TEL AVIV UNIVERSITY אוניברסיטת תל-אביב

RAYMOND AND BEVERLY SACKLER FACULTY OF EXACT SCIENCES
SCHOOL OF MATHEMATICAL SCIENCES

הפקולטה למדעים מדויקים ע"ש ריימונד ובברלי סאקלר
בית הספר למדעי המתמטיקה

Graph Powers and Related Extremal Problems

Thesis submitted for the degree "Doctor of Philosophy"

by

Eyal Lubetzky

under the supervision of

Professor Noga Alon

Submitted to the Senate of Tel Aviv University

June 2007

This work was carried out under the supervision of

Professor Noga Alon.

Acknowledgments

First and foremost, I would like to thank my advisor, Professor Noga Alon. Noga has been a tremendous source of knowledge and inspiration for me throughout these past 5 years. He has taught me the fundamentals of Graph Theory and Combinatorics, and has been a role model for me in every aspect of scientific research. His keen insight and meticulous expertise in numerous fields in Mathematics and Computer Science have turned each conversation with him into a unique learning experience. It has been a privilege to be Noga's student, and I am grateful for the time we spent together.

I am indebted to Professor Benny Sudakov, whose many invaluable advice during my graduate studies helped me focus my research interests and set the goals I wish to achieve as a Mathematician. In addition, several of the projects in this thesis have benefited from Benny's sharp and useful remarks, and I was lucky to enjoy his vast knowledge in Mathematics, over many constructive discussions.

It is a pleasure to thank Professor Itai Benjamini for his never-ending supplies of intriguing and novel ideas for research, as well as for his great enthusiasm in developing them. I wish to thank Professor Simon Litsyn for our illuminating talks, which sparked my interest in Coding Theory. I would also like to thank Professor Michael Krivelevich, whose skills of assessing the primary obstacles in a problem and then tackling them have taught me a lot. I am more than grateful for the fruitful conservations we had when I was facing career decisions.

Many thanks to my dear friends and colleagues Dr. Amit Singer and Uri Stav, whom it has been a pleasure to work with. I want to express my appreciation to my other coauthors during the course of my studies, Tomer

Amiaz, Gidi Amir, Ori Gurel-Gurevich, Simi Haber and Sasha Sodin. I also want to thank my fellow graduate students and friends Sonny Ben-Shimon, Danny Hefetz, Danny Vilenchik and Oded Schwartz, for a work atmosphere with never a dull moment.

Special thanks are due to the Charles Clore Scholars Programme, which generously supported me in the final year of my studies.

Finally, I wish to thank my dear parents and family for their constant encouragement and support over the many years of my education. I feel lucky to have such a giving and loving family, which made this achievement possible. Last but not least, I wish to thank my beloved wife Anat, whom I deeply love and adore. Her advice and support make every task look feasible, and her smile and laugh give everything a new meaning.

E. L.

Tel Aviv, Israel
June 2007

Abstract

The main focus of this work is the study of various parameters of high powers of a given graph, under different definitions of the product operator. The problems we study and the methods used have applications in Extremal Combinatorics, Ramsey Theory, and Coding Theory. This thesis consists of four parts:

In Part I, we study *strong* graph powers, the *Shannon capacity* and problems related to it. This challenging parameter, introduced by Shannon (1956), measures the effective alphabet size in a zero-error transmission over a noisy channel. In Chapter 1 we address the problem of approximating this parameter, and give a probabilistic construction of graphs, whose powers exhibit an arbitrarily complicated behavior in terms of their independence numbers. In particular, this shows that there are graphs, whose capacity cannot be approximated (up to a small power of their number of vertices) by any fixed graph power. Chapter 2 discusses the capacity of a disjoint union of graphs, corresponding to a case where several senders combine their channels. Alon (1998) showed that this capacity may exceed the sum of the individual capacities, and we extend this result as follows. For any family of "privileged" subsets of t senders, we construct a graph for every sender, so that the capacity of the disjoint union of every subset I of these graphs is "large" if I contains a privileged subset, and "small" otherwise. This corresponds to a case where only privileged subsets of senders are allowed to transmit in a high rate. In the process, we obtain an explicit Ramsey construction of t-edge-colorings of the complete graph, where every induced "large" subgraph contains *all* t colors. Chapter 3 deals with *index-coding*, a source-coding problem suggested by Birk and Kol (1998). In this problem, a sender wishes to broadcast codewords of minimal length to a set of receivers; each receiver is interested in a specific block of the input data, and has some prior side-information on other blocks. Bar-Yossef, Birk, Jayram and Kol (2006) characterized the length of an optimal *linear* index-code in terms of the graph modeling the side-information. They proved that in various cases it attains the overall optimum of the problem, and their main conjecture was that linear index-coding is in fact *always* optimal. Using an explicit construction of a Ramsey graph and algebraic upper bounds on its Shannon capacity, we disprove this conjecture in the following strong sense: there are settings where a linear index-code requires $n^{1-o(1)}$ bits, barely improving the n bits required by the naïve protocol, and yet a given non-linear index-code utilizes only $n^{o(1)}$ bits.

Chapter 4 discusses multiple-round index-coding, and relates it to Witsenhausen's rate and colorings of OR graph powers. This provides an alternative proof that linear index-codes are suboptimal (this time, by a multiplicative constant).

Part II is devoted to certain graph powers, which yield dense random-looking graphs, and have applications in Coding Theory and Ramsey Theory. In Chapter 5, we introduce parameters describing the independence numbers and clique numbers in *Xor* powers of a graph, and relate them to problems in Coding Theory. We study the value of these parameters for various families of graphs, and provide general lower and upper bounds for them using tools from Algebra and Spectral Analysis. En route, we prove that large Xor powers of a fixed graph have certain pseudo-random properties, and a natural generalization of the Xor power has useful properties in Ramsey Theory. This generalized graph power is studied in Chapter 6, where we give some tighter bounds on the above coding problems using Delsarte's LP bound, among other ideas. We show that large powers of any nontrivial graph G contain large Ramsey subgraphs; if G is the complete graph, then some power of G matches the bounds of the famous Ramsey construction of Frankl and Wilson (1981), and is in fact a subgraph of a variant of that graph. The mentioned Frankl and Wilson construction is based on set systems with prescribed intersections, motivating our next results.

In Part III, we examine set systems with restricted pairwise intersections. This well studied area is indeed closely related to Ramsey Theory, Coding Theory and Communication Complexity. Two families of subsets of an n-element set, $\mathcal{A}$ and $\mathcal{B}$, are called *ℓ-cross-intersecting* if the intersection of every set in $\mathcal{A}$ with every set in $\mathcal{B}$ contains precisely ℓ elements. The problem of determining the maximal value of $|\mathcal{A}||\mathcal{B}|$ over all ℓ-cross-intersecting pairs of families has attracted a considerable amount of attention, joining a long line of well known problems in Combinatorial Set Theory, with applications in Coding Theory and in Theoretical Computer Science. The best known upper bound on the above was $\Theta(2^n)$, given by Frankl and Rödl (1987), and Ahlswede, Cai and Zhang (1989) provided a construction for an ℓ-cross-intersecting pair of size $\Theta(2^n/\sqrt{\ell})$, and conjectured that it is optimal. However, their conjecture was verified only for the values 0,1,2 of ℓ (the case $\ell = 2$, proved by Keevash and Sudakov (2006), being the latest progress in the study of this problem). In Chapter 7, we settle the conjecture of Ahlswede et al. for every sufficiently large value of ℓ. Furthermore, we obtain the precise structure of all optimal pairs (giving a family of constructions richer than that of Ahlswede et al.).

In Part IV, we consider *tensor* graph powers and related graph isoperimetric inequalities. Chapter 8 discusses the limit of the independence numbers in tensor graph powers, and a related isoperimetric-constant of independent sets in the original graph. We show several connections between these two parameters, and relate them to other long standing open problems involving tensor graph products. One such interesting connection is the relation between these parameters in random graphs, along the *random graph process*. In Chapter 9 we explore the behavior of the isoperimetric-constant in random graphs, and characterize it in every step of the random graph process in terms of the minimal degree of the graph.

Contents

List of Figures

Introduction

The study of graph powers is the analysis of asymptotic properties of a sequence of graphs, generated by repeatedly applying some operator on a fixed graph input. One of the most fundamental and well-known problems in this area is determining the Shannon capacity of a graph, a notoriously challenging graph parameter, introduced by Shannon [97] in 1956. The incentive for this study is zero-error communication over a noisy channel. Shannon introduced a model for a noisy channel, characterized by a graph, and defined the capacity of this graph as an asymptotic quantity which measures the effective alphabet of the channel in zero-error transmission. After establishing some initial properties on the behavior of this parameter, Shannon was able to determine the capacity of all graphs on up to 5 vertices excluding the pentagon, the cycle on 5 vertices. The pentagon was the smallest example where there was a gap between the lower and upper bounds given by Shannon for the capacity, motivating Berge to study such graphs in the 1960's (cf., e.g., [26]).

Berge defined a class of graphs called *perfect graphs*, where in particular, the Shannon capacity is well understood, and made a conjecture on the structure of these graphs. A weaker form of this conjecture, due to Fulkerson [58], was solved by Lovász [79] in 1972, and the *strong perfect graph theorem* was proved by Chudnovsky, Robertson, Seymour and Thomas [39], giving a characterization of all perfect graphs. Despite much effort (cf., e.g., [3], [8], [28], [30], [63], [64], [81], [94], [76]), determining the capacity of non-perfect graphs proved to be a difficult task, and the seemingly simple problem of determining the capacity of the pentagon was solved only in 1979 by Lovász [81], via the celebrated Lovász ϑ-function. Till this day, little is known on the behavior of the Shannon capacity of non-perfect graphs, and the capacity of the cycle on 7 vertices remains unknown.

While many long-standing problems regarding the Shannon capacity are still open, the methods developed over the years in order to deduce bounds on this parameter are interesting on their own account, and proved useful in many other settings. Most notably, the Lovász ϑ-function has many applications in Theoretical and Applied Computer Science, as it can be computed efficiently (in polynomial time, up to arbitrary precision, using Semi-definite Programming), while it is sandwiched between graph parameters which are highly difficult to compute (and are even NP-hard to approximate up to a small power of the number of vertices). See [72] for on excellent survey on this subject. Other useful methods for bounding the Shannon capacity include the algebraic bounds given by Haemers [63],[64] and by Alon [8], which have many applications in Coding Theory and Extremal Combinatorics.

In this thesis we study various capacities of graph powers under different definitions of graph operators, and derive results on problems in Information Theory, Coding Theory, and Ramsey Theory. We concentrate on classical and well-studied graph-power operators, in addition to a newly introduced natural generalization of one of these powers (see [4] for a concise survey on the background and definitions of these different types of graph powers). Our work sheds additional light on the behavior of large powers of a fixed graph, and while there are still many questions in this field awaiting answers, the methods we developed were already useful in settling several open problems in related areas.

The thesis comprises four parts, and in what follows we describe the contents of each of these parts.

Part I: The Shannon capacity of a graph and related problems in Information Theory

In the first part of this thesis, we study problems related to the *strong graph product* and the Shannon capacity of a graph. We begin by stating their formal definitions, as given by C.E. Shannon [97] in his seminal paper from 1956.

Define a channel $\mathcal{C}$ with an input alphabet V and an output alphabet U, as a mapping $V \to P(U)$: each input symbol is associated with a subset of output symbols, where sending the symbol x through the channel may result in each of the symbols $\mathcal{C}(x)$ on the receiver's end. The "noise" of $\mathcal{C}$ is reflected by the fact that sending certain input symbols may result in the same output. It is convenient to model this relationship via *characteristic graph* of the channel, whose vertex set is V, and two vertices are adjacent iff the corresponding input symbols are confusable over $\mathcal{C}$ (that is, xy is an edge iff the output-letter subsets $\mathcal{C}(x)$ and $\mathcal{C}(y)$ intersect).

Shannon was interested in zero-error transmission over $\mathcal{C}$, that is - the sender and receiver agree upon a prescribed set of input letters, enabling the receiver to always recover the sent symbol, without danger of confusion. By the above definitions, such a set of input symbols corresponds to an *independent set* of the characteristic graph G - a set of vertices of G with no edges between them. We denote the cardinality of a maximum independent set by $\alpha(G)$, the *independence number* of G.

As is often the case in Information Theory, one can benefit from transmitting longer words in the above scheme. To this end, Shannon defined G^k, the k-th strong power of G, as the graph whose vertex set is the cartesian k-fold power of G, V^k, where two distinct k-tuples are adjacent iff they are either equal or adjacent in G in each coordinate. According to this definition, each vertex of the k-th power of G corresponds to a k-letter word, and two vertices are adjacent iff the corresponding words are confusable over $\mathcal{C}$ (one coordinate which is distinct and disconnected in G suffices to distinguish between the two words). Hence, a maximum set of k-letter words, which can be transmitted over $\mathcal{C}$ without danger of confusion, corresponds to a maximum independent set of G^k, and has cardinality $\alpha(G^k)$. The *Shannon capacity* of G, $c(G)$, is defined as the limit of the independence numbers of G^k, normalized appropriately: $c(G) = \lim_{k\to\infty} \alpha(G^k)^{1/k}$. The Shannon capacity essentially measures the effective alphabet of the channel in zero-error transmission. For instance, if $c(G) = 7$, then for a sufficiently large word length k, one can send roughly 7^k distinct k-letter words without danger of confusion, and that is optimal; this is analogous to a setting where the input alphabet comprises 7 letters, and there is no confusion whatsoever.

The first two chapters in this part exhibit some of the surprising and nonintuitive properties of the Shannon capacity. In Chapter 1 we discuss the problem of approximating the capacity given a fixed number of graph powers, and demonstrate that graphs can exhibit an arbitrarily complicated behavior in terms of their independence numbers. Chapter 2 focuses on the scenario where several senders combine their individual channels together, and shows that one can manipulate the combined capacities of the different combinations of the senders to be "large" or "small" at will.

In the last two chapters in this part, we use the methods developed in the study of strong graph powers in order to deduce results on *index-coding*, a source-coding problem with motivation in several areas of Information Theory. In Chapter 3 we disprove the main conjecture of Bar-Yossef, Birk, Jayram and Kol [23], which stated that *linear* index coding is always optimal. Using an explicit construction of a Ramsey graph (a graph without large homogenous subgraphs), and algebraic bounds on its Shannon capacity, we show that the gap between the overall optimum and the linear optimum can be essentially the largest possible. In Chapter 4 we relate this problem to *colorings* of strong graph powers and to Witsenhausen's rate [106], yielding that multiple-round index-coding is strictly better, and encouraging the study of the average "rate" of an index-code in sufficiently long transmissions.

The independence numbers of strong graph powers (§1)

The first chapter in this part focuses on the problem of approximating the Shannon capacity of a graph.

Shannon [97] demonstrated graphs where the capacity is attained in the first power, e.g., all perfect graphs (for further results along this line, see the work of Rosenfeld [93] and Ore [90] on "universal graphs"). This corresponds to channels where an optimal zero-error transmission is achieved by repeatedly sending 1-letter messages through the channel. The remarkable Lovász ϑ-function, introduced in [81], provided families of graphs, where the capacity is attained in the second power, e.g., transitive self-complementary graphs, such as the pentagon. In this case, the optimal zero-error transmission is attained by block-coding 2-letter messages repeatedly over the channel. Curiously, there is no known graph whose capacity is attained in *any finite power other than the first or second*, and one may conjecture that the first few powers of G suffice in order to approximate its capacity.

In this work, we provide a probabilistic construction of graphs, whose capacity cannot be approximated (up to a small power of the number of vertices) by any arbitrarily large, yet fixed, sequence of graph powers. The graphs constructed exhibit an arbitrarily complicated behavior in terms of their independence numbers: one can design graphs such that the series of independence numbers of their strong powers repeatedly increases and then stabilizes at arbitrarily chosen positions. The key element in the construction is a random perturbation of an initial graph, whose structure ensures an increase in the series of independence numbers at a desired location. The general result is derived after carefully combining several graphs constructed as above.

We conclude that the Shannon capacity of a graph cannot be approximated by a constant prefix of the series of independence numbers, even if this series demonstrates a sudden increase and thereafter stabilizes. This settles a question raised by Bohman [29].

References: The results of this chapter appear in:

- N. Alon and E. Lubetzky, The Shannon capacity of a graph and the independence numbers of its powers,
  IEEE Transactions on Information Theory 52 (2006), 2172-2176.

Privileged users in zero-error transmission (§2)

The previous chapter demonstrated how the performance of zero-error protocols utilizing word-lengths over a given channel can be quite unpredictable. In the second chapter in this part, we study sums of channels, and show that there are scenarios where seemingly unrelated and independent channels may affect one another. We focus on scenarios where there are multiple senders, and various combinations of these senders wish to cooperate and combine their individual channels together.

A sum of channels, $\mathcal{C} = \sum_i \mathcal{C}_i$, describes a setting where there are $t \geq 2$ senders, each with his own channel $\mathcal{C}_i$, and words can comprise letters from any of the channels. There is no danger of confusion between symbols transmitted over distinct channels, hence this setting corresponds to a disjoint union of the characteristic graphs, $G = \sum_i G_i$. Shannon [97] showed that capacity of the sum of channels is always at least the sum of the capacities (i.e., $c(G) \geq \sum_i c(G_i)$), presented families of graphs where these two quantities are equal, and conjectured that in fact equality holds for all graphs. Intuitively, as the capacity is the effective alphabet size in zero-error transmission, one would indeed expect the combination of the channels to have a zero-error alphabet size which is the sum of the individual alphabets. Surprisingly, this conjecture of Shannon was disproved in 1998 by Alon [8], where it was shown that the capacity of a disjoint union can in fact be larger than any fixed power of the individual capacities.

In this work, we extend the ideas of [8] and prove a stronger result, showing that one may further manipulate the relations between seemingly unrelated channels. Suppose that $\mathcal{F}$ is a family of subsets of $\{1, \ldots, t\}$, thinking of $\mathcal{F}$ as a collection of "privileged" subsets of a group of t senders. Given any such $\mathcal{F}$, we assign a channel $\mathcal{C}_i$ to each sender, such that the combined capacity of a group of senders $X \subset [t]$ is "large" if this group contains some privileged subset (X contains some $F \in \mathcal{F}$) and is "small" otherwise. That is, only privileged subsets of senders are allowed to transmit in a high rate.

For instance, as an analogue to secret sharing, it is possible to ensure that whenever at least k senders combine their channels, they obtain a high capacity, however every group of $k - 1$ senders has a low capacity (and yet is not totally denied of service). The case $k = t = 2$ corresponds to the original conjecture of Shannon.

In the process, we obtain an explicit Ramsey construction of an edge-coloring of the complete graph by t colors, where every "large" induced subgraph contains *all* t colors.

References: The results of this chapter appear in:

- N. Alon and E. Lubetzky, Privileged users in zero-error transmission over a noisy channel,
  Combinatorica, to appear.

Non-linear index coding outperforming the linear optimum (§3)

In the previous chapter, we constructed Ramsey graphs in order to design individual channels to a group of senders, whose combinations satisfy certain properties. In this chapter, we show that a variant of these Ramsey constructions, combined with some additional ideas, has some surprising consequences on *index-coding*. This source-coding problem was introduced by Birk and Kol [27] in 1998, and has applications in Distributed Communication, as well as in other area in Information Theory. The setting of the problem is as follows:

A sender holds a word $x \in \{0,1\}^n$, and wishes to broadcast a codeword to n receivers, $R_1,\ldots,R_n$. The receiver R_i is interested in x_i, and has prior *side information* comprising some subset of the n bits. The server wishes to broadcast a code of minimal word-length, which would always (i.e., for any input word x) allow every receiver to recover the bit he is interested in.

The problem can be reformulated as a graph parameter as follows: the side-information relations are conveniently modeled by a directed graph G on n vertices, where ij is an edge iff R_i knows the bit x_j. An *index code* for G is an encoding scheme which enables each R_i to always reconstruct x_i, given his side information. The minimal word length of an index code was studied by Bar-Yossef, Birk, Jayram and Kol [23]. They introduced a graph parameter, $\mathrm{minrk}_2(G)$, which completely characterizes the length of an optimal *linear* index code for G. The authors of [23] showed that in various cases linear codes attain the optimal word length, and conjectured that linear index coding is in fact *always* optimal.

In this chapter, we disprove the main conjecture of [23] in the following strong sense: for any $\varepsilon > 0$ and sufficiently large n, there is an n-vertex graph G so that every linear index code for G requires codewords of length at least $n^{1-\varepsilon}$ (barely improving the n bits required by the naïve protocol), and yet a given non-linear index code for G has a word length of n^{ε}. This is achieved by an explicit construction, which extends Alon's variant of the celebrated Ramsey construction of Frankl and Wilson.

References: The results of this chapter appear in:

- E. Lubetzky and U. Stav, Non-linear index coding outperforming the

linear optimum,
Proc. of the 48th IEEE FOCS (2007), 161-167.

Index coding and Witsenhausen-type coloring problems (§4)

In this chapter, the final chapter in the first part of this thesis, we relate index-codes for a disjoint union of graphs to the OR product (the complement of a strong graph product) of certain graphs. Using this connection, and based on some classical results in the study of Witsenhausen's rate and colorings of OR graph powers, we obtain that an index-code for $k \cdot G$, a disjoint union of k copies of the graph G, can be strictly shorter than k times the length of an optimal index-code for G. While this result may appear similar to the results of Chapter 2 on the capacity of a sum of channels, this surprising statement is quite different in nature: assuming that every copy of G requires at-least, say, 10 different codewords in an index-code, and that there are k disjoint copies of G each with an independently chosen input word, it is difficult to imagine how fewer than 10^k distinct codewords can suffice for this task...

The above result has two immediate consequences. First, we obtain an alternative proof that linear index coding is suboptimal (this time, by a multiplicative constant). Second, we show that multiple transmissions with the same side-information configuration can be strictly better than the result of repeatedly using the optimal protocol for a single-round index-code. This motivates the study of the "rate" of an index-code of a graph, as the average length of an index-code when performing multiple transmissions with a given side-information graph.

References: The results of this chapter appear in:

- N. Alon, E. Lubetzky and U. Stav, The broadcast rate of a graph.

Part II: Codes and explicit Ramsey graphs

While the previous part of the thesis focused on the classical and well-studied strong graph products, and the related Shannon capacity of a graph, the second part is devoted to a different graph product, which has fascinating random-looking properties, and is connected to problems in Coding Theory and in Ramsey Theory.

Recall that in the k-th *strong* power of a graph, two distinct k-tuples are adjacent iff each of their coordinates is either equal or adjacent in the original graph. This implies that high powers of some fixed graph are quite sparse, as the edge density decreases exponentially in the graph power. While the sizes of maximum independent sets in such graph powers are difficult to analyze, and little is known on their asymptotic behavior, the Shannon capacity, for general graphs, the behavior of other graph parameters becomes trivial due to the graph sparseness. For instance, it is easy and well-known to deduce the structure of all maximum cliques (sets of pairwise adjacent vertices) in strong powers of a graph.

A slightly different definition of the graph power results in much denser graphs, with certain random-looking properties. The adjacency criteria in the k-th power is modified as follows: instead of requiring two distinct k-tuples to be equal or adjacent in every coordinate, they are required to be adjacent in an *odd* number of coordinates. Indeed, a quick look at this definition of this graph power, known as the *Xor* graph power, already reveals some of its unique properties: the edge-density of the k-th Xor power of *any* nonempty regular graph tends to $\frac{1}{2}$ as $k \to \infty$. However, one can show that despite their large density, these graphs do not contain large cliques, and in this sense they are random-looking.

The interesting properties of the Xor graph power were used by Thomason [102] in 1997 to disprove a conjecture of Erdős from 1962, by constructing edge-colorings of the complete-graph with two colors, containing a smaller number of monochromatic copies of K_4 (a complete graph on 4 vertices) than the expected number of such copies in a random coloring. For further information on this problem, see [47],[52],[103].

While the Xor powers of any nontrivial graph are both dense and do not contain large cliques, they do contain large independent sets. Motivated by the search for explicit constructions of Ramsey graphs, we attempt to correct this behavior by inspecting a natural generalization of the Xor powers to other moduli. That is, two k-tuples are adjacent iff the number of their adjacent coordinates is non-zero modulo some integer p. One can show that indeed, choosing large values of p reduces the size of maximum independent sets in high powers of an arbitrary graph, at the cost of larger cliques, and provides a method of producing explicit Ramsey construction.

The two chapters in this part are Chapter 5, which is devoted to the Xor powers and their applications in Coding Theory, and Chapter 6, which studies the generalized p-powers and their applications in Ramsey Theory.

Codes and Xor graph products (§5)

The motivation behind this chapter lies in problems in Coding Theory, as demonstrated by the following two questions:

- What is the maximum possible number, $f_3(n)$, of vectors of length n over $\{0,1,2\}$ such that the Hamming distance between every two is *even*?
- What is the maximum possible number, $g_3(n)$, of vectors in $\{0,1,2\}^n$ such that the Hamming distance between every two is *odd*?

We investigate these questions, and more general ones, by studying Xor powers of graphs, focusing on their independence number and clique number, and by introducing two new parameters of a graph G. Both parameters denote limits of series of either clique numbers or independence numbers of the Xor powers of G (normalized appropriately), and while both limits exist, one of the series grows exponentially as the power tends to infinity, while the other grows linearly. As a special case, it follows that $f_3(n) = \Theta(2^n)$ whereas $g_3(n) = \Theta(n)$.

Unlike the Shannon capacity of a graph, the above mentioned parameters of Xor powers are *non-monotone* with respect to the addition of edges, making the task of determining their values challenging even for complete graphs. In order to obtain general bounds on these parameters, we resort to various tools from Algebra and Spectral Analysis, some of which were developed in the course of the study of the Shannon capacity.

References: The results of this chapter appear in:

- N. Alon and E. Lubetzky, Codes and Xor graph products, **Combinatorica** 27 (2007), 13-33.

Graph p-powers, Delsarte, Hoffman, Ramsey and Shannon (§6)

In this chapter, we attempt to construct explicit Ramsey graphs using graph p-powers, a natural generalization of the Xor power, where the (mod 2) is replaced by (mod p) for some integer p. The bounds we derive in this chapter on the p-powers of a graph improve some of the upper bounds on the Xor capacities, given in the previous chapter, and in particular, settle a conjecture regarding the Xor-powers of cliques up to a factor of 2. For precise bounds on some graphs, we apply Delsarte's remarkable linear programming bound, as well as Hoffman's eigenvalue bound.

While the Xor powers of an arbitrary graph have a logarithmic clique number and a large independence number, we prove that selecting a larger modulo p in the definition of the product operator indeed corrects this behavior: the independence number is reduced at the cost of a poly-logarithmic clique number. We deduce that for any nontrivial graph G, one can point out specific induced subgraphs of large p-powers of G which are "Ramsey", i.e., contain neither a large clique nor a large independent set. This is once again related to the Shannon capacity: we show that the larger the capacity of $\overline{G}$ (the complement of G) is, the larger these subgraphs are. In the special case where G is the complete graph, some p-power of G matches the bounds of the famous Ramsey construction of Frankl and Wilson, and is in fact a subgraph of a variant of that construction. This Ramsey construction of Frankl and Wilson is based on set systems with prescribed pairwise intersections; in the next part, we proceed to investigate this area.

References: The results of this chapter appear in:

- N. Alon and E. Lubetzky, Graph powers, Delsarte, Hoffman, Ramsey and Shannon, **SIAM J. Discrete Math** 21 (2007), 329-348.

Part III: An extremal problem in Finite Set Theory

Our study in the previous part began with problems in Coding Theory, which suggested the study of large Xor powers of a graph, and continued in the natural generalization of these powers to $\pmod{p}$ powers for some integer p. Seeking explicit Ramsey constructions, we proved that large p-powers of certain graphs do not contain large homogenous subgraphs. Optimizing the parameters for this purpose - the initial graph, the power we raise it to and the modulo p in the definition of the graph power - led to a family of graphs which is closely related to the famous Ramsey construction by Frankl and Wilson.

The Frankl and Wilson [56] graph is defined as follows: the vertex set of the graph corresponds to all s-element subsets of a r-element ground set, where $s = p^2 - 1$ and $r = p^3$ for some large prime p; two distinct vertices are adjacent iff their corresponding sets have an intersection congruent to $-1 \pmod{p}$. The Ramsey properties of this graph follow from the following key fact: there cannot be "too many" sets with a cardinality of $k \pmod{p}$, whose pairwise intersections all have cardinalities which are non-congruent to $k \pmod{p}$. Namely, there can be at most $\binom{r}{p-1}$ such sets.

This above statement is a stronger version of a result by Ray-Chaudhuri and Wilson [91], who considered the actual cardinalities of the intersections, without the prime moduli. They show that if $\mathcal{F}$ is a set-system over the ground set $\{1, \ldots, r\}$, and the pairwise intersections of elements of $\mathcal{F}$ have s possible cardinalities, then $\mathcal{F}$ contains at most $\binom{r}{s}$ sets. A well-studied and much more complicated scenario is the one where there are *two* families of sets with prescribed *cross* intersections. Understanding the structure of the extremal families of subsets in this case is difficult even when there is precisely one permitted cardinality for the cross intersection. This problem was introduced by Frankl and Rödl [52] in 1987, and the main result in this part gives the precise structure of these extremal families, thus proving a conjecture of Ahlswede, Cai and Zhang [2] from 1989.

Uniformly cross intersecting families (§7)

Let $\mathcal{A}$ and $\mathcal{B}$ denote two families of subsets of an n-element set. The pair $(\mathcal{A},\mathcal{B})$ is said to be ℓ-cross-intersecting iff $|A \cap B| = \ell$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. When examining the extremal families in this setting, it is clearly possible to extend the size of one family at the expense of the other, hence the interesting quantity to study is the product of the sizes of the two families, $|\mathcal{A}||\mathcal{B}|$, which is also the number of constraints we have. Therefore, let $P_\ell(n)$ denote the maximum value of this quantity over all such pairs:

$$P_\ell(n) = \max \left\{|\mathcal{A}||\mathcal{B}| : \mathcal{A} \text{ and } \mathcal{B} \text{ are } \ell\text{-cross-intersecting over } \{1,\ldots,n\}\right\} .$$

Frankl and Rödl [52] introduced this problem in 1989, and provided the best known upper bound on $P_\ell(n)$, which is $\Theta(2^n)$. For a lower bound, Ahlswede, Cai and Zhang [2] gave in 1989 the following simple construction: take $n \geq 2\ell$, let the family $\mathcal{A}$ consist of the single set $A_1 = \{1,\ldots,2\ell\}$, and let $\mathcal{B}$ comprise all sets whose intersection with A_1 contains precisely ℓ elements. This ℓ-cross-intersecting pair satisfies $|\mathcal{A}||\mathcal{B}| = \binom{2\ell}{\ell}2^{n-2\ell} = \Theta(2^n/\sqrt{\ell})$, and Ahlswede et al. conjectured that this is best possible. That is, they conjectured that $P_\ell(n) = \binom{2\ell}{\ell}2^{n-2\ell}$, and observed that the upper bounds of Frankl and Rödl match this hypothesis for the special cases $\ell = 0$ and $\ell = 1$.

With the upper bound on $P_\ell(n)$ being independent of ℓ as opposed to the lower bound, Sgall [95] asked in 1999 whether or not $P_\ell(n)$ decreases as ℓ grows. The latest progress in the study of this problem was made by Keevash and Sudakov [71] in 2006, where the authors verified the conjecture of Ahlswede, Cai and Zhang for the special case $\ell = 2$.

In this chapter, we confirm the above conjecture of Ahlswede et al. for any sufficiently large ℓ, implying a positive answer to the above question of Sgall as well. By analyzing the linear spaces of the characteristic vectors of $\mathcal{A},\mathcal{B}$ over $\mathbb{R}$, we show that there exists some $\ell_0 > 0$, such that $P_\ell(n) \leq \binom{2\ell}{\ell}2^{n-2\ell}$ for all $\ell \geq \ell_0$. Furthermore, we determine the precise structure of all the pairs of families which attain this maximum (obtaining a family of constructions far more complicated than that of Ahlswede et al.).

References: The results of this chapter appear in:

- N. Alon and E. Lubetzky, Uniformly cross intersecting families.

Part IV: Tensor graph powers and graph isoperimetric inequalities

In the final part of this thesis, we return to the analysis of independence numbers in graph powers, and this time examine the *tensor* graph power, a graph operator which seems to be much better understood than the strong graph power, and yet raises some extremely difficult and challenging problems.

The *tensor* product of two graphs, G and H, is the graph whose vertex set is, as usual, the cartesian product of the vertex sets, and two vertices (u, v) and (u', v') are adjacent iff both uu' are adjacent in G and vv' are adjacent in H. Thus, the k-th tensor power of G has an edge between two k-tuples if there is an edge in each of their coordinates (as opposed to either an edge or *equality* in each of the coordinates in the definition of the strong power).

This graph product has attracted a considerable amount of attention due to a long standing and seemingly naïve conjecture on vertex-colorings of the tensor product of two graphs. The chromatic number of a graph G, denoted by $\chi(G)$, is the minimal number of colors required to *legally* color its vertices (a legal coloring is one where no two adjacent vertices are assigned the same color). The well known conjecture of Hedetniemi [66] from 1966 states that the chromatic number of the tensor product of G and H is equal to $\min\{\chi(G), \chi(H)\}$. See [109] for an extensive survey of this open problem. For further work on colorings of tensor products of graphs, see [9], [61], [74], [100], [101], [108], [110].

While Hedetniemi's conjecture on the chromatic number of a tensor product of two general graphs remains unsolved, the behavior of this parameter in tensor powers of the *same* graph G is rather simple. In that case, it is not difficult to show that the chromatic number of the k-th power of a graph G is always equal to the chromatic number of G. Similarly, one can verify that the clique number of any tensor power of G is equal to the clique number of G. However, the independence numbers of tensor graph powers exhibit a far more interesting behavior. In addition, the methods used to study the independence numbers of tensor powers are interesting on their own account. For instance, in [9] the authors used Fourier Analysis to this end, and in [45] they applied the machinery of noise-stability of functions [88] for this purpose.

In the first chapter in this part, we study the limit of independence numbers (normalized appropriately) in tensor graph powers, that is, the tensor-power analog of the Shannon capacity. This parameter is related to a certain vertex isoperimetric-constant of the initial graph, and we investigate this connection for various families of graphs. In the second chapter in this part we expand our study of graph isoperimetric-constants to the random-graph setting: we obtain a complete characterization of the edge isoperimetric-constant in terms of the minimal degree, along every step of the *random graph process*.

Independent sets in tensor graph powers (§8)

The *independence ratio* of a graph is the ratio of its independence number and number of vertices. The parameter $A(G)$, which was introduced in [37], is the limit of the independence ratios of tensor powers of G. This puzzling parameter takes values only in the range $(0, \frac{1}{2}] \cup \{1\}$, and is lower bounded by a vertex isoperimetric ratio of independent sets of G.

In this chapter, we study the relation between these two parameters further, and ask whether they are essentially equal. We present several families of graphs where equality holds, and discuss the effect the above question has on various open problems related to fractional colorings of tensor graph products.

It is interesting to compare the behavior of the parameter $A(G)$ and the above isoperimetric constant in random graphs and along the random graph process. This is discussed in this chapter, and is related to the topic of the next chapter, where we analyze the (edge) isoperimetric constant of the random graph process.

References: The results of this chapter appear in:

- N. Alon and E. Lubetzky, Independent sets in tensor graph powers, **J. Graph Theory** 54 (2007), 73-87.

The isoperimetric constant of the random graph process (§9)

The *isoperimetric constant* of a graph G on n vertices (sometimes referred to as the *conductance* of G), $i(G)$, is the minimum of $\frac{|\partial S|}{|S|}$, taken over all nonempty subsets $S \subset V(G)$ of size at most $n/2$, where ∂S denotes the set of edges with precisely one end in S.

A *random graph process* on n vertices, $\widetilde{G}(t)$, is a sequence of $\binom{n}{2}$ graphs, where $\widetilde{G}(0)$ is the edgeless graph on n vertices, and $\widetilde{G}(t)$ is the result of adding an edge to $\widetilde{G}(t-1)$, uniformly distributed over all the missing edges.

