lemma 2.3.3(i). We attach a copy of D_r to every edge contained in one of the paths of length $\ell-2$ between x, y, and z and a copy of D_b to every edge contained in one of the paths of length $k-2$. Finally, we add a new vertex v adjacent to x, y, and z, and call the resulting graph G. The construction is illustrated in Figure 5.1 for the case $\ell = 8$ and $k = 6$, showing only the signal edges of the determiners and the edges incident to v. We will now show that $G \to_2 (C_\ell, C_k)$ but $G - v \not\to_2 (C_\ell, C_k)$, implying that any subgraph G' of G that is minimal 2-Ramsey for (C_ℓ, C_k) has to contain v, which in turn proves the proposition.

Consider an arbitrary red/blue-coloring of G. If any copy of D_r or D_b contains a red copy of C_ℓ or a blue copy of C_k, we are done. Otherwise, by property (D2) of D_r and D_b, the paths of length $\ell - 2$ between the vertices x, y, and z must be red and the paths of length $k - 2$ between those vertices must be blue. By the pigeonhole principle, two of the edges incident to v must have the same color; these two edges together with the corresponding red $(\ell - 2)$-path or blue $(k - 2)$-path then form a red copy of C_ℓ or a blue copy of C_k.

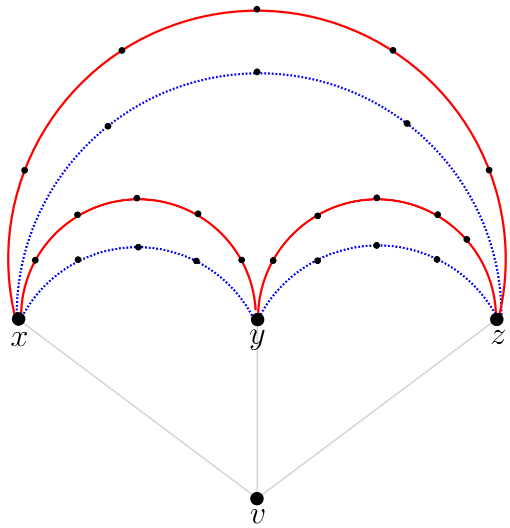

FIGURE 5.1: The graph G in the proof of Proposition 5.1.1 for $\ell = 8$ and $k = 6$ (thick solid/red lines represent edges to which a copy of D_r is attached, dotted/blue lines represent edges to which a copy of D_b is attached).

For the second claim, consider $G - v$ and color each path of length $\ell - 2$ between the vertices x, y, and z red and each path of length $k - 2$ between those vertices blue. Since $k, \ell > 3$, it is easy to see that this partial coloring of $G - v$ is (C_ℓ, C_k)-free. By property (D3) and the safeness of the copies of D_r and D_b, we can extend this coloring to the copies of D_r and D_b so that each determiner has a (C_ℓ, C_k)-safe coloring. By the definition of safeness, this is a (C_ℓ, C_k)-free coloring of $G - v$. □

Note that the construction in Proposition 5.1.1 requires that $k > 3$. The case $k = 3$ is covered by our next construction, dealing with pairs of the form (C_ℓ, K_t). We begin by showing an upper bound on $s_2(C_\ell, K_t)$ in Proposition 5.1.2 and then show a matching lower bound in Proposition 5.1.4. Theorem 1.1.13(ii) then follows immediately from these two propositions. The idea behind the upper bound construction is very similar to the previous one and again relies on the existence of safe determiners, guaranteed by Lemma 2.3.3(ii).

Proposition 5.1.2. *For any integers $\ell \geq 4$ and $t \geq 3$, we have*

$$s_2(C_\ell, K_t) \leq 2(t-1).$$

Proof. To construct a suitable 2-Ramsey graph, we let T be the complete $(t-1)$-partite graph where each vertex class has size two, that is, T consists of $t-1$ vertex classes, each containing two vertices, and two vertices are connected by an edge if and only if they come from different classes. the complete $(t-1)$-partite graph where each independent set contains two vertices. For every pair of vertices in the same class, we add a path of length $\ell-2$; as before, all these paths are vertex-disjoint. Let D_r and D_b be a safe red- and blue-determiner, respectively, as guaranteed by Lemma 2.3.3(ii). We attach a copy of D_b to each edge of T and a copy of D_r to each edge belonging to one of the $t-1$ paths of length $\ell-2$. Finally, we add a new vertex v adjacent to all vertices of T and call the resulting graph G. This construction is illustrated in Figure 5.2 for $\ell = 6$ and $t = 4$, showing only the signal edges of the determiners and the edges incident to v. As in the proof of Proposition 5.1.1, we will show that $G \longrightarrow (C_\ell, K_t)$ but $G - v \not\longrightarrow (C_\ell, K_t)$, which will then imply the claim.

To see the first claim, consider an arbitrary red/blue-coloring of G. If any copy of D_r or D_b contains a red copy of C_ℓ or a blue copy of K_t, then we are done. Hence, we may assume that all determiners have (C_ℓ, K_t)-free colorings, forcing the edges of T to be all blue and the edges in the $(\ell-2)$-paths connecting pairs of vertices from the same class in T to be red (by (D2)). Now, if both edges between v and one of the vertex classes of T are red, there is a red copy of C_ℓ. Otherwise, there is a blue edge from v to each of the $t-1$ vertex classes of T, resulting in an blue copy of K_t.

For the second claim, we color the edges of T blue and the edges of the $(\ell-2)$-paths connecting vertices from the same vertex class of T red. It is not difficult to see that this partial coloring of $G - v$ is (C_ℓ, K_t)-free. Then, using property (D3) and the safeness of the copies of D_r and D_b, we extend this coloring to all determiners so that each one receives a (C_ℓ, K_t)-safe coloring. By the definition of safeness, this results in a (C_ℓ, K_t)-free coloring of the entire graph $G - v$. □

Note that this upper bound for $s_2(C_\ell, K_t)$ does not match the lower bound from (1.1.2), as the latter only implies $s_2(C_\ell, K_t) \geq t$. It turns out that the upper bound is tight, as we will prove in Proposition 5.1.4 below. We will need an auxiliary lemma, which shows that, if G is a graph on fewer than $2(t-1)$ vertices with no t-clique, then there must be at least one vertex common to all $(t-1)$-cliques.

Lemma 5.1.3. *Let $t \geq 3$ be any integer and G be a graph on $n < 2(t-1)$ vertices with $K_{t-1} \subseteq G$. If*

$$\bigcap_{\substack{H \subseteq G \\ H \cong K_{t-1}}} V(H) = \emptyset\,,$$

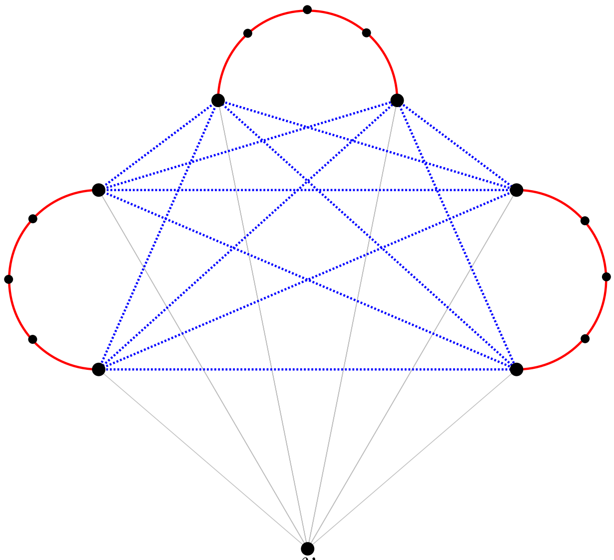

FIGURE 5.2: The graph G in the proof of Proposition 5.1.2 for $\ell = 6$ and $t = 4$ (thick solid/red lines represent edges to which a copy of D_r is attached, dotted/blue lines represent edges to which a copy of D_b is attached).

then $K_t \subseteq G$.

Proof. We proceed by strong induction on t. It is easy to check that the statement is true for $t = 3$. Assume now that $t \geq 3$, and suppose the statement to be true up to t.

Let G be a graph on $n < 2t$ vertices, and let $\mathcal{F} = \{H_0, \ldots, H_m\}$ be a family of distinct t-cliques contained in G whose joint intersection is empty. Suppose additionally that this family is minimal, meaning that every subfamily has a nonempty intersection. Note that we may assume that $m \geq 1$.

Let $S = V(H_1) \cap \ldots \cap V(H_m)$ be the vertex set in the intersection of the t-cliques $H_1, \ldots, H_m$ (without considering H_0). By the minimality of the family $\mathcal{F}$, we know that $|S| > 0$. Further, since G has fewer than $2t$ vertices, it cannot contain two disjoint t-cliques. Therefore, as H_0 is a t-clique and S is another clique disjoint from H_0 in G, it follows that $|S| \leq t - 1$. Write $|S| = t - j$ for some $0 < j < t$.

For $i \in [m]$, let $S_i = V(H_i) \setminus S$. Note that each S_i induces a j-clique. Each vertex in S_i is adjacent to all vertices in S. Therefore, since $|S| = t - j$, if we can find a $(j + 1)$-clique in $G\left[\bigcup_{i=1}^m S_i\right]$, we will have found a $(t + 1)$-clique in G. We consider two possible cases.

<u>*Case 1*</u>*:* Suppose that $\bigcup_{i=1}^m S_i$ has at least $2j$ elements. By definition, both $V(H_0)$ and $\bigcup_{i=1}^m S_i$ have an empty intersection with S, and therefore they are both contained in the set $V(G) \setminus S$ whose size is less than $t + j$. Since $|V(H_0)| = t$ and $|\bigcup_{i=1}^m S_i| \geq 2j$, they must have at least $j + 1$ vertices in common, forming a $(j + 1)$-clique in $G\left[\bigcup_{i=1}^m S_i\right]$.

<u>*Case 2:*</u> Assume next that $\bigcup_{i=1}^{m} S_i$ has fewer than $2j$ elements. Then $G[\bigcup_{i=1}^{m} S_i]$ is a graph on fewer than $2j$ vertices containing a j-clique, namely $G[S_1]$. Since $j < t$ and $\bigcap_{i=1}^{m} S_i = \emptyset$, by the induction hypothesis, it follows that $G\left[\bigcup_{i=1}^{m} S_i\right]$ contains a $(j+1)$-clique. □

We are now ready to prove a lower bound on $s_2(C_\ell, K_t)$ using Lemma 5.1.3.

Proposition 5.1.4. *For any integers $\ell \geq 4$ and $t \geq 3$, we have*

$$s_2(C_\ell, K_t) \geq 2(t-1).$$

Proof. Let G be a minimal 2-Ramsey graph for (C_ℓ, K_t), and suppose v is a vertex of degree at most $2(t-1)-1$ in G, i.e., $|N(v)| < 2(t-1)$. By the minimality of G, there exists a red/blue-coloring φ of the edges of $G - v$ with no red copy of C_ℓ and no blue copy of K_t. If $G[N(v)]$ contains no blue copy of K_{t-1}, then we can extend the coloring φ to G by coloring all edges incident to v blue to obtain a (C_ℓ, K_t)-free coloring of G, a contradiction.

Therefore, assume that there is at least one blue copy of K_{t-1} in $G[N(v)]$. By Lemma 5.1.3, because $G[N(v)]$ has no blue copy of K_t and $|N(v)| < 2(t-1)$, there exists at least one vertex u in the intersection of all blue copies of K_{t-1} in $G[N(v)]$. We now extend φ to G by coloring the edge uv red and all other edges from v to $N(v) \setminus \{u\}$ blue. This coloring does not create a blue copy of K_t and the unique red edge incident to v cannot create a red copy of C_ℓ, again contradicting the fact that G is 2-Ramsey for (C_ℓ, K_t). □

5.2 Proof of Theorems 1.1.14 and 1.1.16

In this section, we use the gadgets constructed in Section 2.6 to prove our main results, Theorems 1.1.14 and 1.1.16. Recall that $\mathcal{T} = \mathcal{T}(q_1, q_2, \ell, t)$ denotes the q-tuple of cycles and cliques as defined in (1.1.3), and that $\mathcal{S}(C_\ell)$ and $\mathcal{S}(K_t)$ denote the cycle-colors $\{1, \ldots, q_1\}$ and clique-colors $\{q_1+1, \ldots, q_1+q_2\}$, respectively, while $\mathcal{S}$ denotes the full color palette $\{1, \ldots, q_1+q_2\}$. We recall the statements below.

Theorem 1.1.14. *For any $t \geq 3$ and any integers $q_1, q_2 \geq 1$, there exists a function $f = f(q_1, q_2, t)$ such that $s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t)) = f(q_1, q_2, t)$ for all $\ell \geq 4$ and*

$$s_{q_2}(K_t) + q_1 \leq f(q_1, q_2, t) \leq s_{q_1+q_2}(K_t). \tag{1.1.4}$$

Theorem 1.1.16. *For all $\ell \geq 4$, $t \geq 3$, $q_2 \geq 1$, and $\varepsilon > 0$, there exists q_0 such that for all $q_1 \geq q_0$, we have*

$$s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t)) \leq (1+\varepsilon)q_1.$$

The idea is to express our function $s_{q_1+q_2}(\mathcal{T})$ in a different way, through a certain packing parameter. This generalizes the ideas of Fox, Grinshpun, Liebenau, Person, and Szabó from [72] discussed in Chapter 3.

5.2.1 Packing parameter

In this section we generalize the packing parameter defined in [72]. Recall that a color pattern on vertex set V is a collection of edge-disjoint graphs $G_1, \ldots, G_m$ on the same vertex set V. A color pattern is H-free if every graph in it is H-free.

Definition 5.2.1. Given positive integers $t \geq 2$ and $q_1, q_2 \geq 0$, let $P_{q_1,q_2}(t)$ be the smallest integer n such that there exists a color pattern $\Gamma_{q_1+1}, \ldots, \Gamma_{q_1+q_2}$ on vertex set $[n]$ such that

(P1) The graph Γ_j is K_{t+1}-free for every $j \in \mathcal{S}(K_t)$.

(P2) For every vertex-coloring $\lambda : [n] \to \mathcal{S}$, we have that (a) two distinct vertices u and w receive the same cycle-color, or (b) there exists a clique-color $j \in \mathcal{S}(K_t)$ such that Γ_j contains a copy of K_t on the vertices of color j.

Note that $P_{q_1,q_2}(t)$ generalizes the parameter $P_q(t)$, defined in [72] and discussed in Chapter 3, since $P_{0,q_2}(t) = P_{q_2}(t)$. Theorem 1.5 in [72] establishes that $s_{q_2}(K_t) = P_{0,q_2}(t-1)$ for all $q_2 \geq 2$ and $t \geq 3$. The following lemma generalizes that theorem and proves that $s_q(\mathcal{T})$ does not depend on ℓ. Setting $f(q_1, q_2, t) = P_{q_1,q_2}(t-1)$ then proves the first part of Theorem 1.1.14.

Lemma 5.2.2. *For all integers $\ell \geq 4$, $t \geq 3$, and $q_1, q_2 \geq 0$, we have*

$$s_{q_1+q_2}(\mathcal{T}(q_1, q_2, \ell, t)) = P_{q_1,q_2}(t-1).$$

Proof. If $q_1 = 0$, we can apply Theorem 1.5 from [72] directly. It is also not difficult to check that $P_{q_1,0}(t-1) = q+1 = s_q(C_\ell)$ [29]. So we may assume $q_1, q_2 \geq 1$. Set $q = q_1 + q_2$ and $\mathcal{T} = \mathcal{T}(q_1, q_2, \ell, t)$. We divide the proof into two claims.

Claim 5.2.3. $s_q(\mathcal{T}) \leq P_{q_1,q_2}(t-1)$.

Proof. Let $n = P_{q_1,q_2}(t-1)$ and $\Gamma_{q_1+1}, \ldots, \Gamma_q$ be a color pattern on $[n]$ that satisfies (P1) and (P2). For every pair of distinct vertices $u, w \in [n]$ and every cycle-color $i \in \mathcal{S}(C_\ell)$, we add a path $P_i(u, w)$ of length $\ell - 2$ between u and w such that these paths are pairwise internally vertex-disjoint. Finally, we add a new vertex v and connect it to each vertex in $[n]$. Call the resulting graph F.

Assume first that $q_1, q_2 > 1$. Now, let S_c^+ and S_c^- be a safe positive and negative $\mathcal{S}(C_\ell)$-sender for $\mathcal{T}$, respectively, and let S_k^+ and S_k^- be a safe positive and negative $\mathcal{S}(K_t)$-sender for $\mathcal{T}$, respectively; all of these gadgets exist by Theorem 2.2.7. Let $E = \{e_1, \dots, e_q\}$ be a matching of size q. For each pair $i, j \in \mathcal{S}(C_\ell)$ of distinct cycle-colors, join the edges e_i and e_j by a copy of S_c^-. Similarly, for each pair $i, j \in \mathcal{S}(K_t)$ of distinct clique-colors, join the edges e_i and e_j by a copy of S_k^-. For every clique-color $i \in \mathcal{S}(K_t)$ and every edge $f \in E(\Gamma_i)$, join the edges e_i and f by a copy of S_k^+. Finally, for each $i \in \mathcal{S}(C_\ell)$ and for each edge $f \in P_i(u, w)$, join the edges e_i and f by a copy of S_c^+. Call the resulting graph G.

We will show that $G \to_q \mathcal{T}$ but $G - v \not\to_q \mathcal{T}$. We begin with the latter. For this we define a $\mathcal{T}$-free coloring of $G - v$. For all $i \in \mathcal{S}(K_t)$, give all edges of Γ_i color i. For all $i \in \mathcal{S}(C_\ell)$ and every pair of distinct vertices $u, w \in [n]$, color the edges of $P_i(u, w)$ with color i. Finally, for all $i \in [q]$, give e_i color i. By property (S3) and the safeness of each gadget, this coloring can now be extended to the set-senders so that each set-sender receives a $\mathcal{T}$-safe coloring. Suppose there exists a monochromatic cycle in a cycle-color or clique in a clique-color. By the safeness of the coloring of each set-sender, we know that such a monochromatic subgraph has to be contained in $F - v$. But $F - v$ contains no monochromatic copy of K_t in a clique-color by property (P1) of the color pattern. By construction, it is not difficult to see that it also contains no monochromatic copy of C_ℓ in a cycle-color. Hence, this is a $\mathcal{T}$-free coloring of $G - v$, as claimed.

We now prove that $G \to_q \mathcal{T}$. For the sake of contradiction, let $\varphi : E(G) \to \mathcal{S}$ be a $\mathcal{T}$-free q-coloring of the edges of G. In any such coloring, property (S2) of the copies of S_c^- and S_k^- ensures that $\{\varphi(e_1), \dots, \varphi(e_{q_1})\} = \mathcal{S}(C_\ell)$, while $\{\varphi(e_{q_1+1}), \dots, \varphi(e_q)\} = \mathcal{S}(K_t)$. Without loss of generality, we may assume that for any $i \in \mathcal{S}$, we have $\varphi(e_i) = i$. Property (S2) of the copies of S_k^+ and S_c^+ further ensures that, for any $i \in \mathcal{S}(K_t)$, each edge in Γ_i has color i, and for each pair of vertices $u, w \in [n]$ and each $j \in \mathcal{S}(C_\ell)$, the edges of $P_j(u, v)$ receive color j.

Consider now the edges from v to $N(v) = [n]$. These induce a natural vertex-coloring $\lambda : [n] \to \mathcal{S}$ defined by $\lambda(u) = \varphi(vu)$ for each $u \in [n]$. Then by property (P2), it follows that either there are two distinct vertices $u, w \in [n]$ such that $\lambda(u) = \lambda(w) = j$ for some $j \in \mathcal{S}(C_\ell)$, or there exists a clique-color $j \in \mathcal{S}(K_t)$ such that $\Gamma_j[\lambda^{-1}(\{j\})]$ contains a copy of K_{t-1}. In the former case $P_j(u, w)$ forms a monochromatic copy of C_ℓ in color j together with v. In the latter case, the copy of K_{t-1} forms a monochromatic copy of K_t in color j together with v.

It follows that G is q-Ramsey for $\mathcal{T}$, while $G - v$ is not. So any minimal q-Ramsey subgraph of G must contain the vertex v, and therefore $s_q(\mathcal{T}) \le d_G(v) = n = P_{q_1,q_2}(t-1)$.

If $q_1 = 1$ and/or $q_2 = 1$, we use a safe $\mathcal{S}(C_\ell)$-determiner D_c instead of a $\mathcal{S}(C_\ell)$-sender, and/or a safe $\mathcal{S}(K_t)$-determiner D_k instead of a $\mathcal{S}(K_t)$-sender. These gadgets exist by Theorem 2.3.4. If $q_1 = 1$, for each edge $f \in P_1(u, w)$ for some $u, w \in [n]$, we attach a copy of D_c to f. If $q_2 = 1$, for every edge $f \in \Gamma_{q_1+1}$, we attach a copy of D_k to f. The rest of the proof is identical to the case $q_1, q_2 > 1$, using the corresponding properties of set-determiners. $\square$

Claim 5.2.4. $s_q(\mathcal{T}) \ge P_{q_1,q_2}(t-1)$.

Proof. Towards a contradiction, assume that there exists a graph G with a vertex v of degree $n < P_{q_1,q_2}(t-1)$ such that G is minimal q-Ramsey for $\mathcal{T}$. By minimality, there exists a $\mathcal{T}$-free q-coloring φ of the edges of $G-v$. This coloring induces a color pattern $\Gamma_{q_1+1},\ldots,\Gamma_q$ on $N(v)$, where Γ_i consists of all edges in $G[N(v)]$ that receive color j under φ for all $j \in \mathcal{S}(K_t)$; then every Γ_j is K_t-free. Since $|N(v)| < P_{q_1,q_2}(t-1)$ and each Γ_j is K_t-free, by property (P2) there must exist a vertex-coloring $\lambda : N(v) \to \mathcal{S}$ such that no two vertices in $N(v)$ receive the same cycle-color and there is no clique-color j such that $\Gamma_j[\lambda^{-1}(\{j\})]$ contains a copy of K_{t-1}. Now, we extend φ to all of G by setting $\varphi(uv) = \lambda(u)$ for each $u \in N(v)$.

By the properties of λ, this extended coloring has no monochromatic copy of C_ℓ in any color $j \in \mathcal{S}(C_\ell)$ and no monochromatic copy of K_t in any color $j \in \mathcal{S}(K_t)$, contradicting the fact that G is q-Ramsey for $\mathcal{T}$. □

5.2.2 Proof of the bounds in Theorem 1.1.14

We are now ready to complete the proof of our first main result in the multicolor setting. We begin with the lower bound.

Lemma 5.2.5. *For all $\ell \geq 4$, $t \geq 3$, and $q_1, q_2 \geq 1$, we have*

$$s_{q_1+q_2}(\mathcal{T}(q_1,q_2,\ell,t)) \geq s_{q_2}(K_t) + s_{q_1}(C_\ell) - 1 = s_{q_2}(K_t) + q_1. \tag{5.2.1}$$

Proof. Set $q = q_1+q_2$ and $\mathcal{T} = \mathcal{T}(q_1,q_2,\ell,t)$, and suppose that G is a minimal q-Ramsey graph for $\mathcal{T}$ containing a vertex v of degree at most $s_{q_2}(K_t) + s_{q_1}(C_\ell) - 2$. Let $\varphi : E(G-v) \to [q]$ be a $\mathcal{T}$-free q-coloring of $G-v$. Let G' be the subgraph of G containing all edges of $G-v$ with colors $q_1+1,\ldots,q$ and any set of $\min\{s_{q_2}(K_t)-1,\ \deg_G(v)\}$ edges of G incident to v. We know that $G'-v$ is not q_2-Ramsey for K_t, and since $\deg_{G'}(v) < s_{q_2}(K_t)$, it follows that G' itself cannot be q_2-Ramsey for K_t. Thus, we can recolor the edges of G' using the colors $q_1+1,\ldots,q$ so that there is no monochromatic copy of K_t in G'. Now, we can apply the same argument to $G-G'$ to obtain a C_ℓ-free coloring of it with the colors $1,\ldots,q_1$. These two colorings together yield a $\mathcal{T}$-free coloring of G, a contradiction. The last equality follows from the fact $s_q(C_\ell) = q+1$ [29]. □

From the proof of this lower bound it becomes clear that this is actually a generalization of the trivial lower bound given in (1.1.2). We now proceed with the upper bound. For this we take a slightly indirect approach: instead of working directly with the parameter s_q, we show a relation between the two packing parameters.

Lemma 5.2.6. *For all $q_1, q_2 \geq 1$, $t \geq 3$, and $\ell \geq 4$, we have*

$$s_{q_1+q_2}(\mathcal{T}(q_1,q_2,\ell,t)) = P_{q_1,q_2}(t-1) \leq P_{0,q_1+q_2}(t-1) = s_{q_1+q_2}(K_t). \tag{5.2.2}$$

Proof. Again set $q = q_1 + q_2$, and let $n = P_{0,q}(t-1)$. Let $\Gamma_1, \ldots, \Gamma_q$ be a K_t-free color pattern on $[n]$, as guaranteed by the definition of $P_{0,q}(t-1)$ (Definition 5.2.1). Consider only the last q_2 graphs; we claim that this color pattern satisfies properties (P1) and (P2) in the definition of $P_{q_1,q_2}(t-1)$ (Definition 5.2.1). The first property is clear. Now let $\lambda : [n] \to \mathcal{S}$ be any coloring. Then we know that there is some $j \in \mathcal{S}$ such that Γ_j contains a monochromatic copy of K_{t-1} on the vertices of color j. Now, if $j > q_1$, then case (b) from property (P2) occurs. Otherwise, we have $j \leq q_1$, and thus there must be at least $t-1 \geq 2$ vertices of color j, implying that case (a) from property (P2) happens. Hence $P_{q_1,q_2}(t-1) \leq P_{0,q}(t-1)$, and the two equalities follow from Lemma 5.2.2 and the discussion that precedes it. □

5.2.3 Proof of Theorem 1.1.16

We now prove our second main result for multiple colors. In [72, Lemmas 4.2 and 4.4], it was shown that, for all $q \geq 2$ and $t \geq 3$, there exists a color pattern $\Gamma_1, \ldots, \Gamma_q$ on the vertex set $[n]$, for some n, such that:

(i) Γ_i is K_t-free for every $i \in [n]$.
(ii) Any subset of $[n]$ of size n/q contains a copy of K_{t-1} in each color.

The results in [72] include bounds on n in terms of q, which are unnecessary for our purpose. Theorem 1.1.16 follows from the next lemma.

Lemma 5.2.7. *Given $0 < \varepsilon < 1$ and integers $q_2 \geq 1$ and $t \geq 3$, there exists an integer $q_0 \geq 1$ such that, for all $q_1 \geq q_0$, we have*

$$P_{q_1,q_2}(t-1) \leq (1+\varepsilon)q_1.$$

Proof. Let $0 < \varepsilon < 1$, $q_2 \geq 1$, and $t \geq 3$ be fixed. For q_1 large enough, there exists a K_t-free color pattern $\Gamma_1, \ldots, \Gamma_{q^*}$ on $n \in [(1+\varepsilon/2)q_1, (1+\varepsilon)q_1]$ vertices, given by the result in [72], with q^* large enough compared to q_2.

Keeping only the first q_2 graphs in the color pattern, which we denote for convenience by $\Gamma_{q_1+1}, \ldots, \Gamma_{q_1+q_2}$, we claim that they satisfy properties (P1) and (P2). The first one is clear. For the second one, consider a vertex coloring $\lambda : [n] \to [q]$, where $q = q_1 + q_2$. Let C be its largest color class in $\mathcal{S}(K_t)$, with color c. If (a) does not hold, by the pigeonhole principle the color class C has size at least $\frac{n-q_1}{q_2}$. Since q^* is large enough compared to q_2, and by the choice of n, we have $\frac{n-q_1}{q_2} \geq \frac{n}{q^*}$. By property (ii) above, we know that there exists a copy of K_{t-1} in $\Gamma_c[C]$. Therefore, if (a) of (P2) does not hold, then (b) does, implying that $P_{q_1,q_2}(t-1) \leq n \leq (1+\varepsilon)q_1$. □

5.3 Concluding remarks

In this chapter, we studied the parameter s_q in the asymmetric setting for tuples consisting of cliques and cycles. The upper and lower bounds we obtained are strongly dependent on the existing bounds for the symmetric parameter $s_q(K_t)$. As noted in [72], the study of $s_q(K_t)$ appears to be tightly connected to the Erdős-Rogers function, implying that any improvements on our current results would probably be nontrivial. We refer to [72, Section 5] for a more detailed discussion on the relationship between $s_q(K_t)$ and the Erdős-Rogers function.

It would be desirable to study other asymmetric cases of the problem, and a natural place to start is to consider pairs of graphs for which safe determiners are known to exist (including all pairs of 3-connected graphs and the pairs considered by Siggers in [140]).

The multicolor asymmetric setting offers even more room for study, as the existence of gadget graphs is an open problem even in some very natural cases. Our method allows us to construct set-determiners and set-senders for tuples of the form $(C_\ell, \dots, C_\ell, K_s, K_t)$. However, we are not aware of a way to build gadget graphs for asymmetric q-tuples of cliques, with $q > 2$. Since studying Ramsey graphs for cliques is a central theme in Ramsey theory, we believe that resolving the following problem would be of interest.

Problem 1. Construct signal senders for asymmetric q-tuples of the form $(K_{t_1}, \dots, K_{t_q})$.

Once we have the necessary tools, it would be very interesting to investigate the parameter s_q for such tuples.

It would also be desirable to determine if the upper bound in Theorem 1.1.14 holds in other cases. In particular, it was conjectured by Fox, Grinshpun, Liebenau, Person, and Szabó [72] that $s_q(K_{t-1}) \leq s_q(K_t)$ for $q \geq 3$. Perhaps the following asymmetric version would be more approachable.

Problem 2. Show that $s_{q_1+q_2}(\underbrace{K_{t-1}, \dots, K_{t-1}}_{q_1+1 \text{ times}}, \underbrace{K_t, \dots, K_t}_{q_2-1 \text{ times}}) \leq s_{q_1+q_2}(\underbrace{K_{t-1}, \dots, K_{t-1}}_{q_1 \text{ times}}, \underbrace{K_t, \dots, K_t}_{q_2 \text{ times}})$.

Chapter 6

Ramsey non-equivalent graphs

The goal of this chapter is to prove Theorems 1.1.18 and 1.1.20 and Proposition 1.1.19. We begin with the simpler of the two theorems, namely Theorem 1.1.20. Both proofs are inspired by previous work by Rödl and Siggers [128] and Siggers [139–141].

6.1 Cycles

We begin by recalling the statements of Proposition 1.1.19 and Theorem 1.1.20.

Proposition 1.1.19. *For every $\ell \geq 4$, the graphs C_ℓ and $C_\ell \cdot K_2$ are not 2-Ramsey equivalent.*

Theorem 1.1.20. *For every odd $\ell \geq 5$, as $n \to \infty$, there are $2^{\Omega(n^2)}$ graphs on at most n vertices that are minimal 2-Ramsey for C_ℓ but are not 2-Ramsey for $C_\ell \cdot K_2$.*

Throughout this section, let $\ell \geq 4$ be fixed. We begin by describing our basic building block and its properties. Recall that we write P_t for a path on t vertices.

Construction 6.1.1. Let the graph T be defined as follows:

- Let P^1, P^2, P^3 be three copies of the path $P_{\ell-2}$. For each $i \in [3]$, let y_i and z_i be the endpoints of P^i.
- For each $i \in [3]$, create a copy of $P_{\ell-1}$ on a new set of vertices and identify its endpoints with y_i and y_{i+1} (for convenience we set $y_4 = y_1$); let h_i denote the unique edge incident to y_i on this path.
- For each $i \in [3]$, create a copy of $P_{\ell-1}$ on a new set of vertices and identify its endpoints with z_i and z_{i+1} (again we set $z_4 = z_1$).

Let φ be the partial coloring of T in which every edge contained in P^i for $i \in [3]$ is red, and all other edges except for h_1, h_2, and h_3 are blue.

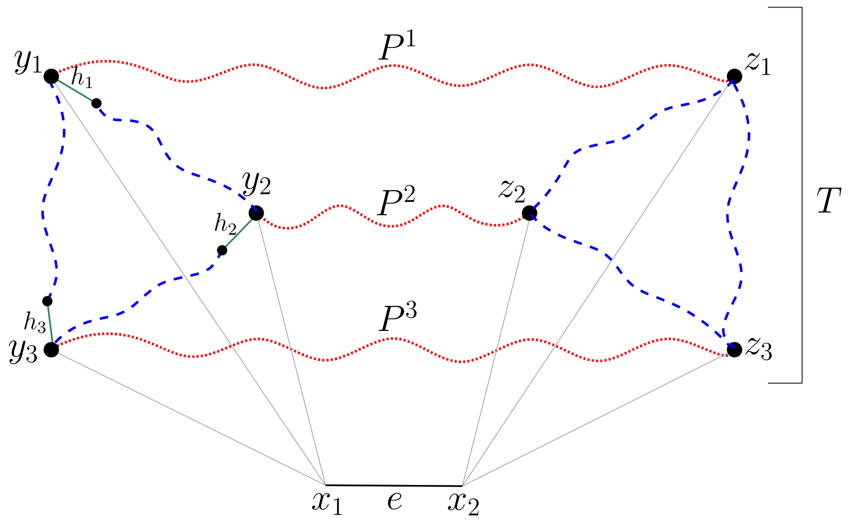

FIGURE 6.1: The graph T' and the coloring φ from Construction 6.1.1 (thick dotted lines represent red paths, thick dashed lines represent blue paths, and thin solid lines represent edges).

Let T' be the graph obtained from T by adding an edge $e = x_1x_2$ disjoint from T and inserting the edges x_1y_i and x_2z_i for each $i \in [3]$.

The graph T' and the partial coloring φ are illustrated in Figure 6.1. We now show some properties of the graphs T and T' and the partial coloring φ.

Lemma 6.1.2. *The girth of T is $3(\ell - 2)$, and therefore T contains no copy of C_ℓ. The girth of T' is ℓ.*

Proof. If a cycle in T contains an internal vertex of P^i for some $i \in [3]$, it must have length at least $2(\ell - 3) + 2(\ell - 2) > \ell$. Otherwise, the graph obtained from T by removing the internal vertices of these three paths has two components, each of which is a cycle of length $3(\ell - 2) > \ell$. For the second part, it suffices to consider only the cycles containing at least one vertex of e. Suppose a cycle in T' contains x_1. Again, the cycle must contain an internal vertex of one of the paths P^i or an internal vertex of a path connecting y_j and y_{j+1} or z_j and z_{j+1} for some $i, j \in [3]$; in both cases the length of the cycle must be at least ℓ. □

Lemma 6.1.3. *Let φ' be a partial coloring of T' extending φ to the edges h_1, h_2, h_3, and e.*

(a) If e is blue under φ', then coloring all edges incident to e red extends φ' to a C_ℓ-free coloring of T'.
(b) If e is red and h_1, h_2, h_3 are all red under φ', then coloring the edges from x_1 to the vertices y_1, y_2, y_3 blue and the edges from x_2 to z_1, z_2, z_3 red extends φ' to a C_ℓ-free coloring of T'.
(c) If e is red and h_1, h_2, h_3 are all blue under φ', then φ' cannot be extended to a C_ℓ-free coloring of T'.

Proof.

(a) We color all edges between e and T red. The blue subgraph of T' is then contained in $T \cup \{e\}$. Since the blue edge e is isolated in this subgraph, it cannot participate in a monochromatic copy of C_ℓ. Since T itself contains no copy of C_ℓ by Lemma 6.1.2, there can be no blue copy of C_ℓ. In addition, the red subgraph of T' consists of three internally vertex-disjoint paths between x_1 and x_2, each of length $\ell - 1$, plus possibly some of the edges h_i as pendent edges. This graph has girth $2(\ell - 1) > \ell$, so it also contains no copy of C_ℓ.

(b) We color the edges from x_1 to the vertices y_1, y_2, y_3 blue and the edges from x_2 to z_1, z_2, z_3 red. Now, the red subgraph of T' is a tree: it consists of three paths of length $\ell - 1$ and the edge e, all meeting at the vertex x_2. So there is no red copy of C_ℓ. Similarly, the blue subgraph of T' is made up of a tree consisting of three paths of length $\ell - 2$ meeting at x_1 and a cycle of length $3(\ell - 2)$, which again contains no copy of C_ℓ.

(c) If there is an $i \in [3]$ such that both x_1y_i and x_2z_i are red, then P^i together with the edges x_1y_i, x_2z_i, and e forms a red copy of C_ℓ. Hence, we may assume that, for each $i \in [3]$, at least one of the edges x_1y_i and x_2z_i is blue. By the pigeonhole principle, there exists a $j \in [2]$ such that there are at least two blue edges incident to x_j. Without loss of generality, we may assume that $j = 1$ and that the blue edges are x_1y_1 and x_1y_2. Then these two edges together with the copy of $P_{\ell-1}$ between y_1 and y_2 yield a blue copy of C_ℓ.

□

As a first step, we will show that there exist safe signal senders for C_ℓ that have C_ℓ-safe colorings satisfying an additional property.

Lemma 6.1.4. *For any $\ell \geq 4$, there exists a safe negative signal sender $S^- = S^-(C_\ell, 2, e, f)$ (resp., a safe positive signal sender $S^+ = S^+(C_\ell, 2, e, f)$) that has a C_ℓ-safe coloring in which all edges incident to a signal edge $h \in \{e, f\}$ receive a color different from that of h.*

Proof. We begin by constructing a weaker gadget, namely a negative signal sender with signal edges e and f that has a C_ℓ-safe coloring in which all edges incident to e receive a color different from that of e. We construct a graph S as follows. Let S_0^+ and S_0^- be a safe positive and a safe negative signal sender for C_ℓ, respectively, which exist by Lemma 2.2.4 and Corollary 2.2.5. We start with a copy of the graph T' and an edge f disjoint from T'. We now connect the edges of $T \subseteq T'$ to f by signal senders: every edge contained in one of the paths P^1, P^2, or P^3 is joined to f by a copy of S_0^+ and every other edge is joined to f by a copy of S_0^-.

We now show that S satisfies the required properties, namely that S is a safe negative signal sender with signal edges e and f and that it has the required special coloring.

We begin by showing property (S2). Consider an arbitrary C_ℓ-free coloring of S. Without loss of generality, assume that f is red. Then by property (S2) of the signal senders S_0^+ and S_0^-, we know that the paths P^1, P^2, and P^3 are monochromatic in red and that the remaining edges of T are all blue. Thus T' has a coloring extending the coloring φ in which the edges h_1, h_2, and h_3 are colored blue. Hence, by Lemma 6.1.3(c), the edge e cannot be red. So e must be blue, as needed.

Property (S3) follows from (S1) and the fact that we can switch the two colors. Now we describe a coloring φ' of S, through which we will verify all remaining properties. To this end, we color the edges h_1, h_2, h_3, and e blue and the edge f red, and we color the remaining edges of T according to the coloring φ. By Lemma 6.1.3(a), we can color the edges incident to e red without creating a monochromatic C_ℓ. Finally, we extend this partial coloring to all copies of S_0^+ and S_0^- using a C_ℓ-safe coloring, which is possible by (S3) because the signal edges of each signal sender are colored compatibly. Since each signal sender was given a safe coloring, it follows that there is no monochromatic C_ℓ anywhere in S, verifying property (S1). Further, all edges incident to e have the opposite color, establishing the required additional property.

Finally, to show that the above coloring is safe at $\{e, f\}$, suppose G is a graph with $V(G) \cap V(S) = V(e) \cup V(f)$ and $E(G) \cap E(S) = \{e, f\}$ and G is given a C_ℓ-free coloring that agrees with φ' on $\{e, f\}$. Suppose there is a monochromatic copy C of C_ℓ in $G \cup S$. We know that each copy of S_0^+ and S_0^- was given a C_ℓ-safe coloring, implying that C is fully contained in the graph $G \cup T' \cup \{f\}$. As in the proof of Corollary 2.2.5, we can conclude that C must contain a vertex $x \in V(T') \setminus e = V(T)$ and a vertex $y \in V(G) \setminus e$. There are two internally vertex-disjoint paths between x and y in C, so C must contain both vertices of e. But then the segments of C connecting x to the vertices of e, which are fully contained in T', together with the edge e form a cycle of length at most $\ell - 1$ in T', contradicting the fact that the girth of T' is ℓ (Lemma 6.1.2). Hence the coloring φ' is C_ℓ-safe at $\{e, f\}$.

We now use a standard construction (see [40]) to obtain the promised gadgets. To construct a safe positive signal sender with the required additional property, we take two copies S_1 and S_2 of the graph S constructed above and denote their signal edges by e_1, f_1 and e_2, f_2, respectively. The graph obtained by identifying the edges f_1 and f_2 is easily seen to satisfy properties (S1), (S2), and (S3). The existence of a C_ℓ-safe coloring with the required additional property follows from the existence of a C_ℓ-safe coloring of S in which all edges incident to e receive a color different from that of e. Similarly, for a safe negative signal sender, we take a third copy S_3 of the graph S with signal edges e_3, f_3, and identify f_1 with f_2 and f_3 with e_1. □

Using the signal senders constructed in the above lemma, we can also construct indicators with a similar property.

Corollary 6.1.5. *For any $\ell \geq 4$ and any graph C_ℓ-free graph F with $m \geq 2$ edges, there exists a safe indicator $I = I(C_\ell, F, 2, e)$ with $O(m)$ vertices such that there are no edges between F and*

e and the C_ℓ-free colorings guaranteed by (I1) and (I3) can be chosen so that all edges incident to e have a color different from that of e.

Proof. Fix a safe indicator $I' = I(C_\ell, F, 2, e')$ with $O(m)$ vertices, as given by Lemma 2.4.2, and a safe positive signal sender $S = S^+(C_\ell, 2, e'', e)$ as in Lemma 6.1.4. Let I be the graph obtained from I' and S by identifying the indicator edge e' of I' with the signal edge e'' of S. The claimed properties of I then follow immediately from the properties of I' and S. □

Before we proceed with the proof of Theorem 1.1.20, we prove Proposition 1.1.19 as a simple application of Lemma 6.1.4.

Proof of Proposition 1.1.19. We construct a graph that is 2-Ramsey for C_ℓ but not 2-Ramsey for $C_\ell \cdot K_2$. Let C be a copy of C_ℓ and e be a disjoint edge. Let S be a safe positive signal sender for C_ℓ satisfying the additional property from Lemma 6.1.4. We then join every edge of C to e by a copy of S and call the resulting graph G.

Now, to see that G is 2-Ramsey for C_ℓ, consider an arbitrary 2-coloring of G and observe that either some copy of S contains a monochromatic copy of C_ℓ, or by (S2) all edges of C have the same color.

For the second claim, we provide a coloring. Assign the color red to all edges of C and to e, and give each copy of S a C_ℓ-safe coloring in which all edges incident to the signal edges are blue. Since every copy of C_ℓ other than C is contained in a signal sender, the only monochromatic copy of C_ℓ under this coloring is C; every edge incident to only one vertex of C is contained in a signal sender and is thus blue. Hence, there is no monochromatic copy of $C_\ell \cdot K_2$. □

We are now ready to prove Theorem 1.1.20. We first present the construction and then justify its properties in a series of lemmas.

Let $\ell \geq 5$ be odd. Fix an integer $m \geq 1$ and m vectors $r_1, \ldots, r_m \in \{0,1\}^m$. We write $r_i(j)$ for the jth coordinate of the vector r_i. Also, fix a safe positive and a safe negative signal sender S^+ and S^- for C_ℓ. Finally, fix a safe indicator $I = I(C_\ell, F, 2, e)$, where F is a matching of size m, as given by Corollary 6.1.5.

Construction 6.1.6. Let $G = G(m, r_1, \ldots, r_m)$ be defined as follows.

- Let $T_1, \ldots, T_m$ be disjoint copies of the graph T from Construction 6.1.1. Let F denote the matching consisting of the copies of h_1 in $T_1, \ldots, T_m$.
- Let $e_1, \ldots, e_{2m}$ be a matching disjoint from the T_i. For each $j \in [2m]$, let $x_{j,1}$ and $x_{j,2}$ denote the endpoints of e_j.

- For all $i, j \in [m]$, add a bipartite graph between T_i and $e_{r_i(j)m+j}$ so that $V(T_i) \cup V(e_{r_i(j)m+j})$ induces a copy of T' in which $x_{r_i(j)m+j,1}$ and $x_{r_i(j)m+j,2}$ correspond to x_1 and x_2, respectively. In other words, depending on the value of $r_i(j)$, we connect T_i and either e_j (if $r_i(j) = 0$) or e_{m+j} (if $r_i(j) = 1$) so that they form a copy of T'.
- Let f be a new edge disjoint from all the T_i and e_j.
- For each copy of T_i, join f to the edges in the copies of P^1, P^2, and P^3 by a copy of S^+ and to the remaining edges of T_i, except for the copies of h_1, h_2, and h_3, by a copy of S^-.
- For all $i \in [m]$, connect the copy of h_1 in T_i to the copies of h_2 and h_3 in T_i by copies of S^+.
- For all $j \in [m]$, join e_j and e_{m+j} by a copy of S^+.
- Let Q_1 and Q_2 be two copies of C_ℓ that are disjoint from the rest of the graph defined so far.
- For each edge h of Q_1, join h and the matching $\{e_1, \ldots, e_m\}$ by a copy of I.
- For each edge h of Q_2, join h and the matching F by a copy of I.

We will show that, for any choice of the $(m+1)$-tuple $(m, r_1, \ldots, r_m)$, the graph $G(m, r_1, \ldots, r_m)$ is 2-Ramsey for C_ℓ but not 2-Ramsey for $C_\ell \cdot K_2$, and if we remove the bipartite graph between any pair (T_i, e_j), the resulting graph is no longer Ramsey for C_ℓ. Further, we will show that the number of vertices of $G(m, r_1, \ldots, r_m)$ is $O(m)$. This is the content of the next four lemmas. For each of these lemmas, assume that the tuple $(m, r_1, \ldots, r_m)$ is fixed and write $G = G(m, r_1, \ldots, r_m)$.

Lemma 6.1.7. *The graph G is 2-Ramsey for C_ℓ.*

Proof. Consider an arbitrary red/blue-coloring of G. By symmetry, we may assume that f is red. If there is a monochromatic copy of C_ℓ in some signal sender or indicator, then we are done. So assume all signal senders and indicators receive C_ℓ-free colorings. Now, by property (S2) of the signal senders S^+ and S^-, we know that each T_i must get a coloring extending φ. If the copies of h_1 in $T_1, \ldots, T_m$ all have the same color, then by property (I2) of the indicators connecting the matching F and the edges of Q_2, all edges of Q_2 have the same color, and hence Q_2 is a monochromatic copy of C_ℓ. Similarly, if the edges $e_1, \ldots, e_m$ all have the same color, then the indicators connecting the matching $\{e_1, \ldots, e_m\}$ and the edges of Q_1 ensure that Q_1 is a monochromatic copy of C_ℓ.

Thus we may assume that there exist $i, j \in [m]$ such that the copy of h_1 in T_i is blue and e_j is red. By property (S2) of the signal sender S^+, we know that the edges h_2 and h_3 in T_i are also blue and that the edge e_{m+j} is also red. The vertices of T_i together with the vertices of either e_j or e_{m+j} form a copy of the graph T', and by Lemma 6.1.3(c) it follows that this copy of T' contains a monochromatic copy of C_ℓ. □

Lemma 6.1.8. *For any $i, j \in [m]$, the graph obtained from G by removing the edges in the bipartite graph between T_i and e_j and e_{m+j} is not 2-Ramsey for C_ℓ.*

Proof. By symmetry, we may assume the edges between T_1 and e_1 and e_{m+1} are removed. Let G' be the resulting graph. Consider the following partial coloring of G'.

- For each $i \in [m]$, color T_i except for the copies of the edge h_1, h_2, h_3 according to the coloring φ.
- Color the copies of h_1, h_2, h_3 in T_1 blue and in every other T_i red.
- Color e_1 and e_{m+1} red and all other e_j blue.
- For all $i \in [m] \setminus \{1\}$, color the bipartite graph between T_i and e_1 and e_{m+1} with the C_ℓ-free coloring given by Lemma 6.1.3(b).
- For all $i \in [m]$ and $j \in [m] \setminus \{1\}$, color the bipartite graph between T_i and e_j and e_{m+j} with the C_ℓ-free coloring given by Lemma 6.1.3(a).
- Color the cycles Q_1 and Q_2 so that they are not monochromatic.
- Color f red.

Now, we can extend this coloring to all signal senders and indicators using a C_ℓ-safe coloring for each gadget. This is possible since the colors so far are chosen so that no indicator has a monochromatic indicator subgraph and any two edges that appear as the signal edges of a signal sender have compatible colors.

We now show that this coloring is indeed C_ℓ-free. Since each signal sender and each indicator was given a C_ℓ-safe coloring, in order to show that the coloring of G' is C_ℓ-free, it suffices to consider the graph $G'' = \bigcup_{i=1}^m T_i \cup \bigcup_{j=1}^{2m} e_j \cup B \cup Q_1 \cup Q_2 \cup f$, where B is the bipartite graph containing all edges between the T_i and the e_j in G', and check that G'' contains no monochromatic copy of C_ℓ.

Suppose for a contradiction that there exists a monochromatic copy C of C_ℓ in G''. First note that the cycles Q_1 and Q_2 are both isolated in G''. Hence, if C contains a vertex of Q_1, then C must be Q_1 itself, which is not monochromatic. Similarly for Q_2.

We now consider two cases depending on the color of C. Suppose first that C is red. Since the graph B between $\bigcup_{i=1}^m V(T_i)$ and $\bigcup_{j=1}^{2m} V(e_j)$ is bipartite and C is not (because ℓ is odd), we know that C must contain one of the edges e_1 or e_{m+1} or some edge belonging to a T_i. The endpoints $x_{1,1}$ and $x_{m+1,1}$ of e_1 and e_{m+1}, respectively, each have degree one in the red subgraph of G'' (by the fact that bipartite graphs between T_i and e_1 and e_{m+1} are all colored according to Lemma 6.1.3(b)), so neither of those edges can be part of a red cycle. So C contains a red edge from some copy of T_i of T. The copies of h_1, h_2, h_3 all have an endpoint of degree at most one in red, so we can conclude that no copy of h_1, h_2, or h_3 can be part of C. Hence C must contain an edge from the copy of one of the paths P^1, P^2, or P^3 in T_i; we may assume C contains an edge from the copy of P^1. But then C must contain the entire copy of P^1 in T_i, which has length $\ell - 3$. The remaining vertices of C can only be from a single edge e_j, but this is impossible as each such edge except for e_1 and e_{m+1} is blue. So C cannot be red. Hence we may assume that C is blue. Now, the blue subgraph of G'' consists of the following components:

- the edges e_j for $j \in [2m] \setminus \{1, m+1\}$;
- m cycles of length $3(\ell - 2)$ containing the copies of z_1, z_2, and z_3 in each T_i for $i \in [m]$;
- a cycle of length $3(\ell - 2)$ containing the copies of y_1, y_2, and y_3 in T_1;
- a tree consisting of paths of length $\ell - 2$ meeting at the endpoint $x_{1,1}$ of e_1;
- a tree consisting of paths of length $\ell - 2$ meeting at the endpoint $x_{m+1,1}$ of e_{m+1}.

None of these graphs contains a copy of C_ℓ. Therefore, we conclude that there is no monochromatic copy of C_ℓ in G'' and hence also in G'. □

Lemma 6.1.9. *The graph G is not 2-Ramsey for $C_\ell \cdot K_2$.*

Proof. We define a coloring of G containing exactly two monochromatic copies of C_ℓ, namely Q_1 and Q_2, in which all edges incident to Q_1 and Q_2 are colored differently from Q_1 and Q_2. Consider the following coloring of G:

- For each $i \in [m]$, color the edges of T_i except for the copies of h_1, h_2, h_3 according to the coloring φ.
- Color the copies of h_1, h_2, h_3 in every T_i for $i \in [m]$ blue.
- Color e_j blue for all $j \in [2m]$.
- For all $i, j \in [m]$, color all edges of the complete bipartite graph between T_i and e_j and e_{m+j} red.
- Color the cycles Q_1 and Q_2 so that each is monochromatic in blue.
- Color f red.

We can now extend this coloring to all signal senders and indicators so that each gadget receives a C_ℓ-safe coloring. In particular, for the copies of I we use the coloring guaranteed by Corollary 6.1.5.

The cycles Q_1 and Q_2 are both monochromatic. We now verify that they are the only monochromatic copies of C_ℓ. Indeed, since each signal sender and each indicator was given a C_ℓ-safe coloring, it suffices to show that the graph $G' = \bigcup_{i=1}^{m} T_i \cup \bigcup_{j=1}^{2m} e_j \cup B \cup f$, where B is the bipartite graph containing all edges between the T_i and the e_j in G, contains no monochromatic copy of C_ℓ. The blue subgraph of G' consists of isolated edges and isolated cycles of length $3(\ell - 2)$, so there is no blue copy of C_ℓ in G'. Suppose there exists a red copy C of C_ℓ in G'. Since ℓ is odd and all copies of h_1, h_2, and h_3 and all e_j are blue, we know that C must contain a red edge from some copy T_i of T. But then C must contain the entire copy of some path P^1, P^2, or P^3 and the vertices of some edge e_j, so it cannot be monochromatic in red. Hence, Q_1 and Q_2 are the only monochromatic copies of C_ℓ in G and they are blue.

Now we consider the full graph G again. By construction, all edges incident to a vertex of Q_1 or Q_2 that are not part of the cycles Q_1 and Q_2 are contained in the copies of I. Since each copy of I was given the coloring guaranteed by Corollary 6.1.5, we know that all those edges must be red, and thus there is no monochromatic copy of $C_\ell \cdot K_2$ in G. □

Lemma 6.1.10. *The graph G has $O(m)$ vertices.*

Proof. Note that each vertex of G belongs to either a signal sender that has f as an edge, or a signal sender with signal edges e_j and e_{m+j} for some $j \in [m]$, or to a signal sender that has a copy of h_1 as a signal edge, or to one of the copies of I with indicator edge in Q_1 or Q_2. The number of vertices in each of S^+, S^-, and T is a constant (depending only on ℓ), and there are at most $me(T)$ signal senders with f as one of their signal edges, m signal senders with signal edges e_j and e_{m+j} for some $j \in [m]$, and $2m$ signal senders that have h_1 as one of their signal edges. Further, by Corollary 6.1.5, we know that each of the 2ℓ copies of I has $O(m)$ vertices. Therefore in total G has $O(m)$ vertices. □

We are now ready to complete the proof of Theorem 1.1.20.

Proof of Theorem 1.1.20. The remainder of the proof is now the same as in [128, 139–141]. By Lemmas 6.1.7 and 6.1.10, we know that for any $m \geq 2$ and for any choice of $r_1, \dots, r_m$, we can construct a graph $G(m, r_1, \dots, r_m)$ that is 2-Ramsey for C_ℓ and has at most Cm vertices for some constant $C = C(\ell) > 0$. Further, for any distinct $(r_1, \dots, r_m) \neq (r'_1, \dots, r'_m)$, the corresponding labeled graphs $G(m, r_1, \dots, r_m)$ and $G(m, r'_1, \dots, r'_m)$ are distinct. By Lemma 6.1.8, we know that removing any of the bipartite graphs between a copy T_i of T and an edge of the matching $\{e_1, \dots, e_{2m}\}$ destroys the Ramsey property of $G(m, r_1, \dots, r_m)$. Hence, if $\tilde{G} \subseteq G(m, r_1, \dots, r_m)$ is minimal 2-Ramsey for C_ℓ, we know that $\tilde{G}$ contains at least one edge from each of these bipartite graphs. Hence, different choices of the vector $(r_1, \dots, r_m)$ result in different labeled *minimal* 2-Ramsey graphs for C_ℓ on at most Cm vertices. The number of choices for $(r_1, \dots, r_m)$ is given by 2^{m^2}. Hence, setting $n = Cm$, there are at least $2^{n^2/C^2}$ labeled minimal 2-Ramsey graphs for C_ℓ that are not 2-Ramsey for $C_\ell \cdot K_2$. Each of these graphs is isomorphic to at most $n! \leq n^n = e^{n \log n}$ others, so there are at least $\frac{2^{n^2/C^2}}{e^{n \log n}} = 2^{\Omega(n^2)}$ non-isomorphic graphs satisfying the desired properties. □

6.2 Cliques

In this section, we prove Theorem 1.1.18.

Theorem 1.1.18. *For every $t \geq 3$, as $n \to \infty$, there are $2^{\Omega(n^2)}$ graphs on at most n vertices that are minimal 2-Ramsey for K_t but are not 2-Ramsey for $K_t \cdot K_2$.*

The strategy is the same as in the previous section: we will again construct $2^{\Omega(n^2)}$ different labeled graphs that are minimal 2-Ramsey for K_t but not 2-Ramsey for $K_t \cdot K_2$; the rest of the proof is the same as the proof of Theorem 1.1.20.

As in the case of cycles, we begin by showing that we can construct signal senders satisfying an additional property.

Lemma 6.2.1. *For any $t \geq 3$, there exists a safe negative signal sender $S^- = S^-(K_t, 2, e, f)$ (resp., a safe positive signal sender $S^+ = S^+(K_t, 2, e, f)$) that has a K_t-safe coloring in which all edges incident to a signal edge $h \in \{e, f\}$ receive a color different from that of h.*

Proof. We begin by constructing a weaker gadget, namely a negative signal sender with signal edges e and f that has a K_t-safe coloring in which all edges incident to e receive a color different from that of e. The rest follows as in the proof of Lemma 6.1.4. We use a construction that is essentially a simplified version of the one given in [83, Lemma 2.6.3].

Let $\mathcal{H}$ be a $(t-1)$-uniform hypergraph with $\chi(\mathcal{H}) \geq 3$, girth at least four, and no isolated vertices, which is known to exist by the work of Erdős and Hajnal [63]. Write $v(\mathcal{H}) = N$ and $V(\mathcal{H}) = \{v_1, \ldots, v_N\}$. Let S^- be a safe negative signal sender for K_t, as guaranteed by Corollary 2.2.3. Construct a graph S as follows:

- Start with N vertex-disjoint copies of K_{t-2}, denoting their vertex sets by $V_1, \ldots, V_N$.
- Add a complete bipartite graph between V_i and V_j whenever v_i and v_j share a hyperedge in $\mathcal{H}$.
- Let e and f be edges that are disjoint from all the V_i and from each other.
- Add edges between the endpoints of e and all vertices in $\bigcup_{i=1}^{N} V_i$.
- Join f to all edges with endpoints in two different V_i by copies of S^-.

We now verify that the graph S is a negative signal sender satisfying the required additional property.

We begin by showing property (S2). Consider an arbitrary K_t-free coloring of S. Without loss of generality, assume that f is red. Suppose for a contradiction that e is also red. By property (S2) of the copies of S^-, we know that all edges with endpoints in two different V_i are blue. Suppose first that there exists an $i \in [N]$ and vertices $u, v \in V_i$ such that the edge uv is blue. We know that there is a hyperedge $h \in \mathcal{H}$ containing v_i. Then taking $t-2$ vertices, one from each set V_j such that $v_j \in h \setminus \{v_i\}$, and the vertices u and v, we find a blue copy of K_t. So we may assume that, for all $i \in [N]$, the $(t-2)$-clique induced by V_i is monochromatic in red.

Now, write $e = x_1x_2$. If all edges from x_1 and x_2 to some V_i are red, then the vertices in $V_i \cup \{x_1, x_2\}$ induce a red copy of K_t. Thus, we may assume that, for all $i \in [N]$, there is a vertex $w_i \in V_i$ such that there is a blue edge from x_1 or x_2 to w_i. We define a coloring on the *vertices* of $\mathcal{H}$ as follows: let v_i have color 1 if there is a blue edge from w_i to x_1 and color 2 otherwise. By assumption, the chromatic number of $\mathcal{H}$ is at least three, so there exists a hyperedge $\{v_{i_1}, \dots, v_{i_{t-1}}\} \in \mathcal{H}$ such that all vertices have the same color; say this color is 1. Then the vertices $w_{i_1}, \dots, w_{i_{t-1}}, x_1$ form a blue copy of K_t.

Next we show properties (S1) and (S3), the safeness of S, and the required additional property by describing a coloring φ of S. Again, note that property (S3) follows from (S1) and the fact that we can switch the two colors. We color f, all edges inside the V_i, and all edges incident to e red; we then color all edges with endpoints in two different V_i and the edge e blue. The signal edges of every copy of S^- are colored differently, so we can extend this coloring to the copies of S^-, giving each signal sender has a K_t-safe coloring.

We now show that φ is K_t-free. Since each copy of S^- received a K_t-safe coloring, we know that it suffices to show that the graph S' obtained from the first four steps in the construction (that is, before adding the signal senders) contains no monochromatic copy of K_t. Suppose there exists a monochromatic copy K of K_t in S'. The red subgraph of $S' - V(e)$ consists of disjoint cliques on $t-2$ vertices, and since e is blue, the red subgraph of S' has clique number $t-1$. Thus, K must be blue. The blue subgraph of S' consists of the edges between the sets V_i and the isolated edge e. Since the girth of $\mathcal{H}$ is at least four, K must be contained within the complete multipartite graph corresponding to a single hyperedge of $\mathcal{H}$, which contains $t-1$ independent sets of size $t-2$. This multipartite graph clearly contains no copy of K_t, verifying property (S1). Note also that all edges incident to the blue edge e are red, verifying the additional property.

Finally, it remains to show that the coloring φ is K_t-safe. Suppose we have a graph G such that $V(G) \cap V(S) = V(e) \cup V(f)$ and $E(G) \cap E(S) = \{e, f\}$ and that G is given a K_t-free coloring that agrees with φ on $\{e, f\}$, that is, f is colored red and e is colored blue. Suppose there is a monochromatic copy K of K_t in $G \cup S$; by the safeness of the coloring of each copy of S^-, we may assume that K is fully contained in $G \cup S'$, where S' is as defined above. As in the proof of Corollary 2.2.3, we can conclude that K must contain a vertex $x \in \bigcup_{i=1}^N V_N$ and a vertex $y \in V(G) \setminus e$. But then x and y cannot be connected by an edge, so they cannot share a copy of K_t. Hence the coloring φ is K_t-safe at $\{e, f\}$. □

As in the cycle case, Lemma 6.2.1 allows us to build safe indicators for K_t satisfying a similar additional property. The proof is the same as that of Corollary 6.1.5

Corollary 6.2.2. *For any $t \geq 3$ and any graph K_t-free graph F with $m \geq 2$ edges, there exists a safe indicator $I = I(K_t, F, 2, e)$ with $O(m)$ vertices such that there are no edges between F and e*

and the K_t-free colorings guaranteed by (I1) and (I3) can be chosen so that all edges incident to e have a color different from that of e.

We are now ready to present the construction proving Theorem 1.1.18. Let $t \geq 3$. Let $m \geq 3$ be an odd integer and $r_1, \ldots, r_m \in \{0,1\}^m$ be fixed. One of the main building blocks is due to Rödl and Siggers [128]. We begin by describing this building block and its properties; the following is a slightly modified version of Construction 3.1 in [128].

Construction 6.2.3. Let $G^* = G^*(m, r_1, \ldots, r_m)$ be defined as follows:

- Let $C : v_1, \ldots, v_{5m}$ be a cycle of length $5m$ with edges $v_i v_{i+1}$ for all $i \in [5m]$, where for convenience we set $v_{5m+1} = v_1$.
- For all $i \in [m]$, let u_1^i, u_2^i, u_3^i, and u_4^i be four new vertices; add the edges $u_1^i u_2^i$, $u_2^i u_3^i$, $u_3^i u_4^i$, and $u_4^i u_1^i$, so that the four vertices form a copy of C_4.
- For all $i, j \in [m]$, add a complete bipartite graph $K_{2,7}$ between $\{u_1^i, u_2^i\}$ and $\{v_{5(j-1)}, \ldots, v_{5j+1}\}$ if $r_i(j) = 1$ and between $\{u_3^i, u_4^i\}$ and $\{v_{5(j-1)}, \ldots, v_{5j+1}\}$ if $r_i(j) = 0$.

The lemma below is a slightly restated version of Lemma 3.3 in [128]; we state the lemma in a more precise way in order to be able to use it directly in our proof; the original proof of Rödl and Siggers establishes this more precise statement.

Lemma 6.2.4.

(a) Suppose that φ is a coloring of G^* in which the cycle C is monochromatic. If for some $i \in [m]$ we have $\varphi(u_1^i u_2^i) = \varphi(u_3^i u_4^i)$ and the color of $u_1^i u_2^i$ is different from that of C, then there is a copy of K_3 in G^* that is monochromatic under φ.

(b) For any edge of G^* of the form $e = u_1^i v_{5j+3}$, the graph $G^* - e$ has a coloring φ not containing a monochromatic triangle such that:
- *The cycle C is monochromatic.*
- *$\varphi(u_1^i u_2^i) = \varphi(u_3^i u_4^i)$ and the color of $u_1^i u_2^i$ differs from that of C.*
- *$\varphi(u_1^{i'} u_2^{i'}) = \varphi(u_3^{i'} u_4^{i'})$ and the color of $u_1^{i'} u_2^{i'}$ is the same as the color of C for all $i' \in [m] \setminus \{i\}$.*

(c) $v(G^*) = 9m$.

We are now ready to define our final graph G. Let S^+ and S^- be a safe positive and a safe negative signal sender for K_t, respectively. Also fix a safe indicator $I = I(K_t, F, 2, e)$, where F is a matching of size m satisfying the additional property given by Corollary 6.2.2.

Construction 6.2.5. Let $G = G(m, r_1, \ldots, r_m)$ be defined as follows.

- Let $G^* = G^*(m, r_1, \ldots, r_m)$ be the graph given by Construction 6.2.3.
- For each $i \in [m]$, connect the edges $u_1^i u_2^i$ and $u_3^i u_4^i$ in G^* by a copy of S^+.

- Let R_1 and R_2 be two copies of K_{t-3} on a new set of $2(t-3)$ vertices.
- Add a complete bipartite graph between $V(R_1) \cup V(R_2)$ and $V(G^*)$.
- Let f be an edge on a new pair of vertices.
- Join f with each edge of the cycle C in G^* by a copy of S^+.
- Connect f to every edge with at least one vertex in $V(R_1)$ by a copy of S^+ and to every edge with at least one vertex in $V(R_2)$ by a copy of S^-.
- Let Q be a copy of K_t on a new set of vertices.
- For each edge h of Q, join the matching $\{u_1^i u_2^i : i \in [m]\}$ and h by a copy of the indicator I.

We proceed as in Section 6.1 to show that G is 2-Ramsey for K_t but not 2-Ramsey for $K_t \cdot K_2$, that G contains certain edges whose removal destroys the Ramsey property of G, and that G has $O(m)$ vertices. This is the content of the next lemmas. Once we have shown the lemmas, Theorem 1.1.18 can be proved in the same way as Theorem 1.1.20, so we omit the details. For each of the next lemmas, assume that the tuple $(m, r_1, \cdots r_m)$ is fixed and write $G = G(m, r_1, \ldots, r_m)$.