There has been much study of the isoperimetric constants of various graphs, such as grid graphs, torus graphs, the n-cube, and more generally, cartesian products of graphs. See, for instance, [35],[36],[42],[68], [87]. In 1988, Bollobás [33] studied the isoperimetric constant of *random d-regular graphs*, and showed that for infinitely many d-regular graphs, its value is at least $\frac{d}{2} - O(\sqrt{d})$. Alon [5] proved in 1997 that this inequality is in fact tight, by showing that any d-regular graph G on a sufficiently large number of vertices satisfies $i(G) \leq \frac{d}{2} - c\sqrt{d}$ (where $c > 0$ is some absolute constant).

In this chapter, we study the isoperimetric constant of general random graphs $\mathcal{G}(n,p)$, $\mathcal{G}(n,M)$, and the random graph process, and show that in these graphs, the ratio between the isoperimetric constant and the minimal degree exhibits an interesting behavior. We prove that, in almost every graph process, $i(\widetilde{G}(t))$ equals the minimal degree of $\widetilde{G}(t)$ in every step, as long as the minimal degree is $o(\log n)$. Furthermore, we show that this result is essentially best possible, by demonstrating that along the period in which the minimum degree is typically $\Theta(\log n)$, the ratio between the isoperimetric constant and the minimum degree falls from 1 to $\frac{1}{2}$, its final value.

References: The results of this chapter appear in:

- I. Benjamini, S. Haber, M. Krivelevich and E. Lubetzky, The isoperimetric constant of the random graph process,
  Random Structures and Algorithms 32 (2008), 101-114.

Part I

The Shannon capacity of a graph and related problems in Information Theory

Chapter 1

The independence numbers of strong graph powers

The results of this chapter appear in [14]

The independence numbers of powers of graphs have been long studied, under several definitions of graph products, and in particular, under the strong graph product. We show that the series of independence numbers in strong powers of a fixed graph can exhibit a complex structure, implying that the Shannon Capacity of a graph cannot be approximated (up to a sub-polynomial factor of the number of vertices) by any arbitrarily large, yet fixed, prefix of the series. This is true even if this prefix shows a significant increase of the independence number at a given power, after which it stabilizes for a while.

1.1 Introduction

Given two graphs, G_1 and G_2, their *strong graph product* $G_1 \cdot G_2$ has a vertex set $V(G_1) \times V(G_2)$, and two distinct vertices (v_1, v_2) and (u_1, u_2) are connected iff they are adjacent or equal in each coordinate (i.e., for $i \in \{1, 2\}$, either $v_i = u_i$ or $v_i u_i \in E(G_i)$). This product is associative and commutative, and we can thus define G^k as the product of k copies of G. In [97], Shannon introduced the parameter $c(G)$, the *Shannon Capacity* of a graph G, which is the limit $\lim_{k\to\infty} \sqrt[k]{\alpha(G^k)}$, where $\alpha(G^k)$ is the independence number of G^k (it is easy to see that this limit exists by super-multiplicativity). The considerable amount of interest that $c(G)$ has received (see, e.g., [3], [8], [28], [30], [63], [64], [81], [94], [76]) is motivated by Information Theory concerns: this parameter represents the effective size of an alphabet, in a communication model where the graph G represents the channel. In other words, we consider a transmission scheme where the input is a set of single letters $V = \{1, \ldots, n\}$, and our graph G has V as its set of vertices, and an edge between each pair of letters, iff they are confusable in transmission (i.e., $(1, 2) \in E(G)$ indicates that sending an input of 1 or an input of 2 might result in the same output). Clearly $\alpha(G)$ is the maximum size of a set of single letters which can be predefined, then sent with zero-error. By definition, $\alpha(G^k)$ represents such a set of words of length k (since two distinct words are distinguishable iff at least one of their coordinates is distinguishable), leading to the intuitive interpretation of $c(G)$ as the effective size of the alphabet of the channel (extending the word length to infinity, while normalizing it in each step).

Consider the series $a_k = a_k(G) = \sqrt[k]{\alpha(G^k)}$, which we call "the *independence series* of G". As observed in [97], the limit $c(G) = \lim_{k\to\infty} a_k$ exists and equals its supremum, and $a_{mk} \geq a_k$ for all integers m, k. Our motivation for the study of the series a_k is the computational problem of approximating $c(G)$. So far, all graphs whose Shannon capacity is known, attain the capacity either at a_1 (the independence number, e.g., perfect graphs), a_2 (e.g., self complementary vertex-transitive graphs) or do not attain it at any a_k (e.g., the cycle C_5 with the addition of an isolated vertex). One might suspect that once the a_k series remains roughly a constant for several consecutive values of k, its value becomes a good approximation to its limit, $c(G)$. This, however, is false. Moreover, it remains false even when restricting ourselves to cases where a_k increases significantly before it stabilizes for a few steps. We thus address the following questions:

1. Is it true that for every arbitrarily large integer k, there is a $\delta = \delta(k) > 0$ and a graph G on n vertices such that the values $\{a_i\}_{i<k}$ are all at least n^δ- far from $c(G)$?

2. Can the series a_k increase significantly (in terms of $n = |V(G)|$) in an arbitrary number of places?

In this chapter we show that the answer to both questions above is positive. The first question is settled by Theorem 1.1.1, proved in section 1.2.

Theorem 1.1.1. *For every fixed $\nu \in \mathbb{N}$ and $\varepsilon > 0$ there exists a graph G on N vertices such that for all $k < \nu$, $a_k \leq c_k \log_2(N)$ (where $c_k = c_{k,\nu}$), and yet $a_\nu \geq N^{\frac{1}{\nu}}$.*

Indeed, for any fixed k, there exists a graph G on N vertices, whose Shannon Capacity satisfies $c(G) > N^\delta \max_{i<k}\{a_i\}$, where $\delta = \frac{1-o(1)}{k}$.

Theorem 1.1.2, proved in section 1.3, settles the second question, and implies the existence of a graph G whose independence series a_k contains an arbitrary number of "jumps" at arbitrarily chosen locations; hence, noticing a significant increase in this series, or noticing that it stabilizes for a while, does not ensure any proximity to $c(G)$.

Theorem 1.1.2. *For every fixed $\nu_1 < \ldots < \nu_s \in \mathbb{N}$ and $\varepsilon > 0$ there exists a graph G such that for all $k < \nu_i$, $a_k < a_{\nu_i}^{\varepsilon}$ ($i \in \{1, \ldots, s\}$).*

For a visualization of Theorem 1.1.2, see Figure 1.1 (notice that this figure is an illustration of the behavior of the independence series, rather than a numerical computation of a specific instance of our construction).

The above theorems imply that the naïve approach of computing the a_k values for some k does not provide even a PSPACE algorithm for approximating $c(G)$. Additional remarks on the complexity of the problem of estimating $c(G)$, as well as several open problems, appear in the final section 1.4.

1.2 The capacity and the initial a_k-s

In this section we prove Theorem 1.1.1, using a probabilistic approach, which is based on the method of [18], but requires some additional ideas.

Let $2 \leq \nu \in \mathbb{N}$; define $N = n\nu$ (n will be a sufficiently large integer) , and let $V(G) = \{0, \ldots, N-1\}$. Let $\mathcal{R}$ denote the equivalence relation on the set of unordered pairs of *distinct* vertices, in which (x, y) is identical to (y, x)

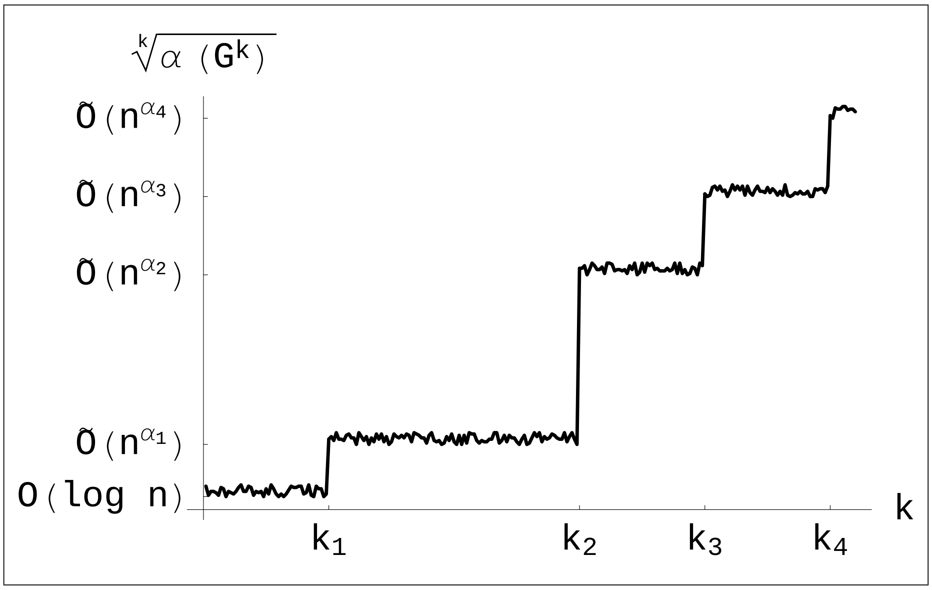

Figure 1.1: An illustration of the independence numbers of powers of the graph constructed in Theorem 1.1.2.

and is equivalent to $(x+kn, y+kn)$ for all $0 \le k \le \nu - 1$, where addition is reduced modulo N. Let $\{\mathcal{R}_1, \ldots, \mathcal{R}_M\}$ denote the different equivalence classes of $\mathcal{R}$. For every $x \ne y$, let $\mathcal{R}(x,y)$ denote the equivalence class of (x,y) under $\mathcal{R}$; then either $|\mathcal{R}(x,y)|$ is precisely ν, or the following equality holds for some $l < \nu$:

$$(x,y) \equiv (y+ln, x+ln) \pmod{N}$$

This implies that $N \mid 2ln$, hence $2l = \nu$. We deduce that if ν is odd, $|\mathcal{R}_i| = \nu$ for all $1 \le i \le M$, and $M = \frac{1}{\nu}\binom{N}{2}$. If ν is even:

$$|\mathcal{R}(x,y)| = \begin{cases} \frac{1}{2}\nu & \text{If } y \equiv x + \frac{1}{2}\nu n, \\ \nu & \text{Otherwise.} \end{cases}$$

and $M = \frac{1}{\nu}\binom{N}{2} + \frac{1}{2\nu}N$, i.e., in case of an even ν there are $N/2$ pairs which belong to n smaller classes, each of which is of size $\frac{1}{2}\nu$, while the remaining edges belong to ordinary edge classes of size ν.

The edges of G are chosen randomly, by starting with the complete graph and excluding a single edge from each equivalence class, uniformly and independently, thus $|E(G)| = \binom{N}{2} - M = \binom{N}{2}\left(\frac{\nu-1}{\nu} + o(1)\right)$.

A standard first moment consideration (c.f., e.g., [19]) shows that $a_1 = \alpha(G) < \lceil 2\log_\nu(N)\rceil$ almost surely. To see this, set $s = \lceil 2\log_\nu(N)\rceil$, and take an arbitrary set $S \subset V(G)$ of size s. If S contains more than one member of some edge class $\mathcal{R}_i$, it cannot be independent. Otherwise, its edge probabilities are independent, and all that is left is examining the lengths of the corresponding edge classes. Assume S contains r pairs which belong to short edge classes: $(x_1, y_1), \ldots, (x_r, y_r)$. If ν is odd, $r = 0$, otherwise $y_i = x_i + \frac{1}{2}\nu n$ for all i, and $x_i \neq x_j \pmod{\frac{1}{2}\nu n}$ for all $i \neq j$ (distinct pairs in S belong to distinct edge classes). It follows that $r \leq \frac{s}{2}$, and we deduce that for each such set S:

$$\Pr[S \text{ is independent}] \leq \left(\frac{1}{\nu}\right)^{\binom{s}{2}-r}\left(\frac{2}{\nu}\right)^{r} \leq \left(\frac{1}{\nu}\right)^{\binom{s}{2}} 2^{s/2}$$

Applying a union bound and using the fact that $\frac{(2\nu)^{s/2}}{s!}$ tends to 0 as N, and hence s, tend to infinity, we obtain:

$$\begin{aligned}\Pr[\alpha(G) \geq s] &\leq \binom{N}{s}\nu^{-\binom{s}{2}}2^{s/2} \leq \frac{2^{s/2}}{s!}\left(N\nu^{-\frac{s-1}{2}}\right)^{s} \\ &= \frac{(2\nu)^{s/2}}{s!}\left(N\nu^{-\frac{s}{2}}\right)^{s} \leq \frac{(2\nu)^{s/2}}{s!} = o(1),\end{aligned}$$

where the $o(1)$ term here, and in what follows, tends to 0 as N tends to infinity.

We next deal with G^k for $2 \leq k < \nu$. Fix a set $S \subset V(G^k)$ of size $s = \lceil c_k \log_2^k(N)\rceil$, where c_k will be determined later. Define S', a subset of S, in the following manner: start with $S' = \phi$, order the vertices of S arbitrarily, and then process them one by one according to that order. When processing a vertex $v = (v_1, \ldots, v_k) \in S$, we add it to S', and remove from S all of the following vertices which contain $v_i + tn \pmod N$ in any of their coordinates, for any $i \in [k]$ and $t \in \{0, \ldots, \nu - 1\}$. In other words, once we add v to S', we make sure that its coordinates modulo n will not appear anywhere else in S'. If S is independent, it has at most $\alpha(G^{k-1})$ vertices with a fixed coordinate, thus $s' = |S'| \geq s/\left(k^2 \cdot \nu \cdot \alpha(G^{k-1})\right)$. Notice that each two distinct vertices $u, v \in S'$ have distinct vertices of G in every coordinate, thus $\mathcal{R}(u_i, v_i)$ is defined for all i; furthermore, for any *other* pair of distinct vertices $u', v' \in S'$, the sets $\{\mathcal{R}(u_1, v_1), \ldots, \mathcal{R}(u_k, v_k)\}$ and $\{\mathcal{R}(u'_1, v'_1), \ldots, \mathcal{R}(u'_k, v'_k)\}$ are disjoint.

We next bound the probability of an edge between a pair of vertices $u \neq v \in S'$. Let k' denote the number of **distinct** pairs of corresponding coordinates of u, v, and let t_l, $1 \leq l \leq M$, be the number of all such distinct pairs whose edge class is $\mathcal{R}_l$ (obviously $\sum_{l=1}^{M} t_l = k'$). For example, when all the corresponding pairs are distinct, we get $k' = k$ and $t_l = |\{1 \leq i \leq k : \mathcal{R}(u_i, v_i) = \mathcal{R}_l\}|$. Notice that, by definition of S', for every i, $v_i \neq u_i + \frac{1}{2}\nu n$, and thus $\mathcal{R}(u_i, v_i)$ is an ordinary edge class. It follows that:

$$\Pr[uv \in E(G^k)] = \prod_{l=1}^{M} \frac{\nu - t_l}{\nu} \tag{1.1}$$

This expression is minimal when $t_l = k'$ for some l, since replacing $t_{l_1}, t_{l_2} > 0$ with $t'_{l_1} = t_{l_1} + t_{l_2}, t'_{l_2} = 0$ strictly decreases its value. Therefore $\Pr[uv \notin E(G^k)] \leq \frac{k'}{\nu} \leq \frac{k}{\nu}$. Notice that, crucially, by the structure of S', as each edge class appears in at most one pair of vertices of S', the events $uv \notin E(G^k)$ are independent for different pairs u, v. Let $A_{S'}$ denote the event that there is an independent set S' of the above form of size $s' = \lceil c' \log_2(N) \rceil$, where $c' = 2k^2$. Applying the same consideration used on S and G to S' and G^k, gives (assuming N is sufficiently large):

$$\begin{aligned}\Pr[A_{S'}] &\leq \binom{N^k}{s'} \left(\frac{k}{\nu}\right)^{\binom{s'}{2}} \leq N^{ks'} 2^{-\frac{1}{2}s'^2 \log_2(\frac{\nu}{k})} \leq \\ &\leq 2^{\left(kc' - \frac{1}{2}\log_2(\frac{\nu}{k}) c'^2\right) \log_2^2(N)} .\end{aligned}$$

Now, our choice of c' should satisfy $c' > \frac{2k}{\log_2(\frac{\nu}{k})}$ for this probability to tend to zero. Whenever $2 \leq k \leq \frac{\nu}{2}$ we get $k \log_2(\frac{\nu}{k}) \geq k > 1$, thus $c' = 2k^2 > \frac{2k}{\log_2(\frac{\nu}{k})}$. For $\frac{\nu}{2} < k < \nu$ we have $1 < \frac{\nu}{k} < 2$ and thus $\log_2(\frac{\nu}{k}) > \frac{\nu}{k} - 1$. Taking any $c' \geq \frac{2k^2}{\nu - k}$ would be sufficient in this case, hence $c' = 2k^2$ will do. Overall, we get that $\Pr[A_{S'}]$ tends to 0 as N tends to infinity.

Altogether, we have shown that for every $2 \leq k < \nu$:

$$\begin{aligned}\alpha(G^k) &\leq k^2\nu\alpha(G^{k-1})2k^2\log_2(N) = \\ &= 2k^4\nu\log_2(N)\alpha(G^{k-1}) \ .\end{aligned}$$

Hence, plugging in the facts that $\alpha(G) \leq 2\log_\nu(N) < 2\log_2(N)$ and $2^{\frac{m}{2}}m! \leq m^m$ for $m \geq 2$, we obtain the following bound for all $k \in \{1, \ldots, \nu - 1\}$:

$$\alpha(G^k) \leq 2^k(k!)^4\nu^{k-1}\log_2^k(N) \leq 2^{-k}k^{4k}\nu^{k-1}\log_2^k(N) \ ,$$

$$a_k \leq \frac{1}{2}k^4\nu\log_2(N) \leq \frac{1}{2}\nu^5\log_2(N) \ .$$

It remains to show that a_ν is large. Consider the following set of vertices in G^ν (with addition modulo N):

$$I = \{ \ \overline{x} = (x, x+n, \ldots, x+(\nu-1)n) \mid 0 \leq x < N \ \} \tag{1.2}$$

Clearly I is independent, since for any $0 \leq x < y < N$, the corresponding coordinates of $\overline{x}, \overline{y}$ form one complete edge class, thus exactly *one* of these coordinates is disconnected in G. This implies that $a_\nu \geq N^{\frac{1}{\nu}}$.

Hence, we have shown that for every value of ν, there exists a graph G on N vertices such that:

$$\left\{\begin{array}{ll} a_i \leq c_i\log_2(N) & (i = 1, \ldots, \nu-1) \\ a_\nu \geq N^{\frac{1}{\nu}} & \end{array}\right. \ , \tag{1.3}$$

as required. ■

We note that a simpler construction could have been used, had we wanted slightly weaker results, which are still asymptotically sufficient for proving the theorem. To see this, take $N = n\nu$ and start with the complete graph K_N. Now order the N vertices arbitrarily in n rows (each of length ν), as (v_{ij}) $(1 \leq i \leq n,\ 1 \leq j \leq \nu)$. For each pair of rows i, i', choose (independently) a single column $1 \leq j \leq \nu$, and remove the edge $v_{ij}v_{i'j}$ from the graph. This gives a graph G with $\binom{N}{2} - \binom{n}{2}$ edges. A calculation similar to the one above shows that with high probability $a_k \leq c_k\log_2(N)$ for $k < \nu$, and yet $\alpha(G^\nu) \geq n$ (as opposed to N in the original construction), hence $a_\nu \geq \left(\frac{N}{\nu}\right)^{\frac{1}{\nu}} \geq \frac{1}{2}N^{\frac{1}{\nu}}$.

1.3 Graphs with an irregular independence series

Theorem 1.1.2 states that there exists a graph G whose independence series exhibits an arbitrary (finite) number of jumps. Our first step towards proving this theorem is to examine the behavior of fixed powers of the form $k \geq \nu$ for the graphs described in the previous section. We show that these graphs, with high probability, satisfy $a_k = (1 + O(\log N))\, N^{\lfloor \frac{k}{\nu} \rfloor \frac{1}{k}}$, for every fixed $k \geq \nu$. The notation $a_k = (1 + O(\log N))\, N^{\alpha}$, here and in what follows, denotes that $N^{\alpha} \leq a_k \leq cN^{\alpha} \log N$ for a fixed $c > 0$. The lower bound of $N^{\lfloor \frac{k}{\nu} \rfloor \frac{1}{k}}$ for a_k can be derived from the cartesian product of the set I, defined in (1.2), with itself, $\lfloor \frac{k}{\nu} \rfloor$ times; the upper bound is more interesting. Fix an arbitrary set S, as before, however, this time, prior to generating S', we first remove from S all vertices which contain among their coordinates a set of the form $\{x, x+n, \ldots, x+(\nu-1)n\}$. This amounts to at most $\binom{k}{\nu}\nu! n\alpha(G^{k-\nu})$ vertices. This step ensures that S will not contain vertices that share a relation, such as the one appearing in the set I defined in (1.2). However, an edge class may still be completely contained in the coordinates of $u, v \in S$, in an interlaced form, for instance: $u = (x, y+n, x+2n, \ldots, x+(\nu-1)n, \ldots)$ and $v = (y, x+n, y+2n, \ldots, y+(\nu-1)n, \ldots)$. This will be automatically handled in generating S', since all vectors v with $x+tn$ in any of their coordinates are removed from S' after processing the vector u. Equation (1.1) remains valid, with $t_i < \nu$ for all i, however now me must be more careful in minimizing its right hand side. We note that for every $0 < t_i, t_j < \nu - 1$, setting $t'_i = \nu - 1, t'_j = t_i + t_j - t'_i$ reduces the product of $\frac{(\nu-t_i)(\nu-t_j)}{\nu^2}$. Therefore, again denoting by k' the number of distinct pairs of corresponding coordinates, we obtain the following bound on the probability of the edge uv:

$$\begin{aligned}\Pr[uv \in E(G^k)] \;&\geq\; \left(\frac{1}{\nu}\right)^{\lfloor \frac{k'}{\nu-1} \rfloor} \frac{\nu - (k' \bmod (\nu-1))}{\nu} \geq \\ &\geq\; \left(\frac{1}{\nu}\right)^{\frac{k'}{\nu-1}} \geq \left(\frac{1}{\nu}\right)^{\frac{k}{\nu-1}} \;. \qquad (1.4)\end{aligned}$$

Thus:

$$\Pr[uv \notin E(G^k)] \leq \mathrm{e}^{-\left(\frac{1}{\nu}\right)^{\frac{k}{\nu-1}}}$$

Now, the same consideration that showed $\alpha(G) \leq 2\log_\nu(N)$ implies that any set S' generated from S in this manner, which is of size $s' \geq 2k\nu^{\frac{k}{\nu-1}}\log(N)$, is almost surely not independent (for the sake of convenience, we set $p = \mathrm{e}^{-\left(\frac{1}{\nu}\right)^{\frac{k}{\nu-1}}}$). Indeed, the probability that there is such an independent set S' is at most:

$$\begin{aligned}\binom{N^k}{s'}p^{\binom{s'}{2}} &\leq \frac{p^{-s'/2}}{s'!}\left(N^k p^{\frac{s'}{2}}\right)^{s'} = \\ &= \frac{p^{-s'/2}}{s'!}\exp\left(k\log(N) - \frac{s'}{2}\nu^{-\frac{k}{\nu-1}}\right)^{s'} \leq \\ &\leq \frac{p^{-s'/2}}{s'!} = o(1)\ .\end{aligned}$$

Thus, almost surely, $|S'| \leq 2k\nu^{\frac{k}{\nu-1}}\log(N)$. Altogether, we have:

$$\alpha(G^k) \leq \binom{k}{\nu}\nu!n\alpha(G^{k-\nu}) + k^2\nu\alpha(G^{k-1}) \cdot 2k\nu^{\frac{k}{\nu-1}}\log(N) =$$

$$= \binom{k}{\nu}(\nu-1)!N\alpha(G^{k-\nu}) + 2k^3\nu^{1+\frac{k}{\nu-1}}\log(N)\alpha(G^{k-1}) \qquad (1.5)$$

For $k = \nu$ and a sufficiently large N, we get

$$N \leq \alpha(G^\nu) \leq N(\nu-1)! + 2\nu^{5+\frac{1}{\nu-1}}\log(N)\left(c_{\nu-1}\log_2(N)\right)^{\nu-1} \leq$$

$$\leq N\log_2(N)\ . \qquad (1.6)$$

Set $d_1 = \ldots = d_{\nu-1} = 0$, $d_\nu = 1$ and $d_k = 4k^3\nu^{1+\frac{k}{\nu-1}}d_{k-1}$ for $k > \nu$. It is easy to verify that $\frac{1}{2}d_k \geq \binom{k}{\nu}(\nu-1)!d_{k-\nu}$, and $\frac{1}{2}d_k \geq 2k^3\nu^{1+\frac{k}{\nu-1}}d_{k-1}$. Hence, by induction, equations (1.5) and (1.6) imply that for all $k \geq \nu$:

$$\alpha(G^k) \leq d_k N^{\lfloor\frac{k}{\nu}\rfloor}\log_2^k(N)$$

By definition of the d_k series,

$$\begin{aligned}d_k &\leq 4^{k-\nu}\left(\frac{k!}{\nu!}\right)^3\nu^{(k-\nu)\left(1+\frac{k}{\nu-1}\right)} \leq \\ &\leq 4^k(k!)^3\nu^{k\left(1+\frac{k-\nu}{\nu-1}\right)} = 4^k(k!)^3\nu^{k\frac{k-1}{\nu-1}}\ .\end{aligned}$$

Hence,

$$1 \leq \frac{a_k}{N^{\lfloor \frac{k}{\nu} \rfloor \frac{1}{k}}} \leq \sqrt{2} k^3 \nu^{\frac{k-1}{\nu-1}} \log_2(N) \ ,$$

as required.

Let us construct a graph whose independence series exhibits two jumps (an easy generalization will provide any finite number of jumps). Take a random graph, G_1, as described above, for some index ν_1 and a sufficiently large number of vertices N_1, and another (independent) random graph, G_2 for some other index $\nu_2 > \nu_1$, on $N_2 = N_1^{\alpha \frac{\nu_2}{\nu_1}}$ vertices (when $\alpha > 1$). Let $G = G_1 \cdot G_2$ be the strong product of the two graphs; note that G has $N = N_1 N_2$ vertices. It is crucial that we do not take G_1 and G_2 with jumps at indices ν_1, ν_2 respectively *separately*, but instead consider the product G of two random graphs constructed as above. We claim that with high probability, G satisfies:

$$a_k(G) = \begin{cases} O(\log N) & k < \nu_1 \\ (1 + O(\log N))\, N_1^{\lfloor \frac{k}{\nu_1} \rfloor \frac{1}{k}} & \nu_1 \leq k < \nu_2 \\ (1 + O(\log N))\, N_2^{\lfloor \frac{k}{\nu_2} \rfloor \frac{1}{k}} & k \geq \nu_2 \end{cases}$$

i.e., for $\nu_1 \leq k < \nu_2$ which is a multiple of ν_1, we have

$$a_k = (1 + O(\log N))\, N^{\frac{1}{\nu_1 + \alpha \nu_2}} \ ;$$

for $k \geq \nu_2$ which is a multiple of ν_2 we have

$$a_k = (1 + O(\log N))\, N^{\frac{\alpha}{\nu_1 + \alpha \nu_2}} \ .$$

Therefore, we get an exponential increase of order α at the index ν_2, and obtain two jumps, as required.

To prove the claim, argue as follows: following the formerly described methods, we filter an arbitrary set $S \subset V(G^k)$ to a subset S', in which every two vertices have a positive probability of being connected, and all such events are independent. This filtering is done as before - only now, we consider the criteria of both G_1 and G_2 when we discard vertices. In other words, if we denote by u^1, u^2 the k-tuples corresponding to G_1^k and G_2^k of a vertex $u \in V(G^k)$, then a vertex $u \in S$ filters out the vertex v from S iff u^1

would filter out v^1 in G_1^k or u^2 would filter out v^2 in G_2^k (or both). Recall that, by the method S' is generated from S, no two vertices in S' share an identical k-tuple of G_1^k or of G_2^k. Hence, two vertices $u, v \in S'$ are adjacent in G^k iff they are adjacent both in G_1^k and in G_2^k. These are two independent events, thus, by (1.1) and (1.4), we get the following fixed lower bound on the probability of u and v being adjacent:

$$\Pr[uv \in E(G^k)] = \Pr[u^1v^1 \in E(G_1^k)] \Pr[u^2v^2 \in E(G_2^k)] \geq \Omega(1)$$

This provides a bound of $O(\log N)$ for the size of S'. Combining this with the increase in the values of $\{a_k\}$ at indices ν_1 and ν_2 ($a_{\nu_i} \geq N_i^{\frac{1}{\nu_i}}$ for $i = 1, 2$) proves our claim.

In order to obtain any finite number of jumps, at indices $\nu_1, \ldots, \nu_s$, simply take a sufficiently large N_1 and set $N_i = N_{i-1}^{\alpha\frac{\nu_i}{\nu_{i-1}}}$ for $1 < i \leq s$, where $\alpha > 1$. By the same considerations used above, with high probability the graph $G = G_1 \cdot \ldots \cdot G_s$ (where G_i is a random graph designed to have a jump at index ν_i almost surely) satisfies $a_{\nu_i} \geq a_k^\alpha$ for all $k < \nu_i$. Hence for every $\varepsilon > 0$ we can choose $\alpha > \frac{1}{\varepsilon}$ and a sufficiently large N_1 so that $a_k < a_{\nu_i}^\varepsilon$ for all $k < \nu_i$. This completes the proof. ■

1.4 Concluding remarks and open problems

We have shown that even when the independence series stabilizes for an arbitrary (fixed) number of elements, or jumps and then stabilizes, it still does not necessarily approximate the Shannon capacity up to any power of $\varepsilon > 0$. However, our constructions require the number of vertices to be exponentially large in the values of the jump indices ν_i. We believe that this is not a coincidence, namely a prefix of *linear* (in the number of vertices) length of the independence series can provide a good approximation of the Shannon capacity. The following two specific conjectures seem plausible:

Conjecture 1.4.1. *For every graph G on n vertices, $\max\{a_k\}_{k\leq n} \geq \frac{1}{2}c(G)$, that is, the largest of the first n elements of the independence series gives a 2-approximation for $c(G)$.*

Conjecture 1.4.2. *For every $\varepsilon > 0$ there exists an $r = r(\varepsilon)$ such that for a sufficiently large n and for every graph G on n vertices, the following is true:* $\max\{a_k\}_{k \leq n^r} \geq (1-\varepsilon)c(G)$.

Our proof of Theorem 1.1.1 shows the existence of a graph whose independence series increases by a factor of N^{δ} at the k-th power, where $\delta = \frac{1-o(1)}{k}$. It would be interesting to decide if there is a graph satisfying this property for a *constant* $\delta > 0$ (independent of k). This relates to a question on channel discrepancy raised in [18], where the authors show that the ratio between the independence number and the Shannon capacity of a graph on n vertices can be at least $n^{\frac{1}{2}-o(1)}$, and ask whether this is the largest ratio possible. Proving Theorem 1.1.1 for a constant $\delta > 0$ will give a negative answer for the following question, which generalizes the channel discrepancy question mentioned above:

Question 1.4.3. *Does* $\max\{a_i\}_{i \leq k}$*, for any fixed $k \geq 2$, approximate $c(G)$ up to a factor of $n^{\frac{1}{k}+o(1)}$ (where $n = |V(G)|$)?*

Although our results exhibit the difficulty in approximating the Shannon capacity of a given graph G, this problem is not even known to be NP-hard (although it seems plausible that it is in fact much harder). We conclude with a question concerning the complexity of determining the value of $c(G)$ accurately for a given graph G:

Question 1.4.4. *Is the problem of deciding whether the Shannon Capacity of a given graph exceeds a given value decidable?*

Chapter 2

Privileged users in zero-error transmission

The results of this chapter appear in [13]

The k-th power of a graph G is the graph whose vertex set is $V(G)^k$, where two distinct k-tuples are adjacent iff they are equal or adjacent in G in each coordinate. The Shannon capacity of G, $c(G)$, is $\lim_{k\to\infty} \alpha(G^k)^{\frac{1}{k}}$, where $\alpha(G)$ denotes the independence number of G. When G is the characteristic graph of a channel $\mathcal{C}$, $c(G)$ measures the effective alphabet size of $\mathcal{C}$ in a zero-error protocol. A sum of channels, $\mathcal{C} = \sum_i \mathcal{C}_i$, describes a setting when there are $t \geq 2$ senders, each with his own channel $\mathcal{C}_i$, and each letter in a word can be selected from any of the channels. This corresponds to a disjoint union of the characteristic graphs, $G = \sum_i G_i$. It is well known that $c(G) \geq \sum_i c(G_i)$, and in [8] it is shown that in fact $c(G)$ can be larger than any fixed power of the above sum.

We extend the ideas of [8] and show that for every $\mathcal{F}$, a family of subsets of $[t]$, it is possible to assign a channel $\mathcal{C}_i$ to each sender $i \in [t]$, such that the capacity of a group of senders $X \subset [t]$ is high iff X contains some $F \in \mathcal{F}$. This corresponds to a case where only privileged subsets of senders are allowed to transmit in a high rate. For instance, as an analogue to secret sharing, it is possible to ensure that whenever at least k senders combine their channels, they obtain a high capacity, however every group of $k-1$ senders has a low capacity (and yet is not totally denied of service). In the process, we obtain an explicit Ramsey construction of an edge-coloring of the complete graph on n vertices by t colors, where every induced subgraph on $\exp\big(\Omega(\sqrt{\log n \log\log n})\big)$ vertices contains all t colors.

2.1 Introduction

A channel $\mathcal{C}$ on an input alphabet V and an output alphabet U maps each $x \in V$ to some $S(x) \subset U$, such that transmitting x results in one of the letters of $S(x)$. The characteristic graph of the channel $\mathcal{C}$, $G = G(\mathcal{C})$, has a vertex set V, and two vertices $x \neq y \in V$ are adjacent iff $S(x) \cap S(y) \neq \emptyset$, i.e., the corresponding input letters are confusable in the channel. Clearly, a maximum set of predefined letters which can be transmitted in $\mathcal{C}$ without possibility of error corresponds to a maximum independent set in the graph G, whose size is $\alpha(G)$ (the independence number of G).

The strong product of two graphs, $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ is the graph, $G_1 \cdot G_2$, on the vertex set $V_1 \times V_2$, where two vertices $(u_1, u_2) \neq (v_1, v_2)$ are adjacent iff for all $i = 1, 2$, either $u_i = v_i$ or $u_i v_i \in E_i$. In other words, the pairs of vertices in both coordinates are either equal or adjacent. This product is associative and commutative, hence we can define G^k to be the k-th power of G, where two vertices $(u_1, \ldots, u_k) \neq (v_1, \ldots, v_k)$ are adjacent iff for all $i = 1, \ldots, k$, either $u_i = v_i$ or $u_i v_i \in E(G)$.

Note that if I, J are independent sets of two graphs, G, H, then $I \times J$ is an independent set of $G \cdot H$. Therefore, $\alpha(G^{n+m}) \geq \alpha(G^n)\alpha(G^m)$ for every $m, n \geq 1$, and by Fekete's lemma (cf., e.g., [76], p. 85), the limit $\lim_{n\to\infty} \alpha(G^n)^{\frac{1}{n}}$ exists and equals $\sup_n \alpha(G^n)^{\frac{1}{n}}$. This parameter, introduced by Shannon in [97], is the Shannon capacity of G, denoted by $c(G)$.

When sending k-letter words in the channel $\mathcal{C}$, two words are confusable iff the pairs of letters in each of their k-coordinates are confusable. Thus, the maximal number of k-letter words which can be sent in $\mathcal{C}$ without possibility of error is precisely $\alpha(G^k)$, where $G = G(\mathcal{C})$. It follows that for sufficiently large values of k, the maximal number of k-letter words which can be sent without possibility of error is roughly $c(G)^k$. Hence, $c(G)$ represents the effective alphabet size of the channel in zero-error transmission.

The sum of two channels, $\mathcal{C}_1 + \mathcal{C}_2$, describes the setting where each letter can be sent from either of the two channels, and letters from $\mathcal{C}_1$ cannot be confused with letters from $\mathcal{C}_2$. The characteristic graph in this case is the disjoint union $G_1 + G_2$, where G_i is the characteristic graph of $\mathcal{C}_i$. Shannon showed in [97] that $c(G_1 + G_2) \geq c(G_1) + c(G_2)$ for every two graphs G_1 and G_2, and conjectured that in fact $c(G_1 + G_2) = c(G_1) + c(G_2)$ for all G_1 and G_2. This was disproved by Alon [8] using an explicit construction of two graphs G_1, G_2 with a capacity $c(G_i) \leq k$, satisfying $c(G_1 + G_2) \geq k^{\Omega(\frac{\log k}{\log \log k})}$.

We extend the ideas of [8] and show that it is possible to construct t graphs, G_i ($i \in [t] = \{1, 2, \ldots, t\}$), such that for every subset $X \subseteq [t]$, the Shannon capacity of $\sum_{i \in X} G_i$ is high iff X contains some subset of a predefined family $\mathcal{F}$ of subsets of $[t]$. This corresponds to assigning t channels to t senders, such that designated groups of senders $F \in \mathcal{F}$ can obtain a high capacity by combining their channels ($\sum_{i \in F} \mathcal{C}_i$), and yet every group of senders $X \subseteq [t]$ not containing any $F \in \mathcal{F}$ has a low capacity. In particular, a choice of $\mathcal{F} = \{F \subset [t] \ : \ |F| = k\}$ implies that every set X of senders has a high Shannon capacity of $\sum_{i \in X} \mathcal{C}_i$ if $|X| \geq k$, and a low capacity otherwise. The following theorem, proved in Section 2.2, formalizes the claims above:

Theorem 2.1.1. *Let $T = \{1, \ldots, t\}$ for some fixed $t \geq 2$, and let $\mathcal{F}$ be a family of subsets of T. For every (large) n it is possible to construct graphs G_i, $i \in T$, each on n vertices, such that the following two statements hold for all $X \subseteq T$:*

1. *If X contains some $F \in \mathcal{F}$, then $c(\sum_{i \in X} G_i) \geq n^{1/|F|} \geq n^{1/t}$.*
2. *If X does not contain any $F \in \mathcal{F}$, then*

$$c(\sum_{i \in X} G_i) \leq \mathrm{e}^{(1+o(1))\sqrt{2 \log n \log \log n}} \ ,$$

where the $o(1)$-term tends to 0 as $n \to \infty$.

As a by-product, we obtain the following Ramsey construction, where instead of forbidding monochromatic subgraphs, we require “rainbow” subgraphs (containing all the colors used for the edge-coloring). This is stated by the next proposition, which is proved in Section 2.3:

Proposition 2.1.2. *For every (large) n and $t \leq \sqrt{\frac{2\log n}{(\log\log n)^3}}$ there is an explicit t-edge-coloring of the complete graph on n vertices, such that every induced subgraph on*

$$\mathrm{e}^{(1+o(1))\sqrt{8\log n \log\log n}}$$

vertices contains all t colors.

This extends the construction of Frankl and Wilson [56] that deals with the case $t = 2$ (using a slightly different construction).

2.2 Graphs with high capacities for unions of predefined subsets

The upper bound on the capacities of subsets not containing any $F \in \mathcal{F}$ relies on the algebraic bound for the Shannon capacity using representations by polynomials, proved in [8]. See also Haemers [63] for a related approach.

Definition 2.1. *Let $\mathbb{K}$ be a field, and let $\mathcal{H}$ be a linear subspace of polynomials in r variables over $\mathbb{K}$. A* ***representation*** *of a graph $G = (V, E)$ over $\mathcal{H}$ is an assignment of a polynomial $f_v \in \mathcal{H}$ and a value $c_v \in \mathbb{K}^r$ to every $v \in V$, such that the following holds: for every $v \in V$, $f_v(c_v) \neq 0$, and for every $u \neq v \in V$ such that $uv \notin E$, $f_u(c_v) = 0$.*

Theorem 2.2.1 ([8])**.** *Let $G = (V, E)$ be a graph and let $\mathcal{H}$ be a space of polynomials in r variables over a field $\mathbb{K}$. If G has a representation over $\mathcal{H}$ then $c(G) \leq \dim(\mathcal{H})$.*

We need the following simple lemma:

Lemma 2.2.2. *Let $T = [t]$ for $t \geq 1$, and let $\mathcal{F}$ be a family of subsets of T. There exist sets $A_1, A_2, \ldots, A_t$ such that for every $X \subseteq T$:*

$$X \textit{ does not contain any } F \in \mathcal{F} \iff \bigcap_{i \in X} A_i \neq \emptyset\ .$$

Furthermore, $|\bigcup_{i=1}^{t} A_i| \leq \binom{t}{\lfloor t/2 \rfloor}$.

Proof of lemma. Let $\mathcal{Y}$ denote the family of all maximal sets Y such that Y does not contain any $F \in \mathcal{F}$. Assign a unique element p_Y to every $Y \in \mathcal{Y}$, and define:

$$A_i = \{p_Y \ : \ i \in Y \ , \ Y \in \mathcal{Y}\} \ . \tag{2.1}$$

Let $X \subsetneq T$, and note that (2.1) implies that $\bigcap_{i \in X} A_i = \{p_Y \ : \ X \subseteq Y\}$. Thus, if X does not contain any $F \in \mathcal{F}$, then $X \subseteq Y$ for some $Y \in \mathcal{Y}$, and hence $p_Y \in \bigcap_{i \in X} A_i$. Otherwise, X contains some $F \in \mathcal{F}$ and hence is not a subset of any $Y \in \mathcal{Y}$, implying that $\bigcap_{i \in X} A_i = \emptyset$.

Finally, observe that $\mathcal{Y}$ is an anti-chain and that $|\bigcup_{i=1}^t A_i| \leq |\mathcal{Y}|$, hence the bound on $|\bigcup_{i=1}^t A_i|$ follows from Sperner's Theorem [98]. ■

Proof of Theorem 2.1.1. Let p be a large prime, and let $\{p_Y \ : \ Y \in \mathcal{Y}\}$ be the first $|\mathcal{Y}|$ primes succeeding p. Define $s = p^2$ and $r = p^3$, and note that, as t and hence $|\mathcal{Y}|$ are fixed, by well-known results about the distribution of prime numbers, $p_Y = (1 + o(1))p < s$ for all Y, where the $o(1)$-term tends to 0 as $p \to \infty$.

The graph $G_i = (V_i, E_i)$ is defined as follows: its vertex set V_i consists of all $\binom{r}{s}$ possible s-element subsets of $[r]$, and for every $A \neq B \in V_i$:

$$(A, B) \in E_i \iff |A \cap B| \equiv s \pmod{p_Y} \text{ for some } p_Y \in A_i \ . \tag{2.2}$$

Let $X \subsetneq T$. If X does not contain any $F \in \mathcal{F}$, then, by Lemma 2.2.2, $\bigcap_{i \in X} A_i \neq \emptyset$, hence there exists some q such that $q \in A_i$ for every $i \in X$. Therefore, for every $i \in X$, if A, B are disconnected in G_i, then $|A \cap B| \not\equiv s \pmod{q}$. It follows that the graph $\sum_{i \in X} G_i$ has a representation over a subspace of the multi-linear polynomials in $|X|r$ variables over $\mathbb{Z}_q$ with a degree smaller than q. To see this, take the variables $x_j^{(i)}$, $i = 1, \ldots, |X|$, $j = 1, \ldots, r$, and assign the following polynomial to each vertex $A \in V_i$:

$$f_A(\overline{x}) = \prod_{u \not\equiv s} (u - \sum_{j \in A} x_j^{(i)}) \ .$$

The assignment c_A is defined as follows: $x_j^{(i')} = 1$ if $i' = i$ and $j \in A$, otherwise $x_j^{(i')} = 0$. As every assignment $c_{A'}$ gives values in $\{0, 1\}$ to all $x_j^{(i)}$, it is possible to reduce every f_A modulo the polynomials $(x_j^{(i)})^2 - x_j^{(i)}$ for all i

and j, and obtain multi-linear polynomials, equivalent on all the assignments $c_{A'}$.

The following holds for all $A \in V_i$:

$$f_A(c_A) = \prod_{u \not\equiv s}(u - s) \not\equiv 0 \pmod{q} ,$$

and for every $B \neq A$:

$$B \in V_i \ , \ (A,B) \notin E_i \implies f_A(c_B) = \prod_{u \not\equiv s}(u - |A \cap B|) \equiv 0 \pmod{q} ,$$
$$B \notin V_i \implies f_A(c_B) = \prod_{u \not\equiv s} u \equiv 0 \pmod{q} ,$$

where the last equality is by the fact that $s \not\equiv 0 \pmod{q}$, as $s = p^2$ and $p < q$. As the polynomials f_A lie in the direct sum of $|X|$ copies of the space of multi-linear polynomials in r variables of degree less than q, it follows from Theorem 2.2.1 that the Shannon capacity of $\sum_{i\in X} G_i$ is at most:

$$|X| \sum_{i=0}^{q-1} \binom{r}{i} \leq t \sum_{i=0}^{q-1} \binom{r}{i} < t \binom{r}{q} .$$

Recalling that $q = (1 + o(1))p$ and writing $t\binom{r}{q}$ in terms of $n = \binom{r}{s}$ gives the required upper bound on $c(\sum_{i\in X} G_i)$.

Assume now that X contains some $F \in \mathcal{F}$, $F = \{i_1, \ldots, i_{|F|}\}$. We claim that the following set is an independent set in $\left(\sum_{i\in X} G_i\right)^{|F|}$:

$$\{(A^{(i_1)}, A^{(i_2)}, \ldots, A^{(i_{|F|})}) \ : \ A \subseteq [r] \ , \ |A| = s\} ,$$

where $A^{(i_j)}$ is the vertex corresponding to A in V_{i_j}. Indeed, if $(A, A, \ldots, A)$ and $(B, B, \ldots, B)$ are adjacent, then for every $i \in F$, $|A \cap B| \equiv s \pmod{p_Y}$ for some $p_Y \in A_i$. However, $\bigcap_{i\in F} A_i = \emptyset$, hence there exist $p_Y \neq p'_Y$ such that $|A \cap B|$ is equivalent both to $s \pmod{p_Y}$ and to $s \pmod{p'_Y}$. By the Chinese Remainder Lemma, it follows that $|A \cap B| = s$ (as $|A \cap B| < p_Y p'_Y$), thus $A = B$. Therefore, the Shannon capacity of $\sum_{i\in X} G_i$ is at least $\binom{r}{s}^{1/|F|} = n^{1/|F|}$. ■

2.3 Explicit construction for rainbow Ramsey graphs

Proof of Proposition 2.1.2. Let p be a large prime, and let $p_1 < \ldots < p_t$ denote the first t primes succeeding p. We define r, s as in the proof of Theorem 2.1.1: $s = p^2$, $r = p^3$, and consider the complete graph on n vertices, K_n, where $n = \binom{r}{s}$, and each vertex corresponds to an s-element subset of $[r]$. The fact that $t \leq \sqrt{\frac{2\log n}{(\log\log n)^3}}$ implies that $t \leq (\frac{1}{2} + o(1))\frac{p}{\log p}$, and hence, by the distribution of prime numbers, $p_t < 2p$ (with room to spare) for a sufficiently large value of p.

We define an edge-coloring γ of K_n by t colors in the following manner: for every $A, B \in V$, $\gamma(A, B) = i$ if $|A \cap B| \equiv s \pmod{p_i}$ for some $i \in [t]$, and is arbitrary otherwise. Note that for every $i \neq j \in \{1, \ldots, t\}$, $s < p_i p_j$. Hence, if $|A \cap B| \equiv s \pmod{p_i}$ and $|A \cap B| \equiv s \pmod{p_j}$ for such i and j, then by the Chinese Remainder Lemma, $|A \cap B| = s$, and in particular, $A = B$. Therefore, the coloring γ is well-defined.

It remains to show that every large induced subgraph of K_n has all t colors according to γ. Indeed, this follows from the same consideration used in the proof of Theorem 2.1.1. To see this, let G_i denote the spanning subgraph of K_n whose edge set consists of all (A, B) such that $\gamma(A, B) = i$. Each pair $A \neq B$, which is disconnected in G_i, satisfies $|A \cap B| \not\equiv s \pmod{p_i}$. Therefore, G_i has a representation over the multi-linear polynomials in r variables over $\mathbb{Z}_{p_i}$ with a degree smaller than p_i (define $f_A(x_1, \ldots, x_r)$ as is in the proof of Theorem 2.1.1, and take c_A to be the characteristic vector of A). Thus, $c(G_i) < \binom{r}{p_i}$, and in particular, $\alpha(G_i) < \binom{r}{p_i}$. This ensures that every induced subgraph on at least $\binom{r}{p_i} \leq \binom{r}{2p}$ vertices contains an i-colored edge, and the result follows. ■

Chapter 3

Non-linear index coding outperforming the linear optimum

The results of this chapter appear in [83]

The following source coding problem was introduced by Birk and Kol: a sender holds a word $x \in \{0,1\}^n$, and wishes to broadcast a codeword to n receivers, $R_1, \ldots, R_n$. The receiver R_i is interested in x_i, and has prior *side information* comprising some subset of the n bits. This corresponds to a directed graph G on n vertices, where ij is an edge iff R_i knows the bit x_j. An *index code* for G is an encoding scheme which enables each R_i to always reconstruct x_i, given his side information. The minimal word length of an index code was studied by Bar-Yossef, Birk, Jayram and Kol [23]. They introduced a graph parameter, $\mathrm{minrk}_2(G)$, which completely characterizes the length of an optimal *linear* index code for G. The authors of [23] showed that in various cases linear codes attain the optimal word length, and conjectured that linear index coding is in fact *always* optimal.

In this work, we disprove the main conjecture of [23] in the following strong sense: for any $\varepsilon > 0$ and sufficiently large n, there is an n-vertex graph G so that every linear index code for G requires codewords of length at least $n^{1-\varepsilon}$, and yet a non-linear index code for G has a word length of n^{ε}.

This is achieved by an explicit construction, which extends Alon's variant of the celebrated Ramsey construction of Frankl and Wilson.

In addition, we study optimal index codes in various, less restricted, natural models, and prove several related properties of the graph parameter $\mathrm{minrk}(G)$.

3.1 Introduction

Source coding deals with a scenario in which a *sender* has some data string x he wishes to transmit through a broadcast channel to *receivers*. The first and classical result in this area is Shannon's Source Coding Theorem. This has been followed by various scenarios which differ in the nature of the data to be transmitted, the broadcast channel and some assumptions on the computational abilities of the users. Another family of source coding problems, which attracted a considerable amount of attention over the years, deals with the assumption that the receivers possess some prior knowledge on the data string x. It was shown that in some cases even some restricted assumptions on this knowledge may drastically affect the nature of the coding problem.

In this chapter we consider a variant of source coding which was first proposed by Birk and Kol [27]. In this variant, called Informed Source Coding On Demand (ISCOD), each receiver has some prior side information, comprising some subset of the input word x. The sender is aware of the portion of x known to each receiver. Moreover, each receiver is interested in just part of the data. Following [23], we restrict ourselves to the problem which is formalized as follows.

Definition 3.1 (index code). *A sender wishes to send a word $x \in \{0,1\}^n$ to n receivers $R_1, \ldots, R_n$. Each R_i knows some of the bits of x and is interested solely in the bit x_i. An* index code *of length ℓ for this setting is a binary code of word-length ℓ, which enables R_i to recover x_i for any x and i.*

Using a graph model for the side-information, this problem can be restated as a graph parameter. For a directed graph G and a vertex v, let $N_G^+(v)$ be the set of out-neighbors of v in G, and for $x \in \{0,1\}^n$ and $S \subset [n] = \{1, \ldots, n\}$, let $x|_S$ be the restriction of x to the coordinates of S.

Definition 3.2 (**ℓ(G)**)**.** *The setting of Definition 3.1 is characterized by the directed* side information graph *G on the vertex set $[n]$, where (i, j) is an edge iff R_i knows the value of x_j. An* index code *of length ℓ for G is a function $E : \{0, 1\}^n \to \{0, 1\}^\ell$ and functions $D_1, \ldots, D_n$, so that for all $i \in [n]$ and $x \in \{0, 1\}^n$, $D_i(E(x), x|_{N_G^+(i)}) = x_i$. Denote the minimal length of an index code for G by $\ell(G)$.*

Example: Suppose that every receiver R_i knows in advance the whole word x, except for the single bit x_i he wishes to recover. The corresponding side information graph G is the complete graph K_n (that is, (i, j) is an edge for all $i \neq j$). By broadcasting the XOR of all the bits of x, each receiver can easily compute its missing bit:

$$\begin{aligned} &E(x) = \bigoplus_{i=1}^{n} x_i \ , \\ &D_i(E(x), x|_{\{j:j\neq i\}}) = E(x) \oplus (\bigoplus_{j\neq i} x_j) = x_i \ . \end{aligned}$$

In this case the code has length $\ell = 1$ and E is a linear function of x over $GF(2)$.

The problem of Informed Source Coding On Demand (ISCOD) was presented by Birk and Kol [27]. They were motivated by various applications of distributed communication such as satellite communication networks with caching clients. In such applications, the clients have limited storage and maintain part of the transmitted information. Subsequently, the clients receive requests for arbitrary information blocks, and may use a slow backward channel to advise the server of their status. The server, playing the role of the sender in Definition 3.1, then broadcasts a single transmission to all clients (the receivers). As observed by Birk and Kol [27], when the sender has only partial knowledge of the side information (e.g., the *number* of missing blocks for each user), an erasure correcting code such as Reed-Solomon Code performs well. This is also the case if every user is expected to be able to decode the whole information. The authors of [27] present some bounds and heuristics for obtaining efficient encoding schemes, as well as protocols for implementing the above scenario. See [27] and [23] for more details on the relation between the source coding problem, as formulated above, and the

ISCOD problem, as well as the communication complexity of the indexing function, random access codes and network coding.

Bar-Yossef, Birk, Jayram and Kol [23] further investigated index coding. They showed that this problem is different in nature from the well-known source coding problems previously studied by Witsenhausen in [106]. Their main contribution is an upper bound on $\ell(G)$, the optimal length of an index code (Definition 3.2). The upper bound is a graph parameter denoted by $\mathrm{minrk}_2(G)$, which is also shown to be the length of the optimal *linear* index code. It is shown in [23] that in several cases linear codes are in fact optimal, e.g., for directed acyclic graphs, perfect graphs, odd cycles and odd anti-holes. An information theoretic lower bound on $\ell(G)$ is obtained: it is at least the size of a maximal acyclic induced subgraph of G. This lower bound holds even for the relaxed problem of *randomized* index codes, where the sender is allowed to use (public) random coins during encoding, and the receivers are expected to decode their information correctly with high probability over these coin flips. Nevertheless, they show that in some cases the lower bound is not tight.

Having proved that the upper bound $\ell(G) \leq \mathrm{minrk}_2(G)$ is tight for several natural graph families and under some relaxed restrictions on the code ("semi-linearly-decodable"), the authors of [23] conjectured that the length of the optimal index code is in fact equal to $\mathrm{minrk}_2(G)$. That is, they conjectured that linear index coding is always optimal, and concluded that this was the main open problem to be investigated.

Before stating the main results of this chapter, we review the definition of $\mathrm{minrk}_2(G)$ and other related graph theoretic parameters.

3.1.1 Definitions, notations and background

Let $G = (V, E)$ be a directed graph on the vertex set $V = [n]$. The adjacency matrix of G, denoted by $A_G = (a_{ij})$, is the $n \times n$ binary matrix where $a_{ij} = 1$ iff $(i, j) \in E$. An *independent set* of G is a set of vertices which have no edges between them, and the *independence number* of G, $\alpha(G)$, is the cardinality of a maximum independent set. The *chromatic number* of G, $\chi(G)$, is the minimum number of independent sets whose union is all of V. Let $\overline{G}$ denote the *graph complement* of G. A *clique* of G is an independent set of $\overline{G}$ (i.e., a set of vertices such that all edges between them belong to G), and the *clique number* of G, $\omega(G)$, is the cardinality of a maximum clique. Without being formal, a graph G is called "Ramsey" if both $\alpha(G)$ and $\omega(G)$ are "small".

In [23], a binary $n \times n$ matrix $A = (a_{ij})$ was said to "fit" G if A has 1-s on its diagonal, and 0 in all the indices i, j where $i \neq j$ and $(i, j) \notin E$. The parameter $\mathrm{minrk}_2(G)$ was defined to be the minimal possible rank over $GF(2)$ of a matrix which fits G.

To extend this definition to a general field, let $A = (a_{ij})$ be an $n \times n$ matrix over some field $\mathbb{F}$. We say that A *represents* the graph G over $\mathbb{F}$ if $a_{ii} \neq 0$ for all i, and $a_{ij} = 0$ whenever $i \neq j$ and $(i, j) \notin E$. The *minrank* of a directed graph G with respect to the field $\mathbb{F}$ is defined by

$$\mathrm{minrk}_{\mathbb{F}}(G) = \min\{\mathrm{rank}_{\mathbb{F}}(A) : A \text{ represents } G \text{ over } \mathbb{F}\} \ . \tag{3.1}$$

For the common case where $\mathbb{F}$ is a finite field, we abbreviate:

$$\mathrm{minrk}_{p^k}(G) = \mathrm{minrk}_{GF(p^k)}(G) \ .$$

The notion of $\mathrm{minrk}(G)$ for an undirected graph G was first considered in the context of graph capacities by Haemers [63],[64]. The Shannon capacity of the graph G, denoted by $c(G)$, is a notoriously challenging parameter, which was defined by Shannon in [97], and remains unknown even for simple graphs, such as C_7, the cycle on 7 vertices. Lower bounds for $c(G)$ are given in terms of independence numbers of certain graphs, and in particular, $\alpha(G) \leq c(G)$. Haemers showed that for all $\mathbb{F}$, $\mathrm{minrk}_{\mathbb{F}}(G)$ is sandwiched between $c(G)$ and $\chi(\overline{G})$, the chromatic number of the complement of G, altogether giving

$$\alpha(G) \leq c(G) \leq \mathrm{minrk}_{\mathbb{F}}(G) \leq \chi(\overline{G}) \ . \tag{3.2}$$

While $\mathrm{minrk}_{\mathbb{F}}(G)$ can prove to be difficult to compute, the most useful upper bound for $c(G)$ is $\vartheta(G)$, the Lovász ϑ-function, which was introduced in the seminal paper [81] to compute $c(C_5)$. The matrix-rank argument was thereafter introduced by Haemers to answer some questions of [81], and has since been used (under some variants) in additional settings to obtain better bounds than those provided by the ϑ-function (cf., e.g., [8]).

3.1.2 New results

The main result of this chapter is an improved upper bound on the length of index codes, which is shown to strictly improve upon the $\mathrm{minrk}_2(G)$ bound. This disproves the main conjecture of [23] regarding the optimality of linear index coding, as stated by the following theorem.

Theorem 3.1.1. *For any $\varepsilon > 0$ and any sufficiently large n, there is an n-vertex graph G so that:*

1. *Any linear index code for G requires $n^{1-\varepsilon}$ bits, that is,*

$$\mathrm{minrk}_2(G) \geq n^{1-\varepsilon} \ .$$

2. *There exists a non-linear index code for G using n^{ε} bits, that is,*

$$\ell(G) \leq n^{\varepsilon} \ .$$

Moreover, the graph G is undirected and can be constructed explicitly.

Note that this in fact disproves the conjecture of Bar-Yossef et al. in the following strong sense: the ratio between an optimal code and an optimal linear code over $GF(2)$ can be $n^{1-o(1)}$. The essence of the new upper bound lies in the fact that, in some cases, linear codes over other fields[1] may yield significantly better coding schemes. This notion is incorporated in the following upper bound on $\ell(G)$, which is a simple extension of a result of [23] (the special case $\mathbb{F} = GF(2)$).

Theorem 3.1.2. *Let G be a graph, and let A be a matrix which represents G over some field $\mathbb{F}$ (not necessarily finite). Then:*

$$\ell(G) \leq \lceil \ \log_2 |\{Ax : x \in \{0,1\}^n\}| \ \rceil \ .$$

In particular, the following holds:

$$\ell(G) \leq \min_{\mathbb{F}\,:\,|\mathbb{F}|<\infty} \lceil \ \mathrm{minrk}_{\mathbb{F}}(G) \log_2 |\mathbb{F}| \ \rceil \ . \tag{3.3}$$

[1] The term “linear codes over GF(p)” is used to describe a coding scheme, in which the input word is encoded into a sequence of linear functionals of its symbols over $GF(p)$, which are subsequently used for the decoding. The protocol for transmitting these functionals need not be linear.

Indeed, for some graphs the minimum of (3.3) is attained when $\mathbb{F} \neq GF(2)$, in which case the linear code over $GF(2)$ is suboptimal. Proposition 3.2.2 (Section 3.2) provides a construction of such graphs, and is the main ingredient in the proof of Theorem 3.1.1. This proposition, which may be of independent interest, states that for any pair of finite fields with distinct characteristics, $\mathbb{F}$ and $\mathbb{K}$, the gap between $\mathrm{minrk}_{\mathbb{F}}$ and $\mathrm{minrk}_{\mathbb{K}}$ can be $n^{1-o(1)}$. Theorem 3.1.1 is then obtained as a corollary of Theorem 3.1.2 and a special case of Proposition 3.2.2.

As a corollary, Proposition 3.2.2 yields that $\mathrm{minrk}_{\mathbb{F}}(G)/\vartheta(G)$ (where ϑ is the Lovász ϑ-function and $|V(G)| = n$) is in some cases (roughly) at least $\sqrt{n}$, whereas in other cases it is (roughly) at most $1/\sqrt{n}$. This addresses another question of [23] on the relation between these two parameters.

An additional corollary of Proposition 3.2.2 states that $\ell(G)$ may be much smaller than the upper bounds of Theorem 3.1.2. That is, there are graphs for which *all* linear codes are suboptimal.

Corollary 3.1.3. *For any $\varepsilon > 0$ and a sufficiently large n, there is a graph G on n vertices so that $\ell(G) \leq n^{\varepsilon}$, and yet $c(G) \geq \sqrt{n}$. In particular, for any field $\mathbb{F}$, $\mathrm{minrk}_{\mathbb{F}}(G) \geq c(G) \geq \sqrt{n}$.*

We also extend the main construction of Proposition 3.2.2 and give, for any prescribed set of finite fields $\{\mathbb{F}_i\}$ and an additional finite field $\mathbb{K}$ of a distinct characteristic, a construction of a graph G so that $\mathrm{minrk}_{\mathbb{F}_i}(G)$ is "large" for all i, whereas $\mathrm{minrk}_{\mathbb{K}}(G)$ is "small". Thus, one cannot hope for a tight upper bound of the type of Theorem 3.1.2 in which only a finite set of fields is considered.

Proposition 3.1.4. *For any fixed t, let $\mathbb{F}_1, \ldots, \mathbb{F}_t$ denote finite fields, and let $\mathbb{K}$ denote a finite field of a distinct characteristic. For any $\varepsilon > 0$ and a sufficiently large n, there is an explicit construction of a graph G on n vertices, so that $\mathrm{minrk}_{\mathbb{K}}(G) \leq n^{\varepsilon}$, whereas for all $i \in [t]$, $\mathrm{minrk}_{\mathbb{F}_i}(G) \geq n^{(1-\varepsilon)/t}$.*

In the second part of this chapter, we revisit the problem definition. It is shown that the restricted problem given in Definition 3.1 captures many other cases arising from the original distributed applications, which motivated the study of Informed Source Coding On Demand. In particular, we suggest appropriate models and reductions for cases in which multiple users are interested in the same bit, there are multiple rounds of transmission and the transmitted words are over a large alphabet. These models are obtained as natural extensions of the original problem, and exhibit interesting relations to the original parameters $\ell(G)$ and $\text{minrk}(G)$.

3.1.3 Techniques

A key element in the proof of the main result is an extended version of the Ramsey graph constructed by Alon [8], which is a variant of the well-known Ramsey construction of Frankl and Wilson [56]. This graph, $G_{p,q}$ for some large primes p, q, was used by Alon in order to disprove an old conjecture of Shannon [97] on the Shannon capacity of a union of graphs.

Using some properties of the minrk parameter, one can show that the graph $G_{p,q}$ has a "small" minrk_p and a "large" minrk_q, implying that the optimal linear index code over $GF(p)$ may be significantly better than the one over $GF(q)$. However, it is imperative in the above construction that both p and q will be large, whereas we are interested in the case $q = 2$, corresponding to minrk_2. To this end, we extend the above construction of [8] to prime-powers, using some well known results on congruencies of binomial coefficients.

In order to generalize the main result and obtain a graph which has a "short" linear index code over some field $\mathbb{K}$, yet "long" linear index codes over a given set of fields $\mathbb{F}_1, \ldots, \mathbb{F}_t$, we follow the approach of [1] and [89], and consider a graph product of previously constructed graphs. En route, we derive several properties of the minrank parameter, which may be of independent interest.

The proofs of the results throughout the chapter combine arguments from Linear Algebra and Number Theory along with some additional ideas, inspired by the theory of graph capacities under various definitions of graph products.

3.1.4 Organization

The rest of the chapter is organized as follows. Section 3.2 contains a description of the basic construction, and the proof of the main result (Theorem 3.1.1). In Section 3.3, we study the various extensions of the original problem. Section 3.5 contains some concluding remarks and open problems.

3.2 Non-linear index coding schemes

We begin with the proof of Theorem 3.1.2, which is a simple extension of a result in [23].

Proof of Theorem 3.1.2. Let $V = [n]$ denote the vertex set of G, $A = (a_{ij})$ denote a matrix which represents G over some field $\mathbb{F}$ (not necessarily finite), and $S = \{Ax : x \in \{0,1\}^n\} \subset \mathbb{F}^n$. For some arbitrary ordering of the elements of S, the encoding of $x \in \{0,1\}^n$ is the label of Ax, requiring a word-length of $\lceil \log_2 |S| \rceil$ bits. For decoding, the i-th receiver R_i examines $(Ax)_i$, and since the diagonal of A does not contain zero entries by definition, we have:

$$a_{ii}^{-1}(Ax)_i = a_{ii}^{-1}\sum_j a_{ij}x_j = x_i + a_{ii}^{-1} \sum_{j\in N_G^+(i)} a_{ij}x_j \ , \tag{3.4}$$

where the last equality is by the fact that A represents G. As R_i knows $\{x_j : j \in N_G^+(i)\}$, this allows R_i to recover x_i. Therefore, indeed $\ell(G) \leq \lceil \log_2 |S| \rceil$.

To conclude the proof, note that in case $\mathbb{F}$ is finite, we have $|S| \leq |\mathbb{F}|^{\mathrm{rank}_{\mathbb{F}}(A)}$, implying (3.3). Furthermore, in this case it is possible to use a linear code utilizing the same word-length. The sender transmits a binary-encoding of the inner-products $(u_1 \cdot x, \ldots, u_r \cdot x) \in \mathbb{F}^r$, where $\{u_1, \ldots, u_r\}$ is a basis for the rows of A over $\mathbb{F}$. ■

Remark 3.2.1: As proved for the case $\mathbb{F} = GF(2)$ in [23], it is possible to show that the above bound is tight for the case of linear codes over $\mathbb{F}$. That is, the length of an optimal linear index code over a finite field $\mathbb{F}$ is $\lceil \mathrm{minrk}_{\mathbb{F}} \log_2 |\mathbb{F}| \rceil$.

We now turn to the main ingredient in the proof of Theorem 3.1.1. Here and in what follows, all logarithms are in the natural base unless stated otherwise.

Proposition 3.2.2. *Let $\mathbb{F}$ and $\mathbb{K}$ denote two finite fields with distinct characteristics. There is an explicit construction of a family of graphs $G = G(n)$ on n vertices, so that*

$$\mathrm{minrk}_{\mathbb{F}}(G) \leq \exp\left(\sqrt{(2+o(1))\log n \log\log n}\right) = n^{o(1)} \ , \tag{3.5}$$

and yet:

$$\mathrm{minrk}_{\mathbb{K}}(G) \geq n/\exp\left(\sqrt{(2+o(1))\log n \log\log n}\right) = n^{1-o(1)} \ , \tag{3.6}$$

where the $o(1)$-terms tend to 0 as $n \to \infty$.

Proof. We first consider the case $\mathbb{F} = GF(p)$ and $\mathbb{K} = GF(q)$ for distinct primes p and q. Let $\varepsilon > 0$, and let k denote a (large) integer satisfying[2]

$$q^l < p^k < (1+\varepsilon)q^l \ , \text{ where } l = \lfloor k \log_q p \rfloor \ . \tag{3.7}$$

Define:

$$s = p^k q^l - 1 \quad\text{and}\quad r = p^{3k} \ . \tag{3.8}$$

The graph G on $n = \binom{r}{s}$ vertices[3] is defined as follows. Its vertices are all s-element subsets of $[r]$, and two vertices are adjacent iff their corresponding sets have an intersection whose cardinality is congruent to -1 modulo p^k:

$$V(G) = \binom{[r]}{s} \ , \ E(G) = \left\{(X,Y) \in V^2 \ : \ X \neq Y \ , |X \cap Y| \equiv -1 \pmod{p^k}\right\} \ . \tag{3.9}$$

For some integer d to be determined later, define the *inclusion matrix* M_d to be the $\binom{r}{s} \times \binom{r}{d}$ binary matrix, indexed by all s-element and d-element subsets of $[r]$, where $(M_d)_{A,B} = 1$ iff $B \subset A$, for all $A \in \binom{[r]}{s}$ and $B \in \binom{[r]}{d}$.

[2] It is easy to verify that there are infinitely many such integers k, as p, q are distinct primes, and hence the set $\{k \log_q p \pmod 1\}_{k\in\mathbb{N}}$ is dense in $[0,1]$.

[3] By well known properties of the density of prime numbers, and standard graph theoretic arguments, proving the assertion of the proposition for these values of n in fact implies the result for any n.

Notice that the $n \times n$ matrix $M_d(M_d)^T$ satisfies the following for all $A, B \in V$ (not necessarily distinct):

$$(M_d(M_d)^T)_{A,B} = \left| \left\{ X \in \binom{[r]}{d} : X \subset (A \cap B) \right\} \right| = \binom{|A \cap B|}{d} . \tag{3.10}$$

Define $P = M_{p^k-1}(M_{p^k-1})^T$ and $Q = M_{q^l-1}(M_{q^l-1})^T$. We claim that P represents G over $GF(p)$ whereas Q represents $\overline{G}$ over $GF(q)$. To see this, we need the following simple observation, which is a special case of Lucas's Theorem (cf., e.g., [31]) on congruencies of binomial coefficients. It was used, for instance, in [22] for constructing low-degree representations of OR functions modulo composite numbers, as well as in [56].