Lemma 6.2.6. *The graph G is 2-Ramsey for K_t.*

Proof. Consider an arbitrary red/blue-coloring of G. Without loss of generality, assume that f is red. If there is a monochromatic copy of K_t in some signal sender or indicator, we are done. So we may assume that each signal sender and each indicator was given a K_t-free coloring. Thus, by property (S2) of each signal sender, all edges of the cycle C are red, every edge incident to a vertex of R_1 is red, every edge incident to a vertex of R_2 is blue, and, for all $i \in [m]$, the two edges $u_1^i u_2^i$ and $u_3^i u_4^i$ have the same color. If $u_1^i u_2^i$ is blue for some $i \in [m]$, then Lemma 6.2.4(a) implies that there is a monochromatic copy T of K_3 inside G^*. Then, depending on the color of T, either $V(T) \cup V(R_1)$ or $V(T) \cup V(R_2)$ induces a monochromatic copy of K_t. Hence, we may assume that $u_1^i u_2^i$ is red for all $i \in [m]$. The copies of the indicator I then force the t-clique Q to be monochromatic (by property (I2)). □

Lemma 6.2.7. *For any $i, j \in [m]$, the graph obtained from G by removing the edge $e = u_1^i v_{5j+3}$ is not 2-Ramsey for K_t.*

Proof. We describe a K_t-free coloring of $G - e$. We begin by taking the coloring of $G^* - e$ guaranteed by Lemma 6.2.4(b); we may assume that the cycle C is red. We assign the color red to the edge f and all edges incident to $V(R_1)$ and the color blue to all edges incident to $V(R_2)$. Since $G^* - e$ contains no monochromatic copy of K_3, there can be no monochromatic copy of K_t in this partial coloring of $G - e$. We next color the clique Q so that it is not monochromatic. Now, we can extend this partial K_t-free coloring of $G - e$ to all signal senders and indicators, giving each of them a K_t-safe coloring. This is possible since the signal edges of each signal sender are colored compatibly and the matching $\{u_1^i u_2^i : i \in [m]\}$ is not monochromatic (by Lemma 6.2.4(b)). By the safeness of the coloring of each gadget graph, it follows that there is no monochromatic copy of K_t in all of $G - e$. □

Lemma 6.2.8. *The graph G is not 2-Ramsey for $K_t \cdot K_2$.*

Proof. We begin by defining a coloring of G that contains a single monochromatic copy of K_t. We color the cycle C, the edge f, all edges incident to $V(R_1)$, each cycle $u_1^i u_2^i u_3^i u_4^i$ for $i \in [m]$, and the clique Q red. Also, we color all edges in the bipartite graph between $V(C)$ and $\bigcup_{i=1}^{m} \{u_1^i, u_2^i, u_3^i, u_4^i\}$ and all edges incident to $V(R_2)$ blue. Now, we extend this coloring to all signal senders and indicators, giving each gadget a K_t-safe coloring. Further, for the copies of I we use the coloring guaranteed by Corollary 6.2.2; this ensures that all edges incident to $V(Q)$ that are not part of the clique Q are blue.

We now claim that Q is the only monochromatic copy of K_t under this coloring. Since each signal sender and each indicator was given a K_t-safe coloring, we restrict our attention to the graph $G' = G^* \cup R_1 \cup R_2 \cup B \cup f \cup Q$, where B is the bipartite graph between $V(R_1) \cup V(R_2)$ and $V(G^*)$. Note that the red subgraph of $G' - (V(R_1) \cup V(Q))$ consists of the edge f, m isolated cycles of length four, and one isolated cycle of length $5m$, and this graph does not contain a copy of K_3. Then G' cannot contain a red copy of K_t other than Q. Similarly, the blue subgraph of $G' - V(R_2)$ is bipartite, so it contains no copy of K_3, which implies that G' contains no blue copy of K_t.

Hence, Q is the only monochromatic copy of K_t and all edges with exactly one endpoint in $V(Q)$ are blue, so there is no monochromatic copy of $K_t \cdot K_2$ in G. □

Lemma 6.2.9. *The graph G has $O(m)$ vertices.*

Proof. Observe that every vertex of G is contained in the graph G^*, or in a signal sender with signal edges $u_1^i u_2^i$ and $u_3^i u_4^i$ for some $i \in [m]$, or in a signal sender that has f as a signal edge, or in one of the copies of I. There are m signal senders with signal edges $u_1^i u_2^i$ and $u_3^i u_4^i$ for some $i \in [m]$; additionally, the cycle C has $5m$ edges and there are $2\binom{t-3}{2} + 2(t-3)(9m)$ edges incident to $V(R_1) \cup V(R_2)$, so there are $2\binom{t-3}{2} + 2(t-3)(9m) + 5m$ signal senders that have f as a signal edge. Finally, there are $\binom{t}{2}$ indicators with indicator edges in Q. We know that each signal sender has constantly many (depending only on t) vertices, each indicator has $O(m)$ vertices (by Corollary 6.2.2), and the graph G^* has $O(m)$ vertices (by Lemma 6.2.4(c)). Hence, all in all, there are $O(m)$ vertices in G. □

6.3 Concluding remarks

Many related problems remain open. We highlight a couple of them here.

Recall from the introduction that Clemens, Liebenau, and Reding [46] established the non-equivalence of K_t and $K_t \cdot K_2$ also for more than two colors. We believe that the analog of Theorem 1.1.18 should also hold in the multicolor setting, but we have been unable to show it. In particular, we have not been able to generalize Lemma 6.2.1 to more than two colors. Similarly, it should be possible to generalize Theorem 1.1.20 to the multicolor setting.

As a second direction, it would be interesting to explore other pairs of graphs H_1 and H_2 and determine whether a result similar to Theorems 1.1.18 and 1.1.20 is true, that is, whether there are many graphs that are minimal q-Ramsey for one but not q-Ramsey for the other. We consider pairs of 3-connected graphs to be a natural starting point here. Clemens, Liebenau, and Reding [46] showed that, if H_1 and H_2 are non-isomorphic 3-connected graphs, then H_1 and H_2 are not q-Ramsey equivalent for any $q \geq 2$. In addition, Siggers [140] proved that, if H is a non-complete 3-connected non-bipartite graph, then $\mathcal{M}_2(H)$ contains $2^{\Omega(n^2)}$ graphs on at most n vertices as $n \to \infty$ (recall from the introduction that, if H is bipartite, then no such result can hold).

Part II

Orthogonal Latin squares

Chapter 7

Enumerating orthogonal Latin squares

In this chapter, we turn our attention to orthogonal Latin squares and prove Theorem 1.2.3 and Corollaries 1.2.4 and 1.2.5. Before proving these results in Section 7.2, we introduce gerechte designs and discuss an equivalent formulation of the notion of mutually orthogonal Latin squares or gerechte desings that will be more convenient for our proof in Section 7.1; we also introduce our main tool, entropy, in the same section. We then provide explicit constructions of Latin squares with many orthogonal mates in Section 7.3, and close with some further remarks and open problems in Section 7.4. The results presented in this chapter are joint work with Shagnik Das and Tibor Szabó; the text is adapted, with mostly minor modifications, from [30] (`https://doi.org/10.1007/s10623-020-00771-6`, license: `https://creativecommons.org/licenses/by/4.0/`).

7.1 Designs and tools

In this section we will introduce the frameworks of gerechte designs and orthogonal arrays, in which we will prove a generalization of Theorem 1.2.3. We will also review some definitions and results regarding entropy that we shall require in our proofs.

7.1.1 Gerechte designs

Gerechte designs, defined below, are a special class of Latin squares introduced by Behrens in [14].

Definition 7.1.1. Let $[n]^2 = R_1 \sqcup \cdots \sqcup R_n$ be a partition of $[n]^2$ into n regions R_i such that $|R_i| = n$ for all $i \in [n]$. A *gerechte design of order n* with respect to this partition is a Latin square with the additional property that each symbol appears exactly once in each region R_i.

There are several natural examples of gerechte designs. For instance, if one takes the regions to be the n rows (or columns) of the $n \times n$ grid, a gerechte design is simply a Latin square. If $n = m^2$, and one partitions the grid into n subsquares of dimension $m \times m$, the corresponding gerechte designs are known as *Sudoku squares of order n*. Finally, given a Latin square L, define the regions $R_t = \{(i, j) : L(i, j) = t\}$ for all $t \in [n]$. A gerechte design with respect to this partition is an orthogonal mate of L.

It is natural to study orthogonality between Latin squares that are gerechte designs with respect to the same partition and, more generally, to consider systems of mutually orthogonal gerechte designs. Bailey, Cameron and Connelly [6] generalized the function $N(n)$ to the setting of gerechte designs, giving upper bounds on the size of a set of mutually orthogonal gerechte designs that are tight for some orders n.

The counting questions concerning Latin squares discussed in the introduction can also be generalized to gerechte designs, and our method will allow us to derive bounds in this broader setting. For this, note that an $n \times n$ square with entries in $[n]$ is a Latin square if and only if it is orthogonal (in the sense of Definition 1.2.1) to the square S_n, given by $S_n(i, j) = i$ for all $i, j \in [n]$, and its transpose. Similarly, it is not difficult to show that an $n \times n$ square with entries in $[n]$ is a gerechte design with respect to the regions $R_1, \ldots, R_n$ if and only if it is orthogonal to the squares S_n, S_n^T, and B, where B is given by $B(i, j) = t$ if $(i, j) \in R_t$. Note that, while the squares S_n and S_n^T are orthogonal to each other, the square B need not be orthogonal to either (that is, B need not be a Latin square).

7.1.2 Orthogonal arrays and nearly orthogonal arrays

When adding a square to a set of mutually orthogonal gerechte designs, we need to ensure three properties: that it is a Latin square, that it respects the regions of the design, and that it is orthogonal to the previous squares. For our proof, it will be helpful to use an equivalent but more symmetric formulation of mutually orthogonal gerechte designs, where these three properties all take the same form. We begin in the setting of mutually orthogonal Latin squares.

Definition 7.1.2. Let $x = (x_1, \ldots, x_{n^2})$ and $y = (y_1, \ldots, y_{n^2})$ be vectors in $[n]^{n^2}$. We say that x and y are *orthogonal* if, for all pairs $(s, t) \in [n]^2$, there exists a unique index ℓ such that $x_\ell = s$ and $y_\ell = t$. An *orthogonal array $OA(n, d)$* is an $n^2 \times d$ array A with entries in $[n]$ such that all pairs of its columns are orthogonal.

We note that in the literature orthogonal arrays are often defined more generally and Definition 7.1.2 describes what is known as an *orthogonal array with strength two and index one*. For the sake of simplicity, we omit the general definition and refer the reader to [89] for more about orthogonal arrays.

Given a k-MOLS $(L_1, \ldots, L_k)$ of order n, we can construct an orthogonal array $OA(n, k+2)$ by taking, for all $(i, j) \in [n]^2$, the vectors $[i, j, L_1(i, j), L_2(i, j), \ldots, L_k(i, j)]$ as rows of the orthogonal array (and ordering them lexicographically). Similarly, given an $n^2 \times (k+2)$ orthogonal array A, we can construct a k-MOLS of order n by setting $L_j(A(\ell, 1), A(\ell, 2)) = A(\ell, j+2)$ for all $1 \leq \ell \leq n^2$ and $1 \leq j \leq k$ (in fact, any two columns of the orthogonal array can be used to coordinatize the Latin squares; here we use the first two). Notice that distinct sequences of mutually orthogonal Latin squares correspond to distinct orthogonal arrays with first two columns $v_1 = [1, \ldots, 1, 2, \ldots, 2, \ldots, n, \ldots, n]^T$ and $v_2 = [1, 2, \ldots, n, 1, 2, \ldots, n, \ldots, 1, 2, \ldots, n]^T$, and hence the number of k-MOLS of order n is the same as the number of orthogonal arrays $OA(n, k+2)$ with first columns v_1 and v_2.

We now extend these ideas to mutually orthogonal gerechte designs. Let v_3 be a vector in $[n]^{n^2}$ with each integer in $[n]$ appearing n times. Note that v_3 determines a partition of the elements of $[n^2]$ (and thus $[n]^2$, after we fix a linear ordering of this set) into n equally-sized regions. From the equivalence between mutually orthogonal Latin squares and orthogonal arrays and the discussion at the end of Section 7.1.1, we can conclude that an $OA(n, k+2)$, whose first two columns are v_1 and v_2, and in which all other columns are also orthogonal to v_3, is equivalent to k mutually orthogonal gerechte designs with respect to the partition determined by v_3. For notational convenience, we add the column v_3 to the array and call the resulting structure an $n^2 \times (k+3)$ *nearly orthogonal array*.

Definition 7.1.3. Given $n \in \mathbb{Z}_{\geq 1}$ and $d \geq 3$, a *nearly orthogonal array* $NOA(n, d)$ is an $n^2 \times d$ array A with symbols $[n]$ such that:

(a) The first column is v_1 and the second column is v_2, as defined above.
(b) Each symbol in $[n]$ appears exactly n times in the third column v_3.
(c) For all $i \geq 4$, the ith column v_i is orthogonal to all other columns in A.

Again, it follows that the number of nearly orthogonal arrays $NOA(n, k+3)$ is equal to the number of sets of k mutually orthogonal gerechte designs with respect to the partition defined by v_3.

7.1.3 Entropy

The proof of our main result is based on entropy. This method has previously given good asymptotic upper bounds for similar problems; for instance, it is used in [123] to prove Brégman's Theorem on the permanent of a matrix (which yields an asymptotically tight upper bound on the number of Latin squares), in [109] to show an upper bound on the number of Steiner triple systems, later shown to be tight in [97], and in [80] to provide a simpler proof of Taranenko's result on the maximum number of transversals in a Latin square, also shown to be tight in the same paper; see also [113] for some further applications. In this section, we review some basic facts about entropy that will be used in our proof. For more on entropy, see [53].

Let X be a discrete random variable taking values in a given finite set $\mathcal{S}$, and let $p(x) = \mathbb{P}[X = x]$ for all $x \in \mathcal{S}$. The *(base e) entropy* of X is given by

$$H(X) = -\sum_{x\in\mathcal{S}} p(x)\log p(x) = -\mathbb{E}[\log p(X)],$$

where we adopt the convention that $0\log 0 = 0$. The entropy of X can be seen as a measure of the amount of information the random variable encodes. It is not difficult to show that

$$H(X) \leq \log|R(X)|, \tag{7.1.1}$$

where $R(X) = \{x \in \mathcal{S} : p(x) > 0\}$ is the range of the random variable, with equality if and only if X is uniformly distributed over $R(X)$.

This definition can be extended to multiple random variables in the natural way. We define the *joint entropy* of two random variables X and Y to be

$$H(X,Y) = -\sum_{x,y} p(x,y)\log p(x,y) = -\mathbb{E}[\log p(X,Y)],$$

where $p(x,y) = \mathbb{P}[X = x, Y = y]$ denotes the joint distribution of X and Y.

The *conditional entropy* of X given Y is defined to be

$$H(X|Y) = \mathbb{E}_{y\sim Y}[H(X|Y = y)] = \sum_{y} \mathbb{P}[Y = y]H(X|Y = y).$$

Conditional entropy gives us a way to measure how much additional information we expect to learn from X once we know the value of Y. It is a simple exercise to show that the joint entropy and the conditional entropy of several random variables satisfy the following equality, known as the chain rule:

$$H(X_1,\ldots,X_n) = \sum_{i=1}^{n} H(X_i|X_1,\ldots,X_{i-1}).$$

We end this section by outlining the basic idea behind counting proofs based on entropy. Suppose we want to obtain a bound on the size of a set $\mathcal{S}$. We sample an element $X \in \mathcal{S}$ uniformly at random. By the above discussion, we have $H(X) = \log|\mathcal{S}|$, and so an upper bound on the entropy $H(X)$ yields an upper bound on $|\mathcal{S}|$. To bound $H(X)$, we break up the random variable X into simpler random variables; the chain rule then allows us to consider these new random variables one at a time.

7.2 Proofs of our results

We now use the material from the previous section to prove Theorem 1.2.3 and its corollaries.

7.2.1 Bounding the number of extensions

In the language of orthogonal arrays, Theorem 1.2.3 is a statement about the number of ways to extend an orthogonal array by one column. We will in fact prove the following more general result, bounding the number of ways to extend a nearly orthogonal array by one column. Indeed, by inserting a copy of the first column in the third column (and reordering the rows if needed), one obtains a nearly orthogonal array from an orthogonal array.

Theorem 7.2.1. *Given $n \in \mathbb{Z}_{\geq 1}$ and $d \geq 3$, let A be a nearly orthogonal array $NOA(n, d)$. For each row $\ell \in [n^2]$, define*

$$\begin{aligned} r_\ell &= |\{s \neq \ell : A(s,1) = A(\ell,1) \text{ and } A(s,3) = A(\ell,3)\}|, \text{ and} \\ c_\ell &= |\{s \neq \ell : A(s,2) = A(\ell,2) \text{ and } A(s,3) = A(\ell,3)\}|. \end{aligned}$$

Then the logarithm of the number of ways to extend A to a nearly orthogonal array with $d+1$ columns is at most

$$\sum_{\ell=1}^{n^2} \int_0^1 \log(1 + (r_\ell + c_\ell)t^{d-1} + (n - r_\ell - c_\ell - 1)t^d)\, \mathrm{d}t.$$

Observe that in the gerechte design setting, for a cell $\ell \in [n]^2$, r_ℓ counts the number of other cells in the same row and region as ℓ, while c_ℓ counts the number of cells sharing the same column and region.

Before proving Theorem 7.2.1, we quickly derive Theorem 1.2.3.

Theorem 1.2.3. *For $0 \leq k \leq n-2$, the logarithm of the number of ways to extend a k-MOLS of order n to a $(k+1)$-MOLS is at most*

$$n^2 \int_0^1 \log(1 + (n-1)t^{k+2})\, \mathrm{d}t.$$

Proof of Theorem 1.2.3. As previously mentioned, a Latin square is a gerechte design with respect to the partition of the cells into their rows. A k-MOLS is thus equivalent to an $NOA(n, k+3)$ with $v_3 = v_1$, and an extension to a $(k+1)$-MOLS corresponds to adding a column to obtain an $NOA(n, k+4)$.

We can thus apply Theorem 7.2.1 with $d = k + 3 \geq 3$. For our choice of v_3, we have $r_\ell = n - 1$ and $c_\ell = 0$ for all $\ell \in [n^2]$. Substituting in these values, the bound on the number of extensions is

$$\sum_{\ell=1}^{n^2} \int_0^1 \log(1 + (n-1)t^{k+2})\,\mathrm{d}t = n^2 \int_0^1 \log(1 + (n-1)t^{k+2})\,\mathrm{d}t,$$

as required. □

We now proceed to the proof of the general theorem.

Proof of Theorem 7.2.1. Let A be as given, and let $\mathcal{S}$ denote the set of column vectors that are valid extensions for A. Our goal is to bound $|\mathcal{S}|$. We can assume $\mathcal{S} \neq \emptyset$, otherwise we are done. Let $X \in \mathcal{S}$ be chosen uniformly at random. Then $H(X) = \log|\mathcal{S}|$, and so it suffices to bound the entropy of X. We will expose the coordinates of X one at a time, using the chain rule to express the total entropy $H(X)$ as the sum of the conditional entropies from each successive reveal.

For $\ell \in [n^2]$, we denote the ℓth coordinate of X by X_ℓ and, given a permutation π of $[n^2]$, we reveal the coordinates in the order $X_{\pi(1)}, X_{\pi(2)}, \ldots, X_{\pi(n^2)}$. The chain rule then gives

$$\begin{aligned}
\log|\mathcal{S}| = H(X) &= \sum_{j=1}^{n^2} H(X_{\pi(j)} | X_{\pi(s)} : s < j) \\
&= \sum_{j=1}^{n^2} \mathbb{E}_{(x_{\pi(s)} \sim X_{\pi(s)} : s<j)} [H(X_{\pi(j)} | X_{\pi(s)} = x_{\pi(s)} : s < j)]. \qquad (7.2.1)
\end{aligned}$$

Given $x \in [n]^{n^2}$, let $R_{\pi(j)}(\pi, x) = R(X_{\pi(j)} | X_{\pi(s)} = x_{\pi(s)} : s < j)$ denote the range of this conditional random variable, that is,

$$R_{\pi(j)}(\pi, x) = \{y \in [n] : \exists Y \in \mathcal{S} : Y_{\pi(j)} = y \text{ and } \forall s < j, Y_{\pi(s)} = x_{\pi(s)}\},$$

and let $N_{\pi(j)}(\pi, x) = |R_{\pi(j)}(\pi, x)|$ be the size of this range. Note that $R_{\pi(j)}(\pi, x)$, and hence also $N_{\pi(j)}(\pi, x)$, only depends on the first $j - 1$ coordinates of x with respect to π; for $s \geq j$, the values $x_{\pi(s)}$ can be chosen arbitrarily without changing the range of the random variable.

Thus, by (7.1.1), we can bound the conditional entropy by $H(X_{\pi(j)} | X_{\pi(s)} = x_{\pi(s)} : s < j) \leq \log\big(N_{\pi(j)}(\pi, x)\big)$ for all $x \in [n]^{n^2}$. Substituting this into (7.2.1) and reordering the sum gives

$$\log|\mathcal{S}| \leq \sum_{j=1}^{n^2} \sum_{x \in [n]^{n^2}} \mathbb{P}[X = x] \log\big(N_{\pi(j)}(\pi, x)\big) = \sum_{\ell=1}^{n^2} \mathbb{E}_X[\log(N_\ell(\pi, x))].$$

This bound holds for any permutation π, and thus it holds when we average over the choice of π. As our calculations are more convenient in the continuous setting, we sample the uniformly random permutation of $[n^2]$ by choosing a vector $\alpha = (\alpha_\ell)_{\ell=1}^{n^2}$, with α_ℓ sampled uniformly at random from $[0,1]$ for all ℓ (all choices are made independently), and defining $\pi_\alpha = \pi$ to be such that $\alpha_{\pi(1)} > \alpha_{\pi(2)} > \cdots > \alpha_{\pi(n^2)}$. We then have

$$\begin{aligned}\log|\mathcal{S}| &\leq \mathbb{E}_\alpha\left[\sum_{\ell=1}^{n^2} \mathbb{E}_X[\log(N_\ell(\pi,x))]\right] = \sum_{\ell=1}^{n^2} \mathbb{E}_X[\mathbb{E}_\alpha[\log(N_\ell(\pi,x))]] \\ &= \sum_{\ell=1}^{n^2} \mathbb{E}_X\big[\mathbb{E}_{\alpha_\ell}[\mathbb{E}_{\alpha|\alpha_\ell}[\log(N_\ell(\pi,x))]]\big] \leq \sum_{\ell=1}^{n^2} \mathbb{E}_X\big[\mathbb{E}_{\alpha_\ell}[\log(\mathbb{E}_{\alpha|\alpha_\ell}[N_\ell(\pi,x)])]\big],\end{aligned}$$

where the last inequality follows from Jensen's inequality and the concavity of $y \mapsto \log y$. It therefore suffices to show that, for all $\ell \in [n^2]$ and all $x \in \mathcal{S}$, we have

$$\mathbb{E}_{\alpha_\ell}[\log(\mathbb{E}_{\alpha|\alpha_\ell}[N_\ell(\pi,x)])] \leq \int_0^1 \log(1 + (r_\ell + c_\ell)t^{d-1} + (n - r_\ell - c_\ell - 1)t^d)\,\mathrm{d}t. \tag{7.2.2}$$

We first estimate the inner expectation $\mathbb{E}_{\alpha|\alpha_\ell}[N_\ell(\pi,x)] = \mathbb{E}_\alpha[N_\ell(\pi,x)|\alpha_\ell]$. By the linearity of expectation, this is equal to $\sum_{y\in[n]} \mathbb{P}[y \in R_\ell(\pi,x)|\alpha_\ell]$. Unfortunately, it is not straightforward to determine whether or not $y \in R_\ell(\pi,x)$, and so we shall instead use a simple necessary condition that we call *availability*.

Recall that for the column x to be orthogonal to the ith column of A, the pairs $(A(s,i), x_s)$ must be distinct for all $s \in [n^2]$. Therefore, if for some symbol $y \in [n]$ there is some column $i \in [d]$ and previously exposed coordinate s such that $A(s,i) = A(\ell,i)$ and $x_s = y$, we cannot also have $x_\ell = y$. In this case we declare y *unavailable*, and observe that we must have $y \notin R_\ell(\pi,x)$. Otherwise, if there is no such column i and coordinate s, we say y is *available*. We now seek to compute the probability that a symbol y is available.

Fix a symbol $y \in [n]$. If y is the true value of the entry in the ℓth coordinate of x, then y cannot possibly have been ruled out by the previously exposed entries, and is thus available with probability 1.

Now suppose $y \in [n]$ is not the true value of x_ℓ. For each $i \in [d]$, since x is orthogonal to the ith column of A, there must be a unique entry $s_i(y) \neq \ell$ such that $x_{s_i(y)} = y$ and $A(s_i(y), i) = A(\ell, i)$. In order for y to be available, ℓ must be exposed before the entries in the set $S(y) = \{s_i(y) : i \in [d]\}$.

To find the probability of y being available, then, we need to compute the size of $S(y)$. Suppose for distinct columns $1 \leq i < j \leq d$ we had $s_i(y) = s_j(y)$. It then follows that $A(s_i(y), i) = A(\ell, i)$ and $A(s_i(y), j) = A(\ell, j)$, and thus the ith and jth columns cannot be orthogonal. Since A is nearly orthogonal, the only possibilities are $i \in \{1,2\}$ and $j = 3$ (by definition, all columns after the third column are orthogonal to all others, and the first two columns are orthogonal by construction).

Therefore $|S(y)| = d$, unless either $s_1(y) = s_3(y)$ or $s_2(y) = s_3(y)$. Note that these cannot happen simultaneously, as we have ruled out $s_1(y) = s_2(y)$, and thus in these cases we have $|S(y)| = d - 1$. There are r_ℓ choices of $s \neq \ell$ for which $A(s,1) = A(\ell,1)$ and $A(s,3) = A(\ell,3)$, and hence r_ℓ values y for which $s_1(y) = s_3(y)$. By orthogonality of x with the first column of A, these values are all distinct. Similarly, there are c_ℓ choices for y with $s_2(y) = s_3(y)$.

To summarize, there is one choice of y that is available with probability 1, there are $r_\ell + c_\ell$ choices of y that are available only if the ℓth coordinate is exposed before some fixed set of $d-1$ other coordinates, and the remaining $n - r_\ell - c_\ell - 1$ choices of y are available only if the ℓth coordinate precedes some d other coordinates.

A coordinate s is revealed after ℓ if $\alpha_s < \alpha_\ell$, which occurs with probability α_ℓ. Moreover, these events are independent for distinct coordinates, and so the probabilities in the latter two cases are α_ℓ^{d-1} and α_ℓ^d respectively. This gives

$$\begin{aligned}\mathbb{E}_{\alpha|\alpha_\ell}[N_\ell(\pi,x)] &= \sum_{y\in[n]} \mathbb{P}[y \in R_\ell(\pi,x)] \\ &\leq \sum_{y\in[n]} \mathbb{P}[y \text{ is available}] = 1 + (r_\ell + c_\ell)\alpha_\ell^{d-1} + (n - r_\ell - c_\ell - 1)\alpha_\ell^d.\end{aligned}$$

Since α_ℓ is uniformly distributed over $[0,1]$, substituting this into $\mathbb{E}_{\alpha_\ell}[\log(\mathbb{E}_{\alpha|\alpha_\ell}[N_\ell(\pi,x)])]$ results in (7.2.2), completing the proof. □

7.2.2 Estimating the integral

In order to apply Theorem 1.2.3, we need to understand the asymptotics of the bound it provides. In this next lemma, we show how to estimate the integral from the theorem.

Lemma 7.2.2. *Let* $2 \leq d \leq n$ *and* $I_d = \int_0^1 \log(1 + (n-1)t^d)\,\mathrm{d}t$. *Then*

$$I_d \leq \log\left(\frac{n-1}{e^d}\right) + \frac{d}{(n-1)^{1/d}} + \frac{3}{d(n-1)^{1/d}}.$$

Proof. Set $t_0 = (n-1)^{-1/d}$. Note that $(n-1)t^d < 1$ if and only if $t < t_0$. We have

$$\begin{aligned}I_d &= \int_0^1 \log(1 + (n-1)t^d)\,\mathrm{d}t \\ &= \int_0^{t_0} \log(1 + (n-1)t^d)\,\mathrm{d}t + \int_{t_0}^1 \log((n-1)t^d)\,\mathrm{d}t + \int_{t_0}^1 \log\left(1 + \frac{1}{(n-1)t^d}\right)\mathrm{d}t.\end{aligned}$$

We estimate the three integrals in turn:

$$\int_0^{t_0} \log(1+(n-1)t^d)\,\mathrm{d}t \le \int_0^{t_0} (n-1)t^d\,\mathrm{d}t = \frac{n-1}{d+1}t^{d+1}\Big|_0^{t_0} = \frac{t_0}{d+1},$$

where for the inequality we use the fact that $\log(1+x) \le x$ for all $x > -1$, and in the final equality we use $t_0^d = (n-1)^{-1}$,

$$\begin{aligned}\int_{t_0}^1 \log((n-1)t^d)\,\mathrm{d}t &= \int_{t_0}^1 \log(n-1) + d\log t\,\mathrm{d}t \\ &= (1-t_0)\log(n-1) + d(t\log t - t)\Big|_{t_0}^1 \\ &= (1-t_0)\log(n-1) + (t_0-1)d + t_0\log(n-1) \\ &= \log(n-1) + (t_0-1)d,\end{aligned}$$

where the penultimate equality again follows from $t_0^d = (n-1)^{-1}$, and

$$\int_{t_0}^1 \log\left(1+\frac{1}{(n-1)t^d}\right)\mathrm{d}t \le \int_{t_0}^1 \frac{1}{(n-1)t^d}\,\mathrm{d}t = \frac{-1}{(n-1)(d-1)} + \frac{t_0}{d-1}.$$

Hence, we have

$$\begin{aligned}I_d &\le \frac{t_0}{d+1} + \log(n-1) + (t_0-1)d - \frac{1}{(n-1)(d-1)} + \frac{t_0}{d-1} \\ &\le \log\left(\frac{n-1}{e^d}\right) + \frac{d}{(n-1)^{1/d}} + \frac{3}{d(n-1)^{1/d}},\end{aligned}$$

where we ignore the negative term and bound $\frac{1}{d+1} + \frac{1}{d-1}$ by $\frac{3}{d}$. □

Corollary 1.2.4 now follows easily from Theorem 1.2.3 and Lemma 7.2.2.

Corollary 1.2.4. *For every fixed $k \in \mathbb{Z}_{\ge 1}$, the maximum number of ways to extend a k-MOLS of order n to a $(k+1)$-MOLS is*

$$\left((1+o(1))\frac{n}{e^{k+2}}\right)^{n^2}.$$

Proof of Corollary 1.2.4. The lower bound comes from the average number of extensions of a k-MOLS, computed in (1.2.3). For the upper bound, Theorem 1.2.3 asserts that the logarithm of the number of extensions of a k-MOLS of order n is, in the notation of Lemma 7.2.2, at most $n^2 I_{k+2}$. By the lemma, this is bounded by

$$n^2\left(\log\left(\frac{n-1}{e^{k+2}}\right) + \frac{k+2}{(n-1)^{1/(k+2)}} + \frac{3}{(k+2)(n-1)^{1/(k+2)}}\right) \le n^2\left(\log\left(\frac{n-1}{e^{k+2}}\right) + \frac{k+4}{(n-1)^{1/(k+2)}}\right).$$

Since k is fixed as n tends to infinity, this is

$$n^2\left(\log\left(\frac{n-1}{e^{k+2}}\right)+o(1)\right)=n^2\log\left((1+o(1))\frac{n-1}{e^{k+2}}\right)=n^2\log\left((1+o(1))\frac{n}{e^{k+2}}\right),$$

giving the desired upper bound. □

Finally, we deduce our upper bound on the number of large sets of mutually orthogonal Latin squares.

Corollary 1.2.5. *As $n \to \infty$,*

(a) $\log L^{(k)}(n) \le \left(k\log n - \binom{k+2}{2} + 1 + k^2 n^{-1/(k+2)}\right)n^2$ *if $k = o(\log n)$,*
(b) $\log L^{(k)}(n) \le (c(\beta)+o(1))kn^2\log n$ *if $k = \beta\log n$, for fixed $\beta > 0$,*
(c) $\log L^{(k)}(n) \le \left(\frac{1}{2}+o(1)\right)(\log k - \log\log n)n^2\log^2 n$ *if $k = \omega(\log n)$,*

where in (b) we define $c(\beta) = 1 - \beta^{-1}\int_0^\beta x(1-e^{-1/x})\,\mathrm{d}x \in [0,1]$.

Proof of Corollary 1.2.5. We can build a k-MOLS by starting with the empty 0-MOLS, and extending it by one Latin square at a time. Theorem 1.2.3 bounds the number of possible extensions at each step, and so, in the notation of Lemma 7.2.2, we have

$$\log L^{(k)}(n) \le n^2\sum_{d=2}^{k+1} I_d. \tag{7.2.3}$$

We shall prove each part of the corollary by estimating this sum appropriately.

(i) By Lemma 7.2.2, we have

$$I_d \le \log(n-1) - d + \frac{d+2}{(n-1)^{1/d}}.$$

Hence, summing over d, we obtain

$$\sum_{d=2}^{k+1} I_d \le k\log(n-1) - \left(\binom{k+2}{2}-1\right) + \binom{k+4}{2}(n-1)^{-1/(k+2)},$$

from which the bound follows. Note that when $k = o(\log n)$, the final summand is a lower order term compared to k^2, and we can be generous in our estimation.

(ii) When $k = \Omega(\log n)$, that final term above becomes significant. While the previous upper bound remains valid, we obtain a better result through more careful calculation.

Rearranging the bound in Lemma 7.2.2 gives

$$\begin{aligned} I_d &\leq \log(n-1) - d\left(1-(n-1)^{-1/d}\right) + \frac{3}{d(n-1)^{1/d}} \\ &\leq \log(n-1) - d\left(1-e^{-\frac{\log(n-1)}{d}}\right) + \frac{3}{d}. \end{aligned}$$

Therefore we have

$$\sum_{d=2}^{k+1} I_d \leq k\log(n-1) - \sum_{d=2}^{k+1} d\left(1-e^{-\frac{\log(n-1)}{d}}\right) + \sum_{d=2}^{k+1}\frac{3}{d}.$$

The second sum, an error term, is at most $3\log(k+1)$. For the first sum, by making the substitution $x = \frac{d}{\log(n-1)}$, we observe that this is related to the estimation of the integral $\int x\left(1-e^{-1/x}\right)\mathrm{d}x$ by the Riemann sum with step size $1/\log(n-1)$. More precisely, we have

$$\frac{1}{\log(n-1)}\sum_{d=2}^{k+1}\frac{d}{\log(n-1)}\left(1-e^{-\frac{\log(n-1)}{d}}\right) = \int_{\frac{2}{\log(n-1)}}^{\frac{k+1}{\log(n-1)}} x\left(1-e^{-1/x}\right)\mathrm{d}x + o(1).$$

Making the necessary substitutions and letting n tend to infinity gives the claimed bound.