Observation 3.2.3. *For every prime p and integers i, j, e with $i < p^e$, $\binom{j+p^e}{i} \equiv \binom{j}{i} \pmod{p}$.*

Consider some $A \in V$; since s satisfies both $s \equiv (p^k - 1) \pmod{p^k}$ and $s \equiv (q^l - 1) \pmod{q^l}$, combining (3.10) with Observation 3.2.3 gives

$$P_{A,A} = \binom{s}{p^k - 1} \equiv 1 \pmod{p} , \quad \text{and} \quad Q_{A,A} = \binom{s}{q^l - 1} \equiv 1 \pmod{q} .$$

Thus, indeed the diagonal entries of P and Q are non-zero; it remains to show that their (A, B)-entries are 0 wherever A, B are distinct non-adjacent vertices. To this end, take $A, B \in V$ so that $A \neq B$ and $AB \notin E(G)$; by (3.9), $|A \cap B| \not\equiv -1 \pmod{p^k}$, hence

$$P_{A,B} = \binom{|A \cap B|}{p^k - 1} \equiv 0 \pmod{p} ,$$

the last equivalence again following from Observation 3.2.3, as $\binom{x}{p^k-1} = 0$ for all $x \in \{0, \ldots, p^k - 2\}$. Finally, suppose that $A, B \in V$ satisfy $A \neq B$ and $AB \notin E(\overline{G})$. That is, $AB \in E(G)$, hence $|A \cap B| \equiv -1 \pmod{p^k}$. The Chinese Remainder Theorem now implies $|A \cap B| \not\equiv -1 \pmod{q^l}$, otherwise we would get $|A \cap B| = s$ and $A = B$. Since $\binom{x}{q^l-1} = 0$ for all $x \in \{0, \ldots, q^l - 2\}$, we get

$$Q_{A,B} = \binom{|A \cap B|}{q^l - 1} \equiv 0 \pmod{q} .$$

Altogether, P represents G over $GF(p)$, and Q represents $\overline{G}$ over $GF(q)$. Therefore, $\mathrm{minrk}_p(G)$ is at most $\mathrm{rank}_p(P) \leq \mathrm{rank}_p(M_{p^k-1})$, and similarly, $\mathrm{minrk}_q(\overline{G})$ is at most $\mathrm{rank}_q(Q) \leq \mathrm{rank}_q(M_{q^l-1})$. As M_d has $\binom{r}{d}$ columns, $n = \binom{p^{3k}}{p^k q^l}$ and $q^l < p^k < (1+\varepsilon)q^l$, a straightforward calculation now gives:

$$\mathrm{minrk}_p(G) \leq \binom{r}{p^k - 1} < \exp\left(\sqrt{(1+\varepsilon+o(1))2\log n \log\log n}\right) ,$$
$$\mathrm{minrk}_q(\overline{G}) \leq \binom{r}{q^l - 1} < \exp\left(\sqrt{(1+\varepsilon+o(1))2\log n \log\log n}\right) .$$

The next simple claim relates $\mathrm{minrk}_q(G)$ and $\mathrm{minrk}_q(\overline{G})$:

Claim 3.2.4. *For any graph G on n vertices and any field $\mathbb{F}$, $\mathrm{minrk}_{\mathbb{F}}(G) \cdot \mathrm{minrk}_{\mathbb{F}}(\overline{G}) \geq n$.*

Proof. We use the following definition of graph product due to Shannon [97]: $G_1 \times G_2$, the *strong graph product* of G_1 and G_2, is the graph whose vertex set is $V(G_1) \times V(G_2)$, where two distinct vertices $(u_1, u_2) \neq (v_1, v_2)$ are adjacent iff for all $i \in \{1, 2\}$, either $u_i = v_i$ or $(u_i, v_i) \in E(G_i)$.

As observed by Haemers [63], if A_1 and A_2 represent G_1 and G_2 respectively over $\mathbb{F}$, then the tensor product $A_1 \otimes A_2$ represents $G_1 \times G_2$ over $\mathbb{F}$. To see this, notice that the diagonal of $A_1 \otimes A_2$ does not contain zero entries, and that if $(u_1, u_2) \neq (v_1, v_2)$ are disconnected vertices in $G_1 \times G_2$, then by definition $(A_1)_{(u_1,v_1)}(A_2)_{(u_2,v_2)} = 0$, since in this case for some $i \in \{1, 2\}$ we have $u_i \neq v_i$ and $u_i v_i \notin E(G_i)$. Letting A_1 and A_2 denote matrices which attain $\mathrm{minrk}_{\mathbb{F}}(G)$ and $\mathrm{minrk}_{\mathbb{F}}(\overline{G})$ respectively, the above discussion implies that:

$$\mathrm{minrk}_{\mathbb{F}}(G \times \overline{G}) \leq \mathrm{rank}(A_1 \otimes A_2) = \mathrm{minrk}_{\mathbb{F}}(G) \cdot \mathrm{minrk}_{\mathbb{F}}(\overline{G}) .$$

However, the set $\{(u, u) : u \in V(G)\}$ is an independent-set of $G \times \overline{G}$, since for $u \neq v$, either $uv \in E(G)$ and $uv \notin E(\overline{G})$ or vice versa. Therefore, (3.2) gives $\mathrm{minrk}_{\mathbb{F}}(G \times \overline{G}) \geq \alpha(G \times \overline{G}) \geq n$, completing the proof of the claim. ■

This concludes the proof of the proposition for the case $\mathbb{F} = GF(p)$, $\mathbb{K} = GF(q)$, where p, q are two distinct primes. The generalization to the case of prime-powers is an immediate consequence of the next claim, whose proof is given in Section 3.4.

Claim 3.2.5. *Let G be a graph, p be a prime and k be an integer. The following holds:*

$$\frac{1}{k}\operatorname{minrk}_p(G) \leq \operatorname{minrk}_{p^k}(G) \leq \operatorname{minrk}_p(G)\ . \tag{3.11}$$

This concludes the proof of Proposition 3.2.2. ■

Remark 3.2.6: Alon's Ramsey construction [8] is the graph on the vertex set $V = \binom{[r]}{s}$, where $r = p^3$ and $s = pq - 1$ for some large primes $p \sim q$, and two distinct vertices A, B are adjacent iff $|A \cap B| \equiv -1 \pmod{p}$. Our construction allows p and q to be large prime-powers $p^k \sim q^l$. Note that the original construction by Frankl and Wilson [56] had the parameters $r = q^3$ and $s = q^2 - 1$ for some prime-power q, and two distinct vertices A and B are adjacent iff $|A \cap B| \equiv -1 \pmod{q}$.

Proof of Theorem 3.1.1. In order to derive Theorem 3.1.1 from Theorem 3.1.2 and Proposition 3.2.2, apply Proposition 3.2.2, setting $\mathbb{F} = GF(p)$ and $\mathbb{K} = GF(2)$, where $p > 2$ is any fixed (odd) prime. Let $\varepsilon > 0$; for any sufficiently large n, the graph obtained above satisfies

$$\begin{aligned}&\operatorname{minrk}_2(G) \geq n/\exp(O(\sqrt{\log n \log\log n})) \geq n^{1-\varepsilon}\ , \qquad \text{and yet}\\ &\ell(G) \leq \lceil \operatorname{minrk}_p(G)\log_2(p)\rceil \leq \exp(O(\sqrt{\log n \log\log n})) \leq n^{\varepsilon}\ .\end{aligned}$$ ■

3.3 The problem definition revisited

Call the problem of finding the optimal index code, as defined in Definition 3.1, **Problem 3.1**. At first glance, Problem 3.1. seems to capture only very restricted instances of the source coding problem for ISCOD, and its motivating applications in communication. Namely, the main restrictions are:

(1) Each receiver requests exactly one data block.

(2) Each data block is requested only once.

(3) Every data block consists of a single bit.

In [27], where Definition 3.1 was stated, it is proved that the source coding problem for ISCOD can be reduced to a similar one which satisfies restriction (1). This is achieved by replacing a user that requests $k > 1$ blocks by k users, all having the same side information, and each requesting a different block. On the other hand, restriction (2) appeared in [27] to simplify the problem and to enable the side-information to be modeled by a directed graph.[4] Restriction (3) is stated assuming a larger block size does not dramatically effect the nature of the problem. In what follows, we aim to reconsider the last two restrictions.

3.3.1 Shared requests

Problem 3.2: The generalization of Problem 3.1 to $m \geq n$ receivers, each interested in a single bit (i.e., we allow several users to ask for the same bit).

In this case, the one-to-one correspondence between message bits and receivers no longer holds, thus the directed side information graph seems unsuitable. However, it is still possible to obtain bounds on the optimal linear and non-linear codes using slightly different models.

Let $\mathcal{P}_2$ denote an instance of Problem 3.2, and let $\ell(\mathcal{P}_2)$ denote the length of an optimal index code in this setting. It is convenient to model the side-information of $\mathcal{P}_2$ using a binary $m \times n$ matrix, where the ij entry is 1 iff the i-th user knows the j-th bit (if $m = n$, this matrix is the adjacency matrix of the side information graph). With this in mind, we extend the notion of representing the side-information graph as follows: an $m \times n$ matrix B *represents* $\mathcal{P}_2$ over $\mathbb{F}$ iff for all i and j:

- If the i-th receiver is interested in the bit x_j, then $B_{ij} \neq 0$.
- If the i-th receiver is neither interested in nor knows the bit x_j, then $B_{ij} = 0$.

Notice that in the special case $m = n$, the above definition coincides with the usual definition of representing the side-information graph. Let $\mathrm{minrk}_{\mathbb{F}}(\mathcal{P}_2)$

[4] It followed the observation that if the same block is requested by several receivers, then most of the communication saving comes from transmitting this block once (*duplicate elimination*).

denote the minimum rank of a matrix B that represents $\mathcal{P}_2$ over $\mathbb{F}$. It is straightforward to verify that a result analogous to Theorem 3.1.2 and Remark 3.2.1 holds for the extended notion of matrix representation:

Theorem 3.3.1. *Let $\mathcal{P}_2$ denote an instance of Problem 3.2. Then the length of an optimal linear code is* $\operatorname{minrk}_2(\mathcal{P}_2)$*, and the upper bounds of Theorem 3.1.2 on arbitrary index codes hold for $\mathcal{P}_2$ as well.*

Next, given $\mathcal{P}_2$, define the following two directed m-vertex graphs G_{ind} and G_{cl}. Both vertex sets correspond to the m users, where each set of users interested in the same bit forms an independent set in G_{ind} and a clique in G_{cl}. In the remaining cases, in both graphs (v_i, v_j) is an edge iff the i-th user knows the bit in which the j-th user is interested (for $m = n$, both graphs are equal to the usual side-information graph defined in Definition 3.2). The following simple claim provides additional bounds on $\ell(\mathcal{P}_2)$; we omit the details of its proof.

Claim 3.3.2. *If $\mathcal{P}_2$ denotes an instance of Problem 3.2, and G_{ind} and G_{cl} are defined as above, then:*

1. $\ell(G_{\mathrm{cl}}) \leq \ell(\mathcal{P}_2)$*, and in addition,* $\operatorname{minrk}_{\mathbb{F}}(G_{\mathrm{cl}}) \leq \operatorname{minrk}_{\mathbb{F}}(\mathcal{P}_2)$ *for all* $\mathbb{F}$.
2. $\ell(\mathcal{P}_2) \leq \ell(G_{\mathrm{ind}})$*, and in addition,* $\operatorname{minrk}_{\mathbb{F}}(\mathcal{P}_2) \leq \operatorname{minrk}_{\mathbb{F}}(G_{\mathrm{ind}})$ *for all* $\mathbb{F}$.

3.3.2 Larger alphabet and multiple rounds

Suppose the data string x is over a possibly larger alphabet, e.g., $\{0,1\}^b$ for some $b \geq 1$:

Problem 3.3: The generalization of Problem 3.1, where each input symbol $x_i \in \{0,1\}^b$ comprises a *block* of b bits. Every user is interested in a single block, and knows a subset of the other blocks.

It is not difficult to see that this case can be reduced to Problem 3.1 by considering the graph $G[b]$, defined as follows. For some integer b, let $G[b]$ denote the *b-blow-up* of G (with independent sets), that is, the graph on the vertex set $V(G) \times [b]$, where (u, i) and (v, j) are adjacent iff $uv \in E(G)$.

Indeed, Problem 3.3 reduces to Problem 3.1 with side information graph $G[b]$, by assigning a receiver to each of the data bits. Therefore, this extension is in fact a special case of the original seemingly restricted problem. It is not difficult to verify that $\ell(G[b]) \leq b \cdot \ell(G)$; the next remark shows this bound is sometimes tight:

Remark 3.3.3: If an undirected graph G satisfies $\ell(G) = \alpha(G)$ (this holds, e.g., for all graphs satisfying $\alpha(G) = \chi(\overline{G})$, and namely for perfect graphs), then $\ell(G[b]) = b \cdot \ell(G)$, as

$$b \cdot \ell(G) \geq \ell(G[b]) \geq \alpha(G[b]) = b \cdot \alpha(G) = b \cdot \ell(G) \ .$$

It seems plausible that there are graphs for which $\ell(G[b]) < b \cdot \ell(G)$. That is, transmission of a block may strictly improve upon independent transmissions. In this case, it would be interesting to study the “rate” of an index code defined by $\lim_{b\to\infty} \frac{\ell(G[b])}{b}$ (the limit exists by sub-additivity).

Another interesting extension of the problem is the scenario of multiple rounds:

Problem 3.4: The generalization of Problem 3.1 to $t \geq 1$ rounds: the sender wishes to transmit t words $x^1, \ldots, x^t \in \{0,1\}^n$, with respective side information graphs $G_1, \ldots, G_t$. Receiver R_i is always interested in the i-th bit of the input words, $x_i^1, \ldots, x_i^t$.

If the side information graph is constant (i.e., $G_i = G_1$ for all i), then Problem 3.4 can (again) be reduced to Problem 3.1 by an independent set blow-up of the side information graph. Hence, in this case it also coincides with Problem 3, where the block size b corresponds to the number of rounds t. Nevertheless, even for general graphs $\{G_i\}$, a reduction to Problem 3.1 is possible: let $G = G_1 \circ \cdots \circ G_t$ denote the directed graph on the vertex set $V(G) = [n] \times [t]$, where for all $i_1, i_2 \in [n]$ and $k_1, k_2 \in [t]$, $((i_1, k_1), (i_2, k_2))$ is an edge of G iff $(i_1, i_2) \in E(G_{k_2})$. Again, it is straightforward to see that $\ell(G)$ is precisely the solution for Problem 3.4.

Interestingly, in this case there are series of graphs such that independent transmissions consume significantly more communication. For instance, for every n there are n-vertex graphs G_1 and G_2 such that $\ell(G_1) + \ell(G_2) = 2n$, whereas $\ell(G_1 \circ G_2) = n + 1$. The example is given in Section 3.4.

3.4 Proofs for remaining results

3.4.1 Proof of Corollary 3.1.3

Let $\varepsilon > 0$, and let G be the graph constructed by Proposition 3.2.2 for $\mathbb{F} = GF(2)$, $\mathbb{K} = GF(3)$, and a sufficiently large n such that $\mathrm{minrk}_2(G) \leq n^{\varepsilon/2}$ and $\mathrm{minrk}_3(\overline{G}) \leq n^{\varepsilon/2}$. Let H denote the graph $G + \overline{G}$, that is, the disjoint union of G and its complement. We claim that

$$\ell(H) < 3n^{\varepsilon/2} \ , \text{ and yet } \ c(H) \geq \sqrt{2n} = \sqrt{|V(H)|} \ .$$

To see this, observe that in order to obtain an index code for a given graph, one may always arbitrarily partition the graph into subgraphs and concatenate their individual index codes:

Observation 3.4.1. *For any graph G and any partition of G to subgraphs $G_1, \ldots, G_r$ (that is, G_i is an induced subgraph of G on some V_i, and $V = \cup_i V_i$), we have $\ell(G) \leq \sum_{i=1}^{r} \ell(G_i)$.*

In particular, in our case, by combining the above with Theorem 3.1.2, we have

$$\ell(H) \leq \ell(G) + \ell(\overline{G}) \leq n^{\varepsilon/2} + \lceil \log_2 3n^{\varepsilon/2} \rceil < 3n^{\varepsilon/2}.$$

Finally, label the vertices of G as $\{v_1, \ldots, v_n\}$ and the corresponding vertices of $\overline{G}$ as $\{v'_1, \ldots, v'_n\}$. Following the arguments of the proof of Claim 3.2.4, it is easy to verify that the set of vertices $\{(v_i, v'_i) : i \in [n]\} \cup \{(v'_i, v_i) : i \in [n]\}$ is an independent set of size $2n$ in $G \times \overline{G} + \overline{G} \times G$, which is an induced subgraph of $H \times H$. Therefore, $c(H) \geq \sqrt{2n}$. ■

Remark 3.4.2: A standard argument gives a slight improvement in the above lower bound on $c(H)$, to $c(H) \geq 2\sqrt{n}$. See, e.g., [8] (proof of Theorem 2.1) for further details.

3.4.2 Proof of Proposition 3.1.4

We need the following definition:

Definition 3.3 (Lexicographic graph product)**.** *The* Lexicographic graph product *of G and H, denoted by $G \cdot H$, is the graph whose vertex set is $V(G) \times V(H)$, where (u_1, v_1) is adjacent to (u_2, v_2) iff either $u_1u_2 \in E(G)$, or $u_1 = u_2$ and $v_1v_2 \in E(H)$.*

The above product is associative, giving meaning to $G_1 \cdot \ldots \cdot G_k$, where two distinct k-tuples are adjacent iff there is an edge in the graph corresponding to the first coordinate where they differ. As noted by Abbott [1], this product satisfies $\alpha(G \cdot H) = \alpha(G)\alpha(H)$ and $\omega(G \cdot H) = \omega(G)\omega(H)$ for any two graphs G and H. This suggests that this product may be used to construct large Ramsey graphs as a high graph power of a small (fixed) Ramsey graph (cf., e.g., [89]).

Notice that $\mathrm{minrk}_{p^e}(G) \le \mathrm{minrk}_{p^d}(G)$ for any prime p and integers $e > d$. Therefore, we can assume without loss of generality that all the $\mathbb{F}_i$-s are fields with pairwise distinct characteristics. Let G_i denote the graph obtained by applying Proposition 3.2.2 on $\mathbb{K}$ and $\mathbb{F}_i$, so that:

$$\mathrm{minrk}_{\mathbb{K}}(G_i) \le n^{\varepsilon} \text{ and } \mathrm{minrk}_{\mathbb{F}_i}(G_i) \ge n^{1-\varepsilon} ,$$

and let $G = G_1 \cdot \ldots \cdot G_t$. The required result will now follow from the next simple claim:

Claim 3.4.3. *The following holds for any two graphs G and H and field $\mathbb{F}$:*

$$\max\{\mathrm{minrk}_{\mathbb{F}}(G), \mathrm{minrk}_{\mathbb{F}}(H)\} \le \mathrm{minrk}_{\mathbb{F}}(G \cdot H) \le \mathrm{minrk}_{\mathbb{F}}(G)\,\mathrm{minrk}_{\mathbb{F}}(H) . \tag{3.12}$$

Proof. For the left inequality, observe that $G \cdot H$ contains an induced subgraph isomorphic to G, and an induced subgraph isomorphic to H. To see this, consider $\{(u, w) : u \in V(G)\}$ for some fixed vertex $w \in V(H)$, and then $\{(w, u) : u \in V(H)\}$ for some fixed $w \in V(G)$.

For the right inequality, suppose $|V(G)| = n$ and $|V(H)| = m$, and let A_G and A_H denote the adjacency matrices of G and H respectively. We claim that, if B_G and B_H are matrices which attain $\mathrm{minrk}_{\mathbb{F}}(G)$ and $\mathrm{minrk}_{\mathbb{F}}(H)$ respectively, then $B = B_G \otimes B_H$ represents $G \cdot H$ over $\mathbb{F}$. To see this, first note that for all $(u, v) \in V(G \cdot H)$ we have

$$B_{(u,v),(u,v)} = (B_G)_{u,u}(B_H)_{v,v} \neq 0 .$$

Second, suppose that $(u_1, v_1) \neq (u_2, v_2)$ are two non-adjacent vertices of $G \cdot H$. If $u_1 \neq u_2$, then by the definition of $G \cdot H$, $u_1 u_2 \notin E(G)$, hence $(B_G)_{u_1,u_2} = 0$. Otherwise, $v_1 \neq v_2$, and again, by the definition of $G \cdot H$, $v_1 v_2 \notin E(H)$, giving $(B_H)_{v_1,v_2} = 0$. Altogether, we deduce that in this case

$$B_{(u_1,v_1),(u_2,v_2)} = (B_G)_{u_1,u_2} (B_H)_{v_1,v_2} = 0 \ ,$$

and $\mathrm{rank}(B) = \mathrm{rank}(B_G)\,\mathrm{rank}(B_H)$, as required. ■

This completes the proof of Proposition 3.1.4. ■

Remark 3.4.4: The proof of Proposition 3.1.4 in fact holds even when $t = o(\sqrt{\frac{\log n}{\log\log n}})$.

3.4.3 Proof of Claim 3.2.5

The statement $\mathrm{minrk}_{p^k}(G) \leq \mathrm{minrk}_p(G)$ follows immediately from the fact that any matrix A which represents G over $GF(p)$ also represents G over $GF(p^k)$, and in addition satisfies $\mathrm{rank}_{p^k}(A) \leq \mathrm{rank}_p(A)$.

To show that $\mathrm{minrk}_p(G) \leq k\,\mathrm{minrk}_{p^k}(G)$, let $V = [n]$ denote the vertex set of G, and let $A = (a_{ij})$ denote a matrix which represents G over $GF(p^k)$ with rank $r = \mathrm{minrk}_{p^k}(G)$. As usual, we represent the elements of $GF(p^k)$ as polynomials of degree at most $k-1$ over $GF(p)$ in the variable x. Since the result of multiplying each row of A by a non-zero element of $GF(p^k)$ is a matrix of rank r which also represents G over $GF(p^k)$, assume without loss of generality that $a_{ii} = 1$ for all $i \in [n]$. By this assumption, the $n \times n$ matrix $B = (b_{ij})$, which contains the free coefficients of the polynomials in A, represents G over $GF(p)$. To complete the proof, we claim that $\mathrm{rank}_p(B) \leq kr$. This follows from the simple fact that, if $\{u_1, \ldots, u_r\}$ is a basis for the rows of A over $GF(p^k)$, then the set $\bigcup_{i=1}^{r} \{u_i, x \cdot u_i, \ldots, x^{k-1} \cdot u_i\}$ spans the rows of A when viewed as kn-dimensional vectors over $GF(p)$. ■

3.4.4 The parameters $\mathrm{minrk}_p(G)$ and $\mathrm{minrk}_p(G[k])$

Claim 3.4.5. *Let G be a graph, p be a prime and k be an integer. The following holds:*

$$\mathrm{minrk}_p(G) \leq \mathrm{minrk}_p(G[k]) \leq k\,\mathrm{minrk}_{p^k}(G) \ . \tag{3.13}$$

Proof. The inequality $\text{minrk}_p(G) \le \text{minrk}_p(G[k])$ follows from the fact that G is an induced subgraph of $G[k]$ (for instance, consider the set of vertices $\{(u,1) : u \in V\}$), hence any matrix which represents $G[k]$ over $GF(p)$ contains a principal submatrix which represents G over $GF(p)$.

It remains to prove that $\text{minrk}_p(G[k]) \le k\,\text{minrk}_{p^k}(G)$. Set $V = [n]$, and let $A = (a_{ij})$ be a matrix which represents G over $GF(p^k)$, whose rank over $GF(p^k)$ is $r = \text{minrk}_{p^k}(G)$. As usual, let us represent the elements of $GF(p^k)$ as polynomials of degree at most $k-1$ over $GF(p)$ in the variable x. As before, assume without loss of generality that $a_{ii} = 1$ for all $i \in [n]$. Next, replace each row $A_i = (a_{i1} \ldots a_{in})$ of A by k rows $\{A_i, x \cdot A_i, \ldots, x^{k-1} \cdot A_i\}$. Treating each element of $GF(p^k)$ as a k-tuple over $GF(p)$, translate the above $kn \times n$ matrix over $GF(p^k)$ to a $kn \times kn$ matrix over $GF(p)$, B, with all the diagonal entries equaling 1. Clearly, the matrix B represents $G[k]$, and furthermore, having included the multiples of each row of A by $1, x, \ldots, x^{k-1}$, it follows that the rank of B over $GF(p)$ is at most kr, as required. ■

3.4.5 The parameters $\text{minrk}(G)$ and $\vartheta(G)$

Consider the n-vertex graph G constructed in Proposition 3.2.2 for $\mathbb{F} = GF(p)$ and $\mathbb{K} = GF(q)$, where p and q are two distinct primes: it satisfies $\text{minrk}_p(G) \le n^{o(1)}$ and $\text{minrk}_q(\overline{G}) \le n^{o(1)}$. Clearly, G is vertex transitive (that is, its automorphism group is closed under all vertex substitutions), as we can always relabel the elements of the ground set $[r]$. By [81] (Theorem 9), every vertex transitive graph G on n vertices satisfies $\vartheta(G)\vartheta(\overline{G}) = n$.

Assume without loss of generality that $\vartheta(G) \ge \sqrt{n} \ge \vartheta(\overline{G})$ (otherwise, switch the roles of p and q and of G and $\overline{G}$). As $\text{minrk}_p(G) \le n^{o(1)}$ and $\text{minrk}_p(\overline{G}) \ge n^{1-o(1)}$, we deduce that

$$\vartheta(G) \ge n^{\frac{1}{2}-o(1)} \cdot \text{minrk}_p(G) \quad \text{and yet} \quad \text{minrk}_p(\overline{G}) \ge n^{\frac{1}{2}-o(1)} \cdot \vartheta(\overline{G}) \ .$$

3.4.6 Example for the benefit of multiple-round index-coding

As a warmup, consider the following case. We have two receivers, R_1 and R_2, and two rounds for transmitting the binary words $x = x_1x_2$ and $y = y_1y_2$.

Suppose that in the first round receiver R_1 knows x_2 and in the second transmission receiver R_2 knows y_1. In this case, each round - if transmitted separately - requires 2 bits to be transmitted. Yet, if the server transmits the 3 bits

$$x_1 \oplus y_1 \ , \ x_2 \oplus y_2 \ , \ x_1 \oplus y_2 \ ,$$

then both receivers can reconstruct their missing bits (and moreover, reconstruct all of x and y).

This in fact is a special case of the following construction. We define a pair of graphs G_1, G_2 such that $\ell(G_1) = \ell(G_2) = n$ and yet only $\ell(G_1 \circ G_2) = n+1$ bits need to transmitted for consecutive transmissions. This is stated in the next claim, where the *transitive tournament graph* on n vertices is isomorphic to the directed graph on the vertex set $[n]$, where (i, j) is an edge iff $i < j$.

Claim 3.4.6. *Let G_1 denote the transitive tournament graph on n vertices, and let G_2 denote the graph obtained from G_1 by reversing all edges. Then $\ell(G_1) + \ell(G_2) = 2n$, and yet $\ell(G_1 \circ G_2) = n + 1$.*

Proof. Without loss of generality, assume that $E(G_1) = \{(i, j) : i < j\}$ and $E(G_2) = \{(i, j) : i > j\}$. Since G_1 and G_2 are both acyclic, the fact that $\ell(G_1) = \ell(G_2) = n$ follows from the lower bound of [23] ($\ell(G)$ is always at least the size a maximum induced acyclic subgraph of G).

Recall that by definition, $G_1 \circ G_2$ is the disjoint union of G_1 and G_2, with the additional edges $\{((i, 1), (j, 2)) : j < i\}$ and $\{((i, 2), (j, 1)) : j > i\}$. Therefore, $G_1 \circ G_2$ has an induced acyclic graph of size $n + 1$: for instance, the set $\{(i, 1) : i \in [n]\} \cup \{(n, 2)\}$ induces such a graph. We deduce that $\ell(G_1 \circ G_2) \geq n + 1$.

To complete the proof of the claim, we give an encoding scheme for $G_1 \circ G_2$ which requires the transmission of $n + 1$ bits, hence $\ell(G_1 \circ G_2) \leq n + 1$. Denote the two words to be transmitted by $x = x_1 \ldots x_n$ and $y = y_1 \ldots y_n$. The coding scheme is linear: by transmitting $x_i \oplus y_i$ for $i \in [n]$ and $\oplus_{i \in [n]} x_i$, it is not difficult to see that each receiver is able to decode its missing bits (in fact, each receiver can reconstruct all the bits of x and y). ■

3.5 Concluding remarks and open problems

In this chapter we introduced constructions of graphs for which linear index coding is suboptimal (Theorem 3.1.1), thus disproving the main conjecture of [23]. The new non-linear codes justified a new upper bound (Theorem 3.1.2) on the minimal length of the code $\ell(G)$, in which linear codes over arbitrary fields are considered.

On the other hand, the best current lower bound on $\ell(G)$ (due to [23]) is known not to be tight. Hence, the main question for further work is trying to close the gap between the bounds on $\ell(G)$.

In addition, we showed that more general scenarios of index coding, as presented in [27], can be reduced to the main problem, which recently attracted attention. In this context, one may save on communication when transmitting t binary words at once, rather than transmitting these words independently. Following the discussion in Section 3.3, it may be interesting to study an appropriate definition of "index code rate" of a given side-information graph.

Chapter 4

Index coding and Witsenhausen-type coloring problems

The results of this chapter appear in [16]

The problem of Informed Source Coding on Demand, introduced by Birk and Kol [27], describes a setting where a server wishes to transmit n data blocks (of t bits each) via broadcast to n receivers; each receiver is interested in a specific block of the input data and may have side information on other blocks. The goal of the sender is to use a code of minimal word-length, while allowing every receiver to recover his desired block. This problem can be formulated as a graph parameter: let G be a directed graph on the vertex set $[n]$, where ij is an edge iff the i-th receiver knows the j-th block, and let $\beta_t(G)$ denote the length of a minimal binary code for G when the blocks are of size t.

The above problem was studied in the case where each block consists of a single bit (namely $t = 1$, see [23], [27] and [83]). In this chapter, we consider the general, and more natural, case. We provide general bounds on $\beta_t(G)$ and show that in some cases usage of large data blocks may strictly improve upon the trivial extension from the binary case, thus answering a question of [83]. This motivates the study of a new graph parameter, *the broadcast rate* of a graph G, defined by $\beta(G) = \lim_{t\to\infty} \beta_t(G)/t$. En route, we show that -

surprisingly - an optimal code for a disjoint union of graphs can be strictly better than a concatenation of the optimal codes for the individual graphs, even when each of these graphs is a copy of C_5.

The proofs are based on a relation between this problem and some results in the study of Witsenhausen's rate, OR graph products, colorings of graph powers and some properties of Cayley graphs.

4.1 Introduction

4.1.1 Background and definitions

Source coding deals with a scenario in which a *sender* has some data string x he wishes to transmit through a broadcast channel to *receivers*. In this chapter we consider a variant of source coding which was first proposed by Birk and Kol [27]. In this variant, called Informed Source Coding On Demand (ISCOD), each receiver has some prior side information, comprising some subset of the input string x. The sender is aware of the portion of x known to each receiver. Moreover, each receiver is interested in just part of the data. Following Bar-Yossef, Birk, Jayram and Kol [23], we use the formalization of **Index code** problem as given in Definition 3.1.

For an example of the applications which motivate the study of this problem, consider a central server and a collection of caching clients in a satellite transmission network. Each of the clients has limited storage, and a slow backward channel to the server. The server holds a data string, comprising a large number of blocks, whereas each client can only store a relatively small number of data blocks at any given point. During the transmission of the data, the clients opt to hold certain data blocks, and subsequently receive the actual user request for specific data blocks. The backward channel can be used to notify the server which blocks are present at the client end, and which are required. Finally, given the state of all clients (namely, the side information of each client), the server must retransmit an encoding of its data, in the most economical possible way, which would allow each client to recover its required blocks (from the new transmission and its prior side information).

Following [23], the problem in Definition 3.1 can be restated as a graph parameter, as described in Definition 3.2, by modeling the side information state via a directed graph on n vertices. We next generalize this notion to depend on the input block size. To this end, we identify the i-th vertex of the graph both as the i-th receiver, R_i, and as the i-th block of input, x_i, and place an edge (i, j) whenever R_i knows x_j. This is formulated in the next definition, where the minimum possible length of an index code is expressed as a parameter of the corresponding side information graph. Here and in what follows, for a directed graph G and a vertex v, let $N_G^+(v)$ be the set of out-neighbors of v in G, and for $x = x_1 \dots x_n$ and $S \subset [n] = \{1, \dots, n\}$, let $x|_S$ be the restriction of x to the coordinates of S

Definition 4.1 ($\beta_\mathbf{t}(\mathbf{G})$)**.** *Let G be a directed* side information graph *G on the vertex set $[n]$, where (i, j) is an edge iff R_i knows the value of x_j. An* index code *of length ℓ for G is a function $E : \{0, 1\}^{n \cdot t} \to \{0, 1\}^\ell$ and functions $D_1, \dots, D_n$, so that for all $i \in [n]$ and $x \in \{0, 1\}^{n \cdot t}$, $D_i(E(x), x|_{N_G^+(i)}) = x_i$. We denote the minimum possible length of such a code, for blocks of length t, by $\beta_t(G)$.*

Notice that, according to the notation of Definition 3.2 given in the previous chapter, $\ell(G)$ is the special case of $\beta_t(G)$ where $t = 1$.

4.1.2 Preliminaries

The following two trivial examples exhibit some of the properties of $\beta_t(G)$ which we will later review. Throughout the chapter, an undirected side information graph corresponds to the directed graph, where each (formerly undirected) edge appears in both directions.

Consider, first, the case where $G = K_n$ is the complete graph. In this case, broadcasting the XOR of the n data blocks enables each user to reconstruct its missing block. The length of this code is t. It is also easy to show that this is the optimal coding scheme for this graph, and hence $\beta_t(K_n) = t$.

On the other hand, assume $G = E_n$ is an edgeless graph, thus none of the receivers has any prior side information. A straightforward counting argument implies that $\beta_t(E_n) = t \cdot n$ (one cannot improve upon the naïve protocol of retransmitting the entire input word x).

Before reviewing the basic facts and results of index codes, we define the following graph theoretic parameters. An independent set in $G = (V, E)$ is a set of vertices which induces an edgeless graph. The *chromatic number* of G, $\chi(G)$, is the minimum number of independent sets whose union is all of V (each such set is referred to as a *color class*). Let $\overline{G}$ denote the *graph complement* of G. We denote by $G + H$ the disjoint union of the graphs G and H, and denote by $k \cdot G$ the disjoint union of k copies of the graph G.

An immediate property of the parameter $\beta_t(G)$ is monotonicity with respect to the removal of vertices as well as edges:

(i) If G_2 is obtained from G_1 by removing some of the vertices (i.e., G_2 is an induced subgraph of G_1) then $\beta_t(G_1) \geq \beta_t(G_2)$.

(ii) If G_2 is obtained from G_1 by removing some of the edges then $\beta_t(G_1) \leq \beta_t(G_2)$.

(iii) $\beta_t(G + H) \leq \beta_t(G) + \beta_t(H)$ and hence $\beta_t(k \cdot G) \leq k\beta_t(G)$.

Combining item (i) with the second example, we obtain that

$$\beta_t(G) \geq t \cdot \alpha(G) \ , \tag{4.1}$$

where $\alpha(G)$ is the cardinality of a maximum independent set [1]. In addition, items (ii), (iii) together with the first example show that

$$\beta_t(G) \leq t \cdot \chi(\overline{G}) \ . \tag{4.2}$$

See [27], [23] and [83] for several other examples and properties of index coding schemes and details on the distributed application which motivates this problem.

In all the previous works on index coding, the parameter $\beta_1(G)$ (denoted $\ell(G)$ there) was studied. It was observed in [83] that the parameter $\beta_t(G)$ can always be reduced to a parameter $\beta_1(G')$ by considering a graph G' which is the *t-blow-up* of G (with independent sets). This graph, denoted by $G[t]$, has

[1] A slightly more sophisticated argument is used in [23] to show that for a directed graph G: $\beta_1(G) \geq \mathrm{MAIS}(G)$ where $\mathrm{MAIS}(G)$ denotes the maximum number of vertices in an induced acyclic subgraph of G.

the vertex set $V(G) \times [t]$, and an edge from (u, i) to (v, j) iff $uv \in E(G)$. By definition, an index code for $G[t]$ with block size $t = 1$ is also an index code for G with block size t, and vice versa. To see this, simply split each R_i into t receivers, each interested in recovering a single bit in the i-th block, and notice that each of these new receivers knows all the bits that R_i originally knew.

Another simple fact is that one can always encode each of the t bits in the blocks independently, and hence

$$\beta_t(G) \leq t \cdot \beta_1(G) \ .$$

More generally, the above reduction immediately yields sub-additivity of $\beta_t(G)$ for any graph G, namely $\beta_{t+s}(G) \leq \beta_t(G) + \beta_s(G)$ for any G, s and t.