(iii) When $k = \omega(\log n)$, the above integral has an infinite domain, but we shall show that it still converges. First, we estimate $e^{-1/x}$ to observe that

$$x\left(1-e^{-1/x}\right) = 1 - \frac{1}{2x} + O(x^{-2}),$$

where the asymptotics are as x tends to infinity. Hence, when β tends to infinity,

$$\int_0^\beta x\left(1-e^{-1/x}\right)\mathrm{d}x = \int_0^\beta 1 - \frac{1}{2x} + O(x^{-2})\,\mathrm{d}x = \beta - \tfrac{1}{2}\log(\beta) + O(1).$$

The result then follows by substituting this into the statement of part (ii) with $\beta = \frac{k}{\log n}$; since the integrand $x\left(1-e^{-1/x}\right)$ is bounded and monotone increasing for large x, the Riemann sum remains a good approximation of the integral when $\beta \to \infty$.

□

7.3 Explicit constructions

Corollary 1.2.4 establishes the existence of Latin squares with several orthogonal mates. Given the numerous applications of orthogonal Latin squares, however, it is of great interest to have explicit constructions of such squares. For instance, in the closely related problem of counting transversals in Latin squares, Taranenko [145] showed that a Latin square of order n can have at most $\left((1+o(1))\frac{n}{e^2}\right)^n$ transversals. Glebov and Luria [80] later proved that Taranenko's bound is tight via a probabilistic construction. Recent results of Eberhard, Manners, and Mrazović [58] and Eberhard [57] give a constructive proof of the theorem of Glebov and Luria, providing explicit examples of Latin squares attaining this bound (in a very precise sense). They showed that the Cayley table of any abelian group G where $\sum_{g\in G} g = 0$ has $\left(\frac{2\pi n^2}{\sqrt{e}}+o(1)\right)\left(\frac{n}{e^2}\right)^n$ transversals.

To see the relation between transversals and orthogonal mates, observe that the n translates of any transversal in a Cayley table partition the Latin square. For each such partition into transversals, we can construct $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ distinct orthogonal mates by assigning distinct symbols in $[n]$ to the n transversals. The results of Eberhard, Manners, and Mrazović thus imply that Cayley tables of abelian groups have at least $\left(\sqrt{\frac{8\pi^3 n^5}{e}}+o(1)\right)\left(\frac{n^2}{e^3}\right)^n$ orthogonal mates. This lower bound is much smaller than the upper bound we would like to match, because this simple argument only counts orthogonal mates of a very special type. Here we describe a construction of MacNeish [114] that allows us to significantly improve this bound, even if we still fall slightly short of the true maximum number of orthogonal mates given by Corollary 1.2.4.

The *Kronecker product* of two Latin squares L_1 and L_2 of order n_1 and n_2 respectively is the Latin square $L_1 \otimes L_2$ of order $n_1 n_2$ given by $(L_1 \otimes L_2)((i_1, j_1), (i_2, j_2)) = (L_1(i_1, i_2), L_2(j_1, j_2))$. (Of course, the row and column indices and the symbols of $L_1 \otimes L_2$ can be seen as elements of $[n_1 n_2]$ after fixing an arbitrary bijection from $[n_1] \times [n_2]$.) For a Latin square L, we write $L^{\otimes k}$ to denote the k-fold product $\underbrace{L \otimes \cdots \otimes L}_{k \text{ times}}$.

Proposition 7.3.1. *Let L_1 and L_2 be Latin squares of order n_1 and n_2 that have q_1 and q_2 orthogonal mates respectively. Then the number of orthogonal mates of $L_1 \otimes L_2$ is at least $q_1 q_2^{n_1^2} \frac{(n_1 n_2)!}{n_1!(n_2!)^{n_1}}$.*

Proof. We will show how orthogonal mates of L_1 and L_2 can be combined in several ways to produce orthogonal mates of the product $L_1 \otimes L_2$. For this, it is again useful to view an orthogonal mate as an ordered partition of $L_1 \otimes L_2$ into disjoint transversals.

Further observe that $L_1 \otimes L_2$ can be partitioned into n_1^2 blocks of the form $L_1(i_1, i_2) \otimes L_2$ for $i_1, i_2 \in [n_1]$. Each of these is isomorphic to L_2, and thus admits q_2 orthogonal mates.

There are q_1 orthogonal mates of L_1, and thus $\frac{q_1}{n_1!}$ *unordered* partitions of L_1 into disjoint transversals, say $\{T_1, \ldots, T_{n_1}\}$. In the product $L_1 \otimes L_2$, this partitions the blocks into n_1 disjoint sets.

Let T_j be one of the transversals in this decomposition of L_1. The corresponding blocks $T_j \otimes L_2 = \{L_1(i_1, i_2) \otimes L_2 : (i_1, i_2) \in T_j\}$ then have all distinct symbols from $[n_1]$ in the first coordinate, and hence cover each symbol in $[n_1] \times [n_2]$ precisely n_2 times. To get a transversal of $L_1 \otimes L_2$, we can choose a transversal in each block $L_1(i_1, i_2) \otimes L_2$ and stitch them together. Furthermore, if we partition each block into transversals, stitching them together gives a partition of $T_j \otimes L_2$ into transversals of $L_1 \otimes L_2$.

There are $q_2^{n_1}$ ways to choose orthogonal mates for each of the n_1 blocks in $T_j \otimes L_2$. Here we keep the ordering, as that tells us which transversals in different blocks should be stitched together. This gives us an ordered partition of $T_j \otimes L_2$ into n_2 transversals of $L_1 \otimes L_2$, and so there are $\frac{q_2^{n_1}}{n_2!}$ unordered partitions of this set of blocks into transversals.

Making these choices for each T_j, we obtain a total of $\frac{q_1}{n_1!}\left(\frac{q_2^{n_1}}{n_2!}\right)^{n_1}$ partitions of $L_1 \otimes L_2$ into $n_1 n_2$ disjoint transversals, each of which can easily be shown to be distinct. To obtain an orthogonal mate, we can order these transversals arbitrarily, and thus obtain $q_1 q_2^{n_1^2} \frac{(n_1 n_2)!}{n_1!(n_2!)^{n_1}}$ mates, as claimed. □

In particular, this implies that powers of a single Latin square have many orthogonal mates.

Corollary 7.3.2. *Let L be a Latin square of order m with q orthogonal mates. Then $L^{\otimes k}$ is a Latin square of order m^k with at least $q^{\frac{m^{2k}-1}{m^2-1}}$ orthogonal mates.*

Proof. We proceed by induction. The statement is clearly true for $k = 1$. Suppose it holds for some $k \geq 1$. Then, by Proposition 7.3.1, with $L_1 = L^{\otimes k}$, $L_2 = L$, $n_1 = m^k$, $n_2 = m$, $q_1 = \frac{m^{2k}-1}{m^2-1}$, and $q_2 = q$, we know the number of mates of $L^{\otimes(k+1)}$ is at least

$$q^{\frac{m^{2k}-1}{m^2-1}} q^{m^{2k}} \frac{(m^{k+1})!}{(m^k)!(m!)^{m^k}} \geq q^{\frac{m^{2(k+1)}-1}{m^2-1}}.$$

□

If we take L to be the Cayley table of $\mathbb{Z}_3$, then we have $q = 6$. The k-fold Kronecker product of the Cayley table gives the Cayley table of the product group $\mathbb{Z}_3^k$, which by Corollary 7.3.2 has at least $(6^{1/8})^{3^{2k}-1}$ orthogonal mates. In the next corollary, we show that the constant in the base of the exponent can be made arbitrarily large at the cost of having a slightly less explicit construction.

Corollary 7.3.3. *For any $C > 0$, there are infinitely many orders n for which we can efficiently produce Latin squares with at least C^{n^2} orthogonal mates.*

Proof. Let m be such that $\frac{m}{2e^3} > C$. By Corollary 1.2.4, provided m is sufficiently large, there is a Latin square L of order m with $\left((1 + o(1))\frac{m}{e^3}\right)^{m^2} > C^{m^2}$ orthogonal mates; we can find such a square with a (finite) exhaustive search. By Corollary 7.3.2, we know that the Latin square $L^{\otimes k}$ of order $n = m^k$ has at least $C^{\frac{m^2(m^{2k}-1)}{m^2-1}} \geq C^{n^2}$ orthogonal mates. □

7.4 Concluding remarks and open questions

By bounding the number of extensions of a set of mutually orthogonal Latin squares, we obtained upper bounds on the number of k-MOLS when k grows with n. The obvious question is how tight these bounds are — can we find corresponding lower bounds? The constructions of Donovan and Grannell [56], valid for infinitely many values of n when $k \leq \sqrt{n}$, give lower bounds of the form $\log L^{(k)}(n) = \Omega\big(\gamma(k,n)n^2 \log n\big)$, where $\gamma(k,n) = \max\left\{\frac{\log k}{k^2 \log n}, \frac{1}{k^4}\right\}$. This is considerably smaller than our upper bounds in Corollary 1.2.5, and it would be of great interest to narrow the gap. One might hope to extend the lower bounds of Keevash [98], which were tight for constant k, but, as he notes in his paper, it is unclear how his methods could be used when k grows.

Aside from the enumeration of k-MOLS, there are several other related open problems, and we elaborate on these possible directions of study below.

Orthogonal mates We have bounded the maximum number of orthogonal mates a Latin square can have, but it is natural to ask if it is typical for a Latin square to have any orthogonal mates at all. Computational results in this direction are given in [35], [59], and [117]. His study of squares of small order led van Rees [148] to conjecture that, as $n \to \infty$, the proportion of Latin squares without orthogonal mates tends to one. On the other hand, having studied slightly larger orders, Wanless and Webb [150] suggested that the opposite may be true.

In (1.2.4), we saw that the results of Luria [113] and Keevash [98] imply that the average Latin square of order n has $\left((1+o(1))\frac{n}{e^3}\right)^{n^2}$ orthogonal mates. By Theorem 1.2.3 and Corollary 1.2.4, the same expression describes the maximum number of mates a square can have. Given the number of Latin squares (see (1.2.1)), it follows by double-counting orthogonal pairs that at least $\left((1+o(1))\frac{n}{e^2}\right)^{n^2}$ squares must have an orthogonal mate. Unfortunately, due to the lower order difference compared to (1.2.1) being in the base of the exponent, this falls short of resolving the question of whether or not most Latin squares have orthogonal mates.

Some evidence that this may not be straightforward to resolve is provided in [43], where Cavenagh and Wanless showed that, for almost all even n, there are at least $n^{(1-o(1))n^2}$ Latin squares of order n without a transversal, let alone an orthogonal mate. However, Ferber and Kwan [68] studied the analogous question in Steiner triple systems, and showed that almost all Steiner triple systems are almost resolvable. In the context of Latin squares, they suggest that their methods would show that almost all Latin squares have $(1-o(1))n$ disjoint transversals. Still, some new ideas would be needed to find the n disjoint transversals that form an orthogonal mate.

In Section 7.3 we showed, for any given $C > 0$, that we can, for infinitely many n, construct Latin squares of order n with at least C^{n^2} orthogonal mates. Given the existence of Latin squares with many more, namely $n^{(1+o(1))n^2}$, orthogonal mates, it is natural to seek better constructions.

Problem 3. Is there an explicit construction of a Latin square of order n with at least $n^{\Omega(n^2)}$ orthogonal mates?

In our product construction in Section 7.3, we only considered orthogonal mates consisting of very special kinds of transversals (those built within blocks, using transversals of the two factor squares). It is likely that these product squares have a much larger number of orthogonal mates, perhaps even close to the maximum possible.

We have also been vague with regards to what we mean by an explicit construction. As is customary in computer science, by *explicit* we mean there is an algorithm that constructs the Latin square in question in time polynomial in n. One can go further, and call a construction *strongly explicit* if each individual entry of the Latin square can be determined in polylogarithmic time. One can verify that our construction in the previous section is indeed strongly explicit. Yet one feels somewhat cheated, as in the first step of the construction we perform an exhaustive search to find an initial Latin square with many orthogonal mates (whose existence is guaranteed by random methods, see Corollary 1.2.4). It would be desirable to find constructions that are also "morally explicit" in the sense that they can be described mathematically, and in particular avoid any initial brute-force search. In this direction, it would be natural to investigate whether the Cayley tables of abelian groups G with $\sum_{g \in G} g = 0$ give examples of such Latin squares (cf. [57, 58]).

Affine and projective planes As mentioned earlier, $(n-1)$-MOLS of order n correspond to affine, and hence projective, planes of order n. Before we proceed, we briefly review the relevant concepts from the theory of affine and projective planes (for simplicity's sake, we omit the axiomatic definitions; for more background on the topic, see for example [9]). For an integer $n \geq 2$, a *projective plane of order n* consists of a set of points $\mathcal{P}$ and a set $\mathcal{L}$ of subsets of $\mathcal{P}$, called *lines*, such that $|\mathcal{P}| = n^2 + n + 1$, $|\ell| = n + 1$ for every $\ell \in \mathcal{L}$, and, for any distinct points in $\mathcal{P}$, there exists a unique line in $\mathcal{L}$ containing both. It is not difficult to check that any two lines in $\mathcal{L}$ intersect nontrivially and that $|\mathcal{L}| = n^2 + n + 1$. An *affine plane of order n* is a structure consisting of a set of points $\mathcal{P}'$ and a set of lines $\mathcal{L}'$ such that $|\mathcal{P}'| = n^2$, $|\ell| = n$ for every $\ell \in \mathcal{L}'$, and, for any distinct points in $\mathcal{P}'$, there exists a unique line in $\mathcal{L}'$ containing both. It is well known that in an affine plane we have $|\mathcal{L}'| = n^2 + n$ and the set $\mathcal{L}'$ can be partitioned into $n + 1$ *parallel classes*, each containing n lines, such that two lines intersect in a point if and only if they come from different parallel classes. Further, it is well known that an affine plane of order n can be extended to a projective plane of order n in a unique way as follows: for each parallel class, we extend the lines to a common new point, and add a line at infinity consisting of the new points. Similarly, we can obtain an affine plane of order n from a projective plane of order n with point set $\mathcal{P}$ and line set $\mathcal{L}$ by choosing an arbitrary line $\ell \in \mathcal{L}$ and setting $\mathcal{P}' = \mathcal{P} \setminus \ell$ and

$\mathcal{L}' = \mathcal{L} \setminus \{\ell\}$; as a result, a projective plane corresponds to at most $n^2 + n + 1$ different affine planes depending on the choice of ℓ.

We now briefly explain the connection between $(n-1)$-MOLS and affine planes. Given an $(n-1)$-MOLS $(L_1, \dots, L_{n-1})$ of order n, we add to it the two squares $L_n = S_n$ and $L_{n+1} = S_n^T$, where the square S_n is as defined in Section 7.1.1. Then let $\mathcal{P}' = \{(i,j) : i,j \in [n]\}$ and, for each $s \in [n+1]$, let the lines of the sth parallel class be given by $\{(i,j) \in \mathcal{P}' : L_s(i,j) = t\}$ for each $t \in [n]$. This constriction can also be reversed; note that each affine plane corresponds to at most $(n+1)!(n!)^{n+1} = e^{(1+o(1))n^2 \log n}$ distinct $(n-1)$-MOLS, since we can permute the parallel classes, and the lines within parallel classes.

Our knowledge of lower bounds for the number of $(n-1)$-MOLS of order n is even direr. It is known that such a system exists if n is a prime power, and it is conjectured that no such system exists for any other value of n. Further, it is believed that there is a unique (up to isomorphism) projective plane when n is a prime. As a step towards proving these conjectures, one could seek to bound the number of affine/projective planes from above, a problem raised by Hedayat and Federer [88].

In our definition of $L^{(k)}(n)$, we do not account for isomorphism. Thus, given a single $(n-1)$-MOLS, we can permute the symbols within each square of the corresponding $(n-1)$-MOLS to obtain $(n!)^{n-1}$ distinct $(n-1)$-MOLS. This gives a lower bound of $L^{(n-1)}(n) \geq (n!)^{n-1} = e^{(1-o(1))n^2 \log n}$ whenever n is a prime power. We remark that, for certain prime powers n, Kantor [93] and Kantor and Williams [94] provide algebraic constructions of superpolynomially many non-isomorphic projective planes of order n, but this contributes a lower order term in the above bound.

For an upper bound, Corollary 1.2.5 yields $L^{(n-1)}(n) \leq e^{\left(\frac{1}{2}+o(1)\right)n^2 \log^3 n}$. However, since a projective plane corresponds to a maximum possible set of mutually orthogonal Latin squares, it has a very restricted structure, and we can take advantage of this to obtain a better upper bound. Given a projective plane Π_n of order n, a subset H of its lines is called a *defining set* if Π_n is the unique projective plane containing H — that is, the lines in H determine the remaining lines in Π_n. Building on the work of Kahn [92], Boros, Szőnyi and Tichler [26] showed that every projective plane admits a small defining set.

Theorem 7.4.1 (Boros, Szőnyi and Tichler, 2005)**.** *Every projective plane of order n (for n sufficiently large) contains a defining set of size at most $22n \log n$.*

This immediately improves our upper bound.

Corollary 7.4.2. $L^{(n-1)}(n) \leq e^{(22+o(1))n^2 \log^2 n}$.

Proof. By Theorem 7.4.1, each projective plane of order n contains a set H of $22n\log n$ lines that determine the remaining ones uniquely. Each line is a subset of size $n+1$ of the n^2+n+1 points. Thus, there are $\binom{n^2+n+1}{n+1} = e^{(1+o(1))n\log n}$ possible lines and at most $\left(e^{(1+o(1))n\log n}\right)^{22n\log n}$ possible sets H and hence projective planes of order n.

As discussed earlier, a projective plane yields at most n^2+n+1 different affine planes, and each affine plane corresponds to at most $(n+1)!(n!)^{n+1} = e^{(1+o(1))n^2\log n}$ distinct $(n-1)$-MOLS. This contributes a lower order term, and so we can also bound $L^{(n-1)}(n) \le e^{(22+o(1))n^2\log^2 n}$. □

It would be of great interest to remove the extra logarithmic factor in the exponent of the upper bound, and thus reduce it log-asymptotically to the lower bound.

Conjecture 7.1.

$$L^{(n-1)}(n) = e^{O\left(n^2\log n\right)}.$$

As we obtain this many $(n-1)$-MOLS from a single projective plane, this would provide qualitative evidence in favor of the non-existence conjectures concerning projective planes. Note that a stronger result than Conjecture 7.1 (and a possible avenue of attack) would be to improve Theorem 7.4.1, which is not known to be tight. The best known lower bound for Theorem 7.4.1 is only linear in n, and, if every projective plane were to indeed contain a defining set of $O(n)$ lines, that would imply $L^{(n-1)}(n) = e^{\Theta(n^2\log n)}$.

Sudoku squares As mentioned in Section 7.1.1, Sudoku squares are a special class of gerechte designs of order n, where $n = m^2$, with the array partitioned into $m\times m$ subsquares in the natural way. After Golomb [81] asked about the existence of a pair of orthogonal Sudoku squares of order 9 (corresponding to the popular puzzle), systems of k *mutually orthogonal Sudoku squares (k-MOSS)* have been studied by several authors. This research has primarily sought to determine the largest k for which a k-MOSS of order n can exist; we refer the reader to [6, 95, 96, 110, 111, 122] for constructions and results in this direction.

One may ask the same counting questions as before for this restricted class of Latin squares, and these are relatively less well-studied. The number of Sudoku squares of order n is known to be $\left((1+o(1))\frac{n}{e^3}\right)^{n^2}$; the upper bound is shown independently by Luria [113] (using entropy) and Berend [15] (using Brégman's Theorem), while the matching lower bound is due to Keevash [98].

To enumerate k-MOSS for fixed k, we extend an idea of Keevash, defining a 4-uniform hypergraph H with vertices $V(H) = \{x_1, x_2, y_1, y_2, z_1^{(1)}, z_2^{(1)}, \dots, z_1^{(k)}, z_2^{(k)}\}$ and edges $\{x_1, x_2, y_1, y_2\}$, $\{x_1, x_2, z_1^{(i)}, z_2^{(i)}\}$, $\{y_1, y_2, z_1^{(i)}, z_2^{(i)}\}$, $\{x_1, y_1, z_1^{(i)}, z_2^{(i)}\}$ and $\{z_1^{(i)}, z_2^{(i)}, z_1^{(j)}, z_2^{(j)}\}$ for all $1 \le i < j \le k$. Letting $H(\sqrt{n})$ be the $(2k+4)$-partite 4-uniform hypergraph obtained by blowing each vertex up into $\sqrt{n}$ new vertices, it follows that a k-MOSS is equivalent to a decomposition of $H(\sqrt{n})$ into copies of H. For fixed k, the results of Luria and Keevash show there are $\left((1+o(1))\frac{n^k}{e^{\binom{k+3}{2}-3}}\right)^{n^2}$ such decompositions.

Our results allow us to bound the number of ways of extending a k-MOSS by an additional Sudoku square. Since each cell shares its row (or column) with $\sqrt{n}-1$ other cells from the same subsquare, we have $r_\ell = c_\ell = \sqrt{n}-1$ for each $\ell \in [n^2]$. Applying Theorem 7.2.1 shows that our upper bounds on k-MOSS coincide with our upper bounds on $(k+1)$-MOLS. In particular, the bound in (a) is once again tight, as it matches the average number of extensions of a k-MOSS.

Corollary 7.2. *(a) For all fixed k, the maximum number of extensions of a k-MOSS of order n to a $(k+1)$-MOSS is $\left((1+o(1))\frac{n}{e^{k+3}}\right)^{n^2}$.*

(b) For $k = k(n) \geq 0$, the logarithm of the number of k-MOSS is at most

(i) $\left((k+1)\log n - \binom{k+3}{2} + 3 + (k+1)^2 n^{-1/(k+3)} + o(1)\right)n^2$ *if* $k = o(\log n)$,

(ii) $(c(\beta)+o(1))kn^2\log n$ *if* $k = \beta\log n$, *for fixed* $\beta > 0$,

(iii) $\left(\frac{1}{2}+o(1)\right)(\log k - \log\log n)n^2\log^2 n$ *if* $k = \omega(\log n)$,

where $c(\beta)$, as in Corollary 1.2.5, is defined to be $1-\beta^{-1}\int_0^\beta x(1-e^{-1/x})\,\mathrm{d}x$.

Proof. A k-MOSS corresponds to a $NOA(n, k+3)$ with $r_\ell = c_\ell = \sqrt{n}-1$. Substituting these parameters into Theorem 7.2.1, the logarithm of the number of extensions of a k-MOSS is at most

$$\begin{aligned}
&n^2\int_0^1 \log\Big(1+2(\sqrt{n}-1)t^{k+2}+(n-2\sqrt{n}+1)t^{k+3}\Big)\,\mathrm{d}t\\
&= n^2\int_0^1 \log\Big(1+(n-1)t^{k+3}\Big)\,\mathrm{d}t + n^2\int_0^1 \log\left(1+\frac{2(\sqrt{n}-1)(1-t)t^{k+2}}{1+(n-1)t^{k+3}}\right)\mathrm{d}t.
\end{aligned}$$

The first integral is simply I_{k+3}, as evaluated in Lemma 7.2.2.

To bound the second integral, observe that

$$\begin{aligned}
\int_0^1 \log\left(1+\frac{2(\sqrt{n}-1)(1-t)t^{k+2}}{1+(n-1)t^{k+3}}\right)\mathrm{d}t &\leq \int_0^1 \frac{2(\sqrt{n}-1)(1-t)t^{k+2}}{1+(n-1)t^{k+3}}\,\mathrm{d}t\\
&\leq \int_0^1 \frac{(2\sqrt{n-1})t^{k+2}}{1+(n-1)t^{k+3}}\,\mathrm{d}t\\
&= \frac{2}{(k+3)\sqrt{n-1}}\log\Big(1+(n-1)t^{k+3}\Big)\Big|_0^1 = \frac{2\log n}{(k+3)\sqrt{n-1}}.
\end{aligned}$$

Thus, even if we sum up over all $k \in [n]$, the contribution from this second integral is a lower order error term. Hence our upper bound on the logarithm of the number of extensions of a k-MOSS is $n^2(I_{k+3}+o(1))$, and therefore we obtain the same enumeration as when extending a $(k+1)$-MOLS. □

Aside from the general lower bound of Keevash [98], we are not aware of any lower bounds on the number of k-MOSS. It would therefore be interesting to find lower bounds on the number of k-MOSS when k grows with n.

Problem 4. How tight are the upper bounds in Corollary 7.2(b)? That is, for $k = k(n)$ that grows with n, can we show the existence of many distinct k-MOSS?

Part III

Hyperplane coverings

Chapter 8

Subspace coverings with multiplicities over the binary field

This chapter represents joint work with Anurag Bishnoi, Shagnik Das, and Tamás Mészáros; the text is taken from [18, 19] (https://doi.org/10.1017/S0963548323000123, license: https://creativecommons.org/licenses/by/4.0/) with some modifications.

The objective of this chapter is to prove Theorem 1.3.2.

Theorem 1.3.2. *Let $k \geq 1$ and $n \geq d \geq 1$.*

(a) If $n \leq \lfloor \log_2 k \rfloor + d + 1$, then $f(n, k, d) = 2^d k - \left\lfloor \frac{k}{2^{n-d}} \right\rfloor$.
(b) If $k \geq 2$ and $n \geq \lfloor \log_2 k \rfloor + d + 1$, then $n + 2^d k - d - \log_2(2k) \leq f(n, k, d) \leq n + 2^d k - d - 2$.
(c) If $k \geq 2$ and $n > 2^{2^d k - d - k + 1}$, then $f(n, k, d) = n + 2^d k - d - 2$.

We begin with the proof of part (a), dealing with large multiplicities, in Section 8.1. In the process, we show a recursive bound that will allow us to restrict our attention to the $d = 1$ case in several of the following proofs, which will greatly simplify our presentation. We prove part (c) in Section 8.2, handling the case where the dimension of the ambient space grows quickly. A key step in the proof is showing the intuitive, yet surprisingly not immediate, fact that $f(n, k, d)$ is strictly increasing in n, as a result of which we will also be able to deduce the bounds in (b). In Section 8.3 we discuss the transition between the two extreme ranges.

Before we proceed, we introduce some definitions and notation that we will use in our proofs. To start with, it will be convenient to have some notation for affine hyperplanes. Given a nonzero vector $\vec{u} \in \mathbb{F}_2^n$, let $H_{\vec{u}}$ denote the hyperplane $\{\vec{x} : \vec{u} \cdot \vec{x} = 1\}$. The *(Hamming) weight* of a vector $\vec{x} \in \mathbb{F}_2^n$ is the number of nonzero entries it contains.

Next, it will sometimes be helpful to specify how many times the origin is covered. Hence, given integers $n \geq d \geq 1$ and $k > s \geq 0$, we say that a (k, d)-cover in $\mathbb{F}_2^n$ is a $(k, d; s)$-*cover* if it covers the origin exactly s times. We write $g(n,k,d;s)$ for the minimum possible size of a $(k,d;s)$-cover in $\mathbb{F}_2^n$ and call a cover *optimal* if it has this minimum size. Clearly, we have $f(n,k,d) = \min_{0 \le s < k} g(n,k,d;s)$, so any knowledge about this more refined function directly translates to our main focus of interest.

8.1 Covering with large multiplicity

In this section we prove Theorem 1.3.2(a), handling the case of large multiplicities.

8.1.1 The lower bound

To start with, we prove a general lower bound, valid for all choices of parameters, that follows from a simple double-counting argument. This establishes the lower bound of Theorem 1.3.2(a).

Lemma 8.1.1. *Let n,k,d,s be integers such that $n \ge d \ge 1$ and $k > s \ge 0$. Then*

$$g(n,k,d;s) \ge 2^d k - \left\lfloor \frac{k-s}{2^{n-d}} \right\rfloor .$$

In particular, $f(n,k,d) \ge 2^d k - \left\lfloor \frac{k}{2^{n-d}} \right\rfloor$.

Proof. Let $\mathcal{H}$ be an optimal $(k,d;s)$-cover of $\mathbb{F}_2^n$, so that we have $g(n,k,d;s) = |\mathcal{H}|$. We double-count the pairs $(\vec{x}, S)$ with $\vec{x} \in \mathbb{F}_2^n$, $S \in \mathcal{H}$, and $\vec{x} \in S$. On the one hand, every affine subspace $S \in \mathcal{H}$ contains 2^{n-d} points, and so there are $2^{n-d}|\mathcal{H}|$ such pairs. On the other hand, since every nonzero point is covered at least k times and the origin is covered s times, there are at least $(2^n - 1)k + s$ such pairs. Thus $(2^n-1)k + s \le 2^{n-d}|\mathcal{H}|$, and the claimed lower bound follows from solving for $|\mathcal{H}|$ and observing that $|\mathcal{H}| = g(n,k,d;s)$ is an integer. The bound on $f(n,k,d)$ is obtained by noticing that our lower bound on $g(n,k,d;s)$ is increasing in s, and is therefore minimized when $s = 0$. □

8.1.2 The upper bound construction

To prove the upper bound of Theorem 1.3.2(a), we must construct small (k,d)-covers. As a first step, we introduce a recursive method for $(k,d;s)$-covers that allows us to reduce to the $d = 1$ case.

Lemma 8.1.2. *For integers $n \ge d \ge 2$ and $k > s \ge 0$ we have*

$$g(n,k,d;s) \le g(n-d+1,k,1;s) + 2k(2^{d-1} - 1),$$

and, therefore,

$$f(n,k,d) \le f(n-d+1,k,1) + 2k(2^{d-1}-1).$$

Proof. We first deduce the recursive bound on $g(n,k,d;s)$. Let $S_0 \subseteq \mathbb{F}_2^n$ be an arbitrary $(n-d+1)$-dimensional (vector) subspace, and let $S_1, \dots, S_{2^{d-1}-1}$ be its affine translates, that, together with S_0, partition $\mathbb{F}_2^n$. For every $1 \le i \le 2^{d-1}-1$, partition $S_i \cong \mathbb{F}_2^{n-d+1}$ further into two subspaces, thereby obtaining a total of $2(2^{d-1}-1)$ affine subspaces of dimension $n-d$. We start by taking k copies of each of these affine subspaces. This gives us a multiset of $2k(2^{d-1}-1)$ subspaces, which cover every point outside S_0 exactly k times and leave the points in S_0 completely uncovered.

It thus remains to cover the points within S_0 appropriately. Since $(n-d)$-dimensional subspaces have relative codimension 1 in S_0, this reduces to finding a $(k,1;s)$-cover within $S_0 \cong \mathbb{F}_2^{n-d+1}$. By definition, we can find such a cover consisting of $g(n-d+1,k,1;s)$ subspaces. Adding these to our previous multiset gives a $(k,d;s)$-cover of $\mathbb{F}_2^n$ of size $g(n-d+1,k,1;s)+2k(2^{d-1}-1)$, as required.

Since $f(n,k,d) = \min_s g(n,k,d;s)$, and the recursive bound on $g(n,k,d;s)$ holds for each s, it naturally carries over to $f(n,k,d)$, giving $f(n,k,d) \le f(n-d+1,k,1) + 2k(2^{d-1}-1)$. □

Armed with this preparation, we can now resolve the problem for large multiplicities.

Proof of Theorem 1.3.2(a). The requisite lower bound, of course, is given by Lemma 8.1.1.

For the upper bound, we start by reducing to the case $d=1$. Indeed, suppose we already know the bound for $d=1$; that is, $f(n,k,1) \le 2k - \left\lfloor \frac{k}{2^{n-1}} \right\rfloor$ for all $k \ge 2^{n-2}$. Now, given some $n \ge d \ge 2$ and $k \ge 2^{n-d-1}$, by Lemma 8.1.2 we have

$$f(n,k,d) \le f(n-d+1,k,1)+2k(2^{d-1}-1) \le 2k - \left\lfloor \frac{k}{2^{n-d+1-1}} \right\rfloor + 2k(2^{d-1}-1) = 2^d k - \left\lfloor \frac{k}{2^{n-d}} \right\rfloor,$$

as required.

Hence, it suffices to prove the bound in the hyperplane case. We begin with the lowest multiplicity covered by part (a), namely $k = 2^{n-2}$. Consider the family $\mathcal{H}_0 = \{H_{\vec{u}} : \vec{u} \in \mathbb{F}_2^n, u_n = 1\}$, where we recall that $H_{\vec{u}} = \{\vec{x} : \vec{u}\cdot\vec{x} = 1\}$. Note that we then have $|\mathcal{H}_0| = 2^{n-1} = 2k = 2k - \left\lfloor \frac{k}{2^{n-1}} \right\rfloor$, and none of these hyperplanes covers the origin. Given nonzero vectors $\vec{x} = (\vec{x}', x)$ and $\vec{u} = (\vec{u}', 1)$ with $\vec{x}', \vec{u}' \in \mathbb{F}_2^{n-1}$ and $x \in \mathbb{F}_2$, we have $\vec{u}\cdot\vec{x} = 1$ if and only if $\vec{u}'\cdot\vec{x}' = 1 - x$. If $\vec{x}' \ne \vec{0}$, precisely half of the choices for $\vec{u}'$ satisfy this equation; if $\vec{x}' = \vec{0}$ (and thus necessarily $x = 1$), the equation is satisfied by all choices of $\vec{u}'$. Thus each nonzero point is covered at least 2^{n-2} times, and hence $\mathcal{H}_0$ is a $(2^{n-2}, 1)$-cover of the desired size.

To extend the above construction to the range $2^{n-2} \leq k < 2^{n-1}$, one can simply add an arbitrary choice of $k - 2^{n-2}$ pairs of parallel hyperplanes. The resulting family will have $2^{n-1} + 2\left(k - 2^{n-2}\right) = 2k = 2k - \left\lfloor \frac{k}{2^{n-1}} \right\rfloor$ elements, every nonzero point is covered at least k times, and the origin is covered $k - 2^{n-2} < k$ times.