In various cases such equality holds. For example, whenever G is a perfect graph and hence satisfies $\alpha(G) = \chi(\overline{G})$, it follows that the inequalities (4.1) and (4.2) are in fact equalities, and in particular $\beta_t(G) = t \cdot \beta_1(G) = \beta_1(G[t])$. It is asked in [83] whether, for some graphs, one can benefit from using a unified scheme that encodes the entire blocks at once. In other words, whether for some graph G and some t, $\beta_t(G) < t \cdot \beta_1(G)$.

In this chapter we answer this question in the affirmative. This justifies the following definition:

Definition 4.2 (**Broadcast rate**)**.** *The* broadcast rate *of a graph G is the asymptotic average communication that is required for a single round when allowing a unified transmission for multiple rounds:*

$$\beta(G) := \lim_{t\to\infty} \frac{\beta_t(G)}{t} \ .$$

The above limit exists and equals the infimum by sub-additivity and Fekete's Lemma. Note that the motivation suggests that the quantity $\beta_t(G)$ is of interest mainly for large values of t, leading naturally to the study of the limit $\beta(G)$

4.1.3 New results

Our main technical result in this chapter is Theorem 4.1.1 below, which addresses the size of optimal index codes for a disjoint union of graphs. The case where the side information graph is $k \cdot G$, a disjoint union of k copies of some graph G, describes the setting where there are k separate input words (one for each graph), and receivers corresponding to one graph have no information on any of the input words corresponding to the remaining graphs. The theorem provides an upper bound on the size of the optimal index code in this case, which will demonstrate the counterintuitive behavior that $\beta_1(k \cdot G)$ exhibits.

It will be convenient to address the more precise notion of the *number of codewords* in an index code. We say that $\mathcal{C}$, an index code for G, is *optimal*, if it contains the minimum possible number of codewords (in which case, $\beta_1(G) = \lceil \log_2 |\mathcal{C}| \rceil$). Moreover, let us denote by γ the maximal cardinality of a set of input-strings in $\{0,1\}^n$, which is unconfusable. That is, all the strings can be encoded by the same codeword in an index code for G, and for any pair of strings there is no coordinate in which the strings differ, but the side information of this coordinate is identical in both strings.

Theorem 4.1.1. *Let G be a directed side information graph on n vertices, and let γ be defined as above. The following holds for any integer k:*

$$\left(\frac{2^n}{\gamma}\right)^k \leq |\mathcal{C}| \leq \left\lceil \left(\frac{2^n}{\gamma}\right)^k kn \log 2 \right\rceil \tag{4.3}$$

where $\mathcal{C}$ is an optimal index code for $k \cdot G$. In particular,

$$\lim_{k \to \infty} \frac{\beta_1(k \cdot G)}{k} = n - \log_2 \gamma \ .$$

Index codes for a disjoint union of graphs

Clearly, $\beta_1(k \cdot G) \leq k \cdot \beta_1(G)$, as one can always obtain an index code for $k \cdot G$ by taking the k-fold concatenation of an optimal index code for G. Furthermore, this bound is tight for all perfect graphs. Hence, the smallest graph where $\beta_1(k \cdot G)$ may possibly be smaller than $k \cdot \beta_1(G)$ is C_5, the cycle on 5 vertices - the smallest non-perfect graph. Indeed, in this case index codes for $k \cdot C_5$ can be significantly better than those obtained by treating each copy of C_5 separately. This is stated in the next corollary.

Corollary 4.1.2. *The following holds: The index code for $k \cdot C_5$ comprising index codes of the individual C_5 copies is suboptimal: $\beta_1(C_5) = 3$, whereas $\beta_1(k \cdot C_5)/k = 5 - \log_2 5 + o(1) \approx 2.68 + o(1)$, with the $o(1)$-terms tending to 0 as $k \to \infty$.*

Therefore, there is a graph G with an optimal index code $\mathcal{C}$, so that much less than $|\mathcal{C}|^k$ words suffice to establish an index code for $k \cdot G$, although each of the k copies of G has no side information on any of the bits corresponding to the remaining copies.

Broadcast rate

The main corollary of Theorem 4.1.1 is the following theorem, which provides a general bound on the broadcast rate of a graph.

Theorem 4.1.3. *Let G be a side information graph on n vertices, and let γ be as in Theorem 4.1.1. The following holds:*

$$\alpha(G) \leq \beta(G) \leq n - \log_2 \gamma \ . \tag{4.4}$$

We note that the lower bound on the broadcast rate of a general directed side information graph G is in fact $\beta(G) \geq \text{MAIS}(G)$, where $\text{MAIS}(G) \geq \alpha(G)$ is the maximum number of vertices in an induced acyclic subgraph of G.

The following is a corollary of Theorem 4.1.3 and Corollary 4.1.2. This shows that there exists a graph G for which $\beta(G) < \beta_1(G)$.

Corollary 4.1.4. *For the special case $G = C_5$, we have*

$$2 \leq \beta(C_5) \leq 5 - \log_2 5 \approx 2.678 \ ,$$

whereas $\beta_1(C_5) = 3$.

In addition, we deduce several properties of linear and non-linear index coding schemes. In particular, some of the results of [83] can be reproved in a stronger form using the new construction. The details follow in Section 4.3.3.

4.1.4 Methods and organization

The basic idea behind the proof of the main technical result is an interesting relation between optimal index codes of a disjoint union of graphs, and vertex-colorings of OR graph products. The products are applied to the confusion graph (introduced in [23]), which is an auxiliary graph whose vertex set consists of all possible input data strings. This connection is then translated, using probabilistic bounds on the fractional chromatic number, into a probabilistic construction of index coding schemes. While the best known upper bounds on index coding were all based on explicit linear coding schemes, our new construction is probabilistic and inherently non-linear.

Using this connection, we obtain several rather surprising results, both on the broadcast rate and on index coding of a disjoint union of graphs. These results are achieved by combining the construction with an analysis of the auxiliary graphs.

The rest of this chapter is organized as follows. In Section 4.2 we prove Theorem 4.1.1, using a connection between index codes for disjoint unions of graphs and the chromatic number of OR graph products. Section 4.3 contains applications of this theorem. The first application is for disjoint union of graphs, namely the proof of Corollary 4.1.2. Then the application for the broadcast rate is discussed, proving Theorem 4.1.3 and Corollary 4.1.4. The applications section ends with brief comments on non-linear and linear index coding schemes. Section 4.4 is devoted to concluding remarks and open problems.

4.2 Optimal index codes for a disjoint union of graphs

Throughout this section the length t of the blocks considered is 1.

Proof of Theorem 4.1.1. The OR graph product is equivalent to the complement of the *strong* product[2], which was thoroughly studied in the investi-

[2] Namely, the OR product of G and H is the complement of the strong product of $\overline{G}$ and $\overline{H}$.

gation of the Shannon capacity of a graph, a notoriously challenging graph parameter introduced by Shannon [97].

Definition 4.3 (**OR graph product**)**.** *The* OR graph product *of G and H, denoted by $G \vee H$, is the graph on the vertex set $V(G) \times V(H)$, where (u,v) and (u',v') are adjacent iff either $uu' \in E(G)$ or $vv' \in E(H)$ (or both). Let $G^{\vee k}$ denote the k-fold OR product of a graph G.*

The size of an optimal index code for a given directed graph may be restated as a problem of determining a chromatic number of a graph, as observed by Bar-Yossef et al. [23]. We need the following definition:

Definition 4.4 (**Confusion graph**)**.** *Let $G = ([n], E)$ be a directed side information graph. The* confusion graph *of G, $\mathfrak{C}(G)$, is the undirected graph whose vertex set is $\{0,1\}^n$, and two vertices $x, y \in \{0,1\}^n$ are adjacent iff for some $i \in [n]$, $x_i \neq y_i$ and yet $x|_{N_G^+(i)} = y|_{N_G^+(i)}$.*

In other words, $\mathfrak{C}(G)$ is the graph whose vertex set is all possible input-words, and two vertices are adjacent iff they cannot be encoded by the same codeword in an index code for G (otherwise, the decoding of at least one of the receivers would be ambiguous). Hence, every index code for G is equivalent to a legal vertex coloring of $\mathfrak{C}(G)$, where each color class corresponds to a distinct codeword. Consequently, if $\mathcal{C}$ is an optimal index code for G, then $|\mathcal{C}| = \chi(\mathfrak{C}(G))$.

Let G and H denote directed graphs on the vertex-sets $[m]$ and $[n]$ respectively, and consider an index code for their disjoint union, $G + H$. As there are no edges between G and H, such an index code cannot encode two input-words $x, y \in \{0,1\}^{m+n}$ by the same codeword iff this forms an ambiguity either with respect to G or with respect to H (or both). Hence:

Observation 4.2.1. *For any two directed graphs G and H, the two graphs $\mathfrak{C}(G+H)$ and $\mathfrak{C}(G) \vee \mathfrak{C}(H)$ are isomorphic.*

Thus, the number of codewords in an optimal index code for $k \cdot G$ is equal to $\chi(\mathfrak{C}(G)^{\vee k})$. The chromatic numbers of strong powers of a graph, as well as those of OR graph powers, have been thoroughly studied. In the former case, they correspond to the Witsenhausen rate of a graph (see [106]). In the latter case, the following was proved by McEliece and Posner [86], and also by Berge and Simonovits [25]:

$$\lim_{k\to\infty}\left(\chi(H^{\vee k})\right)^{1/k} = \chi_f(H)\ , \tag{4.5}$$

where $\chi_f(H)$ is the *fractional chromatic number* of the graph H, defined as follows. A legal vertex coloring corresponds to an assignment of $\{0,1\}$-weights to independent-sets, such that every vertex will be "covered" by a total weight of at least 1. A fractional coloring is the relaxation of this problem where the weights belong to $[0,1]$, and χ_f is the minimum possible sum of weights in such a fractional coloring.

To obtain an estimate on the rate of the convergence in (4.5), we will use the following well-known properties of the fractional chromatic number and OR graph products (cf. [18],[80],[75] and also [51]):

(i) For any graph H, $\chi_f(H^{\vee k}) = \chi_f(H)^k$.

(ii) For any graph H, $\chi_f(H) \leq \chi(H) \leq \lceil\ \chi_f(H)\log|V(H)|\ \rceil$. [This is proved by selecting $r = \lceil\ \chi_f(H)\log|V(H)|\ \rceil$ independent sets, chosen randomly and independently according to the weight distribution, dictated by the optimal weight-function achieving χ_f, and by showing that the expected number of uncovered vertices is less than 1.]

(iii) For any vertex transitive graph H (that is, a graph whose automorphism group is transitive), $\chi_f(H) = |V(H)|/\alpha(H)$ (cf., e.g., [60]).

In order to translate (ii) to the statement of (4.3), notice that γ, as defined in Theorem 4.1.1 is precisely $\alpha(\mathfrak{C}(G))$. In addition, the graph $\mathfrak{C}(G)$ is indeed vertex transitive (as it is a Cayley graph of Z_2^n), and combining the above facts we obtain that:

$$\chi_f\left(\mathfrak{C}(G)^{\vee k}\right)^{1/k} = \frac{2^n}{\alpha(\mathfrak{C}(G))} = \frac{2^n}{\gamma}\ .$$

Plugging the above equation into (ii), and recalling that $\chi\left(\mathfrak{C}(G)^{\vee k}\right)$ is the size of the optimal index code for $k\cdot G$, complete the proof of the theorem. ■

Remark 4.2.2: The right-hand-side of (4.3) can be replaced by

$$\left(\frac{2^n}{\gamma}\right)^k \lceil 1 + k \log \gamma \rceil \ .$$

To see this, combine the simple fact that $\alpha(G^{\vee k}) = \alpha(G)^k$ with the bound $\chi(H) \leq \lceil \chi_f(H)(1 + \ln \alpha(H)) \rceil$ given in [80] (which can be proved by choosing $\lceil \chi_f(H) \log \alpha(H) \rceil$ independent sets randomly as before, leaving at most $\lceil \chi_f(H) \rceil$ uncovered vertices, to be covered separately).

4.3 Applications

4.3.1 Index-coding for disjoint unions of graphs

Recall that for any perfect graph $\beta_1(G) = \alpha(G)$. Since the disjoint union of perfect graphs is perfect as well, and its independence number is the sum of the independence numbers of the individual graphs, we conclude that

$$\beta_1(k \cdot G) = k \cdot \beta_1(G) \ \text{ for any perfect graph } G \text{ and integer } k.$$

Therefore, the smallest example where $\beta_1(k \cdot G)$ might be nontrivial is C_5, the smallest non-perfect graph. Indeed, in this case it is possible to do better than $k \cdot \beta_1(C_5)$ in an index code for $k \cdot C_5$:

Proof of Corollary 4.1.2. One can verify that the following is a maximum independent set of size 5 in $\mathfrak{C}(C_5)$:

$$\{00000, 01100, 00011, 11011, 11101\} \ .$$

In the formulation of Theorem 4.1.1, $\gamma = 5$, and the theorem now implies that $\beta_1(k \cdot C_5)/k$ tends to $5 - \log_2 5$ as $k \to \infty$. On the other hand, one can verify[3] that $\chi(\mathfrak{C}(C_5)) = 8$, hence $\beta_1(C_5) = 3$. ■

Remark 4.3.1: Using the upper bound of (4.3) in its alternate form, as stated in Remark 4.2.2, we obtain that $\beta_1(k \cdot C_5) < k \cdot \beta_1(C_5)$ already for $k = 15$.

[3] This fact can be verified by a computer assisted proof, as stated in [23].

4.3.2 Broadcast rate

We now turn to the main application of Theorem 4.1.1, discussing the broadcast rate of a graph G. We first prove the general bound on the broadcast rate of an arbitrary graph.

Proof of Theorem 4.1.3. To show the lower bound $\beta(G) \geq \alpha(G)$, notice that $\alpha(G[t]) = t \cdot \alpha(G)$ for any integer t, hence $\beta_1(G[t]) \geq t \cdot \alpha(G)$ and $\beta(G) \geq \alpha(G)$.

For the upper bound $\beta(G) \leq n - \log_2 \gamma$, notice that $G[t]$ contains the subgraph $t \cdot G$ (for each $i \in [t]$, the graph $V(G) \times \{i\}$ is isomorphic to G). Thus, by the monotonicity of index coding with respect to addition of edges (Item (ii) of the monotonicity property), $\beta_1(G[t]) \leq \beta_1(t \cdot G)$, and Theorem 4.1.1 now provides the required bound. ■

As in the case of a disjoint union of graphs, any perfect graph satisfies $\beta(G) = \beta_1(G) = \alpha(G)$. Therefore, once again, the smallest example where the broadcast rate can be strictly smaller than β_1 is C_5 (the smallest non-perfect graph). Indeed, that proves to be the case.

Proof of Corollary 4.1.4. In the special case $G = C_5$, recalling that $\alpha(G) = 2$ whereas $\gamma = 5$ (as stated in the proof of Corollary 4.1.2) gives:

$$2 \leq \beta(C_5) \leq 5 - \log_2 5 \approx 2.678 ,$$

as required. ■

4.3.3 Linear vs. Non-linear index coding

An index coding scheme is linear over $GF(2)$ if every bit in it is a linear function of the input word x. The authors of [23] proved that the minimum possible length of such a scheme can be expressed as the minimum possible rank of an appropriate matrix associated with the side information graph. They further conjectured that no index code can outperform the best linear coding scheme. This has been disproved in [83] where the authors showed that sometimes it is better to view the input word x as a word over a larger field and use a linear encoding over that field. It is also possible to split x

into several pieces, and apply in each piece a linear encoding function over a different field, thus obtaining some further savings.

Our results here show that inherently non-linear encodings are sometimes better. In particular, when the side information graph is a disjoint union of many copies of C_5, the non-linear scheme discussed in the previous subsections can be shown to be better than any hybrid of linear schemes over any collection of fields. This is proved by expressing the minimum possible length of such a hybrid by the minimum possible sum of ranks of matrices defined appropriately over the corresponding fields, and by showing that the non-linear scheme is better.

4.4 Concluding remarks and open problems

- In this chapter, we have shown that for large values of k and for every graph G, $\beta_1(k \cdot G) = (n - \log_2 \alpha(\mathfrak{C}(G)) + o(1))\, k$, where the $o(1)$-term tends to 0 as $k \to \infty$.

- Our results also imply that encoding the entire block at once can be strictly better than concatenating the optimal index code for G with a single bit block. This justifies the definition of the broadcast rate of G, $\beta(G)$, as the optimal *asymptotic average* number of bits required for a single bit of index coding for G.

  In the above case of C_5, $2 \leq \beta(C_5) \leq 2.678$, and it would be interesting to determine the index coding rate of C_5 precisely. It would be further interesting to determine the index coding rate of additional families of graphs, and in particular, to decide if there exists a family of graphs G on n vertices where $\beta(G) = O(1)$ whereas $\beta_1(G) = \omega(1)$.

  Moreover, it should be noted that the parameter $\beta_t(G)$ is computable (though maybe not efficiently) for any graph G and any t. It would be interesting to find out whether the parameter $\beta(G)$ is also computable.

- We have also shown that $\beta_1(k \cdot C_5)/k \approx 2.678$ for large values of k, whereas $\beta_1(C_5) = 3$. Hence, the optimal index code for a disjoint union of k copies of a graph G can be strictly better than the concatenation of k optimal index codes for G (benefit is gained already for $k = 15$). This is surprising, considering the lack of mutual information between receivers which correspond to distinct copies.

Part II

Codes and explicit Ramsey Constructions

Chapter 5

Codes and Xor graph products

The results of this chapter appear in [10]

What is the maximum possible number, $f_3(n)$, of vectors of length n over $\{0,1,2\}$ such that the Hamming distance between every two is even? What is the maximum possible number, $g_3(n)$, of vectors in $\{0,1,2\}^n$ such that the Hamming distance between every two is odd? We investigate these questions, and more general ones, by studying Xor powers of graphs, focusing on their independence number and clique number, and by introducing two new parameters of a graph G. Both parameters denote limits of series of either clique numbers or independence numbers of the Xor powers of G (normalized appropriately), and while both limits exist, one of the series grows exponentially as the power tends to infinity, while the other grows linearly. As a special case, it follows that $f_3(n) = \Theta(2^n)$ whereas $g_3(n) = \Theta(n)$.

5.1 Introduction

The *Xor product* of two graphs, $G = (V, E)$ and $H = (V', E')$, is the graph whose vertex set is the Cartesian product $V \times V'$, where two vertices (u, u') and (v, v') are connected iff either $uv \in E$, $u'v' \notin E'$ or $uv \notin E$, $u'v' \in E'$, i.e., the vertices are adjacent in precisely one of their two coordinates. This product is commutative and associative, and it follows that for any $n \geq 1$, the product of $G_1, \ldots, G_n$ is the graph whose vertex set is $\prod V(G_i)$, where two vertices are connected iff they are adjacent in an *odd* number of coordinates. Throughout this chapter, let $G \cdot H$ denote the Xor product of G and H, and let G^n denote the Xor product of n copies of G.

The Xor graph product was studied in [102], where the author used its properties to construct edge colorings of the complete graph with two colors, containing a smaller number of monochromatic copies of K_4 than the expected number of such copies in a random coloring. See also [47],[52],[103] for more about this problem.

Examine K_3, the complete graph on 3 vertices. Each vertex of K_3^n can be naturally represented by a vector in $\{0,1,2\}^n$, and two vertices are connected in K_3^n iff their representing vectors differ in an odd number of coordinates, or equivalently, have an odd Hamming distance. Thus, a set of vectors in $\{0,1,2\}^n$, in which every two vectors have an even Hamming distance, represents an independent set in K_3^n; similarly, a set of vectors of $\{0,1,2\}^n$ in which each pair has an odd Hamming distance represents a clique in K_3^n, and hence:

$$f_3(n) = \alpha(K_3^n)\ ,$$
$$g_3(n) = \omega(K_3^n)\ ,$$

where $\alpha(G)$ denotes the independence number of G and $\omega(G)$ denotes the clique number of G. Studying the series of independence numbers and the series of clique numbers of powers of a fixed graph G provides several interesting questions and results. Both series, when normalized appropriately, converge, however one has an exponential growth while the other grows linearly.

In section 5.2 we show that the series of independence numbers, when normalized, converges to its supremum, which we denote by $x_\alpha(G)$:

$$x_\alpha(G) = \lim_{n\to\infty} \sqrt[n]{\alpha(G^n)} = \sup_n \sqrt[n]{\alpha(G^n)}$$

We calculate this parameter for several families of graphs and multi-graphs, and study some of its properties.

In section 5.3 we show, this time using a linear normalization, that the series $\omega(G^n)/n$ converges as well. We denote its limit by $x_\omega(G)$:

$$x_\omega(G) = \lim_{n\to\infty} \frac{\omega(G^n)}{n} = \sup_n \frac{\omega(G^n)-2}{n+1}$$

Determining the value of x_α and x_ω for K_3 and for a general complete graph K_r gives the asymptotic behavior of $f_3(n)$ and $g_3(n)$, and similarly, of $f_r(n)$ and $g_r(n)$, defined analogously with r replacing the alphabet size of 3. For a general G, it seems that merely approximating x_α and x_ω can be extremely difficult. Both parameters are non-monotone with respect to the addition of edges to the graph, and we use combinatorial ideas, tools from linear algebra and spectral techniques in order to provide bounds for them for different graphs.

5.2 Independence numbers of Xor powers

5.2.1 The independence series and x_α

We begin with an immediate observation: for every two graphs G and H, and every two independent sets $I \subset V(G)$ and $J \subset V(H)$, $I \times J$ is an independent set of $G \cdot H$. Therefore, the function $f(n) = \alpha(G^n)$ is super-multiplicative: $f(m+n) \geq f(m)f(n)$, and by Fekete's lemma (c.f., e.g., [76], p. 85), we deduce that

$$\exists \lim_{n\to\infty} \sqrt[n]{f(n)} = \sup_n \sqrt[n]{f(n)}$$

Let $x_\alpha(G)$ denote this limit.

We note that the definition of the Xor product and of x_α applies to multi-graphs as well: indeed, since only the parity of the number of edges between two vertices dictates their adjacency, we can assume that there are no multiple edges, however there may be (self) loops in the graph. The function $f(n) = \alpha(G^n)$ remains super-multiplicative (notice that an independent set I of G^n can never contain a vertex $v = (v_1, \ldots, v_n)$ with an odd number of coordinates $\{v_{i_j}\}$, which have loops). However, in the single scenario where every vertex of G has a loop, $\alpha(G) = 0$ and we cannot apply Fekete's lemma (indeed, in this case, $f(2n+1) = 0$ and $f(2n) \geq 1$ for all n). In all other cases, $x_\alpha(G)$ is well defined. Furthermore, if we negate the adjacency matrix of G, obtaining the multi-graph complement $\overline{G}$ (u and v are adjacent in $\overline{G}$ iff they are disconnected in G, including the case $u = v$), we get $x_\alpha(G) = x_\alpha(\overline{G})$, as long as $x_\alpha(\overline{G})$ is also defined. To see this fact, take the even powers $2k$

of the independence series, in which two vertices are adjacent in G^{2k} iff they are adjacent in $\overline{G}^{2k}$.

Proposition 5.2.1. *For every multi-graph $G = (V, E)$ satisfying $\alpha(G) > 0$, $x_\alpha(G)$ is well defined. Furthermore, if in addition $\alpha(\overline{G}) > 0$, where $\overline{G}$ is the multi-graph-complement of G, then $x_\alpha(G) = x_\alpha(\overline{G})$.*

5.2.2 General bounds for x_α

It is obvious that $x_\alpha(G) \leq |V(G)|$, and this upper bound is tight, for instance, for the edgeless graph. For the lower bound, the following simple fact holds:

Claim 5.2.2 (Uniform lower bound)**.** *Let $G = (V, E)$ be a multi-graph satisfying $\alpha(G) > 0$. Then:*

$$x_\alpha(G) \geq \sqrt{|V|} \tag{5.1}$$

Proof. Let $I \subset V(G^2)$ denote the set $\{(v, v) \mid v \in V\}$. Clearly, I is an independent set of G^2 of size $|V|$, thus $x_\alpha(G) \geq |V|^{\frac{1}{2}}$ (and similarly, for all k we get an explicit independent set of size $|V|^k$ in G^{2k}). ■

For a better understanding of the parameter $x_\alpha(G)$, we next show several infinite families of graphs which attain either the lower bound of (5.1) or the upper bound of $|V(G)|$. While, trivially, the edgeless graph G on n vertices satisfies $x_\alpha(G) = n$, it is interesting that complete bipartite graphs also share this property:

Claim 5.2.3. *Let $K_{m,n}$ denote the complete bipartite graph with color classes of sizes m, n, where $m \geq n$. Then for every $k \geq 1$, $K_{m,n}^k$ is a complete bipartite graph with color classes W_0, W_1 of sizes:*

$$|W_0| = \frac{1}{2}\left((m+n)^k + (m-n)^k\right) \quad , \quad |W_1| = \frac{1}{2}\left((m+n)^k - (m-n)^k\right)$$

Therefore, $x_\alpha(K_{m,n}) = m + n$.

Proof. Let $G = K_{m,n}$, $m \geq n$, and denote its color classes by U_0, U_1, where $|U_0| = m$. For every vertex $v = (v_1, \ldots, v_k) \in V(G^k)$, define a vector $w_v \in \{0, 1\}^k$, in the following manner: $(w_v)_i = 0$ iff $v_i \in U_0$. By the definition of the Xor product (recall that G is a complete bipartite graph), the following holds for every $u, v \in V(G^k)$:

$$uv \notin E(G^k) \iff |\{1 \leq i \leq k \mid (w_u)_i \neq (w_v)_i\}| = 0 \pmod 2$$

Equivalently, performing addition and dot-product over $GF(2^k)$:

$$uv \notin E(G^k) \iff (w_u + w_v) \cdot \underline{1} = 0 \tag{5.2}$$

Let W_0 denote the set of all vertices in $v \in V(G^k)$ such that the Hamming weight of w_v is even, and let W_1 denote the set of all those whose corresponding vectors have an odd Hamming weight. In other words, we partition the vertices of G^k into two sets, according to the parity of the number of times a coordinate was taken from U_0. Notice that:

$$|W_0| = \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k}{2i} n^{2i} m^{k-2i} = \frac{1}{2}\left((m+n)^k + (m-n)^k\right) \ ,$$

and similarly:

$$|W_1| = \frac{1}{2}\left((m+n)^k - (m-n)^k\right)$$

To see that G^k is a complete bipartite graph with color classes W_0, W_1, argue as follows: take $u, v \in W_i$ $(i \in \{0,1\})$; clearly, we have:

$$(w_u + w_v) \cdot \underline{1} = w_u \cdot \underline{1} + w_v \cdot \underline{1} = i + i = 0 \ ,$$

hence, by (5.2), W_0 and W_1 are both independent sets. Next, for every $u \in W_0$ and $v \in W_1$, we have:

$$(w_u + w_v) \cdot \underline{1} = 0 + 1 = 1 \ ,$$

implying that u and v are adjacent. This completes the proof. ■

The previous claim shows that $x_\alpha(K_{n,n}) = 2n = n x_\alpha(K_2)$. This is a special case of the following property of x_α:

Claim 5.2.4. *Let $G = (V, E)$ be a graph on the vertex set $V = [n]$. We define the r-blow-up of G, $G[r]$, as the n-partite graph whose color groups are*

$(V_1,\ldots,V_n)$, where for all i, $|V_i| = r$, and two vertices $x \in V_i$ and $y \in V_j$ are connected iff $ij \in E$. Then:

$$x_\alpha(G[r]) = r \cdot x_\alpha(G)$$

Furthermore, every maximum independent set of $G[r]^k$ is an r-blow-up of a maximum independent set of G^k.

Proof. Let $T : V(G[r]) \to V(G)$ be the mapping from each vertex in $G[r]$ to its corresponding vertex in G (i.e., if $x \in V_i$, then $T(x) = i$), and define $T^{\circ k} : V(G[r]^k) \to V(G^k)$ by

$$T^{\circ k}(v_1,\ldots,v_k) = (T(v_1),\ldots,T(v_k))$$

Then, by the definition of $G[r]$, $T^{\circ k}(G[r]^k)$ is isomorphic to G^k, and furthermore, a set I is independent in $G[r]^k$ iff $T^{\circ k}(I)$ is independent in G^k. This implies that every maximum independent set of $G[r]^k$ can be obtained by taking a maximum independent set of G^k and expanding each coordinate in each of the r possible ways. In particular:

$$\alpha(G[r]^k)^{\frac{1}{k}} = \left(r^k \alpha(G^k)\right)^{\frac{1}{k}} = r \cdot \alpha(G^k)^{\frac{1}{k}}$$

and the desired result follows. ■

A simple algebraic consideration provides an example for a family of multi-graphs which attain the lower bound - the Hadamard multi-graphs (see , e.g., [76] for further information on Sylvester-Hadamard matrices):

Claim 5.2.5. *Let H_{2^n} be the multi-graph whose adjacency matrix is the Sylvester-Hadamard matrix on 2^n vertices: two (not necessarily distinct) vertices u and v, represented as vectors in $GF(2^n)$, are adjacent iff their dot product equals 1. Then: $x_\alpha(H_{2^n}) = 2^{n/2}$*

Proof. Let $H = H_{2^n}$. Notice that exactly 2^{n-1} vertices have loops, and in particular there is a non-empty independent set in H and x_α is defined. Examine H^k; by definition, $u = (u_1,\ldots,u_k)$ and $v = (v_1,\ldots,v_k)$ are adjacent in H^k iff $\sum_i u_i \cdot v_i = 1 \pmod 2$. This implies, by the definition of the Hadamard multi-graph, that:

$$H_{2^n}^k = H_{2^{nk}}$$

We are thus left with showing that $H = H_{2^n}$ satisfies $\alpha(H) \leq \sqrt{|H|}$, and this follows from the fact that an independent set in H is a self-orthogonal set of vectors in $GF(2^n)$, hence the rank of its span is at most $n/2$ and thus:

$$\alpha(H) \leq 2^{n/2} = \sqrt{|H|}\ ,$$

as needed. ■

Note that the result above is also true for multi-graphs whose adjacency matrix is a general-type Hadamard matrix, H_n; this can be proved using spectral analysis, in a way similar to the treatment of strongly-regular graphs in the next subsection. As another corollary of the analysis of strongly-regular graphs in the next subsection, we will show that the Paley graph P_q, defined there, has q vertices and satisfies $x_\alpha(P_q) \leq \sqrt{q}+1$, hence there exists a family of simple graphs which roughly attain the general lower bound on x_α.

5.2.3 Properties of x_α and bounds for codes

The normalizing factor applied to the independence series when calculating x_α depends only on the current graph power, therefore restricting ourselves to an induced subgraph of a graph G immediately gives a lower bound for $x_\alpha(G)$. It turns out that x_α cannot drastically change with the addition of a single vertex to the graph - each added vertex may increase x_α by at most 1. However, x_α is non-monotone with respect to the addition of edges. The next few claims summarize these facts.

Claim 5.2.6. *Let $G = (V, E)$ be a multi-graph, and let H be an induced subgraph on $U \subset V$, satisfying $\alpha(H) > 0$. Then:*

$$x_\alpha(H) \leq x_\alpha(G) \leq x_\alpha(H) + |V| - |U|$$

Proof. The first inequality is trivial, since we can always restrict our choice of coordinates in independent sets of G^k to vertices of U. In order to prove the second inequality, it is enough to prove the case of $|U| = |V| - 1$. Denote by v the single vertex of $V \setminus U$, and assume that v does not have a loop. Let I be a maximum independent set of G^k. For every pattern of i appearances of v in the coordinates of vertices of I, the set of all vertices of I containing this pattern (and no other appearances of v) is an independent set. This set remains independent in H^{k-i}, after omitting from each of these vertices its i appearances of v, hence its size is at most $\alpha(H^{k-i})$. Since $x_\alpha(H)$ is the supremum of $\sqrt[n]{\alpha(H^n)}$, we get the following bound for I:

$$|I| \leq \sum_{i=0}^{k} \binom{k}{i} \alpha(H^{k-i}) \leq \sum_{i=0}^{k} \binom{k}{i} x_\alpha(H)^{k-i} = (x_\alpha(H)+1)^k \ .$$

Taking the k-th root gives $x_\alpha(G) \leq x_\alpha(H)+1$.

We are left with the case where v has a loop. If H has no loops, then every vertex of I must have an even number of appearances of v in its coordinates (as an independent set cannot contain loops). Hence, every pattern of i appearances of v in the coordinates of vertices of I still represents an independent set in H^{k-i}, and the calculation above is valid. In fact, it gives that

$$|I| \leq \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k}{i} \alpha(H^{k-2i}) = \frac{1}{2}\left((x_\alpha(H)+1)^k + (x_\alpha(H)-1)^k\right) < (x_\alpha+1)^k \ .$$

If H does contain loops, then $\alpha(\overline{H}) > 0$, and we can apply the previous argument to $\overline{G}$ with respect to $\overline{H}$ and v (which does not have a loop in $\overline{G}$), obtaining:

$$x_\alpha(G) = x_\alpha(\overline{G}) \leq x_\alpha(\overline{H}) + 1 = x_\alpha(H) + 1 \ ,$$

where the last equality holds since $\alpha(H) > 0$, guaranteeing that at least one vertex of H does not have a loop. ∎

Notice that, by the last claim, we can apply the vertex-exposure Martingale on the random graph $\mathcal{G}_{n,\frac{1}{2}}$, and obtain a concentration result for x_α (see for example [19], Chapter 7):

Corollary 5.2.7. *Almost surely, that is, with probability that tends to 1 as n tends to infinity, the random graph $G = \mathcal{G}_{n,\frac{1}{2}}$ satisfies*

$$|x_\alpha(G) - \mathbb{E}x_\alpha(G)| \leq O(\sqrt{n})$$

A counterexample for edge-addition monotonicity exists already when $|V| = 3$, as the next claim shows.

Claim 5.2.8. *x_α is non-monotone with respect to the addition of edges.*

Proof. Let $G = (V, E)$ be the graph on three vertices $V = \mathbb{Z}_3$ and one edge $E = \{(0, 1)\}$. We show that $x_\alpha(G) = 2$, thus if we *remove* the single edge (creating the empty graph on 3 vertices) or *add* the edge $(1, 2)$ (creating the complete bipartite graph $K_{1,2}$) we increase x_α to a value of 3. In fact, up to an automorphism of the graph G in each coordinate, there is exactly one maximum independent set of G^k, which is $\{(v_1, \ldots, v_k) : v_i \in \{0, 2\}\}$.

The proof is by induction on k, stating that every maximum independent set of G^k is the Cartesian product of either $\{0, 2\}$ or $\{1, 2\}$ in each of the coordinates (it is obvious that this set is indeed independent). The case $k = 1$ is trivial. For $k > 1$, let I be a maximum independent set of G^k, and notice that by the construction of the independent set above, we have $|I| = \alpha(G^k) \geq 2^k$. Let A_i ($i \in \mathbb{Z}_3$) be the set of vertices of I whose first coordinate is i. We denote by A'_i the set of vertices of G^{k-1} formed by omitting the first coordinate from A_i. Since $A_i \subset I$ is independent, so is A'_i for every i. However, every vertex of A'_0 is adjacent to every vertex of A'_1 (again since I is independent).

Note that, by induction, $|A_i| = |A'_i| \leq 2^{k-1}$. Clearly, this implies that if either A_0 or A_1 are empty, we are done, and I is the Cartesian product of a maximum independent set $I' \subset G^{k-1}$ of size 2^{k-1}, with either $\{0, 2\}$ or $\{1, 2\}$. Indeed, if for instance A_1 is empty, then both A'_0 and A'_2 are maximum independent sets of G^{k-1} (otherwise, the size of I would be strictly less than 2^k), with the same automorphism of G in each coordinate (otherwise I would not be independent - consider the two vertices which contain 2 in all coordinates except the one where the automorphism is different).

Assume therefore that $A_0, A_1 \neq \emptyset$. By a similar argument, $A_2 \neq \emptyset$, otherwise $|I| \geq 2^k$ would imply that both A'_0 and A'_1 are maximum independent sets in G^{k-1} (of size 2^{k-1} each), and by induction, both contain the vector $\underline{2}$, contradicting the independence of I. We therefore have:

$$|I| = \sum_i |A_i| = \sum_i |A'_i| < (|A'_0| + |A'_2|) + (|A'_1| + |A'_2|) \leq 2 \cdot 2^{k-1} = 2^k$$

The last inequality is by the fact that $A'_2 \cap A'_0 = A'_2 \cap A'_1 = \emptyset$, since, for instance, all vertices in A'_0 are adjacent to all vertices in A'_1 but disconnected from all vertices in A_2'. We therefore obtained a contradiction to the fact that $|I| \geq 2^k$. ■

We next prove a general upper bound for x_α of regular graphs. As a corollary, this will determine $x_\alpha(K_3)$ and give the asymptotic behavior of the function $f_3(n)$, mentioned in the abstract.

Theorem 5.2.9. *Let G be a loopless nontrivial d-regular graph on n vertices, and let $d = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ denote the eigenvalues of G. Then:*

$$x_\alpha(G) \leq \max\{|n-2d|, 2|\lambda_2|, 2|\lambda_n|\}$$

Proof. We use spectral analysis to bound the independence numbers of powers of the graph G. Denote by $A = A_G$ the adjacency matrix of G, and let $B = B_G = (-1)^A$, i.e.:

$$B_{ij} \overset{def}{=} \begin{cases} -1 & ij \in E(G) \\ 1 & ij \notin E(G) \end{cases}$$

Notice that $B_{G \cdot H} = B_G \otimes B_H$, where $\otimes$ denotes the tensor-product:

$$(B_G \otimes B_H)_{(u,v),(u',v')} = B_{G u,v} \cdot B_{H u',v'} = \begin{cases} -1 & (u,v)(u',v') \in E(G \cdot H) \\ 1 & (u,v)(u',v') \notin E(G \cdot H) \end{cases}$$

Our aim in using B_G is to obtain expressions for the eigenvalues of A_{G^k}, and then use the following bound, proved by Hoffman: every regular graph H with eigenvalues $\mu_1 \geq \ldots \geq \mu_m$ satisfies:

$$\alpha(H) \leq \frac{-|H|\mu_m}{\mu_1 - \mu_m} \tag{5.3}$$

(see [67], [81]). Recall that the eigenvalues of A are:

$$\lambda(A) = \{d = \lambda_1, \ldots, \lambda_n\}$$

By definition, $B_G = J_n - 2A_G$, where J_n is the all 1-s matrix of order n, and fortunately, the single non-zero eigenvalue of J_n (the eigenvalue n) corresponds to an eigenvector of $\underline{1}$, which is also an eigenvector of A (with the eigenvalue d). Thus, if we denote the spectrum of B by Λ:

$$\Lambda = \lambda(B) = \{n - 2d, -2\lambda_2, \ldots, -2\lambda_n\}$$

Define $\Lambda^k = \{\mu_1\mu_2 \ldots \mu_k : \mu_i \in \Lambda\}$. As usual with tensor-products (c.f., e.g., [9]), we use the fact that:

$$\lambda(B^{\otimes k}) = \{\lambda_{i_1}\lambda_{i_2} \cdot \ldots \cdot \lambda_{i_k} \mid \lambda_{i_j} \in \lambda(B)\} = \Lambda^k$$

Returning to A_{G^k}, we have $A_{G^k} = \frac{1}{2}(J_{n^k} - B_{G^k})$, and $\underline{1}$ is an eigenvector of B_{G^k} corresponding to the eigenvalue $(n-2d)^k$. Hence, $\underline{1}$ is an eigenvector of A_{G^k} with an eigenvalue of:

$$\lambda_M = \frac{n^k - (n-2d)^k}{2}$$

Since this is the regularity degree of G^k, by the Perron-Frobenius theorem it is also its largest eigenvalue. The remaining eigenvalues of A_{G^k} are $\left\{-\frac{1}{2}\mu : \mu \in \Lambda^k, \mu \neq (n-2d)^k\right\}$. Hence, if we define:

$$\beta(k) = \max\left\{\Lambda^k \setminus \{(n-2d)^k\}\right\}$$

then the minimal eigenvalue of A_{G^k}, λ_m, equals $-\frac{1}{2}\beta(k)$. Applying (5.3) gives:

$$\alpha(G^k) \leq \frac{-n^k\lambda_m}{\lambda_M - \lambda_m} = \frac{\beta(k)}{1 - (1-\frac{2d}{n})^k + \beta(k)/n^k} \tag{5.4}$$

Examine the right hand side of (5.4). The term $\left(1 - \frac{2d}{n}\right)^k$ tends to zero as k tends to infinity, since G is simple and hence $1 \leq d \leq n-1$. Considering $\beta(k)$, notice that for sufficiently large values of k, in order to obtain the maximum of $\Lambda^k \setminus \{(n-2d)^k\}$, one must choose the element of Λ whose absolute value is maximal with plurality at least $k-2$ (the remaining two choices of elements should possibly be used to correct the sign of the product, making sure the choice made is not the one corresponding to the degree of G^k). Therefore, if we set $r = \max\{|n-2d|, 2|\lambda_2|, 2|\lambda_n|\}$, we get $\beta(k) = \Theta(r^k)$. To bound r, we use the following simple argument, which shows that

$$\lambda = \max\{|\lambda_2|, \ldots, |\lambda_n|\} \leq \frac{n}{2}$$

(equality is precisely in the cases where G is complete bipartite with $d = \frac{n}{2}$). Indeed, the square of the adjacency matrix A of G has the values d on its diagonal (as G is d-regular), hence:

$$d^2 + \lambda^2 \leq \sum_i \lambda_i^2 = \mathrm{tr}(A^2) = nd ,$$

implying that:

$$\lambda \leq \sqrt{d(n-d)} \leq \frac{n}{2}$$

Therefore, either $r = 2\lambda \leq n$ or $r = |n - 2d| < n$, and in both cases we obtain that $\beta(k)/n^k = O(1)$. Taking the k-th root in (5.4), gives:

$$x_\alpha(G) \leq \lim_{k\to\infty} \sqrt[k]{\beta(k)} = r \ ,$$

as required. ■

Note that the above proof in fact provides upper bounds for the independence numbers of every power k of a given regular graph G (not only for the asymptotic behavior as k tends to infinity) by calculating $\beta(k)$ and applying (5.4).

Corollary 5.2.10. *For the complete graphs K_3 and K_4,*

$$x_\alpha(K_3) = x_\alpha(K_4) = 2 \ .$$

Proof. It is easy and well known that the eigenvalues of the complete graph K_n on $n \geq 2$ vertices are: $\{n-1, -1, \ldots, -1\}$. By Theorem 5.2.9, we have, for every $n \geq 2$:

$$x_\alpha(K_n) \leq \max\{n-2, 2\}$$

For $n = 3$, this implies $x_\alpha(K_3) \leq 2$, and for $n \geq 4$ this implies $x_\alpha(K_n) \leq n-2$. The lower bounds for K_3 and K_4 follow from the fact that $x_\alpha(K_2) = 2$.

We note that (5.4) gives the following bounds on $\alpha(K_n^k)$ for every $k \geq 1$:

$$\alpha(K_3^k) \leq \frac{2^k}{1 - \left(-\frac{1}{3}\right)^k + \left(\frac{2}{3}\right)^k} \ ,$$

$$\alpha(K_n^k) \leq \frac{2(n-2)^{k-1}}{1 - (\frac{2-n}{n})^k + \frac{2}{n}(\frac{n-2}{n})^{k-1}} \ , \quad n \geq 4 \ , \quad 2 \nmid k \ ,$$

$$\alpha(K_n^k) \leq \frac{2(n-2)^{k-1}}{1 - (\frac{2-n}{n})^k + \frac{4}{n^2}(\frac{n-2}{n})^{k-2}} \ , \quad n \geq 4 \ , \quad 2 \mid k \ .$$

■

Recalling the motivation of the codes considered in the introduction, the last claim implies that

$$f_3(n) = \Theta(2^n)$$
$$f_4(n) = \Theta(2^n)$$

In other words, extending the alphabet from 3 letters to 4 does not increase the maximal asymptotic size of the required code, and both cases are asymptotically equivalent to using a binary alphabet. However, adding additional letters to the alphabet does increase this asymptotic size, as it is immediate by Claim 5.2.2 that $f_5(n)$ is at least $\Omega(\sqrt{5}^n)$. Using a simple probabilistic argument (similar to the one used in [9]), we can derive an upper bound for $x_\alpha(K_5)$ from the result on K_4 :

Claim 5.2.11. *Let G be a vertex transitive graph, and let H be an induced subgraph of G. Then:*

$$x_\alpha(G) \leq x_\alpha(H)\frac{|G|}{|H|}$$

Combining this with Corollary 5.2.10, we get:

Corollary 5.2.12. *For all $m < n$, $x(K_n) \leq \frac{x_\alpha(K_m)}{m}n$, and in particular, $\sqrt{5} \leq x_\alpha(K_5) \leq \frac{5}{2}$.*

Proof of claim. Let I be a maximum independent set of G^k, and denote by $\sigma_1, \sigma_2, \ldots, \sigma_k$ random automorphisms of G, chosen independently and uniformly out of all the automorphisms of G. The permutation τ, which maps $v = (v_1, \ldots, v_k) \in G^k$ to $(\sigma_1(v_1), \ldots, \sigma_k(v_k))$, is an automorphism of G^k, and moreover, if we fix a vertex v in G^k, then $\tau(v)$ is uniformly distributed over all the vertices of G^k. Let S be an induced copy of H^k in G^k, and notice that by the properties of τ,

$$\mathbb{E}|\tau(S) \cap I| = |I|\frac{|S|}{|G^k|} = |I|\left(\frac{|H|}{|G|}\right)^k$$

On the other hand, I is an independent set, therefore $|\tau(S) \cap I| \leq \alpha(H^k) \leq (x_\alpha(H))^k$. Choose an automorphism τ for which this random variable attains at least its expected value of $\mathbb{E}|\tau(S) \cap I|$, and it follows that:

$$|I| \leq \left(x_\alpha(H)\frac{|G|}{|H|}\right)^k$$

■

While the best upper bound we have for K_n, when $n \geq 5$, is $n/2$, the last corollary, as well as some simple observations on the first few powers of complete graphs, lead to the following conjecture:

Conjecture 5.2.13. *For every $n \geq 4$, the complete graph on n vertices satisfies $x_\alpha(K_n) = \sqrt{n}$.*

It seems possible that the Delsarte linear programming bound (c.f., e.g., [85]) may provide improved upper bounds for $\alpha(K_n^k)$ when $n \geq 4$, but it does not seem to supply a proof of the last conjecture.

As another corollary of Theorem 5.2.9, we can derive bounds for x_α of strongly-regular graphs. Recall that a strongly-regular graph G with parameters (n, d, λ, μ) is a d-regular graph on n vertices, where the co-degree (the number of common neighbors) of every two adjacent vertices is λ, and the co-degree of every two non-adjacent vertices is μ. The eigenvalues of such a graph are d and the solutions to the quadratic equation $x^2 + (\mu - \lambda)x + (\mu - k) = 0$ (c.f., e.g. [60], Chapter 10). As an example, we consider the Paley graphs:

Corollary 5.2.14. *The Paley graph P_q (where q is a prime power, $q = 1 \pmod 4$) satisfies $\sqrt{q} \leq x_\alpha(P_q) \leq \sqrt{q} + 1$.*

Proof. Recall that P_q has a vertex set $V(P_q) = GF(q)$ and $i, j \in V$ are connected iff $i - j$ is a quadratic residue in $GF(q)$. It is easy to check that P_q is a $(q, \frac{q-1}{2}, \frac{q-5}{4}, \frac{q-1}{4})$ strongly regular graph (c.f., e.g., [60]). Hence, its largest eigenvalue is $\frac{q-1}{2}$, and its remaining eigenvalues are the solutions of the equation $x^2 + x - \frac{q-1}{4} = 0$, i.e., $\{\frac{-1 \pm \sqrt{q}}{2}\}$. By Theorem 5.2.9:

$$x_\alpha(P_q) \leq \max\{1, \sqrt{q} + 1\} = \sqrt{q} + 1$$

■

We conclude this section with another example of an extremal problem on codes, which can easily be translated to the terms of x_α: let $\tilde{f}_3(n)$ be the maximum size of a set of words over $\mathbb{Z}_3^n$, where for every two not necessarily distinct words u, v, the Hamming weight of their sum $u + v$ (addition is performed modulo 3) is even. Determining $\tilde{f}_3(n)$ asymptotically becomes relatively simple, once the problem is translated to the problem of determining $x_\alpha(H)$ for an appropriate multi-graph H. This graph H has a vertex set $V = \mathbb{Z}_3$, where 0 is connected to both 1 and -1, and there are loops on the vertices $1, -1$. It is easy to confirm that a maximum independence set in H^n corresponds to a code of maximum size, meeting the requirements mentioned above. This is an induced subgraph of H_4, the Hadamard graph on 4 vertices (assign the vertices $\{0, 1, -1\}$ the values $\{11, 01, 10\}$ respectively), hence $x_\alpha(H) \leq x_\alpha(H_4) = 2$. The lower bound is immediate, and therefore, $\tilde{f}_3(n) = \Theta(2^n)$.

5.3 Clique numbers of Xor powers

5.3.1 The clique series and x_ω

In the previous section, we examined independent sets in Xor powers of graphs; the behavior of cliques in Xor powers of graphs proves to be significantly different.

Theorem 5.3.1. *For every graph $G = (V, E)$, the limit of $\frac{\omega(G^n)}{n}$ as n tends to infinity exists. Let $x_\omega(G)$ denote this limit. Then:*

$$0 \leq x_\omega(G) = \sup_n \frac{\omega(G^n) - 2}{n + 1} \leq |V|$$

Proof. Let G and H denote two simple graphs, and let $\{v_1, \ldots, v_r\}$ and $\{u_1, \ldots, u_s\}$ be maximum cliques in G and H respectively. The following set is a clique in the graph $G \cdot H \cdot K_2$, where the vertex set of K_2 is $\{0, 1\}$:

$$\{v_2, \ldots, v_r\} \times \{u_1\} \times \{0\} \quad \cup \quad \{v_1\} \times \{u_2, \ldots, u_s\} \times \{1\} \tag{5.5}$$

Thus, the following inequality applies to every two simple graphs G and H:

$$\omega(G \cdot H \cdot K_2) \geq \omega(G) + \omega(H) - 2 \tag{5.6}$$

Note that there are graphs G and H for which equation (5.6) is tight. For example, take both G and H to be powers of K_2. The graph K_2^n is triangle free (recall that by Claim 5.2.3, K_2^n is bipartite), therefore, $\omega(K_2^{k+l}) = 2 = \omega(K_2^k) + \omega(K_2^l) - 2$.

Consider a graph G, and define $g(n) = \omega(G^n)$. If G contains no edges, then each of its powers is an edgeless graph, and $g(n) = 1$ for all n. Otherwise, it contains a copy of K_2, hence equation (5.6) implies that for every $m, n \geq 1$:

$$g(m+n+1) \geq g(m) + g(n) - 2$$

Defining, for every $n \geq 1$,

$$\hat{g}(n) = g(n-1) - 2$$

gives:

$$\hat{g}(m+n) = g(m+n-1) - 2 \geq g(m-1) + g(n-1) - 4 = \hat{g}(m) + \hat{g}(n)$$

Therefore, the function $\hat{g}$ is super-additive, and by Fekete's lemma, the limit of the series $\frac{\hat{g}(n)}{n}$ exists and equals its supremum. We note that this applies for edgeless graphs as well, where this limit equals 0. Denote this limit by x_ω:

$$x_\omega(G) = \lim_{n\to\infty} \frac{\omega(G^n)}{n} = \sup_n \frac{\omega(G^n) - 2}{n+1} \tag{5.7}$$

It remains to show that $x_\omega(G) \leq |V|$. We first need the following definition: A function $f : V \to \mathbb{Z}_2^k$ (for some $k \geq 1$) will be called a **proper representation** of G, if there is a $b_f \in \{0,1\}$, such that for every (not necessarily distinct) $u, v \in V$, $uv \in E$ iff $f(u) \cdot f(v) = b_f$. The dimension of the representation, $\dim(f)$, is defined to be $\dim(f(V))$ in $\mathbb{Z}_2^k$.

The upper bound for x_ω is given by the following lemma:

Lemma 5.3.2. *If $G = (V, E)$ has a proper representation f, then $x_\omega(G) \leq \dim(f)$.*

Proof. Let $x \circ y$ denote the concatenation of the vectors x and y. By the definition of the Xor product, for every two graphs G and H, if g is a proper representation of G and h is a proper representation of H, then $g \circ h$, which maps each vector $(u, v) \in V(G \cdot H)$ to $g(u) \circ h(v)$, is a proper representation of $G \cdot H$, with $b_{g\circ h} = b_g + b_h + 1 \pmod 2$. Clearly, $\dim(g \circ h) \leq \dim(g) + \dim(h)$.

Suppose f is a proper representation of G of dimension d, and let g denote the k-fold concatenation of f. Allowing $\dim(g)$ to be at most $kd+1$ we may assume that $b_g = 0$ (by adding a new coordinate of 1 to all vectors if necessary). Let S be a maximum clique in G^k, $|S| = s$. We define B to be the matrix whose s columns are $\{g(v) : v \in S\}$. Since S is a clique, and g is a proper representation of G^k with $b_g = 0$, then $B^tB = I$. The rank of B^tB is thus s, hence:

$$s = \mathrm{rank}(B^tB) \leq \mathrm{rank}(B) \leq \dim(g) \leq kd + 1$$

We conclude that for every k, $\frac{\omega(G^k)}{k} \leq d + \frac{1}{k}$, and the result follows. ■

To prove that $x_\omega(G) \leq |V|$, it suffices to show that there exists a proper representation for every G (the dimension of the span of n vectors can never exceed n). Set $|V| = n$ and $|E| = m$, and examine the function $f : V \to \mathbb{Z}_2^m$, which maps each vertex v to its corresponding row in the incidence matrix of G. For every $u \neq v \in V$, either $uv \in E$, in which case there is a single index at which $f(u) = f(v) = 1$, or $uv \notin E$ and there is no such index. Hence $f(u) \cdot f(v) = 1$ iff $uv \in E$ (and in particular, this applies to the dot product in $\mathbb{Z}_2^m$ as well). All that remains in order to turn f into a proper representation of G (with $b_f = 1$) is to adjust the values of $f(u) \cdot f(u)$ to 0 for every $u \in V$. Note that $f(u) \cdot f(u)$ is precisely the degree of u modulo 2, hence the vertices which requires adjusting are precisely those of odd degree. Let $S = \{v_1, \ldots, v_s\}$ denote the set of vertices of odd degree (clearly, s is even). We adjust the representation as follows: add s new coordinates to all vectors. For every $u \notin S$, set all of its new coordinates to 0. For v_i, $1 \leq i \leq s$, set the i-th new coordinate to 1 and the remaining new coordinates to 0. In this manner, we reversed the parity of the v_i vectors, while preserving the dot product of v_i and v_j, guaranteeing this is a proper representation of G. This completes the proof of Theorem 5.3.1. ■

Remark 5.3.3: Lemma 5.3.2 can give better upper bounds for various graphs, by constructing proper representations of dimension strictly smaller than $|V|$. For instance, for every Eulerian graph $G = (V, E)$, the incidence matrix is a proper representation of G (there is no need to modify the parity of any of the vertices, since the degrees are all even). Since each column has precisely two occurrences of the value 1, the sum of all rows is 0 in $GF(2)$, hence the rank of the matrix is at most $|V|-1$. More generally, if G has k Eulerian connected components, then $x_\omega(G) \leq |V|-k$ (by creating a dependency in each set of rows corresponding to an Eulerian component). Finally, since the matrix whose rows are the vectors of the proper representation, B, satisfies either $BB^t = A$ or $BB^t = A+J$ (operating over $GF(2)$), where A is the adjacency matrix of G), then every proper representation f satisfies $\dim(f) \geq \min\{\mathrm{rank}(A), \mathrm{rank}(A+J)\}$ over $GF(2)$. In particular, if both A and $A+J$ are of full rank over $GF(2)$, then there cannot exist a proper representation which gives a better bound than $|V|$.

We now wish to extend our definition of x_ω to multi-graphs. Recall that without loss of generality, there are no parallel edges, hence a clique in a multi-graph G is a set where every two distinct vertices are adjacent, however, it contains no loops. We note that if we were to examine sets in G, where each two vertices are adjacent, and in addition, each vertex has a loop, then this notion would be equivalent to independent sets in the multi-graph complement $\overline{G}$, and would thus be treated by the results in the previous section.

Notice that equation (5.6) remains valid, by the same argument, when G and H are multi-graphs. It therefore follows that if a graph G satisfies $\omega(G) \geq 2$, or equivalently, if there are two adjacent vertices in G, each of which does not have a loop, then x_ω is well defined and satisfies equation (5.7).

If $\omega(G) = 0$, then every vertex of G has a loop, hence $\omega(G^{2n+1}) = 0$ and yet $\omega(G^{2n}) \geq 1$ for every n, thus the series $\frac{g(n)}{n}$ alternates between zero and non zero values. Indeed, it is easy to come up with examples for such graphs where this series does not converge (the disjoint union of 3 loops is an example: the second power, which is exactly the square lattice graph $L_2(3)$, contains a copy of K_3, hence the subseries of even indices does not converge to 0).

If $\omega(G) = 1$, then either the graph is simple (and hence edgeless), or there exist two vertices a and b, such that a has a loop and b does not. In this case, we can modify the clique in (5.5) to use the induced graph on $\{a, b\}$

instead of a copy of K_2:

$$\{v_2,\dots,v_r\}\times\{u_1\}\times\{aba\} \ \cup \ \{v_1\}\times\{u_2,\dots,u_s\}\times\{aab\} \tag{5.8}$$

We can therefore slightly modify the argument used on simple graphs, and obtain a similar result. The function $g(n)$ now satisfies the inequality:

$$g(m+n+3) \geq g(m)+g(n)-2$$

hence we can define $\hat{g}$ as:

$$\hat{g}(n) = g(n-3)-2$$

and obtain the following definition for x_ω:

$$x_\omega(G) = \lim_{n\to\infty}\frac{\omega(G^n)}{n} = \sup_n \frac{\omega(G^n)-2}{n+3} \tag{5.9}$$

Altogether, we have shown that x_ω, the limit of $\frac{g(n)}{n}$, exists for every multi-graph G satisfying $\omega(G) > 0$. Examining the even powers of G, it is clear that two possibly equal vertices u and v are adjacent in G^{2n} iff they are adjacent in $\overline{G}^{2n}$ (where $\overline{G}$ is the multi-graph complement of G, as defined in the previous section). Hence, we obtain the following proposition, analogous to Proposition 5.2.1:

Proposition 5.3.4. *For every multi-graph $G = (V, E)$ satisfying $\omega(G) > 0$, $x_\omega(G)$ is well defined. Furthermore, if in addition $\omega(\overline{G}) > 0$, where $\overline{G}$ is the multi-graph-complement of G, then $x_\omega(G) = x_\omega(\overline{G})$.*

We note that the upper bound of $|V|$ in Theorem 5.3.1 applies to multigraphs as well: Lemma 5.3.2 does not rely on the fact that G has no loops, and in the constructions of proper representations for G, we have already dealt with the scenario of having to modify the value of $f(u_i)\cdot f(u_i)$ for a subset of the vertices $\{u_i\} \subset V$. The loops merely effect the choice of the vertices whose parity we need to modify.

5.3.2 Properties of x_ω and bounds for codes

While defining x_ω in the previous section, we commented that the lower bound of 0 is trivially tight for edgeless graphs. It is interesting to state that $x_\omega(G)$ may be 0 even if the graph G is quite dense: recall that the powers of complete bipartite graphs are complete bipartite (Claim 5.2.3). Therefore, for every $k \geq 1$, $\omega(K_{m,n}^k) = 2$, and $x_\omega(K_{m,n}) = 0$.

It is now natural to ask whether $x_\omega(G) = 0$ holds for every (not necessarily complete) bipartite graph. This is false, as the following example shows: take P_4, the path on 4 vertices, $w - x - y - z$. The set $\{(w,x),(y,y),(z,y)\}$ is a triangle in P_4^2, hence (5.7) implies that $x_\omega(P_4) \geq \frac{1}{3} > 0$. However, adding the edge (w,z) completes P_4 into a cycle $C_4 = K_{2,2}$, which satisfies $x_\omega(K_{2,2}) = 0$ by the discussion above. This proves the following property of x_ω:

Claim 5.3.5. *x_ω is non-monotone with respect to the addition of edges.*

Recall the motivation of examining $g_3(n)$, the maximal number of vectors in $\{0,1,2\}^n$ such that the Hamming distance between every two is odd. We already noted in the introduction that $g_3(n) = \omega(K_3^n)$; it is now clear from the lower and upper bounds we have presented for x_ω that $g_3(n) = \Theta(n)$, and more generally, that when the alphabet is $\{0,\ldots,r-1\}$ for some fixed r, $g_r(n) = \Theta(n)$. The following holds for general complete graphs:

Theorem 5.3.6. *The complete graph K_r ($r \geq 3$) satisfies:*

$$x_\omega(G) = (1 - o(1))\, r \ ,$$

where the $o(1)$-term tends to 0 as r tends to infinity.

Proof. We first prove the following lemma, addressing the case of r being a prime power:

Lemma 5.3.7. *Let $r = p^k$ for some prime $p \geq 3$ and $k \geq 1$. Then:*

$$r - 1 - \frac{r}{r+2} \leq x_\omega(K_r) \leq r - 1$$

Proof. The upper bound of $r-1$ is derived from the remark following Theorem 5.3.1 (r is odd and hence K_r is Eulerian). For the lower bound, argue as follows: let $\mathcal{L}$ denote the set of all lines with finite slopes in the affine plane

$GF(p^k)$. Let $\{x_1, \ldots, x_{p^k}\}$ denote the elements of $GF(p^k)$, and represent each such line $\ell \in \mathcal{L}$, $\ell = ax + b$ by the vector:

$$f(\ell) = (a, ax_1 + b, ax_2 + b, \ldots, ax_{p^k} + b)$$

(i.e., represent ℓ by its slope followed by the y-coordinates of its set of points). Every two distinct lines $\ell_1, \ell_2 \in \mathcal{L}$ are either parallel ($a_1 = a_2$ and $b_1 \neq b_2$) or intersect in precisely one point ($x = (b_1 - b_2)(a_2 - a_1)^{-1}$). In both cases, precisely one coordinate in $f(\ell_1), f(\ell_2)$ is equal, hence the Hamming distance between them is p^k. Since p is odd, the above set of vectors forms a clique of size $|\mathcal{L}| = p^{2k}$ in $K_{p^k}^{p^k+1}$. Equation (5.7) yields:

$$x_\omega(K_{p^k}) \geq \frac{p^{2k} - 2}{(p^k + 1) + 1} = p^k - 1 - \frac{p^k}{p^k + 2} \ ,$$

as required. ■

There exists a $\frac{1}{2} < \Theta < 1$ such that for every sufficiently large n, the interval $[n - n^\Theta, n]$ contains a prime number (see, e.g., [69] for $\Theta = 23/42$). Combining this fact with the lower bound of the above lemma immediately implies the asymptotic result for every sufficiently large r. ■

Remark 5.3.8: Lemma 5.3.7 gives a lower bound of 1.4 for $x_\omega(K_3)$. Using a computer search, we improved this lower bound to 1.7 (compared to the upper bound of 2), by finding a clique of size 19 in K_3^9.

It is not difficult to see that the upper bounds of proper representations, given for cliques, can be extended to complete r-partite graphs, by assigning the same vector to all the vertices in a given color class. This is a special case of the following property, analogous to Claim 5.2.4:

Claim 5.3.9. *Let $G = (V, E)$ be a graph on the vertex set $V = [n]$. The r-blow-up of G, $G[r]$ (see Claim 5.2.4 for the definition) satisfies:*

$$x_\omega(G[r]) = x_\omega(G)$$

Furthermore, every maximum clique of $G[r]^k$ corresponds to a maximum clique of the same size of G^k.

Proof. Define the *pattern* of a vertex $v = (v_1, \ldots, v_k) \in G[r]^k$ to be the vector $w_v = (w_1, \ldots, w_k) \in G^k$, such that every coordinate of v belongs in $G[r]$ to the color class of the corresponding coordinate of w_v in G (i.e., v_i belongs to the independent set of size r which corresponds to w_i in $G[r]$). Let S be a maximum clique of $G[r]^k$; then every vertex $v \in S$ has a unique pattern in S (by definition, two vertices sharing the same pattern are disconnected in every coordinate). Thus, we can fix a vertex in each color class of $G[r]$ (note that this is an induced copy of G in $G[r]$), and without loss of generality, we can assume that these are the only vertices used in every $v \in S$. This completes the proof of the claim. ■

Corollary 5.3.10. *For every complete r-partite graph G, $\frac{r}{2} - 1 \leq x_\omega(G) \leq r$, and in addition, $x_\omega(G) = (1 - o(1))\, r$, where the $o(1)$-term tends to 0 as r tends to infinity.*

We have so far seen that for every graph G on n vertices and a maximum clique of size r, $\Omega(r) \leq x_\omega(G) \leq O(n)$. For complete graphs, $x_\omega(G) = (1 - o(1))r$, and one might suspect that $x_\omega(G)$ cannot be significantly larger than r. The following claim settles this issue, by examining self complementary Ramsey graphs (following the ideas of [18]):

Claim 5.3.11. *For every $n \in \mathbb{N}$ there is a graph G on n vertices, such that $\omega(G) < 2\lceil \log_2(n) \rceil$ and yet $x_\omega(G) \geq \frac{n-5}{3}$.*

Proof. In section 2.2 of [18], the authors prove the following lemma:

Lemma 5.3.12 ([18])**.** *For every n divisible by 4 there is a self complementary graph G on n vertices satisfying $\alpha(G) < 2\lceil \log_2(n) \rceil$.*

Set $n = 4m + r$ $(0 \leq r \leq 3)$, and let G be the disjoint union of a self-complementary graph H on $4m$ vertices, and r isolated vertices. By the lemma,

$$\omega(G) < 2\lceil \log_2(n) \rceil$$

Furthermore, if τ is an isomorphism mapping H to its complement, the set $\{(v, \tau(v)) : v \in V(H)\}$ is a clique of size $4m$ in G^2, since for every $u \neq v$, $uv \in E(G)$ iff $\tau(u)\tau(v) \notin E(G)$. Hence:

$$x_\omega(G) \geq \frac{\omega(G^2) - 2}{3} \geq \frac{n - r - 2}{3} \geq \frac{n-5}{3}$$

■

We note that a slightly weaker result can be proved rather easily and without using the lemma on self-complementary Ramsey graphs, by taking the disjoint union of a Ramsey graph and its complement. The lower bound on x_ω is again derived from a clique in G^2 of the form $\{(v, \tilde{v})\}$ where $\tilde{v}$ is the vertex corresponding to v in the complement graph. This construction gives, for every even $n \in \mathbb{N}$, a graph G on n vertices, satisfying $\omega(G) \leq 2\log_2(n)$ and yet $x_\omega(G) \geq \frac{n/2-2}{3} = \frac{n-4}{6}$.

5.4 Open problems

We conclude with several open problems related to x_α and x_ω:

Question 5.4.1. *Does every complete graph on $n \geq 4$ vertices, K_n, satisfy $x_\alpha(K_n) = \sqrt{n}$?*

Question 5.4.2. *What is the expected value of x_α for the random graph $\mathcal{G}_{n,\frac{1}{2}}$? What is the expected value of x_ω for the random graph $\mathcal{G}_{n,\frac{1}{2}}$?*

Question 5.4.3. *What is the precise value of $x_\omega(K_n)$ for $n \geq 3$?*

Question 5.4.4. *Is the problem of deciding whether $x_\alpha(G) > k$, for a given graph G and a given value k, decidable? Is the problem of deciding whether $x_\omega(G) > k$, for a given graph G and a given value k, decidable?*

Chapter 6

Graph p-powers, Delsarte, Hoffman, Ramsey and Shannon

The results of this chapter appear in [11]

The k-th p-power of a graph G is the graph on the vertex set $V(G)^k$, where two k-tuples are adjacent iff the number of their coordinates which are adjacent in G is not congruent to 0 modulo p. The clique number of powers of G is poly-logarithmic in the number of vertices, thus graphs with small independence numbers in their p-powers do not contain large homogenous subsets. We provide algebraic upper bounds for the asymptotic behavior of independence numbers of such powers, settling a conjecture of [10] up to a factor of 2. For precise bounds on some graphs, we apply Delsarte's linear programming bound and Hoffman's eigenvalue bound. Finally, we show that for any nontrivial graph G, one can point out specific induced subgraphs of large p-powers of G with neither a large clique nor a large independent set. We prove that the larger the Shannon capacity of $\overline{G}$ is, the larger these subgraphs are, and if G is the complete graph, then some p-power of G matches the bounds of the Frankl-Wilson Ramsey construction, and is in fact a subgraph of a variant of that construction.

6.1 Introduction

The k-th Xor graph power of a graph G, $G^{\oplus k}$, is the graph whose vertex set is the cartesian product $V(G)^k$, where two k-tuples are adjacent iff an odd number of their coordinates is adjacent in G. This product was used in [102] to construct edge colorings of the complete graph with two colors, containing a smaller number of monochromatic copies of K_4 than the expected number of such copies in a random coloring.

In [10], the authors studied the independence number, α, and the clique number, ω, of high Xor powers of a fixed graph G, motivated by problems in Coding Theory: cliques and independent sets in such powers correspond to maximal codes satisfying certain natural properties. It is shown in [10] that, while the clique number of $G^{\oplus k}$ is linear in k, the independence number $\alpha(G^{\oplus k})$ grows exponentially: the limit $\alpha(G^{\oplus k})^{\frac{1}{k}}$ exists, and is in the range $[\sqrt{|V(G)|}, |V(G)|]$. Denoting this limit by $x_\alpha(G)$, the problem of determining $x_\alpha(G)$ for a given graph G proves to be extremely difficult, even for simple families of graphs. Using spectral techniques, it is proved in [10] that $x_\alpha(K_n) = 2$ for $n \in \{2, 3, 4\}$, where K_n is the complete graph on n vertices, and it is conjectured that $x_\alpha(K_n) = \sqrt{n}$ for every $n \geq 4$. The best upper bound given in [10] on $x_\alpha(K_n)$ for $n \geq 4$ is $n/2$.

The graph product we introduce in this chapter, which generalizes the Xor product, is motivated by Ramsey Theory. In [49], Erdős proved the existence of graphs on n vertices without cliques or independent sets of size larger than $O(\log n)$ vertices, and that in fact, almost every graph satisfies this property. Ever since, there have been many attempts to provide explicit constructions of such graphs. Throughout the chapter, without being completely formal, we call a graph “Ramsey” if it has neither a “large” clique nor a “large” independent set. The famous Ramsey construction of Frankl and Wilson [56] provided a family of graphs on n vertices, FW_n, with a bound of $\exp\left(\sqrt{(2+o(1))\log n \log\log n}\right)$ on the independence and clique numbers, using results from Extremal Finite Set Theory. Thereafter, constructions with the same bound were produced in [8] using polynomial spaces and in [62] using low degree matrices. Recently, the old Frankl-Wilson record was broken in [21], where the authors provide, for any $\varepsilon > 0$, a polynomial-time algorithm for constructing a Ramsey graph on n vertices without cliques or independent sets on $\exp\left((\log n)^{\varepsilon}\right)$ vertices. The disadvantage of this latest revolutionary construction is that it involves a complicated algorithm, from which it is hard to tell the structure of the resulting graph.