Finally, suppose $k \geq 2^{n-1}$. Then we can write $k = a2^{n-1} + b$ for some $a \geq 1$ and $0 \leq b < 2^{n-1}$. We take $\mathcal{H}_1 = \{H_{\vec{u}} : \vec{u} \in \mathbb{F}_2^n \setminus \{\vec{0}\}\}$ to be the set of all affine hyperplanes avoiding the origin, of which there are $2^n - 1$. Moreover, for each nonzero $\vec{x}$, there are exactly 2^{n-1} vectors $\vec{u}$ with $\vec{u} \cdot \vec{x} = 1$, and so each such point is covered 2^{n-1} times by the hyperplanes in $\mathcal{H}_1$.

Now let $\mathcal{H}$ be the multiset of hyperplanes obtained by taking a copies of $\mathcal{H}_1$ and appending an arbitrary choice of b pairs of parallel planes. Each nonzero point is then covered $a2^{n-1} + b = k$ times, while the origin is only covered $b < 2^{n-1} \leq k$ times, and so $\mathcal{H}$ is a $(k, 1)$-cover. Thus,

$$f(n, k, 1) \leq |\mathcal{H}| = a(2^n - 1) + 2b = 2(a2^{n-1} + b) - a = 2k - \left\lfloor \frac{k}{2^{n-1}} \right\rfloor,$$

proving the upper bound. □

8.2 Covering high-dimensional spaces

In this section we turn our attention to the case where n is large with respect to k, with the aim of proving part (c) of Theorem 1.3.2. Furthermore, the results we prove along the way will allow us to establish the bounds in part (b) as well.

8.2.1 The upper bound construction

In this range, in contrast to the large multiplicity setting, it is the upper bound that is straightforward. This bound follows from the following construction, which is valid for the full range of parameters.

Lemma 8.2.1. *Let n, k, d be positive integers such that $n \geq d \geq 1$ and $k \geq 2$. Then*

$$f(n, k, d) \leq n + 2^d k - d - 2.$$

Proof. We start by resolving the case $d = 1$ and $k = 2$. The construction is the same as the one mentioned in the introduction for the hypercube $\{0, 1\}^n \subseteq \mathbb{R}^n$. More precisely, we consider the family of hyperplanes $\mathcal{H} = \{H_{\vec{e}_i} : i \in [n]\} \cup \{H_{\vec{1}}\}$, where $\vec{e}_i$ is the ith standard basis vector and $\vec{1}$ is the all-one vector. To see that this is a $(2, 1)$-cover of $\mathbb{F}_2^n$, note first that the hyperplanes all avoid the origin. Next, if we have a nonzero vector $\vec{x}$, it is covered by the hyperplanes $\{H_{\vec{e}_i} : i \in [n]\}$

as many times as it has nonzero entries. Thus, all vectors of Hamming weight at least two are covered twice or more. The only remaining vectors are those of weight one, which are covered once by $\{H_{\vec{e}_i} : i \in [n]\}$, but these are all covered for the second time by $H_{\vec{1}}$. Hence $\mathcal{H}$ is indeed a $(2, 1)$-cover, and is of the required size, namely $n + 1$.

Now we can extend this construction to the case $d = 1$ and $k \geq 3$ by simply adding $k - 2$ arbitrary pairs of parallel hyperplanes. The resulting family will be a $(k, 1; k - 2)$-cover (and hence, in particular, a $(k, 1)$-cover) of size $n + 2k - 3$, matching the claimed upper bound.

That leaves us with the case $d \geq 2$, which we can once again handle by appealing to Lemma 8.1.2. In conjunction with the above construction, we have

$$f(n, k, d) \leq f(n - d + 1, k, 1) + 2k(2^{d-1} - 1) \leq n - d + 1 + 2k - 3 + 2k(2^{d-1} - 1),$$

which simplifies to the required $n + 2^d k - d - 2$. □

8.2.2 Recursion, again

The upper bound in Lemma 8.2.1 is strictly increasing in n. Our next step is to show that this behavior is necessary — that is, the higher the dimension, the harder the space is to cover. Although intuitive, this fact turned out to be less elementary than expected, and our proof makes use of the probabilistic method.

Lemma 8.2.2. *Let n, k, d, s be integers such that $n \geq 2$, $n \geq d \geq 1$, and $k > s \geq 0$. Then*

$$g(n, k, d; s) \geq g(n - 1, k, d; s) + 1,$$

and thus

$$f(n, k, d) \geq f(n - 1, k, d) + 1.$$

Proof. Let $\mathcal{H}$ be an optimal $(k, d; s)$-cover of $\mathbb{F}_2^n$. To prove the lower bound on its size, we shall construct from it a $(k, d; s)$-cover $\mathcal{H}'$ of $\mathbb{F}_2^{n-1}$, which must comprise of at least $g(n - 1, k, d; s)$ subspaces. To obtain this cover of a lower-dimensional space, we restrict $\mathcal{H}$ to a random hyperplane $H \subseteq \mathbb{F}_2^n$ that passes through the origin. Since $\mathcal{H}$ is a $(k, d; s)$-cover of all of $\mathbb{F}_2^n$, it certainly covers $H \cong \mathbb{F}_2^{n-1}$ as well.

However, we require $\mathcal{H}'$ to be a $(k, d; s)$-cover of H, which must be built of affine subspaces of codimension d relative to H — that is, subspaces of dimension one less than those in $\mathcal{H}$. Fortunately, when intersecting the subspaces $S \in \mathcal{H}$ with a hyperplane, we can expect their dimension to decrease by one. The exceptional cases are when S is disjoint from H, or when S is contained in H. In the former case, S does not cover any points of H, and can therefore be discarded from $\mathcal{H}'$. In the latter case, we can partition S into two subspaces $S = S_1 \cup S_2$, where each S_i is of codimension d relative to H, and replace S with S_1 and S_2 in $\mathcal{H}'$. By making these changes, we obtain a family $\mathcal{H}'$ of codimension-d subspaces of H. Moreover, these subspaces cover the points of H exactly as often as those of $\mathcal{H}$ do, and thus $\mathcal{H}'$ is a $(k, d; s)$-cover of H.

When building this cover, though, we need to control its size. Let X denote the set of subspaces $S \in \mathcal{H}$ that are disjoint from H, and let Y denote the set of subspaces $S \in \mathcal{H}$ that are contained in H. We then have $|\mathcal{H}'| = |\mathcal{H}| - |X| + |Y|$. The objective, then, is to show that there is a choice of hyperplane H for which $|X| > |Y|$, in which case the cover $\mathcal{H}'$ we build is smaller.

Recall that H was a random hyperplane in $\mathbb{F}_2^n$ passing through the origin, which is to say it has a normal vector $\vec{u}$ chosen uniformly at random from $\mathbb{F}_2^n \setminus \{\vec{0}\}$. To compute the expected sizes of X and Y, we consider the probability that a subspace $S \in \mathcal{H}$ is either disjoint from or contained in H.

Let $S \in \mathcal{H}$ be arbitrary and suppose first that $\vec{0} \in S$. We immediately have $\mathbb{P}(S \in X) = 0$, as in this case $\vec{0} \in S \cap H$, so S and H cannot be disjoint. On the other hand, $\mathbb{P}(S \in Y) = \frac{2^d-1}{2^n-1}$, as we have $S \subseteq H$ exactly when the normal vector $\vec{u}$ is a nonzero element of the d-dimensional orthogonal complement, $S^\perp$, of S in $\mathbb{F}_2^n$.

In the other case, when $\vec{0} \notin S$, we can write S in the form $T + \vec{v}$, where $T \subseteq \mathbb{F}_2^n$ is an $(n-d)$-dimensional subspace such that $\vec{0} \in T$ and $\vec{v} \in \mathbb{F}_2^n \setminus T$. Then S is disjoint from H if and only if $\vec{u} \in T^\perp$ and $\vec{u} \cdot \vec{v} = 1$. Since $\vec{v} \notin T$, these are independent conditions, and so we have $\mathbb{P}(S \in X) = \frac{2^{d-1}}{2^n-1}$. Similarly, in order to have $S \subseteq H$, $\vec{u}$ must be a nonzero vector satisfying $\vec{u} \in T^\perp$ and $\vec{u} \cdot \vec{v} = 0$, and so $\mathbb{P}(S \in Y) = \frac{2^{d-1}-1}{2^n-1}$.

Now, using linearity of expectation, we have

$$\begin{aligned}
\mathbb{E}\left[|X| - |Y|\right] &= \sum_{S \in \mathcal{H}} \left(\mathbb{P}(S \in X) - \mathbb{P}(S \in Y)\right) \\
&= \sum_{S \in \mathcal{H}: \vec{0} \notin S} \left(\frac{2^{d-1}}{2^n - 1} - \frac{2^{d-1} - 1}{2^n - 1}\right) + \sum_{S \in \mathcal{H}: \vec{0} \in S} \left(0 - \frac{2^d - 1}{2^n - 1}\right) \\
&= \frac{|\{S \in \mathcal{H} : \vec{0} \notin S\}| - \left(2^d - 1\right) |\{S \in \mathcal{H} : \vec{0} \in S\}|}{2^n - 1} = \frac{|\mathcal{H}| - 2^d s}{2^n - 1},
\end{aligned}$$

where we used the fact that $\mathcal{H}$ is a $(k, d; s)$-cover, and thus $|\{S \in \mathcal{H} : \vec{0} \in S\}| = s$. We now apply the lower bound on $|\mathcal{H}|$ given by Lemma 8.1.1 to obtain

$$\mathbb{E}\left[|X| - |Y|\right] \geq \frac{2^d k - \left\lfloor \frac{k-s}{2^{n-d}} \right\rfloor - 2^d s}{2^n - 1} = \frac{2^d (k - s) - \left\lfloor \frac{k-s}{2^{n-d}} \right\rfloor}{2^n - 1} > 0.$$

Therefore, there must be a hyperplane H for which $|X| - |Y| \geq 1$. The corresponding cover of H thus has size at most $|\mathcal{H}| - 1$ but, as a $(k, d; s)$-cover of an $(n-1)$-dimensional space, has size at least $g(n-1,k,d;s)$. This gives $|\mathcal{H}|-1 \geq |\mathcal{H}'| \geq g(n-1,k,d;s)$, whence the required bound, $g(n,k,d;s) = |\mathcal{H}| \geq g(n-1,k,d;s)+1$.

Finally, we have $f(n,k,d) = \min_s g(n,k,d;s) \geq \min_s (g(n-1,k,d;s)+1) = f(n-1,k,d)+1$, proving the second part of the lemma. □

While Lemma 8.2.2 will be used in our proof of part (c) of Theorem 1.3.2, it also gives us what we need to prove the bounds in part (b).

Proof of Theorem 1.3.2(b). Lemma 8.2.1 gives us the upper bound, $f(n,k,d) \leq n+2^d k-d-2$, which is in fact valid for all $k \geq 2$ and $n \geq d \geq 1$.

When $n \geq \lfloor \log_2 k \rfloor + d + 1$, we can prove the lower bound, $f(n,k,d) \geq n+2^d k - d - \log_2(2k)$, by induction on n. For the base case, when $n = \lfloor \log_2 k \rfloor + d + 1$, we appeal to Lemma 8.1.1, which gives

$$f(n,k,d) \geq 2^d k - \left\lfloor \frac{k}{2^{n-d}} \right\rfloor = 2^d k = n + 2^d k - d - \lfloor \log_2 k \rfloor - 1 \geq n + 2^d k - d - \log_2(2k).$$

For the induction step we appeal to Lemma 8.2.2. Thus, using the induction hypothesis, for all $n > \lfloor \log_2 k \rfloor + d + 1$ we have

$$f(n,k,d) \geq f(n-1,k,d)+1 \geq n-1+2^d k - d - \log_2(2k) + 1 = n + 2^d k - d - \log_2(2k),$$

completing the proof. □

At this stage, all that remains to be proven from Theorem 1.3.2 is the lower bound of part (c), a task we undertake in the following subsections.

8.2.3 A coding theory connection

In Lemma 8.2.2, we proved a recursive bound on $g(n,k,d;s)$ that is valid for all values of s, the number of times the origin is covered. In this subsection, we establish the promised connection to coding theory, which is the key to our proof. Indeed, as observed in Corollary 8.2.6 below, it allows us to restrict our attention to only two feasible values of s.

We begin with $(k,1;0)$-covers of $\mathbb{F}_2^n$, showing that, in this binary setting, hyperplane covers that avoid the origin are in direct correspondence with linear codes of large minimum distance. In the setting of multiple blocking sets, a similar connection to coding theory was observed by Landjev and Rousseva [107], who used it in combination with the Hamming bound to show that Bruen's bound is far from being tight over $\mathbb{F}_2$ (see Theorems 1 and 7 in [107]). While they only showed that the difference between the minimum size of a multiple blocking set and Bruen's bound exceeds any given constant, we use the same approach to obtain concrete bounds for $g(n, k, 1; 0)$. Before we state the precise result, we briefly review the relevant coding-theoretic terminology (see for example [147, Chapter 20] for more basics related to codes). An *(n-dimensional) linear binary code (of length m) C* is an n-dimensional subspace of $\mathbb{F}_2^m$. The *distance* between two distinct elements $\vec{x}, \vec{y} \in C$ is given by the Hamming weight of $\vec{x} - \vec{y}$; the *minimum distance* of C is the smallest distance between two distinct elements of C. A *generator matrix* for C is an $m \times n$ matrix A such that $C = \{A\vec{x} : \vec{x} \in \mathbb{F}_2^n\}$. We remark that, in order to maintain consistency with earlier papers on hyperplane coverings, we deviate slightly from the standard coding-theoretic notation, where n usually stands for the length of the code, k for its dimension, and d for its minimum distance. In other words, our codes are $[m, n, k]$-codes as opposed to the more standard $[n, k, d]$-codes.

Proposition 8.2.3. *There exists a $(k, 1; 0)$-cover of $\mathbb{F}_2^n$ of cardinality m if and only if there exists an n-dimensional linear binary code of length m and minimum distance at least k.*

Proof. Let $\mathcal{H} = \{H_1, H_2, \dots, H_m\}$ be a $(k, 1; 0)$-cover of $\mathbb{F}_2^n$. Since none of the hyperplanes cover the origin, for each $i \in [m]$, H_i has to be described by the equation $\vec{u}_i \cdot \vec{x} = 1$ for some $\vec{u}_i \in \mathbb{F}_2^n \setminus \{\vec{0}\}$. Let A be the $m \times n$ matrix whose rows are $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m$. We claim that A is a generator matrix for a linear binary code of dimension n, length m, and minimum distance at least k. Since each $\vec{x} \in \mathbb{F}_2^n \setminus \{\vec{0}\}$ is covered by at least k of the hyperplanes, it follows that the vector $A\vec{x}$ has weight at least k, which in turn is equivalent to the vectors in the column space of A having minimum distance at least k. Indeed, any vector $\vec{y}$ in the column space of A can be expressed in the form $A\vec{w}$ for some $\vec{w} \in \mathbb{F}_2^n$. Thus, given two vectors $\vec{y}_1, \vec{y}_2$ in the column space, their difference is of the form $A(\vec{w}_1 - \vec{w}_2)$, where $\vec{x} = \vec{w}_1 - \vec{w}_2$ is nonzero. Hence this difference has weight at least k; i.e., the two vectors $\vec{y}_1$ and $\vec{y}_2$ have distance at least k. The fact that the weight of $A\vec{x}$ is at least $k \geq 1$ for any $\vec{x} \neq 0$ also implies that the kernel of A is trivial; therefore, the dimension of the column space of A, and hence of the binary code generated by A, is n.

Conversely, given a linear binary code of dimension n, length m, and minimum distance at least k, consider a generator matrix for it and let $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m$ be the rows of this matrix. By the same reasoning as above, the hyperplanes $H_{\vec{u}_1}, H_{\vec{u}_2}, \dots, H_{\vec{u}_m}$ form a $(k, 1; 0)$-cover of $\mathbb{F}_2^n$. □

Thus, the problem of finding a small $(k, 1; 0)$-cover of $\mathbb{F}_2^n$ corresponds to finding an n-dimensional linear code of minimum distance at least k and small length. This is a central problem in coding theory and, as such, has been extensively studied. We can therefore leverage known bounds to bound the function $g(n, k, 1; 0)$.

Corollary 8.2.4. *For all $k \geq 2$ and $n \geq 1$,*

$$g(n, k, 1; 0) \geq n + \left\lfloor \frac{k-1}{2} \right\rfloor \log_2 \left(\frac{2n}{k-1} \right).$$

Proof. Let $\mathcal{H}$ be an optimal $(k,1,0)$-cover and let $C \subseteq \mathbb{F}_2^m$ be the equivalent n-dimensional linear binary code of length $m = |\mathcal{H}|$ and minimum distance at least k, as described in Proposition 8.2.3. We can now appeal to the Hamming bound (see e.g. [147, Theorem 20.1]); we include the short proof for completeness. Consider the balls of radius $t = \left\lfloor \frac{k-1}{2} \right\rfloor$ around the 2^n points of C, that is, the set $\mathcal{B}_{\vec{x}} = \{\vec{z} \in \mathbb{F}_2^m : \vec{x} - \vec{z} \text{ has weight at most } t\}$ for each $\vec{x} \in C$; since the code has minimum distance k, these balls must be pairwise disjoint. As each ball has size $\sum_{i=0}^{t} \binom{m}{i}$, and the ambient space has size 2^m, we get

$$2^n \leq \frac{2^m}{\sum_{i=0}^{t} \binom{m}{i}}.$$

We bound the denominator from below by

$$\sum_{i=0}^{t} \binom{m}{i} \geq \binom{m}{t} \geq \left(\frac{m}{t}\right)^t \geq \left(\frac{n}{t}\right)^t = 2^{t \log_2\left(\frac{n}{t}\right)},$$

where the last inequality is valid provided $m \geq n$, as it must be. Thus we conclude

$$g(n,k,1;0) = |\mathcal{H}| = m \geq n + t \log_2\left(\frac{n}{t}\right) \geq n + \left\lfloor \frac{k-1}{2} \right\rfloor \log_2\left(\frac{2n}{k-1}\right). \qquad \square$$

Remark 8.2.5. Although it may seem that some of our bounds might be wasteful, we can deduce an almost matching upper bound when n is large with respect to k by considering a random collection of hyperplanes. We briefly sketch the argument. Let $\vec{u}_1, \ldots, \vec{u}_m \in \mathbb{F}_2^n$ be m vectors obtained by choosing each entry uniformly at random from $\mathbb{F}_2$, independently of all other entries; consider the hyperplanes given by $H_{\vec{u}_i}$ for $i \in [m]$ (for simplicity we allow the empty hyperplane with normal vector $\vec{0}$). It is not difficult to check that, for each $i \in [m]$ and every vector $\vec{y} \in \mathbb{F}_2^n \setminus \{\vec{0}\}$, the probability that $\vec{u}_i \cdot \vec{y} = 1$ is exactly $\frac{1}{2}$. Hence, for a fixed $\vec{y} \in \mathbb{F}_2^n \setminus \{\vec{0}\}$, the probability that $\vec{y}$ is covered fewer than k times is $\sum_{i=0}^{k-1} \binom{m}{i} \frac{1}{2^m} \leq \frac{k}{2^m} \binom{m}{k}$. Taking a union bound over all choices of $\vec{y} \in \mathbb{F}_2^n \setminus \{\vec{0}\}$ and substituting $m = n + k \log_2(3n)$ shows that the probability that some $\vec{y}$ is covered fewer than k times is less than 1, and thus there exists a choice of $\vec{u}_1, \ldots, \vec{u}_m$ that yields a strict k-cover of $\mathbb{F}_2^n$.

In coding theory this construction is well known, resulting in the so-called Gilbert-Varshamov bound for linear codes [79, 149]. The Gilbert-Varshamov bound and the Hamming bound provide lower and upper bounds on the maximum size (dimension) of a code with a given length and minimum distance. Narrowing the gap between these bounds remains an active area of research in coding theory (see for example [76, 91, 118, 137] and the references therein).

The above lower bound can be used to show that if n is large with respect to k and d then every optimal (k,d)-cover has to cover the origin many times. This corollary is critical to our proof of the upper bound.

Corollary 8.2.6. *If $n > 2^{2^d k - k - d + 1}$, then any optimal (k,d)-cover of $\mathbb{F}_2^n$ covers the origin at least $k-2$ times.*

Proof. Let $S_1, \ldots, S_m$ be an optimal (k, d)-cover, and, if necessary, relabel the subspaces so that $S_1, \ldots, S_s$ are the affine subspaces covering the origin. Suppose for a contradiction that $s \le k - 3$, and observe that if we delete the first $k - 3$ subspaces, each nonzero point must still be covered at least thrice, while the origin is left uncovered. That is, $S_{k-2}, S_{k-1}, \ldots, S_m$ forms a $(3, d; 0)$-cover of $\mathbb{F}_2^n$.

For each $k - 2 \le j \le m$, we can then extend S_j to an arbitrary hyperplane H_j that contains S_j and avoids the origin. Then $\{H_{k-2}, H_{k-1}, \ldots, H_m\}$ is a $(3, 1; 0)$-cover, and hence $m - k + 3 \ge g(n, 3, 1; 0)$.

By Corollary 8.2.4, this, together with the assumption $n > 2^{2^d k - k - d + 1}$, implies

$$f(n, k, d) = m \ge g(n, 3, 1; 0) + k - 3 \ge n + \log_2 n + k - 3 > n + 2^d k - k - d + 1 + k - 3 = n + 2^d k - d - 2,$$

which contradicts the upper bound from Lemma 8.2.1. □

Remark 8.2.7. Observe that Corollary 8.2.6 in fact gives us some stability for large dimensions. If $n = 2^{2^d k - k - d + \omega(1)}$, then the above calculation shows that any (k, d)-cover that covers the origin at most $k - 3$ times has size at least $n + 2^d k + \omega(1)$. Thus, when $n = 2^{2^d k - k - d + \omega(1)}$, any (k, d)-cover that is even close to optimal must cover the origin at least $k - 2$ times.

8.2.4 The lower bound

By Corollary 8.2.6, when trying to bound $f(n, k, d) = \min_s g(n, k, d; s)$ for large n, we can restrict our attention to $s \in \{k - 2, k - 1\}$. First we deal with the latter case.

Lemma 8.2.8. *Let n, k, d be positive integers such that $n \ge d \ge 1$. Then*

$$g(n, k, d; k - 1) = n + 2^d k - d - 1.$$

Proof. To prove the statement, we will show that, for all positive integers n, k, d with $n \ge d \ge 1$, we have $g(n + 1, k, d; k - 1) = g(n, k, d; k - 1) + 1$. Combined with the simple observation that $g(d, k, d; k - 1) = 2^d k - 1$ for all $k \ge 1$, since when $d = n$ we are covering with individual points, this fact will indeed imply the desired result.

By Lemma 8.2.2 we know that $g(n + 1, k, d; k - 1) \ge g(n, k, d; k - 1) + 1$. For the other inequality, consider an optimal $(k, d; k-1)$-cover $\mathcal{H}$ of $\mathbb{F}_2^n$. For every $S \in \mathcal{H}$, let $S' = S \times \{0, 1\}$, which is a codimension-d affine subspace of $\mathbb{F}_2^{n+1}$, and let S_0 be any $(n + 1 - d)$-dimensional affine subspace of $\mathbb{F}_2^{n+1}$ that contains the vector $(0, \ldots, 0, 1)$ but avoids the origin. We claim that $\mathcal{H}' = \{S' : S \in \mathcal{H}\} \cup \{S_0\}$ is a $(k, d; k - 1)$-cover of $\mathbb{F}_2^{n+1}$. Indeed, for all $S \in \mathcal{H}$, a point of the form $(\vec{x}, t)$ is covered by S' if and only if $\vec{x}$ is covered by S. Hence, the collection $\{S' : S \in \mathcal{H}\}$

covers $\vec{0}$ exactly $k-1$ times and each point of the form $(\vec{x}, t)$ with $\vec{x} \neq \vec{0}$ at least k times. Finally, the point $(\vec{0}, 1)$ is covered $k-1$ times by the collection $\{S' : S \in \mathcal{H}\}$ and once by the subspace S_0, so it is also covered the correct number of times. Hence $\mathcal{H}'$ is indeed a $(k, d; k-1)$-cover of of size $|\mathcal{H}| + 1$, and so the second inequality follows. □

Remark 8.2.9. Recall that in the special case $d = 1$, the equality $g(n, k, 1; k-1) = n + 2k - 2$ also follows from (the proof of) [133, Theorem 1.5].

The proof of Theorem 1.3.2(c) is now straightforward.

Proof of Theorem 1.3.2(c). The upper bound is given by Lemma 8.2.1. For the lower bound, first observe that for any valid choice of the parameters, we have $g(n, k, d; s+1) \leq g(n, k, d; s) + 1$, as adding any subspace containing the origin to a $(k, d; s)$-cover yields a $(k, d; s+1)$-cover. Then, by Corollary 8.2.6 and Lemma 8.2.8, we obtain

$$f(n, k, d) = \min\{g(n, k, d; k-2), g(n, k, d; k-1)\} \geq g(n, k, d; k-1) - 1 = n + 2^d k - d - 2,$$

as desired. □

8.3 The transition

Parts (a) and (c) of Theorem 1.3.2 determine the function $f(n, k, d)$ exactly in the two extreme ranges of the parameters — when k is exponentially large with respect to n, and when n is exponentially large with respect to k. As remarked upon after the statement of Theorem 1.3.2, we know that in the former case, the bound on k is best possible. For part (c), however, we believe the upper bound of Lemma 8.2.1 should be tight for much smaller values of n as well.

In this section we explore the transition between these two ranges, with an eye towards better understanding when this upper bound becomes tight. As we saw in Lemma 8.1.2, for our upper bounds we can generally reduce to the hyperplane setting, and so we shall focus on the $d = 1$ case in this section. Recall that, to simplify notation, when $d = 1$ we often refer to a $(k, 1)$-cover as a k-cover and write $f(n, k)$ instead of $f(n, k, 1)$.

In this hyperplane setting, the upper bound of Lemma 8.2.1, valid for all $n \geq 1$ and $k \geq 2$, has the simple form $n + 2k - 3$. Given some fixed k, suppose the bound is tight for some n_0; that is, $f(n_0, k) = n_0 + 2k - 3$. The recursion of Lemma 8.2.2 implies $f(n, k) \geq f(n-1, k) + 1$ for all $n \geq 2$, and so these two bounds together imply $f(n, k) = n + 2k - 3$ for all $n \geq n_0$. Hence, for every k, there is a well-defined threshold $n_0(k)$ such that $f(n, k) = n + 2k - 3$ if and only if $n \geq n_0(k)$. Theorem 1.3.2(c) shows $n_0(k) \leq 2^k + 1$, and our goal now is to explore the true behavior of this threshold.

8.3.1 The diagonal case

As a natural starting point, one might ask what lower bound we can provide for $n_0(k)$. From our previous results, in particular Theorem 1.3.2(a), we have seen that $f(n,k)$ behaves differently when k is large compared to n. We therefore know the upper bound of Lemma 8.2.1 is not tight when $k \geq 2^{n-2}$ or, equivalently, we know $n_0(k) > \log_2 k + 2$. However, the following construction, valid when $k \geq 4$, shows that we can improve upon Lemma 8.2.1 for considerably larger values of n as well.

Proposition 8.3.1. *For all $k \geq 4$, we have $f(k,k) \leq 3k - 4$. As a consequence, $n_0(k) \geq k+1$.*

Proof. To prove the upper bound, we must construct a k-cover $\mathcal{H}$ of $\mathbb{F}_2^k$ of size $3k-4$. Letting $\vec{e}_i$ denote the ith standard basis vector and $\vec{1}$ the all-one vector, we take $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3$, where $\mathcal{H}_1 = \left\{H_{\vec{e}_i} : i \in [k]\right\}$, $\mathcal{H}_2 = \left\{H_{\vec{1}-\vec{e}_i} : i \in [k]\right\}$, and $\mathcal{H}_3$ consists of $k-4$ copies of the hyperplane with equation $\vec{1} \cdot \vec{x} = 0$. Then $\mathcal{H}$ has size $3k-4$, while the only planes containing the origin are those in $\mathcal{H}_3$. Thus it only remains to verify that each nonzero point is covered at least k times.

Given a nonzero point $\vec{y} = (y_1, \ldots, y_n)$, let its weight be w. We then see that $\vec{y}$ is covered w times by the planes in $\mathcal{H}_1$. Next, observe that $\left(\vec{1} - \vec{e}_i\right) \cdot \vec{y}$ is equal to w if $y_i = 0$, and is equal to $w-1$ otherwise. Hence, if w is odd, then $\vec{y}$ is covered by $k-w$ planes in $\mathcal{H}_2$, and is thus covered at least k times by $\mathcal{H}$.

On the other hand, if w is even, then $\vec{y}$ is covered w times by the planes in $\mathcal{H}_2$. However, in this case $\vec{1} \cdot \vec{y} = 0$, and so $\vec{y}$ is covered $k-4$ times by $\mathcal{H}_3$ as well. In total, then, $\vec{y}$ is covered $2w + k - 4$ times. As $\vec{y}$ is a nonzero vector of even weight, we must have $w \geq 2$, and hence $\vec{y}$ is covered at least k times in this case as well.

In conclusion, we see that $\mathcal{H}$ forms a k-cover of $\mathbb{F}_2^k$, and thus $f(k,k) \leq |\mathcal{H}| = 3k-4$. As this is smaller than the upper bound of Lemma 8.2.1, it follows that $n_0(k) \geq k+1$. □

8.3.2 Initial values

This still leaves us with a large range of possible values for $n_0(k)$: our lower bound is linear, while our upper bound is exponential. To get a better feel for which bound might be nearer to the truth, we next decided to take a closer look at $f(n,k)$ for small values of the parameters.

To be able to compute a number of these values efficiently, it helped to appeal to our recursive bounds. Lemma 8.2.2 already restricts the behavior of $f(n,k)$ as n changes, showing that the function must be strictly increasing in n. It is also very helpful to understand how $f(n,k)$ responds to changes in k: as the following lemma shows, there is even less flexibility here.

Lemma 8.3.2. *For all $n \geq 1$ and $k \geq 2$, we have $f(n, k-1) + 1 \leq f(n, k) \leq f(n, k-1) + 2$.*

Proof. For the lower bound, observe that, given a k-cover of size $f(n, k)$, removing a hyperplane covering the origin (or, if no such plane exists, an arbitrary plane) leaves us with a $(k-1)$-cover, and thus $f(n, k-1) \leq f(n, k) - 1$.

For the upper bound, given a $(k-1)$-cover of size $f(n, k-1)$, we can add an arbitrary pair of parallel hyperplanes to obtain a k-cover. Thus $f(n, k) \leq f(n, k-1) + 2$. □

Thus, if we know the value of $f(n, k-1)$, there are only two possible values for $f(n, k)$. This becomes even more powerful when used in combination with Lemma 8.2.2, which guarantees $f(n, k) \geq f(n-1, k) + 1$. Hence, in case we have $f(n-1, k) = f(n, k-1) + 1$, the only possible value for $f(n, k)$ is $f(n, k-1) + 2$.

Although this may seem a very conditional statement, this configuration occurs quite frequently, as one can see in Table 8.1 below, and allows us to deduce several values of $f(n, k)$ for free. This observation, together with our previous bounds (and noting that $f(n, 2) = n + 1$), allows us to almost completely determine $f(n, k)$ for $n \leq 6$. We were able to fill in the few outstanding values through a computer search (using SageMath [146] and Gurobi [85]).[1]

n \ k	3	4	5	6	7	8	9	10	11	12	13	14	15	16	···
3	6*	7	9	11	13	14	16	18	20	21	23	25	27	28	···
4	7*	8	10	12	14	15	17	19	21	23	25	27	29	30	···
5	8*	10*	11	13	15	16	18	20	22	24	26	28	30	31	···
6	9*	11*	13*	14	16	18	20	22	23	25	27	29	31	32	···

TABLE 8.1: $f(n, k)$ for $3 \leq n \leq 6$: values in green come from Theorem 1.3.2(a), values in blue are a consequence of the recursive bounds, values in orange follow from Proposition 8.3.1, and values in red were obtained by a computer search. An asterisk denotes values equal to the upper bound of Lemma 8.2.1; that is, where $n \geq n_0(k)$.

8.3.3 The extended Golay code

We see from Table 8.1 that $n_0(k) = k + 1$ for $k \in \{4, 5\}$, leading some credence to the belief that the construction from Proposition 8.3.1 is perhaps indeed the last time (for a fixed k, as n increases) the upper bound from Lemma 8.2.1 can be improved. However, we can once again exploit the coding theory connection of Proposition 8.2.3 to show that this is not always the case.

[1] Some of these values we first proved by hand, via direct case analysis. However, as we do not see any more broadly applicable generalization of the arguments therein, we have omitted these proofs.

The extended binary Golay code is a 12-dimensional code of length 24 and minimum distance 8 (see e.g. [147, Chapter 20] for a definition of this code). By Proposition 8.2.3, this code is equivalent to an $(8, 1; 0)$-cover of $\mathbb{F}_2^{12}$ of size 24, thus implying that $f(12, 8) \leq 24$, whereas the upper bound given by Lemma 8.2.1 is 25. Furthermore, we see in Table 8.1 that $f(6, 8) = 18$. By repeated application of Lemma 8.2.2, we must have $f(12, 8) \geq f(6, 8) + 6$, and thus $f(12, 8) = 24$. Moreover, there must be equality in every step of the recursion, and thus $f(n, 8) = n + 12$ for $6 \leq n \leq 12$.