Relating the above to graph products, the Xor product may be viewed as an operator, $\oplus_k$, which takes a fixed input graph G on n vertices, and produces a graph on n^k vertices, $H = G^{\oplus k}$. The results of [10] imply that the output graph H satisfies $\omega(H) \leq nk = O(\log(|V(H)|))$, and that if G is a nontrivial d-regular graph, then H is d'-regular, with $d' \to \frac{1}{2}|V(H)|$ as k tends to infinity. Thus, $\oplus_k$ transforms any nontrivial d-regular graph into a random looking graph, in the sense that it has an edge density of roughly $\frac{1}{2}$ and a logarithmic clique number. However, the lower bound $\alpha(H) \geq \sqrt{|V(H)|}$, which holds for every even k, implies that $\oplus_k$ cannot be used to produce good Ramsey graphs.

In order to modify the Xor product into a method for constructing Ramsey graphs, one may try to reduce the high lower bound on the independence numbers of Xor graph powers. Therefore, we consider a generalization of the Xor graph product, which replaces the modulo 2 (adjacency of two k-tuples is determined by the parity of the number of adjacent coordinates) with some possibly larger modulo $p \in \mathbb{N}$. Indeed, we show that by selecting a larger p, the lower bound on the independence number, $\alpha(H)$, is reduced from $\sqrt{|V(H)|}$ to $|V(H)|^{1/p}$, at the cost of a polynomial increase in $\omega(H)$. The generalized product is defined as follows:

Definition 6.1. *Let $k, p \in \mathbb{N}$. The k-th p-power of a graph G, denoted by $G^{k_{(p)}}$, is the graph whose vertex set is the cartesian product $V(G)^k$, where two k-tuples are adjacent iff the number of their coordinates which are adjacent in G is not congruent to 0 modulo p, that is:*

$$(u_1, \ldots, u_k)\ (v_1, \ldots, v_k) \in E(G^k) \quad \textit{iff} \quad |\{i : u_i v_i \in E(G)\}| \not\equiv 0 \pmod{p} \ .$$

Throughout the chapter, we use the abbreviation G^k for $G^{k_{(p)}}$ when there is no danger of confusion.

In Section 6.2 we show that the limit $\alpha(G^k)^{\frac{1}{k}}$ exists and is equal to $\sup_k \alpha(G^k)^{\frac{1}{k}}$; denote this limit by $x_\alpha^{(p)}$. A simple lower bound on $x_\alpha^{(p)}$ is

$|V(G)|^{1/p}$, and algebraic arguments show that this bound is nearly tight for the complete graph: $x_\alpha^{(p)}(K_n) = O(n^{1/p})$. In particular, we obtain that

$$\sqrt{n} \leq x_\alpha(K_n) = x_\alpha^{(2)}(K_n) \leq 2\sqrt{n-1}\ ,$$

improving the upper bound of $n/2$ for $n \geq 4$ given in [10], and determining that the behavior of x_α for complete graphs is as stated in Question 4.1 of [10] up to a factor of 2.

For the special case $G = K_n$, it is possible to apply Coding Theory techniques in order to bound $x_\alpha^{(p)}(G)$. The problem of determining $x_\alpha^{(p)}(K_n)$ can be translated into finding the asymptotic maximum size of a code over the alphabet $[n]$, in which the Hamming distance between any two codewords is divisible by p. The related problem for *linear* codes over a field has been well studied: see, e.g., [105] for a survey on this subject. However, as we later note in Section 6.2, the general non-linear case proves to be quite different, and the upper bounds on linear divisible codes do not hold for $x_\alpha^{(p)}(K_n)$. Yet, other methods for bounding sizes of codes are applicable. In Section 6.3 we demonstrate the use of Delsarte's linear programming bound in order to obtain precise values of $\alpha(K_3^{k(3)})$. We show that $\alpha(K_3^{k(3)}) = 3^{k/2}$ whenever $k \equiv 0 \pmod 4$, while $\alpha(K_3^{k(3)}) < \frac{1}{2}3^{k/2}$ for $k \equiv 2 \pmod 4$, hence the series $\alpha(K_3^{k+1(3)})/\alpha(K_3^{k(3)})$ does not converge to a limit.

Section 6.4 gives a general bound on $x_\alpha^{(p)}$ for d-regular graphs in terms of their eigenvalues, using Hoffman's eigenvalue bound. The eigenvalues of p-powers of G are calculated using tensor products of matrices over $\mathbb{C}$, in a way somewhat similar to performing a Fourier transform on the adjacency matrix of G. This method may also be used to derive tight results on $\alpha(G^{k(p)})$, and we demonstrate this on the above mentioned case of $p = 3$ and the graph K_3, where we compare the results with those obtained in Section 6.3 by the Delsarte bound.

Section 6.5 shows, using tools from linear algebra, that indeed the clique number of $G^{k(p)}$ is poly-logarithmic in k, and thus p-powers of graphs attaining the lower bound of $x_\alpha^{(p)}$ are Ramsey. We proceed to show a relation between the Shannon capacity of the complement of G, $c(\overline{G})$, and the Ramsey properties of p-powers of G. Indeed, for any nontrivial graph G, we can point out a large Ramsey induced subgraph of some p-power of G. The larger $c(\overline{G})$ is, the larger these Ramsey subgraphs are. When $G = K_p$ for some prime p, we obtain that $H = K_p^{p^2(p)}$ is a Ramsey graph matching the bound of Frankl-Wilson, and in fact, H contains an induced subgraph which is a modified variant of FW_{N_1} for some N_1, and is contained in another variant of FW_{N_2} for some N_2. The method of proving these bounds on $G^{k(p)}$ provides yet another (simple) proof for the Frankl-Wilson result.

6.2 Algebraic lower and upper bounds on $x_\alpha^{(p)}$

In this section, we define the parameter $x_\alpha^{(p)}$, and provide lower and upper bounds for it. The upper bounds follow from algebraic arguments, using graph representation by polynomials.

6.2.1 The limit of independence numbers of p-powers

The following lemma establishes that $x_\alpha^{(p)}$ exists, and gives simple lower and upper bounds on its range for graphs on n vertices:

Lemma 6.2.1. *Let G be a graph on n vertices, and let $p \geq 2$. The limit of $\alpha(G^{k(p)})^{\frac{1}{k}}$ as $k \to \infty$ exists, and, denoting it by $x_\alpha^{(p)}(G)$, it satisfies:*

$$n^{1/p} \leq x_\alpha^{(p)}(G) = \sup_k \alpha(G^{k(p)})^{\frac{1}{k}} \leq n \ .$$

Proof. Observe that if I and J are independent sets of G^k and G^l respectively, then the set $I \times J$ is an independent set of G^{k+l}, as the number of adjacent coordinates between any two k-tuples of I and between any two l-tuples of J is $0 \pmod p$. Therefore, the function $g(k) = \alpha(G^k)$ is super-multiplicative and strictly positive, and we may apply Fekete's Lemma (cf., e.g., [76], p. 85) to obtain that the limit of $\alpha(G^k)^{\frac{1}{k}}$ as $k \to \infty$ exists, and satisfies:

$$\lim_{k\to\infty} \alpha(G^k)^{\frac{1}{k}} = \sup_k \alpha(G^k)^{\frac{1}{k}} \ . \tag{6.1}$$

Clearly, $\alpha(G^k) \leq n^k$, and it remains to show the lower bound on $x_\alpha^{(p)}$. Notice that the following set is an independent set of G^p:

$$I = \{\ (u, \ldots, u)\ :\ u \in V(G)\} \subset G^p\ ,$$

since for all $u, v \in V(G)$, there are either 0 or p adjacent coordinates between the two corresponding p-tuples in I. By (6.1), we obtain that $x_\alpha^{(p)}(G) \geq |I|^{1/p} = n^{1/p}$. ■

6.2.2 Bounds on $x_\alpha^{(p)}$ of complete graphs

While the upper bound $|V(G)|$ on $x_\alpha^{(p)}(G)$ is clearly attained by an edgeless graph, proving that a family of graphs attains the lower bound requires some effort. The next theorem states that complete graphs achieve the lower bound of Lemma 6.2.1 up to a constant factor:

Theorem 6.2.2. *The following holds for all integer $n, p \geq 2$:*

$$x_\alpha^{(p)}(K_n) \leq 2^{H(1/p)}(n-1)^{1/p} \ , \tag{6.2}$$

where $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function. In particular, $x_\alpha^{(p)}(K_n) = \Theta(n^{1/p})$. In the special case where $n = p = q^r$ for some prime q and $r \geq 1$, the lower bound roughly matches upper bound:

$$p^{\frac{2}{p+1}} \leq x_\alpha^{(p)}(K_p) \leq \left(\mathrm{e}p^2\right)^{1/p} \ .$$

Taking $p = 2$ and noting that $H(\frac{1}{2}) = 1$, we immediately obtain the following corollary for Xor graph products, which determines the asymptotic behavior of x_α for complete graphs:

Corollary 6.2.3. *For all $n \geq 2$, the complete graph on n vertices satisfies*

$$\sqrt{n} \leq x_\alpha(K_n) \leq 2\sqrt{n-1} \ .$$

Proof of Theorem 6.2.2. The upper bound will follow from an argument on polynomial representations, an approach which was used in [8] to bound the Shannon capacity of certain graphs. Take $k \geq 1$, and consider the graph $H = K_n^k$. For every vertex of H, $u = (u_1, \ldots, u_k)$, we define the following polynomial in $\mathbb{R}[x_{i,j}]$, where $i \in [k]$, $j \in [n]$:

$$f_u(x_{1,1}, \ldots, x_{k,n}) = \prod_{t=1}^{\lfloor k/p \rfloor} \left(k - tp - \sum_{i=1}^{k} x_{i,u_i}\right) \ . \tag{6.3}$$

Next, give the following assignment of values for $\{x_{i,j}\}$, x_v, to each $v = (v_1, \ldots, v_k) \in V(H)$:

$$x_{i,j} = \delta_{v_i,j} \ , \tag{6.4}$$

where δ is the Kronecker delta. Definitions (6.3) and (6.4) imply that for every two such vertices $u = (u_1, \ldots, u_k)$ and $v = (v_1, \ldots, v_k)$ in $V(H)$:

$$f_u(x_v) = \prod_{t=1}^{\lfloor k/p \rfloor} \left(k - tp - \sum_{i=1}^{k} \delta_{u_i,v_i} \right) = \prod_{t=1}^{\lfloor k/p \rfloor} \left(|\{i \ : \ u_i \neq v_i\}| - tp \right) \ . \tag{6.5}$$

Notice that, by the last equation, $f_u(x_u) \neq 0$ for all $u \in V(H)$, and consider two distinct non-adjacent vertices $u, v \in V(H)$. The Hamming distance between u and v (considered as vectors in $\mathbb{Z}_n^k$) is by definition $0 \pmod p$ (and is not zero). Thus, (6.5) implies that $f_u(x_v) = 0$.

Recall that for all u, the assignment x_u gives values $x_{i,j} \in \{0, 1\}$ for all i, j, and additionally, $\sum_{j=1}^{n} x_{i,j} = 1$ for all i. Therefore, it is possible to replace all occurrences of $x_{i,n}$ by $1 - \sum_{j=1}^{n-1} x_{i,j}$ in each f_u, and then proceed and reduce the obtained result modulo the polynomials:

$$\bigcup_{i \in [k]} \left(\{x_{i,j}^2 - x_{i,j} : j \in [n]\} \ \cup \ \{x_{i,j} x_{i,l} : j, l \in [n], j \neq l\} \right) \ ,$$

without affecting the value of the polynomials on the above defined substitutions. In other words, after replacing $x_{i,n}$ by $1 - \sum_{j<n} x_{i,j}$, we repeatedly replace $x_{i,j}^2$ by $x_{i,j}$, and let all the monomials containing $x_{i,j} x_{i,l}$ for $j \neq l$ vanish. This gives a set of multi-linear polynomials $\{\tilde{f}_u\}$ satisfying:

$$\begin{cases} \tilde{f}_u(x_u) \neq 0 & \text{for all } u \in V(H) \\ \tilde{f}_u(x_v) = 0 & \text{for } u \neq v \ , \ uv \notin E(H) \end{cases} ,$$

where the monomials of $\tilde{f}_u$ are of the form $\prod_{t=1}^{r} x_{i_t,j_t}$ for some $0 \leq r \leq \lfloor \frac{k}{p} \rfloor$, a set of pairwise distinct indices $\{i_t\} \subset [k]$ and indices $\{j_t\} \subset [n-1]$.

Let $\mathcal{F} = \text{Span}(\{\tilde{f}_u : u \in V(H)\})$, and let I denote a maximum independent set of H. A standard argument shows that $F = \{\tilde{f}_u : u \in I\}$ is linearly independent in $\mathcal{F}$. Indeed, suppose that $\sum_{u \in I} a_u \tilde{f}_u(x) = 0$; then substituting $x = x_v$ for some $v \in I$ gives $a_v = 0$. It follows that $\alpha(H) \leq \dim \mathcal{F}$, and thus:

$$\alpha(H) \leq \sum_{r=0}^{\lfloor k/p \rfloor} \binom{k}{r} (n-1)^r \leq \left(2^{H(1/p)}(n-1)^{1/p}\right)^k \ , \qquad (6.6)$$

where in the last inequality we used the fact that $\sum_{i \leq \lambda n} \binom{n}{i} \leq 2^{nH(\lambda)}$ (cf., e.g., the remark following Corollary 4.2 in [7], and also [19] p. 242). Taking the k-th root and letting k grow to ∞, we obtain:

$$x_\alpha^{(p)}(K_n) \leq 2^{H(1/p)}(n-1)^{1/p} \ ,$$

as required.

In the special case of K_p (that is, $n = p$), note that: $2^{H(\frac{1}{p})} = p^{\frac{1}{p}}(\frac{p}{p-1})^{\frac{p-1}{p}} \leq (\mathrm{e}p)^{\frac{1}{p}}$ and hence in this case $x_\alpha^{(p)}(K_p) \leq (\mathrm{e}p^2)^{1/p}$. If $p = q^r$ is a prime-power we can provide a nearly matching lower bound for $x_\alpha^{(p)}(K_p)$ using a construction of [10], which we shortly describe for the sake of completeness.

Let $\mathcal{L}$ denote the set of all lines with finite slopes in the affine plane $GF(p)$, and write down the following vector w_ℓ for each $\ell \in \mathcal{L}$, $\ell = ax+b$ for some $a, b \in GF(p)$:

$$w_\ell = (a, ax_1 + b, ax_2 + b, \ldots, ax_p + b) \ ,$$

where $x_1, \ldots, x_p$ denote the elements of $GF(p)$. For every two distinct lines ℓ, ℓ', if $\ell \| \ell'$ then $w_\ell, w_{\ell'}$ has a single common coordinate (the slope a). Otherwise, $w_\ell, w_{\ell'}$ has a single common coordinate, which is the unique intersection of ℓ, ℓ'. In any case, we obtain that the Hamming distance of w_ℓ and $w_{\ell'}$ is p, hence $W = \{w_\ell : \ell \in \mathcal{L}\}$ is an independent set in K_p^{p+1}. By (6.1), we deduce that:

$$x_\alpha^{(p)}(K_p) \geq p^{\frac{2}{p+1}} \ ,$$

completing the proof. ■

Remark 6.2.4: The proof of Theorem 6.2.2 used representation of the vertices of K_n^k by polynomials of kn variables over $\mathbb{R}$. It is possible to prove a similar upper bound on $x_\alpha^{(p)}(K_n)$ using a representation by polynomials of k variables over $\mathbb{R}$. To see this, use the natural assignment of $x_i = v_i$ for $v = (v_1, \ldots, v_k)$, denoting it by x_v, and assign the following polynomial to

$u = (u_1, \ldots, u_k)$:

$$f_u(x_1, \ldots, x_k) = \prod_{t=1}^{\lfloor k/p \rfloor} \left(k - tp - \sum_{i=1}^{k} \prod_{\substack{j=1 \\ j \neq u_i}}^{n} \frac{x_i - j}{u_i - j} \right) . \qquad (6.7)$$

The expression $\prod_{j \neq u_i} \frac{x_i - j}{u_i - j}$ is the monomial of the Lagrange interpolation polynomial, and is equal to δ_{x_i, u_i}. Hence, we obtain that $f_u(x_u) \neq 0$ for any vertex u, whereas $f_u(x_v) = 0$ for any two distinct non-adjacent vertices u, v. As the Lagrange monomials yield values in $\{0, 1\}$, we can convert each f_u to a multi-linear combination of these polynomials, $\tilde{f}_u$, while retaining the above properties. Note that there are n possibilities for the Lagrange monomials (determined by the value of u_i), and it is possible to express one as a linear combination of the rest. From this point, a calculation similar to that in Theorem 6.2.2 for the dimension of $\mathrm{Span}(\{\tilde{f}_u : u \in V\})$ gives the upper bound (6.2).

Remark 6.2.5: The value of $\alpha(K_n^{k(p)})$ corresponds to a maximum size of a code C of k-letter words over $\mathbb{Z}_n$, where the Hamming distance between any two codewords is divisible by p. The case of *linear* such codes when $\mathbb{Z}_n$ is a field, that is, we add the restriction that C is a linear subspace of $\mathbb{Z}_n^k$, has been thoroughly studied; it is equivalent to finding a linear subspace of $\mathbb{Z}_n^k$ of maximal dimension, such that the Hamming weight of each element is divisible by p. It is known for this case that if p and n are relatively prime, then the dimension of C is at most k/p (see [104]), and hence the size of C is at most $n^{k/p}$. However, this bound does not hold for the non-linear case (notice that this bound corresponds to the lower bound of Lemma 6.2.1). We give two examples of this:

1. Take $p = 3$ and $n = 4$. The divisible code bound implies an upper bound of $4^{1/3} \approx 1.587$, and yet $x_\alpha^{(3)}(K_4) \geq \sqrt{3} \approx 1.732$. This follows from the geometric construction of Theorem 6.2.2, which provides an independent set of size 9 in $K_3^{4(3)} \subset K_4^{4(3)}$, using only the coordinates $\{0, 1, 2\}$ (this result can be slightly improved by adding an all-3 vector to the above construction in the 12-th power).

2. Take $p = 3$ and $n = 2$. The linear code bound is $2^{1/3} \approx 1.26$, whereas the following construction shows that $\alpha(K_2^{12^{(3)}}) \geq 24$, implying that $x_\alpha^{(3)}(K_2) \geq 24^{1/12} \approx 1.30$. Let $\{v_1, \ldots, v_{12}\}$ denote the rows of a binary Hadamard matrix of order 12 (such a matrix exists by Paley's Theorem, cf. e.g. [65]). For all $i \neq j$, v_i and v_j have precisely 6 common coordinates, and hence, the set $I = \{v_i\} \cup \{\overline{v}_i\}$ (where $\overline{v}_i$ denotes the complement of v_i modulo 2) is an independent set of size 24 in $K_2^{12^{(3)}}$. In fact, I is a maximum independent set of $K_2^{12^{(3)}}$, as Delsarte's linear programming bound (described in Section 6.3) implies that $\alpha(K_2^{12^{(3)}}) \leq 24$.

6.2.3 The value of $x_\alpha^{(3)}(K_3)$

While the upper bound of Theorem 6.2.2 on $x_\alpha^{(p)}(K_n)$ is tight up to a constant factor, the effect of this constant on the independence numbers is exponential in the graph power, and we must resort to other techniques in order to obtain more accurate bounds. For instance, Theorem 6.2.2 implies that:

$$1.732 \approx \sqrt{3} \leq x_\alpha^{(3)}(K_3) \leq 2^{H(\frac{1}{3})} 2^{\frac{1}{3}} = \frac{3}{2^{1/3}} \approx 2.381 \ .$$

In Sections 6.3 and 6.4, we demonstrate the use of Delsarte's linear programming bound and Hoffman's eigenvalue bound for the above problem, and in both cases obtain the exact value of $\alpha(K_3^{k^{(3)}})$ under certain divisibility conditions. However, if we are merely interested in the value of $x_\alpha^{(3)}(K_3)$, a simpler consideration improves the bounds of Theorem 6.2.2 and shows that $x_\alpha^{(3)}(K_3) = \sqrt{3}$:

Lemma 6.2.6. *For any $k \geq 1$, $\alpha(K_3^{k^{(3)}}) \leq 3 \cdot \sqrt{3}^k$, and in particular, $x_\alpha^{(3)}(K_3) = \sqrt{3}$.*

Proof. Treating vertices of K_3^k as vectors of $\mathbb{Z}_3^k$, notice that every two vertices $x = (x_1, \ldots, x_k)$ and $y = (y_1, \ldots, y_k)$ satisfy:

$$\sum_{i=1}^{k} (x_i - y_i)^2 \equiv |\{i : x_i \neq y_i\}| \pmod 3 \ ,$$

and hence if I is an independent set in K_3^k, then:

$$\sum_i (x_i - y_i)^2 \equiv 0 \pmod 3 \text{ for all } x, y \in I \; .$$

Let I denote a maximum independent set of K_3^k, and let $I_c = \{x \in I : \sum_i x_i^2 \equiv c \pmod 3\}$ for $c \in \{0, 1, 2\}$. For every $c \in \{0, 1, 2\}$ we have:

$$\sum_i (x_i - y_i)^2 = 2c - 2x \cdot y \equiv 0 \pmod 3 \text{ for all } x, y \in I_c,$$

and hence $x \cdot y = c$ for all $x, y \in I_c$. Choose c for which $|I_c| \geq |I|/3$, and subtract an arbitrary element $z \in I_c$ from all the elements of I_c. This gives a set J of size at least $|I|/3$, which satisfies:

$$x \cdot y = 0 \text{ for all } x, y \in J \; .$$

Since $\mathrm{Span}(J)$ is a self orthogonal subspace of $\mathbb{Z}_3^k$, its dimension is at most $k/2$, and hence $|J| \leq 3^{k/2}$. Altogether, $\alpha(K_3^k) \leq 3 \cdot \sqrt{3}^k$, as required. ■

6.3 Delsarte's linear programming bound for complete graphs

In this section, we demonstrate how Delsarte's linear programming bound may be used to derive precise values of independence numbers in p-powers of complete graphs. As this method was primarily used on binary codes, we include a short proof of the bound for a general alphabet.

6.3.1 Delsarte's linear programming bound

The linear programming bound follows from the relation between the distance distribution of codes and the Krawtchouk polynomials, defined as follows:

Definition 6.2. *Let $n \in \mathbb{N}$ and take $q \geq 2$. The Krawtchouk polynomials $\mathcal{K}_k^{n;q}(x)$ for $k = 0, \ldots, n$ are defined by:*

$$\mathcal{K}_k^{n;q}(x) = \sum_{j=0}^{k} \binom{x}{j} \binom{n-x}{k-j} (-1)^j (q-1)^{k-j} \; . \tag{6.8}$$

Definition 6.3. *Let C be an n-letter code over the alphabet $\{1,\ldots,q\}$. The distance distribution of C, $B_0, B_1, \ldots, B_n$, is defined by:*

$$B_k = \frac{1}{|C|}|\{(w_1,w_2) \in C^2 : \delta(w_1,w_2) = k\}| \quad (k = 0,\ldots,n)\ ,$$

where δ denotes the Hamming distance.

The Krawtchouk polynomials $\{\mathcal{K}_k^{n;q}(x)\}$ are sometimes defined with a normalizing factor of q^{-k}. Also, it is sometimes customary to define the distance distribution with a different normalizing factor, letting $A_k = \frac{B_k}{|C|}$, in which case A_k is the probability that a random pair of codewords has a Hamming distance k.

The Krawtchouk polynomials $\{\mathcal{K}_k^{n;q} : k = 0,\ldots,n\}$ form a system of orthogonal polynomials with respect to the weight function

$$w(x) = \frac{n!}{\Gamma(1+x)\Gamma(n+1-x)}(q-1)^x\ ,$$

where Γ is the gamma function. For further information on these polynomials see, e.g., [99].

Delsarte [43] (see also [85]) presented a remarkable method for bounding the maximal size of a code with a given set of restrictions on its distance distribution. This relation is given in the next proposition, for which we include a short proof:

Proposition 6.3.1. *Let C be a code of n-letter words over the alphabet $[q]$, whose distance distribution is $B_0,\ldots,B_n$. The following holds:*

$$\sum_{i=0}^{n} B_i \mathcal{K}_k^{n;q}(i) \geq 0 \quad \textit{for all } k = 0,\ldots,n\ . \tag{6.9}$$

Proof. Let $G = \mathbb{Z}_q^n$, and for every two functions $f,g : G \to \mathbb{C}$, define (as usual) their inner product $\langle f,g\rangle$ and their delta-convolution, $f * g$, as:

$$\langle f,g\rangle = \int_G f(x)\overline{g(x)}dx = \frac{1}{|G|}\sum_{T\in G} f(T)\overline{g(T)}\ ,$$

$$(f*g)(s) = \int_G f(x)\overline{g(x-s)}dx\ .$$

Denoting the Fourier expansion of f by: $f = \sum_{S \in G} \widehat{f}(S)\chi_S$, where $\chi_S(x) = \omega^{S \cdot x}$ and ω is the q-th root of unity, it follows that for any $k = 0, \ldots, n$:

$$\sum_{S \in G : |S| = k} \widehat{f}(S) = \frac{1}{|G|} \sum_{i=0}^{n} \mathcal{K}_k^{n;q}(i) \sum_{T \in G : |T| = i} f(T) \ , \tag{6.10}$$

where $|S|$ and $|T|$ denote the Hamming weights of $S, T \in G$. Since the delta-convolution satisfies:

$$\widehat{f * g}(S) = \widehat{f}(S)\overline{\widehat{g}(S)} \ ,$$

every f satisfies:

$$\widehat{f * f}(S) = |\widehat{f}(S)|^2 \geq 0 \ . \tag{6.11}$$

Let f denote the characteristic function of the code C, $f(x) = \mathbf{1}_{\{x \in C\}}$, and notice that:

$$(f * f)(S) = \frac{1}{|G|} \sum_{T \in G} f(T)\overline{f(T-S)} = \frac{1}{|G|} |\{T : T, T - S \in C\}| \ ,$$

and thus:

$$B_i = \frac{|G|}{|C|} \sum_{T : |T| = i} (f * f)(T) \ . \tag{6.12}$$

Putting together (6.10), (6.11) and (6.12), we obtain:

$$0 \leq \sum_{S : |S| = k} \widehat{f * f}(S) = \frac{1}{|G|} \sum_{i=0}^{n} \mathcal{K}_k^{n;q}(i) \sum_{T : |T| = i} (f * f)(T) = \frac{|C|}{|G|^2} \sum_{i=0}^{n} \mathcal{K}_k^{n;q}(i) B_i \ ,$$

as required. ■

Let $F \subset [n]$ be a set of forbidden distances between distinct codewords. Since $|C| = \sum_i B_i$, the following linear program provides an upper bound on the size of any code with no pairwise distances specified by F:

$$\text{maximize } \textstyle\sum_i B_i \text{ subject to the constraints:}$$
$$\left\{ \begin{array}{c} B_0 = 1 \\ B_i \geq 0 \text{ for all } i \\ B_i = 0 \text{ for all } i \in F \\ \sum_{i=0}^{n} B_i \mathcal{K}_k^{n;q}(i) \geq 0 \ \text{ for all } k = 0, \ldots, n \end{array} \right. .$$

By examining the dual program, it is possible to formulate this bound as a minimization problem. The following proposition has been proved in various special cases, (cf., e.g., [44], [77]). For the sake of completeness, we include a short proof of it.

Proposition 6.3.2. *Let C be a code of n-letter words over the alphabet $[q]$, whose distance distribution is $B_0, \ldots, B_n$. Let $P(x) = \sum_{k=0}^{n} \alpha_k \mathcal{K}_k^{n;q}(x)$ denote an n-degree polynomial over $\mathbb{R}$. If $P(x)$ has the following two properties:*

$$\alpha_0 > 0 \quad \textit{and } \alpha_i \geq 0 \quad \textit{for all } i = 1, \ldots, n \ , \tag{6.13}$$

$$P(d) \leq 0 \quad \textit{whenever } B_d > 0 \textit{ for } d = 1, \ldots, n \ , \tag{6.14}$$

then $|C| \leq P(0)/\alpha_0$.

Proof. The Macwilliams transform of the vector $(B_0, \ldots, B_n)$ is defined as follows:

$$B'_k = \frac{1}{|C|} \sum_{i=0}^{n} \mathcal{K}_k^{n;q}(i) B_i \ . \tag{6.15}$$

By the Delsarte inequalities (stated in Proposition 6.3.1), $B'_k \geq 0$, and furthermore:

$$B'_0 = \frac{1}{|C|} \sum_{i=0}^{n} \mathcal{K}_0^{n;q}(i) B_i = \frac{1}{|C|} \sum_{i} B_i = 1 \ .$$

Therefore, as (6.13) guarantees that $\alpha_i \geq 0$ for $i > 0$, we get:

$$\sum_{k=0}^{n} \alpha_k B'_k \geq \alpha_0 \ . \tag{6.16}$$

On the other hand, $B_0 = 1$, and by (6.14), whenever $B_i > 0$ for some $i > 0$ we have $P(i) \leq 0$, thus:

$$\sum_{i=0}^{n} B_i P(i) \leq P(0) \ . \tag{6.17}$$

Combining (6.16) and (6.17) with (6.15) gives:

$$\alpha_0 \leq \sum_{k=0}^{n} \alpha_k B'_k = \frac{1}{|C|} \sum_{i=0}^{n} B_i \sum_{k=0}^{n} \alpha_k \mathcal{K}_k^{n;q}(i) = \frac{1}{|C|} \sum_{i=0}^{n} B_i P(i) \leq \frac{P(0)}{|C|} \ ,$$

and the result follows. ■

We proceed with an application of the last proposition in order to bound the independence numbers of p-powers of complete graphs. In this case, the distance distribution is supported by $\{i : i \equiv 0 \pmod p\}$, and in Section 6.3.2 we present polynomials which satisfy the properties of Proposition 6.3.2 and provide tight bounds on $\alpha(K_3^{k_{(3)}})$.

6.3.2 Improved estimations of $\alpha(K_3^{k_{(3)}})$

Recall that the geometric construction of Theorem 6.2.2 describes an independent set of size p^2 in $K_p^{p+1_{(p)}}$ for every p which is a prime-power. In particular, this gives an independent set of size $3^{k/2}$ in $K_3^{k_{(3)}}$ for every $k \equiv 0 \pmod 4$. Using Proposition 6.3.2 we are able to deduce that indeed $\alpha(K_3^k) = 3^{k/2}$ whenever $k \equiv 0 \pmod 4$, whereas for $k \equiv 2 \pmod 4$ we prove that $\alpha(K_3^k) < \frac{1}{2}3^{k/2}$.

Theorem 6.3.3. *The following holds for any even integer k:*

$$\begin{cases} \alpha(K_3^k) = 3^{k/2} & k \equiv 0 \pmod 4 \\ \frac{1}{3}3^{k/2} \leq \alpha(K_3^k) < \frac{1}{2}3^{k/2} & k \equiv 2 \pmod 4 \end{cases} .$$

Proof. Let k be an even integer, and define the following polynomials:

$$P(x) = \frac{2}{3}3^{k/2} + \sum_{\substack{t=1 \\ t \not\equiv 0 (\text{mod } 3)}}^{k} \mathcal{K}_t^{k;3}(x) \ , \tag{6.18}$$

$$Q(x) = \frac{2}{3}3^{k/2} + \sum_{\substack{t=0 \\ t \equiv 0 (\text{mod } 3)}}^{k} \mathcal{K}_t^{k;3}(x) \ . \tag{6.19}$$

Clearly, both P and Q satisfy (6.13), as $\mathcal{K}_0^{n;q} = 1$ for all n, q. It remains to show that P, Q satisfy (6.14) and to calculate $P(0), Q(0)$. As the following calculation will prove useful later on, we perform it for a general alphabet q and a general modulo p. Denoting the q-th root of unity by $\omega = \mathrm{e}^{2\pi i/q}$, we have:

$$\begin{aligned}\sum_{\substack{t=0\\ t\equiv 0 (\mathrm{mod}\ p)}}^{k} \mathcal{K}_t^{k;q}(s) &= \sum_{\substack{t=0\\ t\equiv 0 (\mathrm{mod}\ p)}}^{k} \sum_{j=0}^{t} \binom{s}{j}\binom{k-s}{t-j}(-1)^j (q-1)^{t-j} \\ &= \sum_{j=0}^{s}\binom{s}{j}(-1)^j \sum_{\substack{l=0\\ j+l\equiv 0 (\mathrm{mod}\ p)}}^{k-s} \binom{k-s}{l}(q-1)^l \\ &= \sum_{j=0}^{s}\binom{s}{j}(-1)^j \sum_{l=0}^{k-s}\binom{k-s}{l}(q-1)^l \frac{1}{q}\sum_{t=0}^{q-1}\omega^{(j+l)t} \\ &= \delta_{s,0}\cdot q^{k-1} + \frac{1}{q}\sum_{t=1}^{q-1}(1+(q-1)\omega^t)^{k-s}(1-\omega^t)^s \ , \qquad (6.20)\end{aligned}$$

where the last equality is by the fact that: $\sum_{j=0}^{s}\binom{s}{j}(-1)^j = \delta_{s,0}$, and therefore the summand for $t=0$ vanishes if $s\neq 0$ and is equal to q^{k-1} if $s=0$. Repeating the above calculation for $t \not\equiv 0 \pmod{p}$ gives:

$$\begin{aligned}\sum_{\substack{t=0\\ t\not\equiv 0 (\mathrm{mod}\ p)}}^{k} \mathcal{K}_t^{k;q}(s) &= \sum_{j=0}^{s}\binom{s}{j}(-1)^j \sum_{l=0}^{k-s}\binom{k-s}{l}(q-1)^l\left(1-\frac{1}{q}\sum_{t=0}^{q-1}\omega^{(j+l)t}\right) \\ &= \delta_{s,0}\cdot(q^k-q^{k-1}) - \frac{1}{q}\sum_{t=1}^{q-1}(1+(q-1)\omega^t)^{k-s}(1-\omega^t)^s \ . \qquad (6.21)\end{aligned}$$

Define:

$$\xi_s = \frac{1}{q}\sum_{t=1}^{q-1}(1+(q-1)\omega^t)^{k-s}(1-\omega^t)^s \ ,$$

and consider the special case $p=q=3$. The fact that $\omega^2=\overline{\omega}$ implies that:

$$\begin{aligned}\xi_s = \frac{2}{3}\mathrm{Re}\left((1+2\omega)^{k-s}(1-\omega)^s\right) &= \frac{2}{3}\mathrm{Re}\left((\sqrt{3}i)^{k-s}(\sqrt{3}\mathrm{e}^{-\frac{\pi}{6}i})^s\right) \\ &= \frac{2}{3}\sqrt{3}^k\cos(\frac{\pi k}{2}-\frac{2\pi s}{3}) \ , \qquad (6.22)\end{aligned}$$

and for even values of k and $s\equiv 0 \pmod 3$ we deduce that:

$$\xi_s = \frac{2}{3}3^{k/2}(-1)^{k/2} \ . \qquad (6.23)$$

Therefore, $\xi_s = \frac{2}{3}3^{k/2}$ whenever $s \equiv 0 \pmod 3$ and $k \equiv 0 \pmod 4$, and (6.21) gives the following for any $k \equiv 0 \pmod 4$:

$$\begin{aligned} P(0) &= \frac{2}{3}3^{k/2} + \frac{2}{3}3^k - \xi_0 = \frac{2}{3}3^k \ , \\ P(s) &= \frac{2}{3}3^{k/2} - \xi_s = 0 \ \text{ for any } 0 \neq s \equiv 0 \pmod 3 \ . \end{aligned}$$

Hence, $P(x)$ satisfies the requirements of Proposition 6.3.2 and we deduce that for any $k \equiv 0 \pmod 4$:

$$\alpha(K_3^k) \leq \frac{P(0)}{\frac{2}{3}3^{k/2}} = 3^{k/2} \ .$$

As mentioned before, the construction used for the lower bound on $x_\alpha^{(p)}(K_3)$ implies that this bound is indeed tight whenever $4 \mid k$.