This result, coupled with the techniques described previously, allows us to extend Table 8.1 to include values for $7 \leq n \leq 12$ and $3 \leq k \leq 10$. These new values are depicted in Table 8.2 below. We see that the equality $n_0(k) = k + 1$ persists for $k = 6, 7$ until the Golay construction comes into existence. In light of Lemma 8.3.2, this ensures $n_0(k) \geq k + 2$ for $8 \leq k \leq 11$.

n \ k	3	4	5	6	7	8	9	10
6	9*	11*	13*	14	16	18	20	22
7	10*	12*	14*	16*	17	19	21	23
8	11*	13*	15*	17*	19*	20	22	24
9	12*	14*	16*	18*	20*	21	23	25
10	13*	15*	17*	19*	21*	22	24	26
11	14*	16*	18*	20*	22*	23	25	27
12	15*	17*	19*	21*	23*	24	26	28
⋮	⋮	⋮	⋮	⋮	⋮			

TABLE 8.2: More values of $f(n, k)$: green represents values coming from Theorem 1.3.2(a), red represents values obtained through computer computations, blue represents values obtained from other values by the recursive bounds, orange represents values obtained by Proposition 8.3.1 and recursion, and cyan represents values obtained by the Golay code construction and its recursive consequences. An asterisk denotes values attaining the upper bound of Lemma 8.2.1; that is, where $n \geq n_0(k)$.

This begs the question of what happens for larger values of k. Does the gap $n_0(k) - k$ continue to grow? Does the threshold return to $k + 1$ at a later point? Unlike the construction in Proposition 8.3.1, the Golay code yields a sporadic construction, which we have not been able to generalize. Furthermore, this is known as a particularly efficient code, and we are not aware of any other code whose parameters lead to an improvement on Proposition 8.3.1. Hence, we are leaning towards the second possibility – not strongly enough, perhaps, to conjecture it as the truth, but enough to pose it as a question.

Question 7. Do we have $n_0(k) = k + 1$ for all $k \geq 12$?

To answer Question 7, we need to determine the value of $f(k + 1, k)$. For an affirmative answer, we need to show $f(k + 1, k) = 3k - 2$, while a negative answer would follow from a construction showing $f(k+1,k) \leq 3k-3$. What could such a construction look like? If we retrace the proof of Theorem 1.3.2(c), we see that any k-cover of $\mathbb{F}_2^{k+1}$ that covers the origin at least $k-2$ times must have size at least $3k-2$. Hence, any construction negating Question 7 must cover the origin at most $k-3$ times.

While this seemingly contradicts Corollary 8.2.6, recall that we needed n to be exponentially large with respect to k to draw that conclusion. Without this condition, the Hamming bound on codes with large distance is not strong enough to provide the requisite lower bound on $f(n,k)$. Indeed, the Gilbert-Varshamov bound, discussed in Remark 8.2.5, shows that a random collection of $k+O(\log_2 k)$ hyperplanes forms a 3-cover of $\mathbb{F}_2^{k+1}$ with high probability. Adding $k-3$ arbitrary pairs of parallel planes then gives a k-cover of size $3k+O(\log_2 k)$ that only covers the origin $k-3$ times. Thus, we can find numerous k-covers that are asymptotically optimal, and we cannot hope for any strong stability when n and k are comparable.

8.4 Concluding remarks

In this chapter, we investigated the minimum number of affine subspaces of a fixed codimension needed to cover all nonzero points of $\mathbb{F}_2^n$ at least k times, while only covering the origin at most $k-1$ times. We were able to determine the answer precisely when k is large with respect to n, or when n is large with respect to k, and provided asymptotically sharp bounds for the range in between these extremes. In this final section, we highlight some open problems and avenues for further research.

Bounding the threshold In the previous section, we raised the question of determining the threshold $n_0(k)$ beyond which the result of Theorem 1.3.2(c) holds. Although our proof requires n to be exponentially large with respect to k, our constructions suggest the threshold might, with limited exceptions, be as small as $k+1$.

It is quite possible that solving Question 7 will require improving the classic bounds on the length of binary codes of large minimum distance, and will therefore perhaps be quite challenging. However, there is plenty of scope to attack the problem from the other direction, and aim to reduce the exponential upper bound on $n_0(k)$.

Our strategy was to prove the lower bound for $g(n,k,1;k-1)$ and $g(n,k,1;k-2)$, using the recursive bounds. By removing hyperplanes covering the origin, we could reduce the remaining cases to $g(n,3,1;0)$, for which, when n is large, the coding theory connection provides a large enough lower bound.

There are two natural ways to improve this argument. The first would be to extend the values s for which we directly prove the lower bound on $g(n,k,1;s)$. For instance, if we could show that $g(n, k, 1; s) \geq n + 2k - 3$ for $s \in \{k-3, k-4\}$ as well, then we could reduce the remaining cases to $g(n, 5, 1; 0)$ instead, for which the Hamming bound gives a stronger lower bound. This would still yield an exponential bound on $n_0(k)$, but with a smaller base.

The second approach concerns our reduction to $g(n, 3, 1; 0)$, where we use the fact that removing a hyperplane from a k-cover leaves us with a $(k-1)$-cover. However, our constructions contain arbitrary pairs of parallel planes, and thus it is possible to remove *two* planes from them and still be left with a $(k-1)$-cover. If we can show that this is true more generally (for instance, for every optimal k-cover of $\mathbb{F}_2^n$ for some range of values of n), it could lead to a linear bound on $n_0(k)$.

Finally, while we have focused on the hyperplane case in Question 7, it would also be worth exploring the corresponding threshold $n_0(k, d)$ for $d \geq 2$. It would be very interesting if there were new constructions that appear in this setting where we cover with affine subspaces of codimension d.

Larger fields In this chapter we worked exclusively over the binary field $\mathbb{F}_2$, but it is also natural to explore these subspace covering problems over larger finite fields, $\mathbb{F}_q$ for $q > 2$. Let us denote the corresponding extremal function by $f_q(n, k, d)$, which is the minimum cardinality of a multiset of $(n-d)$-dimensional affine subspaces that cover all points of $\mathbb{F}_q^n \setminus \{\vec{0}\}$ at least k times, and the origin at most $k-1$ times. The work of Jamison [90] establishes the initial values of this function, showing $f_q(n, 1, d) = (q-1)(n-d) + q^d - 1$. When it comes to multiplicities $k \geq 2$, some of what we have done here can be transferred to larger fields as well.

To start, we can once again resolve the setting where the multiplicity k is large with respect to the dimension n. Indeed, the double-counting lower bound of Lemma 8.1.1 generalizes immediately to this setting, giving $f_q(n, k, d) \geq q^d k - \left\lfloor \frac{k}{q^{n-d}} \right\rfloor$, and one can obtain a matching upper bound by taking multiple copies of every affine subspace.

In the other extreme, where n is large with respect to k, the problem remains widely open. We first note that the reduction to hyperplanes from Lemma 8.1.2 can be extended, giving $f_q(n, k, d) \leq f_q(n-d+1, k, 1) + (q^{d-1} - 1)kq$. Thus, as before, it is best to first focus on the case $d = 1$, and we define $f_q(n, k) := f_q(n, k, 1)$. Then Jamison's result gives $f_q(n, 1) = (q-1)n$.

For an upper bound, let us start by considering 2-covers. It is once again true that if one takes the standard 1-covering by hyperplanes, consisting of all hyperplanes of the form $\{\vec{x} : x_i = c\}$ for some $i \in [n]$ and $c \in \mathbb{F}_q \setminus \{0\}$, the only nonzero vectors that are only covered once are those of Hamming weight 1. However, since the nonzero coordinate of these vectors can take any of $q-1$ different values, it takes a further $q-1$ hyperplanes to cover these again, and so we have $f(n, 2) \leq (q-1)(n+1)$. Now, given a $(k-1)$-cover of $\mathbb{F}_q^n$, one can obtain a k-cover by adding an arbitrary partition of $\mathbb{F}_q^n$ into q parallel planes, and this yields $f_q(n, k) \leq (q-1)(n+1) + q(k-2)$.

This construction is the direct analogue of that from Lemma 8.2.1, and so, as in Theorem 1.3.2(c), we expect it to be tight when n is sufficiently large.

However, the lower bounds are lacking. A simple general lower bound is obtained by noticing that removing $k-1$ hyperplanes from a k-cover leaves us with at least a 1-cover, and so $f_q(n,k) \geq f_q(n,1) + k - 1 = (q-1)n + k - 1$. This remains the best lower bound we know — in particular, even the case of $f_q(n,2)$ is unsolved.

It would of course be very helpful to use some of the machinery we have developed here, and so we briefly explain where the difficulties therein lie. Key to our binary proof was the equivalence with codes of a certain minimum distance, given in Proposition 8.2.3. When working over $\mathbb{F}_q$, unfortunately, that equivalence breaks down. For an n-dimensional linear code with minimum distance k with generator matrix A, we require that, for every nonzero vector $\vec{x} \in \mathbb{F}_q^n$, the vector $A\vec{x}$ has at least k nonzero entries. In the binary setting, this was precisely what we wanted, since $\vec{x}$ was covered by the ith hyperplane if and only if the ith entry of $A\vec{x}$ was nonzero. However, in the q-ary setting, for $\vec{x}$ to be covered by the ith hyperplane, we need the ith entry of $A\vec{x}$ to be equal to a prescribed nonzero value. Hence, while every k-covering of $\mathbb{F}_q^n$ gives rise to a linear q-ary n-dimensional code of minimum distance at least k, the converse is not true. As a result, the coding theoretic bounds, which are of the form $n + O(k \log_q n)$, are not strong enough to give us information here.

Another main tool was the recursion over n, showing that $f(n,k)$ is strictly increasing in n. The same proof goes through here, and we can again show $f_q(n,k) > f_q(n-1,k)$. However, from our bounds, we expect the stronger inequality $f_q(n,k) \geq f_q(n-1,k) + q - 1$ to hold. Intuitively, this is because when we restrict a k-cover of $\mathbb{F}_q^n$ to $\mathbb{F}_q^{n-1} \subseteq \mathbb{F}_q^n$, there are $q-1$ affine copies of $\mathbb{F}_q^{n-1}$ that are lost. However, this does not (appear to) come out of our probabilistic argument.

It would thus be of great interest to develop new tools to handle the q-ary case, as these may also bear fruit when applied to the open problems in the binary setting as well. We believe that new algebraic ideas may be necessary to resolve the following question.

Question 8. For $n \geq n_0(k,q)$, do we have $f_q(n,k) = (q-1)(n+1) + q(k-2)$?

Polynomials with large multiplicity Finally, speaking of algebraic methods, we return to our introductory discussion of the polynomial method. Recall that previous lower bounds in this area have been obtained by considering the more general problem of the minimum degree of a polynomial in $\mathbb{F}[x_1, x_2, \ldots, x_n]$ that vanishes with multiplicity at least k at all nonzero points in some finite grid, and with lower multiplicity at the origin. Sauermann and Wigderson's recent breakthrough, Theorem 1.3.1, resolves this polynomial problem for $n \geq 2k-3$ over fields of characteristic 0, while our results here show that, in the binary setting at least, there is separation between the hyperplane covering and the polynomial problems.

Despite this, we wonder whether the answers to the two problems might coincide in the range where the multiplicity k is large with respect to the dimension n. That is, can the simple double-counting hyperplane lower bound be strengthened to the polynomial setting? We would therefore like to close by emphasizing a question of Sauermann and Wigderson [133], this time over $\mathbb{F}_2$.

Question 9. Given positive integers k, n with $k \geq 2^{n-2}$, let $P \in \mathbb{F}_2[x_1, x_2, \ldots, x_n]$ be a polynomial that vanishes with multiplicity at least k at every nonzero point, and with multiplicity at most $k-1$ at the origin. Must we then have $\deg(P) \geq 2k - \left\lfloor \frac{k}{2^{n-1}} \right\rfloor$?

Chapter 9

Hyperplane coverings over the reals

This chapter represents joint work with Anurag Bishnoi, Shagnik Das, and Yvonne den Bakker; many of the results appear in [17].

Unless otherwise specified, all grids in this chapter are in $\mathbb{R}^2$. Recall that, given sets S_1, S_2, we denote the grid $S_1 \times S_2$ by $\Gamma = \Gamma(S_1, S_2)$, and we always assume that $\vec{0} \in \Gamma$. For convenience, we will often write $\Gamma^- = \Gamma \setminus \{\vec{0}\}$. A point of Γ^- is an interior point if it has no zero coordinates and an axis point otherwise. As usual, the x- and y-axis refer to the lines $y = 0$ and $x = 0$; a vertical line is a line parallel to the y-axis, and a horizontal line is a line parallel to the x-axis. We say that a point is incident to a line, or vice versa, if the line contains the point.

We begin by studying non-square grids in Section 9.1, exploring when the Ball-Serra bound (1.3.1) (see also (1.3.2) for a simplified version) is tight and proving Theorems 1.3.3 and 1.3.4. Next, we shift our attention to square grids in Section 9.2. We reformulate the covering problem as an integer linear program and use duality in linear programming to derive a general method for proving lower bounds. We then use these ideas to prove Proposition 1.3.5, providing general lower and upper bounds for square grids, Proposition 1.3.6 giving a different lower bound for the special case of Δ-bounded grids, and our results about the standard grid, namely Theorems 1.3.7 and 1.3.8.

9.1 The Ball-Serra lower bound

In this section, we will be concerned with the tightness of the Ball-Serra lower bound for $n \times m$ grids. Recall that, for a two-dimensional grid $\Gamma(S_1, S_2)$ with $|S_1| \geq |S_2|$ and $k \geq 2$, the Ball-Serra bound simplifies to

$$h(\Gamma(S_1, S_2), k) \geq k(|S_1| - 1) + (|S_2| - 1).$$

Note that for the next results we do not require that the omitted point $(0,0)$ be a corner of the grid. We begin by showing that, if n is sufficiently large with respect to m and the multiplicity k, then the Ball-Serra bound is always tight, that is, we prove Theorem 1.3.3.

Theorem 1.3.3. *Let $n, m \geq 1$ and $k \geq 2$ be integers such that $n \geq (k-1)(m-1)+1$. Then for any $S_1, S_2 \subseteq \mathbb{R}$ satisfying $|S_1| = n$, $|S_2| = m$, and $0 \in S_1 \cap S_2$, we have $h(\Gamma(S_1, S_2), k) = k(n-1)+(m-1)$, that is, the Ball-Serra bound is tight.*

Proof of Theorem 1.3.3. Write $S_2 = \{0, t_1, \ldots, t_{m-1}\}$. Let $P_1 \cup \cdots \cup P_{m-1}$ be an arbitrary partition of $S_1 \setminus \{0\}$ such that $|P_i| \geq k-1$ for all $i \in [m-1]$; such a partition exists since $n-1 \geq (k-1)(m-1)$.

Now, consider the following collection of lines:

(i) the line $y = t_i$ for all $i \in [m-1]$;
(ii) $k-1$ copies of the line $x = s$ for all $s \in S_1 \setminus \{0\}$;
(iii) the line connecting $(0, t_i)$ and $(s, 0)$ for every $i \in [m-1]$ and $s \in P_i$.

In total, this collection contains $m-1+(k-1)(n-1)+n-1 = k(n-1)+m-1$ lines. It remains to verify that these lines form a valid strict k-cover of Γ. Note first that no line in this collection passes through the point $(0,0)$.

Any interior point of Γ is covered k times by the lines in (i) and (ii). It remains to consider the axis points. A point of the form $(s_1, 0)$ for $s_1 \in S_1 \setminus \{0\}$ is covered $k-1$ times by the lines in (ii) and once by the lines in (iii). Finally, a point $(0, s_2)$ for $s_2 \in S_2 \setminus \{0\}$ is covered once by the lines in (i) and no fewer than $k-1$ times by the lines in (iii). Hence every nonzero point of Γ is covered at least k times, as required. Combined with the Ball-Serra bound (1.3.1), this implies that $h(\Gamma, k) = k(n-1)+(m-1)$. □

We remark that the above construction can be extended to higher dimensions to show that, for any $n_1 \times \cdots \times n_d$ grid $\Gamma(S_1, \ldots, S_d)$ containing the origin, if $n_1 \geq n_2 \geq \cdots \geq n_d$ and $n_1 \geq h(\Gamma(S_2, \ldots, S_d), k-1)+1$, then $h(\Gamma(S_1, \ldots, S_d), k)$ attains the Ball-Serra bound. Indeed, write $S_1 = \{0, s_1, \ldots, s_{n_1-1}\}$ and let $\mathcal{H} = \{H_1, \ldots, H_{n_1-1}\}$ be any collection of n_1-1 hyperplanes in $\mathbb{R}^{d-1}$ containing a strict $(k-1)$-cover of $S_2 \times \cdots \times S_d$; then we can take $k-1$ copies of the hyperplane $x_1 = s$ for all $s \in S_1 \setminus \{0\}$, one copy of the hyperplane $x_i = t$ for all $i \in \{2, \ldots, d\}$ and $t \in S_i \setminus \{0\}$, and the hyperplane spanned by $\{0\} \times H_i$ and $(s_i, 0, \ldots, 0)$ for all $i \in [n_1-1]$. It is not difficult to check that this is indeed a k-cover of $\Gamma(S_1, \ldots, S_d)$ meeting the Ball-Serra lower bound. It is not clear how good this bound on n_1 is, that is, how large n_1 needs to be with respect to the other n_i in order to ensure that the Ball-Serra bound is tight. We show in Theorem 1.3.4 that for certain two-dimensional grids, the bound obtained in Theorem 1.3.3 is indeed best possible. We do not pursue this question further for higher dimensions in this thesis, but we remark that it was shown in [55] that for a grid of the form $\{0, \ldots, n_1 - 1\} \times \{0, \ldots, n_2 - 1\} \times \{0, \ldots, n_3 - 1\}$ with $n_1 \geq n_2 \geq n_3$, the Ball-Serra bound is tight already when $n_1 \geq (k-1)(n_2 - 1) + 1$.

We now turn our attention to the proof of Theorem 1.3.4.

Theorem 1.3.4. *Let $n, m \geq 1$ be integers satisfying $n \geq m$ and $S_1, S_2 \subseteq \mathbb{R}$ be sets with $|S_1| = n$, $|S_2| = m$, and $0 \in S_1 \cap S_2$. Assume further that every line that is not parallel to the x- or y-axis contains at most two points of $\Gamma(S_1, S_2)$. Then, for all $k \geq 2$, we have*

$$(n-1)k + \frac{k}{n+m-2}(m-1)^2 \leq h(\Gamma(S_1, S_2), k)$$

and

$$h(\Gamma(S_1, S_2), k) \leq \left\lceil \frac{(n-1)k}{n+m-2} \right\rceil (n-1) + \left\lceil \frac{(m-1)k}{n+m-2} \right\rceil (m-1) + \left\lceil \frac{(n-1)k}{n+m-2} \right\rceil (m-1).$$

Moreover, if $a = \frac{n-1}{\gcd(n-1, m-1)}$ and $b = \frac{m-1}{\gcd(n-1, m-1)}$ and k is divisible by $a + b$, then

$$h(\Gamma(S_1, S_2), k) = (n-1)k + \frac{k}{n+m-2}(m-1)^2.$$

Proof of Theorem 1.3.4. We first show the lower bound. Let S_1 and S_2 be as in the statement. Fix an optimal strict k-cover $\mathcal{H}$ of $\Gamma(S_1, S_2)$, and suppose it has a_v vertical lines, a_h horizontal lines, and a_d other lines. We use double-counting to obtain a lower bound on the total number of lines $a_v + a_h + a_d = |\mathcal{H}|$.

First, we count the number of point-line incidences $(\ell, \vec{x})$ such that $\ell \in \mathcal{H}$, $\vec{x} \in \Gamma^-$, and $\vec{x} \in \ell$. Observe that each vertical line is incident to m points, each horizontal line to n points, and each other line to at most two points, so the total number of point-line incidences is at most $ma_v + na_h + 2a_d$. On the other hand, each of the $nm - 1$ points in Γ^- is incident to at least k lines in $\mathcal{H}$, yielding the inequality

$$ma_v + na_h + 2a_d \geq (nm-1)k. \tag{9.1.1}$$

Next, we consider the points on the x-axis and count the point-line incidences involving these points. The points on the x-axis are avoided by the horizontal lines; each vertical and each non-axis parallel line contains at most one such point. As each of the $n - 1$ points on the x-axis is covered at least k times, we have

$$a_v + a_d \geq (n-1)k. \tag{9.1.2}$$

Using a similar argument for the points on the y-axis, we obtain

$$a_h + a_d \geq (m-1)k. \tag{9.1.3}$$

Taking the linear combination $\frac{1}{n+m-2}$(9.1.1) + $\frac{n-2}{n+m-2}$(9.1.2) + $\frac{m-2}{n+m-2}$(9.1.3) gives

$$\begin{aligned} a_v + a_h + a_d &\geq \frac{nm-1+(n-1)(n-2)+(m-1)(m-2)}{n+m-2}k \\ &= \frac{(n-1)(n+m-2)+(m-1)^2}{n+m-2}k \\ &= (n-1)k + \frac{k}{n+m-2}(m-1)^2 \\ &= (n-1)k + \frac{bk}{a+b}(m-1), \end{aligned}$$

completing the proof of the lower bound, where for the last step we use the definitions of a and b.

The construction that we use to show the upper bound is similar to the one in the proof of Theorem 1.3.3. Let B be any bipartite (multi)graph with vertex classes $(S_1 \setminus \{0\})$ and $(S_2 \setminus \{0\})$ such that each $s_2 \in S_2 \setminus \{0\}$ has degree $\lceil \frac{ak}{a+b} \rceil$ and each $s_1 \in S_1 \setminus \{0\}$ has degree $\lfloor \frac{bk}{a+b} \rfloor$ or $\lceil \frac{bk}{a+b} \rceil$. We can build such a bipartite (multi)graph as follows. Put $\lceil \frac{ak}{a+b} \rceil$ "half-edges" at each vertex of $S_2 \setminus \{0\}$ and $\lfloor \frac{bk}{a+b} \rfloor$ or $\lceil \frac{bk}{a+b} \rceil$ "half-edges" at each vertex of $S_1 \setminus \{0\}$ in such a way that the number of "half-edges" on each side is the same; this is possible since $\lceil \frac{ak}{a+b} \rceil (m-1) \geq \frac{(m-1)(n-1)k}{n+m-2} = \frac{bk}{a+b}(n-1)$. Finally, pair up the "half-edges" arbitrarily. Now consider the following collection of lines:

(i) $\lceil \frac{ak}{a+b} \rceil$ copies of the line $x = s_1$ for each $s_1 \in S_1 \setminus \{0\}$;
(ii) $\lceil \frac{bk}{a+b} \rceil$ copies of the line $y = s_2$ for each $s_2 \in S_2 \setminus \{0\}$;
(iii) one copy of the line connecting $(s_1, 0)$ and $(0, s_2)$ for all $s_1 \in S_1 \setminus \{0\}$ and $s_2 \in S_2 \setminus \{0\}$ such that $s_1 s_2 \in E(B)$.

To see that this is a valid strict k-cover, observe that every interior point of Γ is covered by at least $\frac{ak}{a+b}$ vertical lines from (i) and $\frac{bk}{a+b}$ horizontal lines from (ii), and is thus covered at least k times in total. A point of the form $(s_1, 0)$ for $s_1 \in S_1 \setminus \{0\}$ is covered $\lceil \frac{ak}{a+b} \rceil$ times by the lines in (i) and at least $\lfloor \frac{bk}{a+b} \rfloor$ times by the lines in (iii) by the choice of the bipartite graph B, so in total it is covered at least k times. Similarly, each point of the form $(0, s_2)$ for $s_2 \in S_2 \setminus \{0\}$ is covered at least k times by the lines in (ii) and (iii). Finally, none of the lines in our collection passes through the origin.

The total number of lines in the cover is

$$\begin{aligned}&\left\lceil \frac{ak}{a+b} \right\rceil (n-1) + \left\lceil \frac{bk}{a+b} \right\rceil (m-1) + \left\lceil \frac{ak}{a+b} \right\rceil (m-1)\\ &= \left\lceil \frac{(n-1)k}{n+m-2} \right\rceil (n-1) + \left\lceil \frac{(m-1)k}{n+m-2} \right\rceil (m-1) + \left\lceil \frac{(n-1)k}{n+m-2} \right\rceil (m-1),\end{aligned}$$

as claimed.

If k is divisible by $a+b$, the above expression becomes $(n-1)k + \frac{bk}{a+b}(m-1)$, where we use the fact that $\frac{ak}{a+b}(m-1) = \frac{bk}{a+b}(n-1)$. □

As remarked in the introduction, Theorem 1.3.4 shows that for the grids satisfying the assumptions of the theorem, the Ball-Serra bound is tight precisely when $n \geq (k-1)(m-1)+1$. We conclude this section by giving an example showing that the Ball-Serra bound can also be tight when $n = (k-1)(m-1)$. Note that the omitted point in this example is not a corner of the grid.

Proposition 9.1.1. *Let $k \geq 3$, $n = 2(k-1)$, $S_1 = \{0, 1, 2, \ldots, n-1\}$, and $S_2 = \{-1, 0, 1\}$. Then $h(\Gamma(S_1, S_2), k) = k(n-1) + 2$, that is, the Ball-Serra bound is tight.*

Proof. If $k = 3$ and $n = 4$, consider the following collection of lines:

(i) $y = 1$ and $y = -1$;
(ii) $x = i$ for $i \in [3]$;
(iii) $x - y = 1$, $x - y = 2$, $x + y = 1$, and $x + y = 2$;
(iv) $\frac{1}{3}x - y = 1$ and $\frac{1}{3}x + y = 1$.

This cover contains $2 + 3 + 4 + 2 = 11$ lines and is illustrated in Figure 9.1, from which it is not difficult to see that this is indeed a strict 3-cover of Γ.

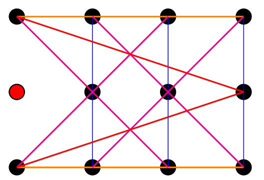

FIGURE 9.1: Strict 3-cover of the grid $\{0, 1, 2, 3\} \times \{-1, 0, 1\}$ with 11 lines

Now suppose $k \geq 4$. Consider the following collection of lines:

(i) $y = -1$ and $y = 1$;
(ii) $k - 4$ copies of the line $x = t$ for each $t \in [n-1]$;
(iii) the lines $x - y = t$ and $x + y = t$ for every $t \in \{1, \ldots, n-2\}$;
(iv) the lines $x + ty = t$ and $x - ty = t$ for all $t \in [k-2]$;
(v) the lines $x - ty = n - 1 - t$ and $x + ty = n - 1 - t$ for all $t \in [k-2]$;

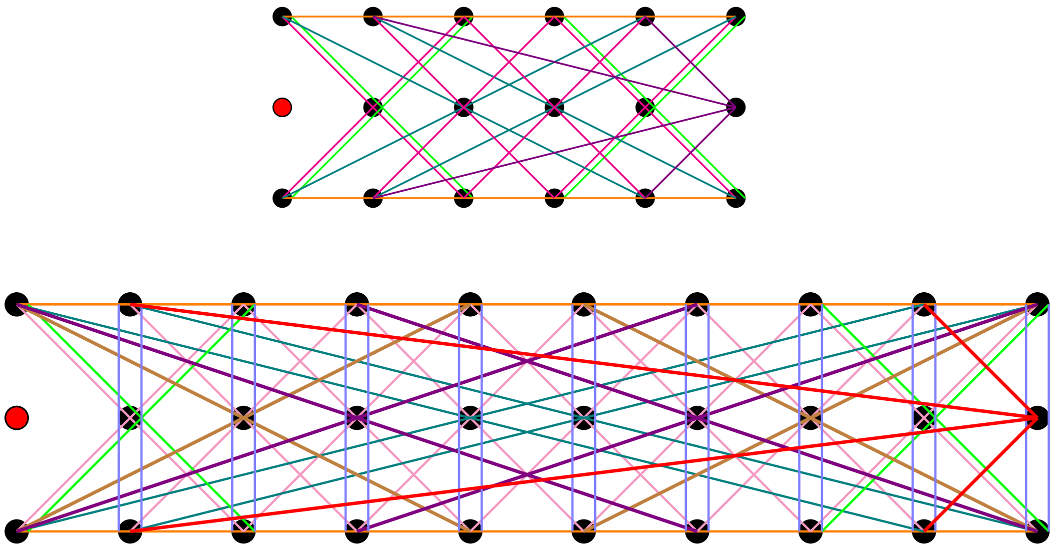

FIGURE 9.2: Strict k-cover of the grid $\{0, 1, 2, \ldots, 2(k-1)-1\} \times \{-1, 0, 1\}$ from the proof of Proposition 9.1.1 for $k = 4$ and $k = 6$

(vi) the four lines connecting $(n-1, 0)$ to $(1, -1), (1, 1), (n-2, -1), (n-2, 1)$.

The construction is illustrated in Figure 9.2 for $k = 4$ and $k = 6$.

We now verify that this is indeed a strict k-cover of Γ. Recalling that $n = 2(k-1)$, we note that the number of lines used is exactly $k(n-1)+2$. Also, observe that no line passes through $(0, 0)$.

It remains to show that each point of Γ^- is covered at least k times. First consider a point of the form $(s, 0)$ for $s \in [n-1]$. Such a point is covered $k-4$ times by the lines in (ii). If $s = n-1$, then the point is covered four more times by the lines in (vi). Otherwise $(s, 0)$ is covered twice more by the lines in (iii) and twice more by the lines in (iv) or (v), depending on whether $s \leq k-2$ or $s \geq k-1$.

Notice that the construction is symmetric around the x-axis, so it suffices to show that a point of the form $(s, 1)$ for $s \in \{0, 1, \ldots, n-1\}$ is covered at least k times. This point is covered once by the horizontal lines in (i). First suppose $s = 0$. The point is then covered $k-2$ times by the lines in (iv) and once more by the lines in (iii). If $s = n-1$, the argument is similar using the lines from (iii) and (v). When $0 < s < n-1$, the point $(s, 1)$ is covered $k-4$ times by the lines in (ii). If $s \in \{1, n-2\}$, then the point is covered once by the lines in (iii), once more by the lines in (vi), and once more by the lines in (iv) or (v) with $t = k-2$. Finally, assume that $s \in \{2, \ldots, n-3\}$. Then the point $(s, 1)$ is covered twice more by the lines in (iii), and once more by a line in (iv) or (v) with $t = \frac{s}{2}$ if s is even and $t = \frac{n-1-s}{2}$ if s is odd (recall that n is always even). $\square$

9.2 Square grids

In the remainder of the chapter, we study square grids in which the origin is a corner. For our lower bounds, we will rely on linear programming and duality; we will use this language in all of our proofs in this section. Before we define the linear program, we remark that, for every strict k-cover of an $n \times n$ grid Γ with $n \geq 2$, there exists a strict k-cover of the same size such that every line contains at least two points of the grid. Indeed, if a line contains only a single point of the grid, we can always exchange it for another line through the same point that contains another point of the grid and still avoids the origin (for example, we can always take either a horizontal or a vertical line). This means that it suffices to restrict our attention to lines containing at least two points of the grid; there are finitely many such lines, and this observation now allows us to reformulate our problem as an integer linear program. We use standard terminology from linear programming and refer the reader to [115] for more background on the topic.

Let $S_1, S_2 \subseteq \mathbb{R}$ be finite sets with $\min S_1 = 0 = \min S_2$ and $\Gamma = \Gamma(S_1, S_2)$. Let $\mathcal{L} = \mathcal{L}(\Gamma)$ be the set of lines containing at least two points of the grid Γ and not containing the origin. Then the problem of determining the minimum size of a strict k-cover of Γ can be expressed via the following integer linear program $\mathcal{I} = \mathcal{I}(\Gamma, k)$. In this linear program, we have a variable $z(\ell)$ for each line $\ell \in \mathcal{L}$, capturing how many times the line ℓ appears in the cover.

$$\begin{aligned}
&\text{minimize} \quad \sum_{\ell \in \mathcal{L}} z(\ell) \\
&\text{subject to} \\
&\qquad \sum_{\substack{\ell \in \mathcal{L}: \\ (x,y) \in \ell}} z(\ell) \geq k \quad \text{for all } (x, y) \in \Gamma^- \\
&\qquad z(\ell) \in \mathbb{Z}_{\geq 0} \quad \text{for all } \ell \in \mathcal{L}
\end{aligned}$$

To use duality, we first consider the linear relaxation of $\mathcal{I}(\Gamma, k)$, which we denote by $\mathcal{P} = \mathcal{P}(\Gamma)$ and formulate in the following way (note that we divide through by k, thus eliminating the dependence on k). The variables are $u(\ell)$ for each line $\ell \in \mathcal{L}$.

$$\begin{aligned}
&\text{minimize} \quad \sum_{\ell \in \mathcal{L}} u(\ell) \\
&\text{subject to} \\
&\qquad \sum_{\substack{\ell \in \mathcal{L}: \\ (x,y) \in \ell}} u(\ell) \geq 1 \quad \text{for all } (x, y) \in \Gamma^- \\
&\qquad u(\ell) \geq 0 \quad \text{for all } \ell \in \mathcal{L}
\end{aligned}$$

Observe that if $(z(\ell) : \ell \in \mathcal{L})$ is a feasible solution of $\mathcal{I}(\Gamma, k)$, then setting $u(\ell) = \frac{1}{k}z(\ell)$ for all $\ell \in \mathcal{L}$ gives a feasible solution of $\mathcal{P}(\Gamma)$. Let $a(\Gamma)$ be the optimal objective value of $\mathcal{P}(\Gamma)$. Then, since $h(\Gamma, k)$ is the optimum of $\mathcal{I}(\Gamma, k)$ and $\mathcal{I}(\Gamma, k)$ has the extra integrality constraints, we obtain

$$ka(\Gamma) \leq h(\Gamma, k). \tag{9.2.1}$$

The dual of $\mathcal{P}(\Gamma)$, which we denote by $\mathcal{D}(\Gamma)$, is given by the following program, where we have a variable $w(x, y)$ for each point $(x, y) \in \Gamma^-$.

$$\begin{aligned} &\text{maximize} \quad \sum_{(x,y)\in\Gamma^-} w(x,y) \\ &\text{subject to} \\ &\qquad \sum_{\substack{(x,y)\in\Gamma^-:\\(x,y)\in\ell}} w(x,y) \leq 1 \quad \text{for all } \ell \in \mathcal{L} \\ &\qquad w(x,y) \geq 0 \quad \text{for all } (x,y) \in \Gamma^- \end{aligned} \tag{9.2.2}$$

By the duality theorem for linear programming (see for example [115, Section 6.1]), we know that the programs $\mathcal{P}(\Gamma)$ and $\mathcal{D}(\Gamma)$ have the same optimal objective value $a(\Gamma)$. Thus, to obtain a lower bound on $h(\Gamma, k)$, it suffices to give a feasible solution to the dual linear program $\mathcal{D}(\Gamma)$. We will make extensive use of this fact. For convenience, given a weighting $(w(x, y) : (x, y) \in \Gamma^-)$ and a subset $S \subseteq \Gamma^-$, we write $w(S) = \sum_{(x,y)\in S} w(x, y)$.

We now use the above formulation to prove Proposition 1.3.5, providing general upper and lower bounds for square grids.

Proposition 1.3.5. *Let $n, k \geq 2$ be integers and $S_1, S_2 \subseteq \mathbb{R}$ be sets satisfying $|S_1| = n = |S_2|$, $\min S_1 = 0 = \min S_2$. Then:*

(a) $h(\Gamma(S_1, S_2), k) \leq \left\lceil \frac{3}{2}k \right\rceil (n - 1)$.
(b) $h(\Gamma(S_1, S_2), k) \geq (4 - 2\sqrt{2} + o(1))k(n - 1)$ as $n \to \infty$.

Proof of Proposition 1.3.5. Let $S_1 = \{0, x_1, \ldots, x_{n-1}\}$ and $S_2 = \{0, y_1, \ldots, y_{n-1}\}$ be subsets of $\mathbb{R}_{\geq 0}$ of size n, where $0 < x_1 < \cdots < x_{n-1}$ and $0 < y_1 < \cdots < y_{n-1}$, and $\Gamma = \Gamma(S_1, S_2)$.

(a) Consider the following assignment of weights on the lines in $\mathcal{L}$:
   (i) $z(\ell) = \left\lceil \frac{k}{2} \right\rceil$ for all $\ell \in \mathcal{L}$ of the form $x = x_i$ or $y = y_i$ for $i \in [n - 1]$;
   (ii) $z(\ell) = \left\lfloor \frac{k}{2} \right\rfloor$, where $\ell \in \mathcal{L}$ is the line connecting $(x_i, 0)$ and $(0, y_i)$ for all $i \in [n - 1]$;
   (iii) $z(\ell) = 0$ for all other $\ell \in \mathcal{L}$.

We now show that this assignment is a feasible solution of $\mathcal{I}(\Gamma, k)$. First note that all $z(\ell)$ are nonnegative integers. Now let $(s_1, s_2) \in \Gamma^-$. If $s_1 s_2 \neq 0$, then $z(x = s_1) + z(y = s_2) = 2\lceil \frac{k}{2} \rceil \geq k$. If $s_1 = 0$ and $s_2 = y_t$, then $z(y = s_2) + z(\ell) = \lceil \frac{k}{2} \rceil + \lfloor \frac{k}{2} \rfloor = k$, where ℓ is the line connecting $(x_t, 0)$ and $(0, y_t)$. A similar argument applies when $s_2 = 0$. Hence, the assignment is indeed a feasible solution to $\mathcal{I}(\Gamma, k)$ with objective value $(n-1)\Big(2\lceil \frac{k}{2} \rceil + \lfloor \frac{k}{2} \rfloor\Big) = \lceil \frac{3}{2}k \rceil (n-1)$, as required. Thus, $h(\Gamma, k) \leq \lceil \frac{3}{2}k \rceil (n-1)$.

(b) We show the required lower bound by constructing a feasible solution to the dual linear program $\mathcal{D}(\Gamma)$ and using (9.2.1). Consider the following assignment of weights to the points of Γ^-:

(i) $w(x, y) = \frac{\sqrt{2}}{2n}$ if $xy \neq 0$;

(ii) $w(x, y) = \frac{2-\sqrt{2}}{2}$ if $x = 0$ and $y \geq y_1$;

(iii) $w(x, y) = \frac{2-\sqrt{2}}{2}$ if $y = 0$ and $x_1 \leq x \leq x_{\lceil (2-\sqrt{2})n \rceil}$;

(iv) $w(x, y) = 0$ otherwise.

To see that this is a feasible solution to $\mathcal{D}(\Gamma)$, we split the lines into two categories. First consider a line with points of nonzero weight from both the x- and y-axes. The x-coordinate of the former point must be at most $x_{\lceil (2-\sqrt{2})n \rceil}$, and hence the line must have fewer than $\lceil (2-\sqrt{2})n \rceil$ interior points. This gives a total weight of at most $2 \cdot \frac{2-\sqrt{2}}{2} + (2-\sqrt{2})n \cdot \frac{\sqrt{2}}{2n} = 1$. Any other line can have at most one positively-weighted axis point and n interior points, which gives a total weight of at most $\frac{2-\sqrt{2}}{2} + n \cdot \frac{\sqrt{2}}{2n} = 1$.

Thus, the desired bound on the weight of each line holds and all weights are nonnegative, implying that the weighting $(w(x, y) : (x, y) \in \Gamma^-)$ is indeed feasible. The objective value satisfies

$$\begin{aligned} w(\Gamma^-) &\geq (n-1)\frac{2-\sqrt{2}}{2} + ((2-\sqrt{2})n - 1)\frac{2-\sqrt{2}}{2} + (n-1)^2 \frac{\sqrt{2}}{2n} \\ &= (3-\sqrt{2})n \cdot \frac{2-\sqrt{2}}{2} + n \cdot \frac{\sqrt{2}}{2} + O(1) \\ &= (4 - 2\sqrt{2} + o(1))n = (4 - 2\sqrt{2} + o(1))(n-1). \end{aligned}$$

Hence, any strict k-cover of Γ must have at least $(4 - 2\sqrt{2} + o(1))k(n-1)$ lines.

□

We now improve the above construction significantly for grids in which the lines connecting two axis points are sparse. Recall that a grid is Δ-bounded if every line containing two axis points contains at most Δ interior points of the grid. As remarked in the introduction, most grids are actually 0-bounded, and we consider Δ-boundedness to be a natural generalization.

Proposition 1.3.6. *Let $n, k \geq 2$ be integers and $S_1, S_2 \subseteq \mathbb{R}$ be sets satisfying $|S_1| = n = |S_2|$, $\min S_1 = 0 = \min S_2$. If $\Gamma(S_1, S_2)$ is Δ-bounded, then*

$$h(\Gamma(S_1, S_2), k) \geq \left[2 - \frac{n-1}{2(n-1)-\Delta}\right] k(n-1).$$

Proof of Proposition 1.3.6. Let Γ be a Δ-bounded grid, where $\Delta \leq n-2$, and consider the following assignment of weights on the points of Γ^-:

$$w(x,y) = \begin{cases} 1 - \frac{n-1}{2(n-1)-\Delta} & \text{if } xy = 0 \\ \frac{1}{2(n-1)-\Delta} & \text{otherwise.} \end{cases}$$

To show that this is a feasible solution of $\mathcal{D}(\Gamma)$, first note that all weights are nonnegative. Now, we classify the lines in $\mathcal{L}$ depending on the number of axis points they contain. If $\ell \in \mathcal{L}$ contains two axis points, then ℓ contains at most Δ interior points, so $w(\ell) \leq 2\left(1 - \frac{n-1}{2(n-1)-\Delta}\right) + \frac{\Delta}{2(n-1)-\Delta} = 1$. If ℓ contains one axis point, then it has at most $n-1$ interior points, and thus $w(\ell) \leq 1 - \frac{n-1}{2(n-1)-\Delta} + \frac{n-1}{2(n-1)-\Delta} = 1$. Finally, if ℓ contains no axis points, then its total weight is at most $\frac{n}{2(n-1)-\Delta} \leq 1$ using the fact that $\Delta \leq n-2$. Hence, the solution is indeed feasible. The total weight is

$$\begin{aligned} w(\Gamma^-) &= 2(n-1) \cdot \left(1 - \frac{n-1}{2(n-1)-\Delta}\right) + (n-1)^2 \cdot \frac{1}{2(n-1)-\Delta} \\ &= \left[2 - \frac{n-1}{2(n-1)-\Delta}\right](n-1), \end{aligned}$$

which, combined with (9.2.1), yields the required result. □

9.2.1 Standard grids

In this section we focus on the most natural square grid, the standard $n \times n$ grid $\{0, 1, 2, \ldots, n-1\}^2$. For convenience, we denote this grid by Γ_n. Our first goal is to prove Theorem 1.3.7.

Theorem 1.3.7. *Let $n, k \geq 2$ be integers and $S_1 = S_2 = \{0, 1, 2, \ldots, n-1\}$. Then, as $n, k \to \infty$, we have*

$$(2 - e^{-1/2} + o(1))k(n-1) \leq h(\Gamma(S_1, S_2), k) \leq (\sqrt{2} + o(1))k(n-1). \tag{1.3.3}$$

We begin by showing that a strict k-cover of the standard grid requires far fewer lines than the upper bound given by Proposition 1.3.5(a), which, as discussed earlier, is asymptotically tight for most grids.

Proof of Theorem 1.3.7 (upper bound). We provide a feasible solution to $\mathcal{I}(\Gamma_n, k)$ with objective value $(\sqrt{2}+o(1))k(n-1)$.

Let $t \in [n-2]$ be any integer. Consider the following assignment of weights on the lines in $\mathcal{L}$:

(i) $z(\ell) = \lceil \frac{i}{n+t-1} k \rceil$ for every line $\ell \in \mathcal{L}$ of the form $x = i$ or $y = i$;
(ii) $z(\ell) = k - \lceil \frac{i}{n+t-1} k \rceil$ for every line $\ell \in \mathcal{L}$ of the form $x + y = i$, where $1 \le i \le n+t-1$;
(iii) $z(\ell) = 0$ for every other line $\ell \in \mathcal{L}$.

We begin by showing that the assignment $(z(\ell) : \ell \in \mathcal{L})$ gives a feasible solution to the integer program $\mathcal{I}(\Gamma_n, k)$. It is clear that $z(\ell) \in \mathbb{Z}_{\ge 0}$ for all $\ell \in \mathcal{L}$. Now consider a point $(s_1, s_2) \in \Gamma_n^-$. If $s_2 = 0$, we have $z(x = s_1) + z(x + y = s_1) = \lceil \frac{s_1}{n+t-1} k \rceil + k - \lceil \frac{s_1}{n+t-1} k \rceil = k$; similarly if $s_1 = 0$, we have $z(y = s_2) + z(x + y = s_2) = k$. So assume that $s_1 s_2 \ne 0$. If $s_1 + s_2 \le n + t - 1$, we have $z(x = s_1) + z(y = s_2) + z(x + y = s_1 + s_2) = \lceil \frac{s_1}{n+t-1} k \rceil + \lceil \frac{s_2}{n+t-1} k \rceil + k - \lceil \frac{s_1+s_2}{n+t-1} k \rceil \ge k$. Finally, if $s_1 + s_2 > n + t - 1$, then $z(x = s_1) + z(y = s_2) = \lceil \frac{s_1}{n+t-1} k \rceil + \lceil \frac{s_2}{n+t-1} k \rceil > k$. Thus $(z(\ell) : \ell \in \mathcal{L})$ is indeed feasible.

Hence, the optimal solution to $\mathcal{I}(\Gamma_n, k)$ is at most

$$\begin{aligned}
\sum_{\ell \in \mathcal{L}} z(\ell) &= \sum_{i=1}^{n-1} z(x = i) + \sum_{i=1}^{n-1} z(y = i) + \sum_{i=1}^{n+t-1} z(x + y = i) \\
&= \sum_{i=1}^{n-1} \left\lceil \frac{i}{n+t-1} k \right\rceil + \sum_{i=1}^{n-1} \left\lceil \frac{i}{n+t-1} k \right\rceil + \sum_{i=1}^{n+t-1} \left(k - \left\lceil \frac{i}{n+t-1} k \right\rceil \right) \\
&\le k \left[2 \sum_{i=1}^{n-1} \frac{i}{n+t-1} + \sum_{i=1}^{n+t-1} \left(1 - \frac{i}{n+t-1} \right) \right] + 2n \\
&\le k \left[\frac{2}{n+t-1} \binom{n}{2} + n + t - \frac{1}{n+t-1} \binom{n+t}{2} \right] + 2n \\
&= k \left[\frac{n(n-1)}{n+t-1} + \frac{n+t}{2} \right] + 2n. \qquad (9.2.3)
\end{aligned}$$

We now want to choose a value of t that minimizes total weight given by (9.2.3). For this, we set $g(t) = \frac{n(n-1)}{n+t-1} + \frac{n+t}{2}$. We then have $g'(t) = \frac{1}{2} - \frac{n(n-1)}{(n+t-1)^2}$, which is zero at $t_+ = -(n-1) + \sqrt{2(n-1)n}$ and $t_- = -(n-1) - \sqrt{2(n-1)n}$. Consider the positive root: we have $g''(t_+) = \frac{2n(n-1)}{(n+t_+-1)^3} > 0$. Thus the function g has a local minimum at t_+.

Choosing $t_0 = \lceil -(n-1) + \sqrt{2(n-1)n} \rceil = (\sqrt{2} - 1 + o(1))(n-1)$, we have $n + t_0 = (\sqrt{2} + o(1))(n-1)$; substituting into (9.2.3), we find that the objective value is at most $k\left(\frac{n(n-1)}{n+t_0-1} + n + t_0 - \frac{n+t_0}{2}\right) + 2n = (\sqrt{2} + o(1))(n-1)k$, as claimed. □

We now turn our attention to the lower bound.

Proof of Theorem 1.3.7 (lower bound). By (9.2.1), it suffices to find a feasible solution to the dual linear program $\mathcal{D}(\Gamma_n)$ that has total weight at least $(2 - e^{-1/2} + o(1))(n-1)$.

Let t be the largest integer such that $\sum_{i=1}^{t} \frac{1}{n-i} \leq \frac{1}{2}$ and consider the following weighting on the points of Γ_n^-:

(i) $w(x,y) = \frac{1}{2}$ if $xy = 0$;
(ii) $w(x,y) = \frac{1}{n-i}$ if $x + y = n - 1 + i$, where $1 \leq i \leq t$;
(iii) $w(x,y) = 0$ for every other point $(x,y) \in \Gamma_n^-$.

We first show that $(w(x,y) : (x,y) \in \Gamma_n)$ gives a feasible solution to the dual linear program $\mathcal{D}(\Gamma_n)$. Clearly $w(x,y) \geq 0$ for all $(x,y) \in \Gamma_n^-$. Now, let $\ell \in \mathcal{L}$ be any line. If ℓ contains two axis points, then any interior point (x,y) on ℓ satisfies $x+y \leq n-1$, and thus has weight zero. It follows that $w(\ell) = 1$. Otherwise, if $(x,y) \mapsto x+y$ is constant on ℓ, then $\ell = \{(x,y) : x+y = n-1+i\}$ for some $i \in \mathbb{Z}$. Then all points on ℓ have weight zero, unless $1 \leq i \leq t$, in which case ℓ contains $n - i$ points of weight $\frac{1}{n-i}$ each. Thus $w(\ell) \leq 1$ in this case. Finally, if $(x,y) \mapsto x+y$ is not constant on ℓ, it must be injective. Then, ℓ contains at most one axis point, which has weight $\frac{1}{2}$, and the weight from the remaining points is at most $\sum_{i=1}^{t} \frac{1}{n-i}$, which by the choice of t is at most $\frac{1}{2}$. So in total we again have $w(\ell) \leq 1$.

To compute the total weight of the grid, observe that each diagonal line of the form $x + y = i$ has weight one if $1 \leq i \leq n - 1 + t$ and zero otherwise. Thus, $w(\Gamma_n^-) = n - 1 + t$.

It remains to estimate t. Note that $t \leq \frac{n}{2}$, since $\sum_{i=1}^{t} \frac{1}{n-i} \geq \sum_{i=1}^{t} \frac{1}{n} = \frac{t}{n}$, and so both $n - 1$ and $n - 1 - t$ go to infinity linearly with n. It is well known that as $n \to \infty$ the partial sums H_n of the Harmonic series satisfy $H_n = \sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + o(1)$, where γ is a constant. Hence,

$$\sum_{i=1}^{t} \frac{1}{n-i} = H_{n-1} - H_{n-1-t} = \log\left(\frac{n-1}{n-1-t}\right) + o(1).$$

Thus, we must have $\log\left(\frac{n-1}{n-1-t}\right) = \frac{1}{2} + o(1)$, or $\log\left(1 - \frac{t}{n-1}\right) = -\frac{1}{2} + o(1)$. This gives $1 - \frac{t}{n-1} = e^{-1/2} + o(1)$, or $t = \left(1 - e^{-1/2} + o(1)\right)(n-1)$, which results in the claimed bound. □

Motivated by the observation that in the proof of the upper bound in Theorem 1.3.7 the only lines with nonzero weight have slope $0, \infty$, and -1, we now consider a variant of our covering problem. We ask: what is the minimum number of lines of slope $0, \infty$, or -1 needed to cover every nonzero point of Γ_n at least k times, while leaving the origin uncovered? We answer this question asymptotically, proving Theorem 1.3.8.

Theorem 1.3.8. *As $n, k \to \infty$, the minimum number of lines of slope $0, \infty$, or -1 needed to cover every nonzero point of Γ_n at least k times, while leaving the origin uncovered, is $\left(\sqrt{2} + o(1)\right)k(n-1)$.*

In a similar fashion as for the general problem, we can show that the answer to this question boils down to solving an integer linear program similar to $\mathcal{I}$. The only necessary modification is to work with the subset $\mathcal{L}' \subseteq \mathcal{L}$ containing only the lines of slope $0, \infty$, or -1 instead of the full set of lines $\mathcal{L}$. We can then consider the linear relaxation of the integer program and its dual to arrive at the equivalent of (9.2.1). Thus, to prove the required lower bound, we provide a feasible solution to the modified dual program with the required objective value; this linear program is obtained from (9.2.2) by replacing the set $\mathcal{L}$ by the subset $\mathcal{L}'$ and we denote it by $\mathcal{D}'$.

Proof of Theorem 1.3.8. The upper bound follows from the above discussion and the fact that in the proof of the upper bound in Theorem 1.3.7 all lines with positive weight belong to $\mathcal{L}'$. We now show the matching lower bound. For convenience, we call the lines of slope -1 diagonals.

Our weighting w will have the following properties for some integer $t \in [n-1]$:

- $w(x, y) = w(y, x)$ for all $(x, y) \in \Gamma_n^-$.
- Every vertical line has weight one (and hence so does every horizontal line).
- On every diagonal, every interior point has the same weight.
- The diagonal $x + y = i$ for $1 \leq i \leq n + t - 1$ has weight one, the diagonal $x + y = n + t$ has weight at most one, and all other diagonals have weight zero.

Let $1 \leq t \leq n-1$ be an integer. Let $\alpha_1, \ldots, \alpha_{n-1}, z \in \mathbb{R}_{\geq 0}$, and consider the following weighting of the points of Γ_n^-:

(i) $w(x, y) = \frac{1}{2} - \alpha_{x+y}$ if $xy = 0$;
(ii) $w(x, y) = \frac{2\alpha_{x+y}}{x+y-1}$ if $2 \leq x + y \leq n - 1$;
(iii) $w(x, y) = \frac{1}{n-i}$ if $x + y = n - 1 + i$ for some $1 \leq i \leq t$;
(iv) $w(x, y) = z$ if $x + y = n + t$;
(v) $w(x, y) = 0$ if $x + y \geq n + t + 1$.

Since each of the diagonals $x + y = i$ for $1 \leq i \leq n + t - 1$ has weight one, the total weight of the grid is at least $n + t - 1$.

We now investigate what relations the parameters $t, \alpha_1, \ldots, \alpha_{n-1}, z$ have to satisfy in order to guarantee that the weighting is feasible and satisfies the above properties.

In order for the weight of each point in the lower triangle to be nonnegative, we must have

$$0 \leq \alpha_i \leq \frac{1}{2} \quad \text{for all } i \in [n-1]. \tag{9.2.4}$$

From the definition of the weighting it is not difficult to see that every line of the form $x+y=i$ for $i \in [n-1]$ has weight at most one. A line of the form $x+y=n-1+i$ for $i \in [n-1]$ contains $n-i$ points, implying the required inequality for every line $x+y=n-1+i$ with $i \in [n-1] \setminus \{t+1\}$. To ensure that the line $x+y=n+t$ also has weight at most one, we need to impose the following constraint on z:

$$0 \leq z \leq \frac{1}{n-t-1}. \tag{9.2.5}$$

As mentioned earlier, we require every vertical line to have weight exactly one, so considering the line $x=n-1$, we obtain:

$$\alpha_{n-1} = \sum_{i=1}^{t} \frac{1}{n-i} + z - \frac{1}{2}. \tag{9.2.6}$$

Now consider the line $x=n-2$. We can express the weight of $x=n-2$ as $w(x=n-2) = w(x=n-1) - (\frac{1}{2} - \alpha_{n-1}) + \frac{2\alpha_{n-1}}{n-2} + (\frac{1}{2} - \alpha_{n-2})$, and since all vertical lines should have weight exactly one, we obtain

$$\alpha_{n-2} = \left(1 + \frac{2}{n-2}\right)\alpha_{n-1}.$$

In a similar fashion, considering the lines $x=i-1$ and $x=i$ for each $2 \leq i \leq n-1$, we obtain the following general relation:

$$\alpha_{i-1} = \left(1 + \frac{2}{i-1}\right)\alpha_i - z\mathbf{1}_{i=t+1} - \frac{1}{n-i}\mathbf{1}_{i \leq t}, \tag{9.2.7}$$

where $\mathbf{1}_A$ is the indicator function of the event A defined as

$$\mathbf{1}_A = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{if } A \text{ is false} \end{cases}.$$

For the line $x+y=1$ to have weight one, we must have

$$\alpha_1 = 0. \tag{9.2.8}$$

Now using (9.2.8) and (9.2.7), we can solve for the other α_i to obtain:

$$\alpha_i = \frac{1}{i(i+1)} \sum_{j=1}^{\min\{t,i\}} \frac{j(j-1)}{n-j} + \mathbf{1}_{i \geq t+1} \frac{t(t+1)}{i(i+1)} z \quad \text{for all } 2 \leq i \leq n-1. \tag{9.2.9}$$

Now, using (9.2.6) and (9.2.9) for $i = n-1$, we can solve for z to obtain:

$$z = \left(\frac{1}{2} - \sum_{j=1}^{t} \frac{1}{n-j}\left(1 - \frac{j(j-1)}{(n-1)n}\right)\right)\frac{n(n-1)}{n(n-1)-t(t+1)} \tag{9.2.10}$$

$$= \frac{(n-1)n\left(\frac{1}{2} - \frac{t(2n+t-1)}{2(n-1)n}\right)}{(n-1)n - t(t+1)}, \tag{9.2.11}$$

where the second equality is obtained as follows:

$$\begin{aligned}
\sum_{j=1}^{t} \frac{1}{n-j}\left(1 - \frac{j(j-1)}{(n-1)n}\right) &= \sum_{j=1}^{t} \frac{1}{n-j}\left(\frac{n^2 - n - j^2 + j}{(n-1)n}\right) \\
&= \frac{1}{n(n-1)} \sum_{j=1}^{t}\left(\frac{(n+j)(n-j) - (n-j)}{n-j}\right) \\
&= \frac{1}{n(n-1)} \sum_{j=1}^{t}(n+j-1) = \frac{1}{n(n-1)}\left(nt - t + \binom{t+1}{2}\right) \\
&= \frac{t(2n+t-1)}{2n(n-1)}.
\end{aligned}$$

From (9.2.5), we know that $0 \le z \le \frac{1}{n+t-1}$. If $n > 1$, we have $z \ge 0$ when $0 \le t \le \frac{1}{2}\left(\sqrt{8n^2 - 8n + 1} - 2n + 1\right)$ and $z \le \frac{1}{n-t-1}$ for $\frac{1}{2}\left(\sqrt{8n^2 - 8n + 1} - 2n - 1\right) \le t < n - 1$. We then choose t to be an integer satisfying $\frac{1}{2}\left(\sqrt{8n^2 - 8n + 1} - 2n - 1\right) \le t \le \frac{1}{2}\left(\sqrt{8n^2 - 8n + 1} - 2n + 1\right)$, so that $t = (\sqrt{2} - 1 + o(1))n$. From here we can conclude that the total weight of the grid is $\left(\sqrt{2} + o(1)\right)(n-1)$.

It remains to verify that $0 \le \alpha_i \le \frac{1}{2}$ for all $1 \le i \le n-1$. Now, using (9.2.7), we obtain:

$$\begin{aligned}
\alpha_i &= \frac{i-1}{i+1}\alpha_{i-1} < \alpha_{i-1} \quad \text{if } t+2 \le i \le n-1 \\
\alpha_i &= \frac{i-1}{i+1}(\alpha_{i-1} + z) \le \frac{i-1}{i+1}\left(\alpha_{i-1} + \frac{1}{n-i}\right) \quad \text{if } i = t+1 \\
\alpha_i &= \frac{i-1}{i+1}\left(\alpha_{i-1} + \frac{1}{n-i}\right) \quad \text{if } 2 \le i \le t.
\end{aligned}$$

Thus, it suffices to show that $\alpha_i \le \frac{1}{2}$ for all $2 \le i \le t+1$. We will do so by showing that for $1 \le i \le t+1$, we have $\alpha_i \le \frac{i-1}{2(n-i)} \le \frac{i}{2(n-i-1)}$ by induction on i. We know that $\alpha_1 = 0$, so the base case is clear. Let $i > 1$ and assume the induction hypothesis; then

$$\begin{aligned}
\alpha_i &\le \frac{i-1}{i+1}\left(\alpha_{i-1} + \frac{1}{n-i}\right) \le \frac{i-1}{i+1}\left(\frac{i-1}{2(n-(i-1)-1)} + \frac{1}{n-i}\right) \\
&= \frac{i-1}{i+1}\left(\frac{i-1}{2(n-i)} + \frac{2}{2(n-i)}\right) = \frac{i-1}{2(n-i)} \le \frac{i}{2(n-i-1)}.
\end{aligned}$$

We have $\frac{i}{2(n-i-1)} \le \frac{1}{2}$ whenever $i \le n-i-1$, which is true since $i \le t+1 = (\sqrt{2}-1+o(1))(n-1) < \frac{1}{2}(n-1)$. □

Unfortunately, with the help of a computer, we found out that the above construction does not produce a feasible solution to the original dual linear program $\mathcal{D}(\Gamma_n)$, as the lines of slope 1 that are close to the origin (for example, the line $x - y = 1$) tend to have weight exceeding one when n is large.

9.3 Concluding remarks

In this chapter, we studied line coverings with multiplicities for two-dimensional real grids. Despite the progress we made, many natural and interesting questions remain open. We highlight several of them below.

In Section 9.1, we investigated for which grids the Ball-Serra bound is tight. We proved that when n is sufficiently large with respect to m and k, the Ball-Serra bound is tight for any $n \times m$ grid. Moreover, we showed that the threshold value for n given by Theorem 1.3.3 is tight for *most* grids but that there are examples of grids for which it is not tight. It would be interesting to understand what the threshold for an evenly spaced grid is.

Question 10. Let Γ be the grid $\{0, 1, 2, \dots, n-1\} \times \{0, 1, 2, \dots, m-1\}$ and $k \ge 2$ be an integer. How large does n need to be with respect to m and k in order for $h(\Gamma, k)$ to meet the Ball-Serra bound?

In Proposition 9.1.1, we present a family of $n \times 3$ grids for which the value of n given by Theorem 1.3.3 is not best possible. It would be interesting to find a construction for general m.

Our main result for standard grids establishes reasonably good asymptotic lower and upper bounds on $h(\Gamma_n, k)$. It would be of interest to close the remaining gap.

Question 11. What is the true asymptotic value of $h(\Gamma_n, k)$?

We tend to believe that $h(\Gamma_n, k) = (\sqrt{2} + o(1))(n-1)k$. In light of Theorem 1.3.8 and the fact that lines of slope 1 appear to be problematic for the weighting given in the proof of this theorem, as an intermediate problem it might be helpful to consider what happens if we want to cover the grid with lines of slope $0, \infty, -1$, and 1.

In our work thus far we observed that the standard grid Γ_n requires many fewer lines to cover than any other grid we considered. Our general lower bound from Proposition 1.3.5(b) is not strong enough to establish this fact and we propose the following problem.

Question 12. Is it true that $h(\Gamma_n, k) \leq h(\Gamma, k)$ for any $n \times n$ grid Γ in which $\vec{0}$ is a corner?

Finally, it would be interesting to generalize this work to higher-dimensional grids over $\mathbb{R}$. Some first results in this direction were shown in [55].

Glossary

We provide a glossary of the notation used throughout the thesis, meant to serve as a quick reference for the reader. The precise definition of each concept can be found in the General terminology and notation section or in the corresponding chapter of the thesis. For the convenience of the reader, the section is divided into four parts, listing first our general notation and then the notation from each of the three parts of the thesis.

General	
$\mathbb{R}$	the real numbers
$\mathbb{F}_q$	the unique field with q elements
$\mathbb{F}$	a general field
$\mathbb{Z}$	the integers
$\mathbb{Z}_{\geq a}$	the set $\{z \in \mathbb{Z} : z \geq a\}$
$\mathbb{Z}_d$	the cyclic group of order d
$[n]$	the set $\{1, 2, \ldots, n\}$
$z = x \pm y$	z is between $x - y$ and $x + y$
$\binom{n}{k}$	the number of ways to pick k elements from a set of size n
log	the natural logarithm
$\log_2$	the binary logarithm
polylog	polynomial function in log
$f = O(g)$	there exists $C > 0$ such that $\lvert f(n)\rvert \leq C\lvert g(n)\rvert$ for all sufficiently large n
$f = \Omega(g)$	$g = O(f)$
$g = \Theta(f)$	$f = O(g)$ and $f = \Omega(g)$
$f = o(g)$	$f(n)/g(n) \to 0$ as $n \to \infty$
$f = \omega(g)$	$f(n)/g(n) \to \infty$ as $n \to \infty$; $g = o(f)$
$f \ll g$	$f = o(g)$
$f \gg g$	$f = \omega(g)$
$\mathbb{P}[A]$	the probability of the event A

$\mathbb{E}[Y]$	the expectation of the random variable Y
w.h.p.	with high probability; with probability tending to 1 as some parameter tends to infinity

Part I

$V(G)$	the set of vertices of a graph G
$v(G)$	the number of vertices of a graph G
$E(G)$	the set of edges of a graph G
$e(G)$	the number of edges of a graph G
$N_G(v), N(v)$	the neighborhood of a vertex v in a graph G
$d_G(v), d(v)$	the degree of a vertex v in a graph G
$\delta(G)$	the minimum degree of a graph G
$\Delta(G)$	the maximum degree of a graph G
$G \cong H$	G is isomorphic to H
$G[U]$	the subgraph of G induced by a vertex set U
$G - v$	the graph $G[V(G) - \{v\}]$
$G - W$	the graph $G[V(G) - W]$ for a set of vertices W
$G - e$	the graph obtained from G by removing an edge e but keeping its endpoints
$G - F$	the graph obtained from G by removing the edges in a set $F \subseteq E(G)$ but keeping their vertices
$G + v$	the graph obtained from G and a new vertex v by connecting v to all vertices of G
$E_G(U, W)$	for two subsets U and W of $V(G)$, the set of edges in G with one endpoint in U and one endpoint in W
$e_G(U, W)$	for two subsets U and W of $V(G)$, the number of edges in G with one endpoint in U and one endpoint in W
$d_{G,p}(U, W)$	for a real number $0 < p \leq 1$ and subsets U and W of $V(G)$, the quantity $\frac{e_G(U,W)}{p\|U\|\|W\|}$ (see Definition 2.6.3)
$E_G(U)$	for a subset U of $V(G)$, the set of edges in $G[U]$
$e_G(U)$	for a subset U of $V(G)$, the number of edges in $G[U]$
$\alpha(G)$	the independence number of G; the largest size of an independent set of G
$\omega(G)$	the clique number of G; the largest size of an clique in G
girth of G	the shortest length of a cycle in G or ∞ if G is acyclic
$G \cup H$	the graph with vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H)$
$\lambda(H)$	the order of the largest component of H (see Theorem 1.1.9)
K_t	the complete graph on t vertices
$K_{s,t}$	the complete bipartite graph with vertex classes of size s and t

P_t	the path on t *vertices*
C_t	the cycle of length t
$\mathrm{Bin}(n, p)$	the binomial distribution with n trials and success probability p
$G(n, p)$	the binomial random graph on vertex set $[n]$ in which each edge is added with probability p, independently of all other edges
$\mathcal{P}_s$	graph property, that is, subset of all graphs on s vertices (see Lemma 4.3.7)
$V(\mathcal{H})$	the set of vertices of a hypergraph $\mathcal{H}$
$v(\mathcal{H})$	the number of vertices of a hypergraph $\mathcal{H}$
$E(\mathcal{H})$	the set of hyperedges of a hypergraph $\mathcal{H}$
$e(\mathcal{H})$	the number of vertices of a hypergraph $\mathcal{H}$
girth of $\mathcal{H}$	the shortest length of a cycle in a hypergraph $\mathcal{H}$, or ∞ if no such cycle exists (see Section 2.6.1)
$\mathcal{H}_{n,p}$	the binomial random h-uniform hypergraph on vertex set $[n]$, in which every h-set is added as an edge with probability p, independently of all other h-sets (see Section 2.6.1)
q-coloring	an edge-coloring of a graph using at most q different colors (usually from the set $[q]$)
$\varphi^{-1}(i)$	the ith color class under φ; the graph formed by all edges that have color i under φ
$\varphi_{\vert F}$	the coloring induced by φ on the subgraph F
$G_1, \ldots, G_q$	a q-color pattern
$G_1[U], \ldots, G_q[U]$	the q-color pattern induced by $G_1, \ldots, G_q$ on a vertex set U
$\varphi^{-1}(1), \ldots, \varphi^{-1}(q)$	the q-color pattern induced by a q-coloring φ
$(H_1, \ldots, H_q)$	a q-tuple of graphs
$(H_1, \ldots, H_q)$-free, H-free coloring	a q-coloring φ of a graph such that $\varphi^{-1}(i)$ is H_i-free (resp., H-free)
$G \to_q (H_1, \ldots, H_q), G \to_q H$	G is q-Ramsey for $(H_1, \ldots, H_q)$ (resp., H)
$G \not\to_q (H_1, \ldots, H_q), G \not\to_q H$	G is not q-Ramsey for $(H_1, \ldots, H_q)$ (resp., H)
$\mathcal{R}_q(H_1, \ldots, H_q), \mathcal{R}_q(H)$	the collection of q-Ramsey graphs for $(H_1, \ldots, H_q)$ (resp., H)
$\mathcal{M}_q(H_1, \ldots, H_q), \mathcal{M}_q(H)$	the collection of minimal q-Ramsey graphs for $(H_1, \ldots, H_q)$ (resp., H)
$r_q(H_1, \ldots, H_q), r_q(H)$	the q-color Ramsey number of $(H_1, \ldots, H_q)$ (resp., H)
$s_q(H_1, \ldots, H_q), s_q(H)$	the smallest minimum degree of a minimal q-Ramsey graph for $(H_1, \ldots, H_q)$ (resp., H)
$\mathcal{T}(q_1, q_2, \ell, t)$	the $(q_1 + q_2)$-tuple consisting of q_1 cycles C_ℓ and q_2 cliques K_t;
$\mathcal{S}, \mathcal{S}(C_\ell), \mathcal{S}(K_t)$	for the tuple $\mathcal{T}(q_1, q_2, \ell, t)$, the color palettes $[q_1 + q_2]$, $[q_1]$, and $\{q_1 + 1, \ldots, q_1 + q_2\}$, respectively
$P_q(t)$	the q-color t-clique packing parameter (see Definition 3.0.1)

$P_{q_1,q_2}(t)$	a packing parameter generalizing $P_q(t)$ (see Definition 5.2.1)
$\tilde{q}(H)$	the simplicity threshold of a graph H
$S^+((H_1,\dots,H_q),q,e,f)$, $S^+(H,q,e,f)$, S^+	a positive signal sender for $(H_1,\dots,H_q)$ (resp., H) in q colors with signal edges e and f
$S^-((H_1,\dots,H_q),q,e,f)$, $S^-(H,q,e,f)$, S^-	a negative signal sender for $(H_1,\dots,H_q)$ (resp., H) in q colors with signal edges e and f
$I(H,F,q,e), I$	an indicator for H in q colors with indicator subgraph F and indicator edge e
$P(H,F,\mathscr{F},q), P$	a pattern gadget for H in q colors with special subgraph F and set of permissible patterns $\mathscr{F}$
$\mathcal{G}(\ell,m,(V_i)_{i=1}^{\ell},M,\varepsilon)$	a collection of graphs that are all blow-ups of C_ℓ satisfying certain regularity properties (see Section 2.6.1)
$\mathcal{F}(\ell,m,(V_i)_{i=1}^{\ell},M,\varepsilon)$	the graphs in $\mathcal{G}(\ell,m,(V_i)_{i=1}^{\ell},M,\varepsilon)$ not containing a copy of C_ℓ (see Section 2.6.1)

Part II

$L(i,j)$	the entry in row i and column j in a Latin square L
k-MOLS of order n	a sequence of k pairwise orthogonal Latin squares of order n
$(L_1,\dots,L_k)$	a k-MOLS
$L(n)$	the number of Latin squares of order n
$L^{(k)}(n)$	the number of k-MOLS of order n
S_n	the $n\times n$ square with entries $S_n(i,j)=i$ for all $i,j\in[n]$
$OA(n,d)$	an $n^2\times d$ orthogonal array with entries in $[n]$ (see Definition 7.1.2)
v_1	the vector $[1,\dots,1,2,\dots,2,\dots,n,\dots,n]$ of length n^2
v_2	the vector $[1,2\dots,n,1,2\dots,n,\dots,1,2\dots,n]$ of length n^2
$NOA(n,d)$	an $n^2\times d$ nearly orthogonal array with entries in n (see Definition 7.1.3)
$A(i,j)$	the entry in row i and column j in a (nearly) orthogonal array A
r_ℓ	$\lvert\{s\neq\ell : A(s,1)=A(\ell,1) \text{ and } A(s,3)=A(\ell,3)\}\rvert$ (see Theorem 7.2.1)
c_ℓ	$\lvert\{s\neq\ell : A(s,2)=A(\ell,2) \text{ and } A(s,3)=A(\ell,3)\}\rvert$ (see Theorem 7.2.1)
$R(X)$	the range of a random variable X (see Section 7.1.3)
$H(X)$	the (base e) entropy of a random variable X (see Section 7.1.3)
$H(X,Y)$	the joint entropy of two random variables X and Y (see Section 7.1.3)
$H(X\mid Y)$	the conditional entropy of a random variable X given a random variable Y (see Section 7.1.3)
I_d	the integral $\int_0^1 \log(1+(n-1)t^d)\,\mathrm{d}t$ (see Lemma 7.2.2)

$L_1 \otimes L_2$	the Kronecker product of two Latin squares (see Section 7.3)
$L^{\otimes k}$	the k-fold Kronecker product $L \otimes \cdots \otimes L$ (see Section 7.3)

Part III

$H_{\vec{u}}$	the hyperplane defined by the equation $\vec{u} \cdot \vec{x} = 1$
$\Gamma(S_1, \ldots, S_n), \Gamma$	the grid $S_1 \times \cdots \times S_n$
Γ^-	$\Gamma \setminus \{\vec{0}\}$
(k, d)-cover of $\Gamma \subseteq \mathbb{F}^n$	a multiset of $(n-d)$-dimensional affine subspaces in $\mathbb{F}^n$ that cover all nonzero points of Γ at least k times while covering $\vec{0}$ fewer times
k-cover	a (k, d)-cover for $d = 1$
$(k, d; s)$-cover	a (k, d)-cover in which the origin is covered exactly s times
strict (k, d)-cover	a (k, d)-cover in which the origin is not covered
strict k-cover	a strict (k, d)-cover for $d = 1$
$f(n, k, d), f(n, k)$	the minimum size of a (k, d)-cover of $\mathbb{F}_2^n$
$g(n, k, d; s)$	the minimum size of a $(k, d; s)$-cover of $\mathbb{F}_2^n$
$h(\Gamma, k)$	minimum size of a strict k-cover of Γ
Δ-bounded grid	a grid in $\mathbb{R}^2$ such that every line containing two axis points contains at most Δ interior points (see Section 1.3.4.2)
Γ_n	the standard grid $\{0, 1, 2, \ldots, n-1\}^2 \subseteq \mathbb{R}^2$

Zusammenfassung

Diese Dissertation besteht aus drei unabhängigen Teilen.

Der erste Teil beschäftigt sich mit Ramseytheorie. Für eine ganze Zahl $q \geq 2$ nennt man einen Graphen *q-Ramsey* für einen anderen Graphen H, wenn jede Kantenfärbung mit q Farben einen einfarbigen Teilgraphen enthält, der isomorph zu H ist. Das zentrale Problem in diesem Gebiet ist die minimale Anzahl von Knoten in einem solchen Graphen zu bestimmen. In dieser Dissertation betrachten wir zwei verschiedene Varianten. Als erstes, beschäftigen wir uns mit dem kleinstmöglichen Minimalgrad eines minimalen (bezüglich Teilgraphen) q-Ramsey-Graphen für einen gegebenen Graphen H. Diese Frage wurde zuerst von Burr, Erdős und Lovász in den 1970er-Jahren studiert. Wir betrachten dieses Problem für einen Zufallsgraphen und untersuchen, wie viele Knoten kleinen Grades ein Ramsey-Graph für gegebenes H enthalten kann. Wir untersuchen auch eine asymmetrische Verallgemeinerung des Minimalgradproblems. Als zweites betrachten wir die Frage, wie sich die Menge aller q-Ramsey-Graphen für H verändert, wenn wir den Graphen H modifizieren. Aufbauend auf den Arbeiten von Fox, Grinshpun, Liebenau, Person und Szabó und Rödl und Siggers beweisen wir, dass bereits der Graph, der aus K_t mit einer hängenden Kante besteht, eine sehr unterschiedliche Menge von 2-Ramsey-Graphen besitzt im Vergleich zu K_t.

Im zweiten Teil geht es um orthogonale lateinische Quadrate. Ein *lateinisches Quadrat der Ordnung n* ist eine $n \times n$-Matrix, gefüllt mit den Zahlen aus $[n]$, in der jede Zahl genau einmal pro Zeile und einmal pro Spalte auftritt. Zwei lateinische Quadrate sind *orthogonal* zueinander, wenn für alle $x, y \in [n]$ genau ein Paar $(i, j) \in [n]^2$ existiert, sodass es $L(i, j) = x$ und $L'(i, j) = y$ gilt. Ein *k-MOLS der Ordnung n* ist eine Menge von k lateinischen Quadraten, die paarweise orthogonal sind. Motiviert von einem bekannten Resultat, welches die Anzahl von lateinischen Quadraten der Ordnung n log-asymptotisch bestimmt, untersuchen wir die Frage, wie viele k-MOLS der Ordnung n es gibt. Dies wurde bereits von Donovan und Grannell und Keevash und Luria studiert. Wir verbessern die beste obere Schranke für einen breiten Bereich von Parametern $k = k(n)$. Zusätzlich bestimmen wir log-asymptotisch zu wie viele anderen lateinischen Quadraten ein lateinisches Quadrat orthogonal sein kann.

Im dritten Teil studieren wir, wie viele Hyperebenen notwendig sind, um die Punkte eines endlichen Gitters zu überdecken, sodass ein bestimmter Punkt maximal $(k-1)$-mal bedeckt ist und alle andere mindestens k-mal. Wir untersuchen diese Anzahl für das Gitter $\mathbb{F}_2^n$ asymptotisch und sogar genau, wenn eins von n und k viel größer als das andere ist. Dies verallgemeinert ein Ergebnis von Jamison für den Fall $k = 1$. Außerdem betrachten wir dieses Problem für Gitter im reellen Vektorraum, wenn der spezielle Punkt überhaupt nicht bedeckt ist. Dies ist durch die Arbeiten von Clifton und Huang und Sauermann und Wigderson motiviert, die den Hyperwürfel $\{0, 1\}^n \subseteq \mathbb{R}^n$ untersucht haben. Wir konzentrieren uns auf zwei-dimensionale Gitter und zeigen, dass schon diese sich sehr unterschiedlich verhalten können.

Selbstständigkeitserklärung

Name: Boyadzhiyska
Vorname: Simona

Ich erkläre gegenüber der Freien Universität Berlin, dass ich die vorliegende Dissertation selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe. Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht. Diese Dissertation wurde in gleicher oder ähnlicher Form noch in keinem früheren Promotionsverfahren eingereicht.

Mit einer Prüfung meiner Arbeit durch ein Plagiatsprüfungsprogramm erkläre ich mich einverstanden.

Datum: ________________ Unterschrift: ________________________________

Bibliography

[1] Research problems. *J. Combinatorial Theory*, 2(1):104–105, 1967.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A*, 29(3):354–360, 1980.

[3] N. Alon and Z. Füredi. Covering the cube by affine hyperplanes. *European J. Combin.*, 14(2):79–83, 1993.

[4] V. Angeltveit and B. D. McKay. $R(5,5) \leq 48$. *J. Graph Theory*, 89(1):5–13, 2018.

[5] M. Axenovich, J. Rollin, and T. Ueckerdt. Conditions on Ramsey nonequivalence. *J. Graph Theory*, 86(2):159–192, 2017.

[6] R. A. Bailey, P. J. Cameron, and R. Connelly. Sudoku, gerechte designs, resolutions, affine space, spreads, reguli, and Hamming codes. *Amer. Math. Monthly*, 115(5):383–404, 2008.

[7] R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes, II. *Proc. London Math. Soc. (3)*, 83(3):532–562, 2001.

[8] S. Ball. On intersection sets in Desarguesian affine spaces. *European J. Combin.*, 21(4):441–446, 2000.

[9] S. Ball. *Finite geometry and combinatorial applications*, volume 82 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2015.

[10] S. Ball and A. Blokhuis. A bound for the maximum weight of a linear code. *SIAM J. Discrete Math.*, 27(1):575–583, 2013.

[11] S. Ball and O. Serra. Punctured Combinatorial Nullstellensätze. *Combinatorica*, 29(5):511–522, 2009.

[12] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *J. Amer. Math. Soc.*, 28(3):669–709, 2015.

[13] J. Bamberg, A. Bishnoi, and T. Lesgourgues. The minimum degree of minimal Ramsey graphs for cliques. *Bull. Lond. Math. Soc.*, 54(5):1827–1838, 2022.

[14] W. U. Behrens. Feldversuchsanordnungen mit verbessertem Ausgleich der Bodenunterschiede (in German). *Zeitschrift für Landwirtschaftliches Versuchs- und Untersuchungswesen*, 2:176–193, 1956.

[15] D. Berend. On the number of Sudoku squares. *Discrete Math.*, 341(11):3241–3248, 2018.

[16] A. Bishnoi, S. Boyadzhiyska, D. Clemens, P. Gupta, T. Lesgourgues, and A. Liebenau. On the minimum degree of minimal Ramsey graphs for cliques versus cycles. *SIAM J. Discrete Math.*, 37(1):25–50, 2023.

[17] A. Bishnoi, S. Boyadzhiyska, S. Das, and Y. den Bakker. Covering grids with multiplicity. *Comb. Theory*, to appear. arXiv:2305.00825.

[18] A. Bishnoi, S. Boyadzhiyska, S. Das, and T. Mészáros. Subspace coverings with multiplicities, 2021. arXiv:2101.11947.

[19] A. Bishnoi, S. Boyadzhiyska, S. Das, and T. Mészáros. Subspace coverings with multiplicities. *Combin. Probab. Comput.*, 32(5):782–795, 2023.

[20] A. Bishnoi, P. L. Clark, A. Potukuchi, and J. R. Schmitt. On zeros of a polynomial in a finite grid. *Combin. Probab. Comput.*, 27(3):310–333, 2018.

[21] A. Blokhuis, A. E. Brouwer, and T. Szőnyi. Covering all points except one. *J. Algebraic Combin.*, 32(1):59–66, 2010.

[22] T. F. Bloom and A. Liebenau. Ramsey equivalence of K_n and $K_n + K_{n-1}$. *Electron. J. Combin.*, 25(3):P3.4, 2018.

[23] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. *Random Structures Algorithms*, 58(2):221–293, 2021.

[24] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[25] B. Bollobás, J. Donadelli, Y. Kohayakawa, and R. H. Schelp. Ramsey minimal graphs. *Journal of the Brazilian Computer Society*, 7(3):27–37, 2001.

[26] E. Boros, T. Szőnyi, and K. Tichler. On defining sets for projective planes. *Discrete Math.*, 303(1-3):17–31, 2005.

[27] S. Boyadzhiyska, D. Clemens, S. Das, and P. Gupta. Ramsey simplicity of random graphs, 2021. arXiv:2109.04140.

[28] S. Boyadzhiyska, D. Clemens, and P. Gupta. Minimal Ramsey graphs with many vertices of small degree, 2020. arXiv:2009.04159.

[29] S. Boyadzhiyska, D. Clemens, and P. Gupta. Minimal Ramsey graphs with many vertices of small degree. *SIAM J. Discrete Math.*, 36(3):1503–1528, 2022.

[30] S. Boyadzhiyska, S. Das, and T. Szabó. Enumerating extensions of mutually orthogonal Latin squares. *Des. Codes Cryptogr.*, 88(10):2187–2206, 2020.

[31] L. Brègman. Some properties of nonnegative matrices and their permanents. *Soviet Math. Dokl.*, 14:945–949, 1973.

[32] A. E. Brouwer and A. Schrijver. The blocking number of an affine space. *J. Combin. Theory Ser. A*, 24(2):251–253, 1978.

[33] A. A. Bruen. Polynomial multiplicities over finite fields and intersection sets. *J. Combin. Theory Ser. A*, 60(1):19–33, 1992.

[34] A. A. Bruen and J. C. Fisher. The Jamison method in Galois geometries. *Des. Codes Cryptogr.*, 1(3):199–205, 1991.

[35] M. Bryant, J. Figler, R. Garcia, C. Mummert, and Y. Singh. The number of mates of Latin squares of sizes 7 and 8. *Congr. Numer.*, 217:53–64, 2013.

[36] S. A. Burr, P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. Ramsey-minimal graphs for star-forests. *Discrete Math.*, 33(3):227–237, 1981.

[37] S. A. Burr, P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. Ramsey-minimal graphs for forests. *Discrete Math.*, 38(1):23–32, 1982.

[38] S. A. Burr, P. Erdős, R. J. Faudree, and R. H. Schelp. A class of Ramsey-finite graphs. In *Proceedings of the Ninth Southeastern Conference on Combinatorics, Graph Theory, and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1978)*, Congress. Numer., XXI, pages 171–180. Utilitas Math., Winnipeg, Man., 1978.

[39] S. A. Burr, P. Erdős, R. J. Faudree, and R. H. Schelp. On the difference between consecutive Ramsey numbers. *Utilitas Math.*, 35:115–118, 1989.

[40] S. A. Burr, P. Erdős, and L. Lovász. On graphs of Ramsey type. *Ars Combin.*, 1(1):167–190, 1976.

[41] S. A. Burr, R. J. Faudree, and R. H. Schelp. On Ramsey-minimal graphs. In *Proceedings of the Eighth Southeastern Conference on Combinatorics, Graph Theory and Computing (Louisiana State Univ., Baton Rouge, La., 1977)*, Congress. Numer., XIX, pages 115–124. Utilitas Math., Winnipeg, Man., 1977.

[42] S. A. Burr, J. Nešetřil, and V. Rödl. On the use of senders in generalized Ramsey theory for graphs. *Discrete Math.*, 54(1):1–13, 1985.

[43] N. J. Cavenagh and I. M. Wanless. Latin squares with no transversals. *Electron. J. Combin.*, 24(2):P2.45, 2017.

[44] S. Chowla, P. Erdős, and E. G. Straus. On the maximal number of pairwise orthogonal Latin squares of a given order. *Canad. J. Math.*, 12:204–208, 1960.

[45] V. Chvátal. Tree-complete graph Ramsey numbers. *J. Graph Theory*, 1(1):93, 1977.

[46] D. Clemens, A. Liebenau, and D. Reding. On minimal Ramsey graphs and Ramsey equivalence in multiple colours. *Combin. Probab. Comput.*, 29(4):537–554, 2020.

[47] A. Clifton and H. Huang. On almost k-covers of hypercubes. *Combinatorica*, 40(4):511–526, 2020.

[48] D. Conlon. A new upper bound for diagonal Ramsey numbers. *Ann. of Math. (2)*, 170(2):941–960, 2009.

[49] D. Conlon. The Ramsey number of dense graphs. *Bull. Lond. Math. Soc.*, 45(3):483–496, 2013.

[50] D. Conlon and A. Ferber. Lower bounds for multicolor Ramsey numbers. *Adv. Math.*, 378:Paper No. 107528, 2021.

[51] D. Conlon, J. Fox, and B. Sudakov. Recent developments in graph Ramsey theory. In *Surveys in combinatorics 2015*, volume 424 of *London Math. Soc. Lecture Note Ser.*, pages 49–118. Cambridge Univ. Press, Cambridge, 2015.

[52] D. Conlon, J. Fox, and B. Sudakov. Short proofs of some extremal results III. *Random Structures Algorithms*, 57(4):958–982, 2020.

[53] T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, Inc., New York, 1991.

[54] G. Damásdi, B. Keszegh, D. Malec, C. Tompkins, Z. Wang, and O. Zamora. Saturation problems in the Ramsey theory of graphs, posets and point sets. *European J. Combin.*, 95:Paper No. 103321, 2021.

[55] Y. den Bakker. *Hyperplane covering problems*. Bachelor's thesis, Technische Universiteit Delft, 2021.

[56] D. M. Donovan and M. J. Grannell. On the number of transversal designs. *J. Combin. Theory Ser. A*, 120(7):1562–1574, 2013.

[57] S. Eberhard. More on additive triples of bijections, 2017. arXiv:1704.02407.

[58] S. Eberhard, F. Manners, and R. Mrazović. Additive triples of bijections, or the toroidal semiqueens problem. *J. Eur. Math. Soc.*, 21(2):441–463, 2019.

[59] J. Egan and I. M. Wanless. Enumeration of MOLS of small order. *Math. Comp.*, 85(298):799–824, 2016.

[60] G. P. Egorychev. The solution of van der Waerden's problem for permanents. *Adv. in Math.*, 42(3):299–305, 1981.

[61] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

[62] P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. The size Ramsey number. *Period. Math. Hungar.*, 9(1-2):145–161, 1978.

[63] P. Erdős and A. Hajnal. On chromatic number of graphs and set-systems. *Acta Math. Acad. Sci. Hungar.*, 17(1):61–99, 1966.

[64] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[65] G. Exoo. A lower bound for $R(5,5)$. *J. Graph Theory*, 13(1):97–98, 1989.

[66] D. I. Falikman. A proof of van der Waerden's conjecture on the permanent of a doubly stochastic matrix (in Russian). *Mat. Zametki*, 29(6):931–938, 1981.

[67] R. J. Faudree and R. H. Schelp. All Ramsey numbers for cycles in graphs. *Discrete Math.*, 8(4):313–329, 1974.

[68] A. Ferber and M. Kwan. Almost all Steiner triple systems are almost resolvable. *Forum Math. Sigma*, 8:E39, 2020.

[69] G. Fiz Pontiveros, S. Griffiths, and R. Morris. The triangle-free process and the Ramsey number $R(3,k)$. *Memoirs of the American Mathematical Society*, 263:v+125, 2020.

[70] J. Folkman. Graphs with monochromatic complete subgraphs in every edge coloring. *SIAM J. Appl. Math.*, 18:19–24, 1970.

[71] J. Fox, A. Grinshpun, A. Liebenau, Y. Person, and T. Szabó. What is Ramsey-equivalent to a clique? *J. Combin. Theory Ser. B*, 109:120–133, 2014.

[72] 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*, 120:64–82, 2016.

[73] J. Fox and K. Lin. The minimum degree of Ramsey-minimal graphs. *J. Graph Theory*, 54(2):167–177, 2007.

[74] J. Fox and B. Sudakov. Two remarks on the Burr-Erdős conjecture. *European J. Combin.*, 30(7):1630–1645, 2009.

[75] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016.

[76] P. Gaborit and G. Zemor. Asymptotic improvement of the Gilbert–Varshamov bound for linear codes. *IEEE Trans. Inform. Theory*, 54(9):3865–3872, 2008.

[77] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger. Small subsets inherit sparse ε-regularity. *J. Combin. Theory Ser. B*, 97(1):34–56, 2007.

[78] S. Gerke and A. Steger. The sparse regularity lemma and its applications. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 227–258. Cambridge Univ. Press, Cambridge, 2005.

[79] E. N. Gilbert. A comparison of signalling alphabets. *The Bell System Technical Journal*, 31(3):504–522, 1952.

[80] R. Glebov and Z. Luria. On the maximum number of Latin transversals. *J. Combin. Theory Ser. A*, 141:136–146, 2016.

[81] S. W. Golomb. Problem 11214. *Amer. Math. Monthly*, 113.3(3):268, 2006.

[82] A. Grinshpun, R. Raina, and R. Sengupta. Minimum degrees of minimal Ramsey graphs for almost-cliques. *J. Graph Theory*, 85(2):349–362, 2017.

[83] A. V. Grinshpun. *Some Problems in Graph Ramsey Theory*. PhD thesis, Massachusetts Institute of Technology, 2015.

[84] H. Guo and L. Warnke. Packing nearly optimal Ramsey $R(3,t)$ graphs. *Combinatorica*, 40(1):63–103, 2020.

[85] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2022.

[86] H. Hàn, V. Rödl, and T. Szabó. Vertex Folkman numbers and the minimum degree of minimal Ramsey graphs. *SIAM J. Discrete Math.*, 32(2):826–838, 2018.

[87] P. E. Haxell, Y. Kohayakawa, and T. Łuczak. The induced size-Ramsey number of cycles. *Combin. Probab. Comput.*, 4(3):217–239, 1995.

[88] A. Hedayat and W. T. Federer. On embedding and enumeration of orthogonal Latin squares. *Ann. Math. Statist.*, 42(2):509–516, 1971.

[89] A. S. Hedayat, N. J. A. Sloane, and J. Stufken. *Orthogonal Arrays: Theory and Applications*. Springer-Verlag, New York, 1999.

[90] R. E. Jamison. Covering finite fields with cosets of subspaces. *J. Combinatorial Theory Ser. A*, 22(3):253–266, 1977.

[91] T. Jiang and A. Vardy. Asymptotic improvement of the Gilbert–Varshamov bound on the size of binary codes. *IEEE Trans. Inform. Theory*, 50(8):1655–1664, 2004.

[92] J. Kahn. On a problem of Erdős and Lovász: random lines in a projective plane. *Combinatorica*, 12(4):417–423, 1992.

[93] W. M. Kantor. Commutative semifields and symplectic spreads. *J. Algebra*, 270(1):96–114, 2003.

[94] W. M. Kantor and M. E. Williams. Symplectic semifield planes and $\mathbb{Z}_4$-linear codes. *Trans. Amer. Math. Soc.*, 356(3):895–938, 2004.

[95] A. D. Keedwell. On Sudoku squares. *Bull. Inst. Comb. Appl.*, 50:52–60, 2007.

[96] A. D. Keedwell. Constructions of complete sets of orthogonal diagonal Sudoku squares. *Australas. J. Combin.*, 47:227–238, 2010.

[97] P. Keevash. Counting Steiner triple systems. In *European Congress of Mathematics*, pages 459–481, 2018.

[98] P. Keevash. Coloured and directed designs. In I. Bárány, G. Katona, and A. Sali, editors, *Building Bridges II: Mathematics of László Lovász*, pages 279–315. Springer Berlin Heidelberg, 2019.

[99] P. Keevash, E. Long, and J. Skokan. Cycle-complete Ramsey numbers. *Int. Math. Res. Not.*, 2021(1):275–300, 2021.

[100] J. H. Kim. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Structures Algorithms*, 7(3):173–207, 1995.

[101] S. Kogan. New results on k-independence of graphs. *Electron. J. Combin.*, 24(2):P2.15, 2017.

[102] Y. Kohayakawa, T. Łuczak, and V. Rödl. On K^4-free subgraphs of random graphs. *Combinatorica*, 17(2):173–213, 1997.

[103] P. Komjáth. Partitions of vector spaces. *Period. Math. Hungar.*, 28(3):187–193, 1994.

[104] G. Kós, T. Mészáros, and L. Rónyai. Some extensions of Alon's Nullstellensatz. *Publ. Math. Debrecen*, 79(3-4):507–519, 2011.

[105] G. Kós and L. Rónyai. Alon's Nullstellensatz for multisets. *Combinatorica*, 32(5):589–605, 2012.

[106] P. Kővári, V. T. Sós, and P. Turán. On a problem of K. Zarankiewicz. *Colloq. Math.*, 3:50–57, 1954.

[107] I. Landjev and A. Rousseva. On the sharpness of Bruen's bound for intersection sets in Desarguesian affine spaces. *Des. Codes Cryptogr.*, 72(3):551–558, 2014.

[108] S. Letzter. Path Ramsey number for random graphs. *Combin. Probab. Comput.*, 25(4):612–622, 2016.

[109] N. Linial and Z. Luria. An upper bound on the number of Steiner triple systems. *Random Structures Algorithms*, 43(4):399–406, 2013.

[110] J. Lorch. Mutually orthogonal families of linear Sudoku solutions. *J. Aust. Math. Soc.*, 87(3):409–420, 2009.

[111] J. Lorch. Orthogonal combings of linear Sudoku solutions. *Australas. J. Combin.*, 47:247–264, 2010.

[112] M. G. Lu. The maximum number of mutually orthogonal Latin squares. *Kexue Tongbao (English Ed.)*, 30(2):154–159, 1985.

[113] Z. Luria. New bounds on the number of n-queens configurations, 2017. arXiv:1705.05225.

[114] H. MacNeish. Euler squares. *Ann. of Math. (2)*, 23(3):221–227, 1922.

[115] J. Matoušek and B. Gärtner. *Understanding and using linear programming*. Springer Berlin, Heidelberg, 2007.

[116] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, pages 195–248. Springer, Berlin, 1998.

[117] B. D. McKay, A. Meynert, and W. Myrvold. Small Latin squares, quasigroups, and loops. *J. Combin. Des.*, 15(2):98–119, 2007.

[118] B. Mounits, T. Etzion, and S. Litsyn. Improved upper bounds on sizes of codes. *IEEE Trans. Inform. Theory*, 48(4):880–886, 2002.

[119] R. Nenadov and A. Steger. A short proof of the Random Ramsey theorem. *Combin. Probab. Comput.*, 25(1):130–144, 2016.

[120] J. Nešetřil and V. Rödl. The Ramsey property for graphs with forbidden complete subgraphs. *J. Combin. Theory Ser. B*, 20(3):243–249, 1976.

[121] J. Nešetřil and V. Rödl. The structure of critical Ramsey graphs. *Acta Math. Acad. Sci. Hungar.*, 32(3-4):295–300, 1978.

[122] R. M. Pedersen and T. L. Vis. Sets of mutually orthogonal Sudoku Latin squares. *College Math. J.*, 40(3):174–180, 2009.

[123] J. Radhakrishnan. An entropy proof of Bregman's theorem. *J. Combin. Theory Ser. A*, 77(1):161–164, 1997.

[124] S. P. Radziszowski. Small Ramsey numbers. *Electron. J. Combin.*, 1:Dynamic Survey 1, 1994.

[125] F. P. Ramsey. On a problem of formal logic. *Proc. London Math. Soc. (2)*, 30(4):264–286, 1929.

[126] V. Rödl and A. Ruciński. Lower bounds on probability thresholds for Ramsey properties. In *Combinatorics, Paul Erdős is eighty*, volume 1, pages 317–346. János Bolyai Math. Soc., Budapest, 1993.

[127] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.*, 8(4):917–942, 1995.

[128] V. Rödl and M. Siggers. On Ramsey minimal graphs. *SIAM J. Discrete Math.*, 22(2):467–488, 2008.

[129] K. Rogers. A note on orthogonal Latin squares. *Pacific J. Math.*, 14:1395–1397, 1964.

[130] V. Rosta. On a Ramsey-type problem of J. A. Bondy and P. Erdős. II. *J. Combin. Theory Ser. B*, 15(1):105–120, 1973.

[131] H. J. Ryser. Permanents and systems of distinct representatives. In *Combinatorial Mathematics and Its Applications, Proc. Conf., Univ. North Carolina, Chapel Hill, NC, 1967*, pages 55–68, Chapel Hill, NC, 1969. Univ. North Carolina Press.

[132] A. Sah. Diagonal Ramsey via effective quasirandomness. *Duke Math. J.*, 172(3):545–567, 2023.

[133] L. Sauermann and Y. Wigderson. Polynomials that vanish to high order on most of the hypercube. *J. Lond. Math. Soc. (2)*, 106(3):2379–2402, 2022.

[134] M. Savery. Chromatic number is Ramsey distinguishing. *J. Graph Theory*, 99(1):152–161, 2022.

[135] W. Sawin. An improved lower bound for multicolor Ramsey numbers and a problem of Erdős. *J. Combin. Theory Ser. A*, 188:Paper No. 105579, 2022.

[136] D. Saxton and A. Thomason. Hypergraph containers. *Invent. Math.*, 201(3):925–992, 2015.

[137] A. Schrijver. New code upper bounds from the Terwilliger algebra and semidefinite programming. *IEEE Trans. Inform. Theory*, 51(8):2859–2866, 2005.

[138] J. B. Shearer. A note on the independence number of triangle-free graphs. *Discrete Math.*, 46(1):83–87, 1983.

[139] M. Siggers. Five cycles are highly Ramsey infinite. *Kyungpook Math. J.*, 52(1):13–20, 2012.

[140] M. Siggers. Non-bipartite pairs of 3-connected graphs are highly Ramsey-infinite. *European J. Combin.*, 36:172–189, 2014.

[141] M. H. Siggers. On highly Ramsey infinite graphs. *J. Graph Theory*, 59(2):97–114, 2008.

[142] J. Spencer. Ramsey's theorem – a new lower bound. *J. Combin. Theory Ser. A*, 18(1):108–115, 1975.

[143] T. Szabó, P. Zumstein, and S. Zürcher. On the minimum degree of minimal Ramsey graphs. *J. Graph Theory*, 64(2):150–164, 2010.

[144] E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

[145] A. A. Taranenko. Multidimensional permanents and an upper bound on the number of transversals in Latin squares. *J. Combin. Des.*, 23(7):305–320, 2015.

[146] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.0-9.5)*, 2020-2022. `https://www.sagemath.org`.

[147] J. H. Van Lint and R. M. Wilson. *A course in combinatorics*. Cambridge University Press, Cambridge, second edition, 2001.

[148] G. H. J. van Rees. Subsquares and transversals in Latin squares. *Ars Combin.*, 29(B):193–204, 1990. Twelfth British Combinatorial Conference (Norwich, 1989).

[149] R. R. Varshamov. Estimate of the number of signals in error correcting codes. *Dokl. Akad. Nauk, SSSR*, 117:739–741, 1957.

[150] I. M. Wanless and B. S. Webb. The existence of Latin squares without orthogonal mates. *Des. Codes Cryptogr.*, 40(1):131–135, 2006.

[151] D. B. West. *Introduction to graph theory*. Prentice hall, Upper Saddle River, NJ, second edition, 2001.

[152] Y. Wigderson. An improved lower bound on multicolor Ramsey numbers. *Proc. Amer. Math. Soc.*, 149(6):2371–2374, 2021.

[153] R. M. Wilson. Concerning the number of mutually orthogonal Latin squares. *Discrete Math.*, 9:181–198, 1974.

[154] C. Zanella. Intersection sets in $\mathrm{AG}(n,q)$ and a characterization of the hyperbolic quadric in $\mathrm{PG}(3,q)$. *Discrete Math.*, 255(1-3):381–386, 2002.