For $k \equiv 2 \pmod 4$ and $s \equiv 0 \pmod 3$ we get $\xi_s = -\frac{2}{3}3^{k/2}$, and by (6.20) we get:

$$\begin{aligned} Q(0) &= \frac{2}{3}3^{k/2} + 3^{k-1} + \xi_0 = 3^{k-1} \ , \\ Q(s) &= \frac{2}{3}3^{k/2} + \xi_s = 0 \ \text{ for any } 0 \neq s \equiv 0 \pmod 3 \ . \end{aligned}$$

Again, $Q(x)$ satisfies the requirements of Proposition 6.3.2 and we obtain the following bound for $k \equiv 2 \pmod 4$:

$$\alpha(K_3^k) \leq \frac{Q(0)}{\frac{2}{3}3^{k/2} + 1} = \frac{3^k}{2 \cdot 3^{k/2} + 3} < \frac{1}{2}3^{k/2} \ .$$

To conclude the proof, take a maximum independent set of size $\sqrt{3}^l$ in K_3^l, where $l = k - 2$, for a lower bound of $\frac{1}{3}3^{k/2}$. ■

6.4 Hoffman's bound on independence numbers of p-powers

In this section we apply spectral analysis in order to bound the independence numbers of p-powers of d-regular graphs. The next theorem generalizes Theorem 2.9 of [10] by considering tensor powers of adjacency matrices whose values are p-th roots of unity.

Theorem 6.4.1. *Let G be a nontrivial d-regular graph on n vertices, whose eigenvalues are $d = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, and let $\lambda = \max\{\lambda_2, |\lambda_n|\}$. The following holds for any $p \geq 2$:*

$$x_\alpha^{(p)}(G) \leq \max\{\sqrt{n^2 - 2\left(1 - \cos(\frac{2\pi}{p})\right) d(n-d)}, \lambda\sqrt{2 - 2\cos\left(\frac{2\pi}{p}\lfloor\frac{p}{2}\rfloor\right)}\} . \tag{6.24}$$

Proof. Let $A = A_G$ denote the adjacency matrix of G, and define the matrices B_t for $t \in \mathbb{Z}_p$ as follows:

$$B_t = J_n + (\omega^t - 1)A , \tag{6.25}$$

where $\omega = \mathrm{e}^{2\pi i/p}$ is the p-th root of unity, and J_n is the all-ones matrix of order n. In other words:

$$(B_t)_{uv} = \omega^{tA_{uv}} = \left\{ \begin{array}{ll} \omega^t & \text{if } uv \in E(G) \\ 1 & \text{if } uv \notin E(G) \end{array} \right. .$$

By the definition of the matrix tensor product $\otimes$, it follows that for all $u = (u_1, \ldots, u_k)$ and $v = (v_1, \ldots, v_k)$ in G^k:

$$(B_t^{\otimes k})_{u,v} = \omega^{t|\{i\ :\ u_i v_i \in E(G)\}|} ,$$

and:

$$\sum_{t=0}^{p-1} (B_t^{\otimes k})_{u,v} = \left\{ \begin{array}{ll} p & \text{if } \ |\{i : u_i v_i \in E(G)\}| \equiv 0 \pmod p \\ 0 & \text{otherwise} \end{array} \right. .$$

Recalling that $uv \in E(G^k)$ iff $|\{i : u_i v_i \in E(G)\}| \not\equiv 0 \pmod p$, we get:

$$A_{G^k} = J_{n^k} - \frac{1}{p}\sum_{t=0}^{p-1} B_t^{\otimes k} = \frac{p-1}{p} J_{n^k} - \frac{1}{p}\sum_{t=1}^{p-1} B_t^{\otimes k} . \tag{6.26}$$

The above relation enables us to obtain expressions for the eigenvalues of G^k, and then apply the following bound, proved by Hoffman (see [67], [81]): every regular nontrivial graph H on N vertices, whose eigenvalues are $\mu_1 \geq \ldots \geq \mu_N$, satisfies:

$$\alpha(H) \leq \frac{-N\mu_N}{\mu_1 - \mu_N} . \tag{6.27}$$

Recall that J_n has a single non-zero eigenvalue of n, corresponding to the all-ones vector $\underline{1}$. Hence, (6.25) implies that $\underline{1}$ is an eigenvector of B_t with an eigenvalue of $n + (\omega^t - 1)d$, and the remaining eigenvalues of B_t are $\{(\omega^t - 1)\lambda_i : i > 1\}$. By well known properties of tensor products, we obtain that the largest eigenvalue of $H = G^k$ (which is its degree of regularity) is:

$$\begin{aligned}
\mu_1 &= n^k - \frac{1}{p}\sum_{t=0}^{p-1}(n + (\omega^t - 1)d)^k = n^k - \frac{1}{p}\sum_{j=0}^{k}\binom{k}{j}(n-d)^{k-j}d^j\sum_{t=0}^{p-1}\omega^{jt} \\
&= n^k - \sum_{\substack{j=0 \\ j\equiv 0 (\mathrm{mod}\ p)}}^{k} \binom{k}{j}(n-d)^{k-j}d^j \ , \qquad (6.28)
\end{aligned}$$

and the remaining eigenvalues are of the form:

$$\mu(\lambda_{i_1}, \ldots, \lambda_{i_s}) = -\frac{1}{p}\sum_{t=1}^{p-1}(n + (\omega^t - 1)d)^{k-s}\prod_{j=1}^{s}(\omega^t - 1)\lambda_{i_j} \ , \qquad (6.29)$$

where $0 < s \le k$ and $1 < i_j \le n$ for all j (corresponding to an eigenvector which is a tensor-product of the eigenvectors of λ_{i_j} for $j = 1, \ldots, s$ and $\underline{1}^{\otimes k-s}$). The following holds for all such choices of s and $\{\lambda_{i_j}\}$:

$$\begin{aligned}
|\mu(\lambda_{i_1}, \ldots, \lambda_{i_s})| &\le \max_{1 \le t \le p-1}\left|(n + (\omega^t - 1)d)^{k-s}\prod_{i=1}^{s}(\omega^t - 1)\lambda_{i_j}\right| \\
&\le \max_{1 \le t \le p-1} |n + (\omega^t - 1)d|^{k-s}(|\omega^t - 1|\lambda)^s \\
&\le \max_{1 \le t \le p-1} \left(\max\{|n + (\omega^t - 1)d|, \lambda|\omega^t - 1|\}\right)^k \ .
\end{aligned}$$

Since for any $1 \le t \le p-1$ we have:

$$\begin{aligned}
|n + (\omega^t - 1)d|^2 &= n^2 - 2\left(1 - \cos(\frac{2\pi t}{p})\right)d(n-d) \\
&\le n^2 - 2\left(1 - \cos(\frac{2\pi}{p})\right)d(n-d) \ , \\
|\omega^t - 1|^2 &= 2 - 2\cos(\frac{2\pi t}{p}) \le 2 - 2\cos\left(\frac{2\pi}{p}\lfloor\frac{p}{2}\rfloor\right) \ ,
\end{aligned}$$

it follows that:

$$|\mu(\lambda_{i_1},\ldots,\lambda_{i_s})| \leq (\max\{\rho_1,\rho_2\})^k \ ,$$

where:

$$\begin{aligned} \rho_1 &= \sqrt{n^2 - 2\left(1-\cos(\tfrac{2\pi}{p})\right)d(n-d)} \\ \rho_2 &= \lambda\sqrt{2-2\cos\left(\tfrac{2\pi}{p}\lfloor\tfrac{p}{2}\rfloor\right)} \end{aligned} \ .$$

By the same argument, (6.28) gives:

$$|\mu_1| \geq n^k - \rho_1^k \ ,$$

and applying Hoffman's bound (6.27), we get:

$$\alpha(G^k) \leq \frac{-n^k\mu_{n^k}}{\mu_1 - \mu_{n^k}} \leq \frac{(\max\{\rho_1,\rho_2\})^k}{1-(\frac{\rho_1}{n})^k + (\frac{\max\{\rho_1,\rho_2\}}{n})^k} \ . \tag{6.30}$$

To complete the proof, we claim that $\max\{\rho_1,\rho_2\} \leq n$, and hence the denominator in the expression above is $\Theta(1)$ as $k \to \infty$. Clearly, $\rho_1 \leq n$, and a simple argument shows that $\lambda \leq n/2$ and hence $\rho_2 \leq n$ as well. To see this, consider the matrix A^2 whose diagonal entries are d; we have:

$$nd = \mathrm{tr}A^2 = \sum_i \lambda_i^2 \geq d^2 + \lambda^2 \ ,$$

implying that $\lambda \leq \sqrt{d(n-d)} \leq \frac{n}{2}$. Altogether, taking the k-th root and letting k tend to ∞ in (6.30), we obtain that $x_\alpha^{(p)}(G) \leq \max\{\rho_1,\rho_2\}$, as required. ■

Examples: For $p = 2,3$ the above theorem gives:

$$\begin{aligned} x_\alpha^{(2)}(G) &\leq \max\{|n-2d|, 2\lambda\} \ , \\ x_\alpha^{(3)}(G) &\leq \max\{\sqrt{n^2-3d(n-d)}, \sqrt{3}\lambda\} \ . \end{aligned}$$

Since the eigenvalues of K_3 are $\{2,-1,-1\}$, this immediately provides another proof for the fact that $x_\alpha^{(3)}(K_3) \leq \sqrt{3}$. Note that, in general, the upper bounds derived in this method for $x_\alpha^{(p)}(K_n)$ are only useful for small values of n, and tend to n as $n \to \infty$, whereas by the results of Section 6.2 we know that $x_\alpha^{(p)}(K_n) = \Theta(n^{1/p})$.

Consider $d = d(n) = \frac{n}{2} + O(\sqrt{n})$, and let $G \sim G_{n,d}$ denote a random d-regular graph on n vertices. By the results of [73], $\lambda = \max\{\lambda_2, |\lambda_n|\} = O(n^{3/4})$, and thus, Theorem 6.4.1 implies that $x_\alpha^{(2)}(G) = O(n^{3/4})$, and that $x_\alpha^{(3)}(G) \leq (1 + o(1))\frac{n}{2}$. We note that one cannot hope for better bounds on $x_\alpha^{(3)}$ in this method, as ρ_1 attains its minimum at $d = \frac{n}{2}$.

Remark 6.4.2: The upper bound (6.24) becomes weaker as p increases. However, if p is divisible by some $q \geq 2$, then clearly any independent set of $G^{k_{(p)}}$ is also an independent set of $G^{k_{(q)}}$, and in particular, $x_\alpha^{(p)}(G) \leq x_\alpha^{(q)}(G)$. Therefore, when applying Theorem 6.4.1 on some graph G, we can replace p by the minimal $q \geq 2$ which divides p. For instance, $x_\alpha^{(4)}(G) \leq x_\alpha^{(2)}(G) \leq \max\{|n - 2d|, 2\lambda\}$, whereas substituting $p = 4$ in (6.24) gives the slightly weaker bound $x_\alpha^{(4)}(G) \leq \{\sqrt{(n-d)^2 + d^2}, 2\lambda\}$.

Remark 6.4.3: In the special case $G = K_n$, the eigenvalues of G are $\{n-1, -1, \ldots, -1\}$, and the general expression for the eigenvalues of G^k in (6.29) takes the following form (note that $\lambda_{i_j} = -1$ for all $1 \leq j \leq s$):

$$\mu(s) = -\frac{1}{p}\sum_{t=1}^{p-1}(1 + (n-1)\omega^t)^{k-s}(1 - \omega^t)^s \ ,$$

and as $s > 0$, we obtain the following from (6.21):

$$\mu(s) = \sum_{\substack{t=0 \\ t \not\equiv 0 (\mathrm{mod}\ p)}}^{k} \mathcal{K}_t^{k;q}(s) \ .$$

Similarly, comparing (6.28) to (6.21) gives:

$$\mu_1 = \sum_{\substack{t=0 \\ t \not\equiv 0 (\mathrm{mod}\ p)}}^{k} \mathcal{K}_t^{k;q}(0) \ .$$

It is possible to deduce this result directly, as K_n^k is a Cayley graph over $\mathbb{Z}_n^k$ with the generator set $S = \{x : |x| \not\equiv 0 \pmod{p}\}$, where $|x|$ denotes the Hamming weight of x. It is well known that the eigenvalues of a Cayley graph are equal to the character sums of the corresponding group elements.

Since for any $k = 0, \ldots, n$ and any $x \in \mathbb{Z}_n^k$ the Krawtchouk polynomial $\mathcal{K}_k^{n;q}$ satisfies:

$$\mathcal{K}_k^{n;q}(|x|) = \sum_{y \in \mathbb{Z}_n^k : |y| = k} \chi_y(x) \ ,$$

the eigenvalue corresponding to $y \in \mathbb{Z}_n^k$ is:

$$\mu(y) = \sum_{x \in S} \chi_x(y) = \sum_{\substack{t=0 \\ t \not\equiv 0 \pmod p}}^{k} \ \sum_{x : |x| = t} \chi_x(y) = \sum_{\substack{t=0 \\ t \not\equiv 0 \pmod p}}^{k} \mathcal{K}_t^{k;q}(|y|) \ .$$

Remark 6.4.4: The upper bound on $x_\alpha^{(p)}$ was derived from an asymptotic analysis of the smallest eigenvalue μ_{n^k} of G^k. Tight results on $\alpha(G^k)$ may be obtained by a careful analysis of the expression in (6.29). To illustrate this, we consider the case $G = K_3$ and $p = 3$. Combining the previous remark with (6.21) and (6.22), we obtain that the eigenvalues of $K_3^{k^{(3)}}$ are:

$$\begin{aligned} \mu_1 &= \frac{2}{3}3^k - \frac{2}{3}\sqrt{3}^k \cos(\frac{\pi k}{2}) \ , \\ \mu(s) &= -\frac{2}{3}\sqrt{3}^k \cos(\frac{\pi k}{2} - \frac{2\pi s}{3}) \ \text{ for } 0 < s \leq k \ . \end{aligned} \tag{6.31}$$

Noticing that $\mu(s)$ depends only on the values of $s \pmod 3$ and $k \pmod 4$, we can determine the minimal eigenvalue of G^k for each given power k, and deduce that:

$$\begin{array}{lll} \alpha(G^k) \leq 3^{k/2} & \text{if } k \equiv 0 & \pmod 4 \\ \alpha(G^k) \leq \dfrac{3^{k+1}}{3 + 2 \cdot 3^{(k+1)/2}} < \dfrac{1}{2} 3^{(k+1)/2} & \text{if } k \equiv 1 & \pmod 2 \\ \alpha(G^k) \leq \dfrac{3^k}{3 + 2 \cdot 3^{k/2}} < \dfrac{1}{2} 3^{k/2} & \text{if } k \equiv 2 & \pmod 4 \end{array} ,$$

matching the results obtained by the Delsarte linear programming bound.

6.5 Ramsey subgraphs in large p-powers of any graph

In order to prove a poly-logarithmic upper bound on the clique sizes of p-powers of a graph G, we use an algebraic argument, similar to the method of representation by polynomials described in the Section 6.2. We note that the same approach provides an upper bound on the size of independent sets. However, for this latter bound, we require another property, which relates the problem to strong graph products and to the Shannon capacity of a graph.

The k-th *strong* power of a graph G (also known as the *and* power), denoted by $G^{\wedge k}$, is the graph whose vertex set is $V(G)^k$, where two distinct k-tuples $u \neq v$ are adjacent iff each of their coordinates is either equal or adjacent in G:

$$(u_1, \ldots, u_k)(v_1, \ldots, v_k) \in E(G^{\wedge k}) \iff$$
$$\text{for all } i = 1, \ldots, k: \ u_i = v_i \text{ or } u_i v_i \in E(G) \ .$$

In 1956, Shannon [97] related the independence numbers of strong powers of a fixed graph G to the effective alphabet size in a zero-error transmission over a noisy channel. Shannon showed that the limit of $\alpha(G^{\wedge k})^{\frac{1}{k}}$ as $k \to \infty$ exists and equals $\sup_k \alpha(G^{\wedge k})^{\frac{1}{k}}$, by super-multiplicativity; this limit is denoted by $c(G)$, the Shannon capacity of G. It follows that $c(G) \geq \alpha(G)$, and in fact equality holds for all perfect graphs. However, for non-perfect graphs, $c(G)$ may exceed $\alpha(G)$, and the smallest (and most famous) example of such a graph is C_5, the cycle on 5 vertices, where $\alpha(C_5) = 2$ and yet $c(C_5) \geq \alpha(C_5^{\wedge 2})^{\frac{1}{2}} = \sqrt{5}$. The seemingly simple question of determining the value of $c(C_5)$ was solved only in 1979 by Lovász [81], who introduced the ϑ-function to show that $c(C_5) = \sqrt{5}$.

The next theorem states the bound on the clique numbers of $G^{k(p)}$, and relates the Shannon capacity of $\overline{G}$, the complement of G, to bounds on independent sets of $G^{k(p)}$.

Theorem 6.5.1. *Let G denote a graph on n vertices and let $p \geq 2$ be a prime. The clique number of $G^{k(p)}$ satisfies:*

$$\omega(G^{k(p)}) \leq \binom{kn + p - 1}{p - 1} , \tag{6.32}$$

and if I is an independent set of both $G^{k(p)}$ and $\overline{G}^{\wedge k}$, then:

$$|I| \leq \binom{kn + \lfloor \frac{k}{p} \rfloor}{\lfloor \frac{k}{p} \rfloor} . \tag{6.33}$$

Moreover, if in addition G is regular then:

$$\omega(G^{k_{(p)}}) \leq \binom{k(n-1)+p}{p-1} \ , \ |I| \leq \binom{k(n-1)+\lfloor\frac{k}{p}\rfloor+1}{\lfloor\frac{k}{p}\rfloor} \ . \tag{6.34}$$

The above theorem implies that if S is an independent set of $\overline{G}^{\wedge k}$, then any independent set I of $G^{k_{(p)}}[S]$, the induced subgraph of $G^{k_{(p)}}$ on S, satisfies inequality (6.33). For large values of k, by definition there exists such a set S of size roughly $c(\overline{G})^k$. Hence, there are induced subgraphs of $G^{k_{(p)}}$ of size tending to $c(\overline{G})^k$, whose clique number and independence number are bounded by the expressions in (6.32) and (6.33) respectively.

In the special case $G = K_n$, the graph $\overline{G}^{\wedge k}$ is an edgeless graph for any k, and hence:

$$\alpha(K_n^{k_{(p)}}) \leq \binom{k(n-1)+\lfloor\frac{k}{p}\rfloor+1}{\lfloor\frac{k}{p}\rfloor} \leq (\mathrm{e}p(n-1)+\mathrm{e}+o(1))^{k/p} \ ,$$

where the $o(1)$-term tends to 0 as $k \to \infty$. This implies an upper bound on $x_\alpha^{(p)}(K_n)$ which nearly matches the upper bound of Theorem 6.2.2 for large values of p.

Proof. Let $g_1 : V(G) \to \mathbb{Z}_p^m$ and $g_2 : V(G) \to \mathbb{C}^m$, for some integer m, denote two representations of G by m-dimensional vectors, satisfying the following for any (not necessarily distinct) $u, v \in V(G)$:

$$\left\{ \begin{array}{ll} g_i(u)\cdot g_i(v) = 0 & \text{if } uv \in E(G) \\ g_i(u)\cdot g_i(v) = 1 & \text{otherwise} \end{array} \right. \quad (i=1,2) \ . \tag{6.35}$$

It is not difficult to see that such representations exist for any graph G. For instance, the standard basis of n-dimensional vectors is such a representation for $G = K_n$. In the general case, it is possible to construct such vectors inductively, in a way similar to a Gram-Schmidt orthogonalization process. To see this, define the lower diagonal $|V(G)| \times |V(G)|$ matrix M as follows:

$$M_{k,i} = \left\{ \begin{array}{cl} -\sum_{j=1}^{i-1} M_{k,j}M_{i,j} & i<k, \ v_iv_k \in E(G) \\ 1-\sum_{j=1}^{i-1} M_{k,j}M_{i,j} & i<k, \ v_iv_k \notin E(G) \\ 1 & i=k \\ 0 & i>k \end{array} \right. .$$

The rows of M satisfy (6.35) for any distinct $u, v \in V(G)$, and it remains to modify the inner product of any vector with itself into 1 without changing the inner products of distinct vectors. This is clearly possible over $\mathbb{Z}_p$ and $\mathbb{C}$ using additional coordinates.

Consider $G^{k_{(p)}}$, and define the vectors $w_u = g_1(u_1) \circ \ldots \circ g_1(u_k)$ for $u = (u_1, \ldots, u_k) \in V(G^k)$, where $\circ$ denotes vector concatenation. By definition:

$$w_u \cdot w_v \equiv k - |\{i : u_i v_i \in E(G)\}| \pmod{p}$$

for any $u, v \in V(G^k)$, and hence, if S is a maximum clique of G^k, then $w_u \cdot w_v \not\equiv k \pmod{p}$ for any $u, v \in S$. It follows that if B is the matrix whose columns are w_u for $u \in S$, then $C = B^t B$ has values which are $k \pmod{p}$ on its diagonal and entries which are not congruent to k modulo p anywhere else. Clearly, $\mathrm{rank}(C) \leq \mathrm{rank}(B)$, and we claim that $\mathrm{rank}(B) \leq kn$, and that furthermore, if G is regular then $\mathrm{rank}(B) \leq k(n-1)+1$. To see this, notice that, as the dimension of $\mathrm{Span}(\{g_1(u) : u \in V\})$ is at most n, the dimension of the span of $\{w_u : u \in G^k\}$ is at most kn. If in addition G is regular, define $z = \sum_{u \in V} g_1(u)$ (assuming without loss of generality that $z \neq 0$), and observe that by (6.35), each of the vectors w_u is orthogonal to the following $k-1$ linearly independent vectors:

$$\{z \circ (-z) \circ \underline{0}^{\circ(k-2)},\ \underline{0} \circ z \circ (-z) \circ \underline{0}^{\circ(k-3)}, \ldots,\ \underline{0}^{\circ(k-2)} \circ z \circ (-z)\}\ . \qquad (6.36)$$

Similarly, the vectors $w'_u = g_2(u_1) \circ \ldots \circ g_2(u_k)$ satisfy the following for any $u, v \in V(G^k)$:

$$w'_u \cdot w'_v = k - |\{i : u_i v_i \in E(G)\}|\ .$$

Let I denote an independent set of $G^{k_{(p)}}$, which is also an independent set of $\overline{G}^{\wedge k}$. By the definition of $\overline{G}^{\wedge k}$, every $u, v \in I$ share a coordinate i such that $u_i v_i \in E(G)$, and combining this with the definition of $G^{k_{(p)}}$, we obtain:

$$0 < |\{i : u_i v_i \in E(G)\}| \equiv 0 \pmod{p} \text{ for any } u, v \in I\ .$$

Therefore, for any $u \neq v \in I$:

$$w'_u \cdot w'_v = k - tp \text{ for some } t \in \{1, \ldots, \lfloor \frac{k}{p} \rfloor\}\ ,$$

and if B' is the matrix whose columns are w'_u for $u \in I$, then $C' = B'^t B'$ has the entries k on its diagonal and entries of the form $k - tp$, $0 < t \leq \lfloor \frac{k}{p} \rfloor$, anywhere else. Again, the definition of g_2 implies that $\mathrm{rank}(C') \leq kn$, and in case G is regular, $\mathrm{rank}(C') \leq k(n-1)+1$ (each w'_u is orthogonal to the vectors of (6.36) for $z = \sum_{u \in V} g_2(u)$).

Define the following polynomials:

$$f_1(x) = \prod_{\substack{j \in \mathbb{Z}_p \\ j \not\equiv k (\mathrm{mod}\ p)}} (j - x) \quad , \quad f_2(x) = \prod_{t=1}^{\lfloor \frac{k}{p} \rfloor} (k - tp - x) \ . \tag{6.37}$$

By the discussion above, the matrices D, D' obtained by applying f_1, f_2 on each element of C, C' respectively, are non-zero on the diagonal and zero anywhere else, and in particular, are of full rank: $\mathrm{rank}(D) = |S|$ and $\mathrm{rank}(D') = |I|$. Recalling that the ranks of C and C' are at most kn, and at most $k(n-1)+1$ if G is regular, the proof is completed by the following simple Lemma of [6]:

Lemma 6.5.2 ([6]). *Let $B = (b_{i,j})$ be an n by n matrix of rank d, and let $P(x)$ be an arbitrary polynomial of degree k. Then the rank of the n by n matrix $(P(b_{i,j}))$ is at most $\binom{k+d}{k}$. Moreover, if $P(x) = x^k$ then the rank of $(P(b_{i,j}))$ is at most $\binom{k+d-1}{k}$.*

■

For large values of k, the upper bounds provided by the above theorem are:

$$\begin{aligned} \omega(H) &\leq \binom{(1+o(1))kn}{p} , \\ \alpha(H) &\leq \binom{(1+o(1))kn}{k/p} . \end{aligned}$$

This gives the following immediate corollary, which states that large p-powers of any nontrivial graph G contain a large induced subgraph without large homogenous sets.

Corollary 6.5.3. *Let G be some fixed nontrivial graph and fix a prime p.*

1. *Let S denote a maximum clique of G, and set $\lambda = \log \omega(G) = \log \alpha(\overline{G})$. For any k, the induced subgraph of $G^{k_{(p)}}$ on S^k, $H = G^{k_{(p)}}[S^k]$, is a graph on $N = \exp(k\lambda)$ vertices which satisfies:*

$$\omega(H) = O(\log^p N) \ , \ \alpha(H) \leq N^{(1+o(1))\frac{\log(np)+1}{p\lambda}} \ .$$

2. *The above formula holds when taking $\lambda = \frac{\log \alpha(\overline{G}^{\wedge \ell})}{\ell}$ for some $\ell \geq 1$ dividing k, S a maximum clique of $\overline{G}^{\wedge \ell}$, and $H = G^{k_{(p)}}[S^{k/\ell}]$. In particular, for sufficiently large values of k, $G^{k_{(p)}}$ has an induced subgraph H on $N = \exp\big((1-o(1))k \log c(\overline{G})\big)$ vertices satisfying:*

$$\omega(H) = O(\log^p N) \ , \ \alpha(H) \leq N^{(1+o(1))\frac{\log(np)+1}{p \log c(\overline{G})}} \ .$$

Remark 6.5.4: In the special case $G = K_n$, where n, p are large and $k > p$, the bound on $\omega(K_n^k)$ is $\binom{(1+o(1))kn}{p}$ whereas the bound on $\alpha(K_n^k)$ is $\binom{(1+o(1))kn}{k/p}$. Hence, the optimal mutual bound on these parameters is obtained at $k = p^2$. Writing $H = K_n^k$, $N = n^k = n^{p^2}$ and $p = n^c$ for some $c > 0$, we get:

$$p = \sqrt{\frac{(2c + o(1)) \log N}{\log \log N}} \ ,$$

and:

$$\begin{aligned} \max\{\omega(H), \alpha(H)\} &\leq ((1+o(1))\mathrm{e}pn)^p \\ &= \exp\left(\left(\frac{1+c}{\sqrt{2c}} + o(1)\right) \sqrt{\log N \log \log N}\right) \ . \end{aligned}$$

The last expression is minimized for $c = 1$, and thus the best Ramsey construction in p-powers of K_n is obtained at $p = n$ and $k = p^2$, giving a graph H on N vertices with no independence set or clique larger than $\exp\big((1+o(1))\sqrt{2 \log N \log \log N}\big)$ vertices. This special case matches the bound of the *FW* Ramsey construction, and is in fact closely related to that construction, as we next describe.

The graph FW_N, where $N = \binom{p^3}{p^2-1}$ for some prime p, is defined as follows: its vertices are the N possible choices of (p^2-1)-element sets of $[p^3]$, and two vertices are adjacent iff the intersection of their corresponding sets is congruent to -1 modulo p. Observe that the vertices of the graph $K_n^{k(p)}$ for $n = p$ and $k = p^2$, as described above, can be viewed as k-element subsets of $[kn]$, where the choice of elements is restricted to precisely one element from each of the k subsets $\{(j-1)n+1, \ldots, jn\}$, $j \in [k]$ (the j-th subset corresponds to the j-th coordinate of the k-tuple). In this formulation, the intersection of two sets corresponds to the number of common coordinates between the corresponding k-tuples. As $k = p^2 \equiv 0 \pmod{p}$, it follows that two vertices in $K_p^{p^2(p)}$ are adjacent iff the intersection of their corresponding sets is not congruent to 0 modulo p. Altogether, we obtain that $K_p^{p^2(p)}$ is an induced subgraph of a slight variant of FW_N, where the differences are in the cardinality of the sets and the criteria for adjacency.

Another relation between the two constructions is the following: one can identify the vertices of $K_2^{p^3(p)}$ with all possible subsets of $[p^3]$, where two vertices are adjacent iff the intersection of their corresponding sets is not congruent to 0 modulo p. In particular, $K_2^{p^3(p)}$ contains all the (p^2-1)-element subsets of $[p^3]$, a variant of FW_N for the above value of N (the difference lies in the criteria for adjacency).

We note that the method of proving Theorem 6.5.1 can be applied to the graph FW_N, giving yet another simple proof for the properties of this well known construction.

Part III

An extremal problem in Finite Set Theory

Chapter 7

Uniformly cross intersecting families

The results of this chapter appear in [15]

Let $\mathcal{A}$ and $\mathcal{B}$ denote two families of subsets of an n-element set. The pair $(\mathcal{A}, \mathcal{B})$ is said to be ℓ-cross-intersecting iff $|A \cap B| = \ell$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Denote by $P_\ell(n)$ the maximum value of $|\mathcal{A}||\mathcal{B}|$ over all such pairs. The best known upper bound on $P_\ell(n)$ is $\Theta(2^n)$, by Frankl and Rödl. For a lower bound, Ahlswede, Cai and Zhang showed, for all $n \geq 2\ell$, a simple construction of an ℓ-cross-intersecting pair $(\mathcal{A}, \mathcal{B})$ with $|\mathcal{A}||\mathcal{B}| = \binom{2\ell}{\ell} 2^{n-2\ell} = \Theta(2^n/\sqrt{\ell})$, and conjectured that this is best possible. Consequently, Sgall asked whether or not $P_\ell(n)$ decreases with ℓ.

In this work, we confirm the above conjecture of Ahlswede et al. for any sufficiently large ℓ, implying a positive answer to the above question of Sgall as well. By analyzing the linear spaces of the characteristic vectors of $\mathcal{A}, \mathcal{B}$ over $\mathbb{R}$, we show that there exists some $\ell_0 > 0$, such that $P_\ell(n) \leq \binom{2\ell}{\ell} 2^{n-2\ell}$ for all $\ell \geq \ell_0$. Furthermore, we determine the precise structure of all the pairs of families which attain this maximum.

7.1 Introduction

Let $\mathcal{A}$ and $\mathcal{B}$ denote two families of subsets of an n-element set. We say that the pair $(\mathcal{A}, \mathcal{B})$ is ℓ-cross-intersecting iff $|A \cap B| = \ell$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Let $P_\ell(n)$ denote the maximum possible value of $|\mathcal{A}||\mathcal{B}|$ over all ℓ-cross-intersecting pairs $(\mathcal{A}, \mathcal{B})$. We are interested in finding the precise value of $P_\ell(n)$, and in characterizing all the extremal pairs $\mathcal{A}, \mathcal{B}$ which achieve this maximum.

The study of the maximal size of a single family of sets $\mathcal{F} \subset 2^{[n]}$, with specified pairwise intersections of its members, has received a considerable amount of attention over the years. For instance, the Erdős-Ko-Rado Theorem [50], one of the most fundamental theorems in Combinatorial Set Theory, gives a tight upper bound $|\mathcal{F}| \leq \binom{n-t}{k-t}$ in case $|F \cap F'| \geq t$ for all $F, F' \in \mathcal{F}$, $|F| = k$ for all $F \in \mathcal{F}$ and n is sufficiently large. The case where there is no restriction on the size of the sets of $\mathcal{F}$ is treated by Katona's Theorem [70]. In both cases, there is a unique (up to a relabeling of the elements of $[n]$) family of sets which achieves the upper bound. For further results of this nature, see, e.g, [53], [54], [56], [91], as well as [20].

A well known conjecture of Erdős [48] stated that if $\mathcal{F} \subset 2^{[n]}$ is a family satisfying $|F \cap F'| \neq \lfloor \frac{n}{4} \rfloor$ for all $F, F' \in \mathcal{F}$, then $|\mathcal{F}| < (2-\varepsilon)^n$ for some $\varepsilon > 0$. This was proved by Frankl and Rödl [55], by considering the corresponding variant on two families: it is shown in [55], that if $\mathcal{A}, \mathcal{B} \subset 2^{[n]}$ and $|A \cap B| \neq l$, where $\eta n \leq l \leq (\frac{1}{2} - \eta)n$ for some $\eta < \frac{1}{4}$, then $|\mathcal{A}||\mathcal{B}| \leq (4 - \varepsilon(\eta))^n$. The authors of [55] studied several additional problems related to cross-intersections of two families of sets, and among their results, they provided the following upper bound on $P_\ell(n)$, which was later reproved in [2]:

$$\left\{ \begin{array}{ll} P_0(n) \leq 2^n & \\ P_\ell(n) \leq 2^{n-1} & \text{for } \ell \geq 1 \end{array} \right. . \tag{7.1}$$

The argument which gives the upper bound of 2^n is simple: consider the characteristic vectors of the sets in $\mathcal{A}, \mathcal{B}$ as vectors in $\mathbb{Z}_2^n$. Notice that the intersection of two sets is equal to the inner product of the two corresponding vectors modulo 2. Therefore, if ℓ is even, then the families $\mathcal{A}, \mathcal{B}$ belong to two orthogonal linear spaces, giving $|\mathcal{A}||\mathcal{B}| \leq 2^n$. Otherwise, we may add an additional coordinate of 1 to all vectors, and repeat (carefully) the above argument, gaining a slight improvement: $|\mathcal{A}||\mathcal{B}| \leq 2^{n-1}$. Similar ideas are used to show that the upper bound 2^{n-1} holds for even values of $\ell > 0$ as well, by performing the analysis over $GF(p)$ for some prime $p > 2$ instead of over $\mathbb{Z}_2$.

As part of their study of questions in Coding Theory, Ahlswede, Cai and Zhang [2] gave the following simple construction of an ℓ-cross-intersecting pair: for $n \geq 2\ell$, let $\mathcal{A}$ contain a single 2ℓ-element set, A, and let $\mathcal{B}$ contain all the sets which contain precisely ℓ elements of A. This gives:

$$|\mathcal{A}||\mathcal{B}| = \binom{2\ell}{\ell} 2^{n-2\ell} = (1+o(1))\frac{2^n}{\sqrt{\pi\ell}} , \tag{7.2}$$

where the $o(1)$-term tends to 0 as $\ell \to \infty$. The upper bound (7.1) implies that this construction achieves the maximum of $P_\ell(n)$ for $\ell \in \{0,1\}$, and the authors of [2] conjectured that this in fact holds for all ℓ.

As the upper bound (7.1) is independent of ℓ, compared to the above lower bound of $\Theta(2^n/\sqrt{\ell})$, Sgall [95] asked whether or not $P_\ell(n)$ is bounded from above by some decreasing function of ℓ. One of the motivations of [95] was a relation between problems of restricted cross-intersections of two families of sets and problems in Communication Complexity; see [95] for more details.

In [71], the authors verified the above conjecture of [2] for the case $\ell = 2$, by showing that $P_2(n) \leq 3 \cdot 2^{n-3}$. However, for any $\ell > 2$ the best known upper bound on $P_\ell(n)$ remained 2^{n-1}.

The following theorem confirms the above conjecture of [2] for all sufficiently large values of ℓ, and thus provides also a positive answer to the above question of Sgall.

Theorem 7.1.1. *There exists some $\ell_0 > 0$ such that, for all $\ell \geq \ell_0$, every ℓ-cross-intersecting pair $\mathcal{A}, \mathcal{B} \subset 2^{[n]}$ satisfies:*

$$|\mathcal{A}||\mathcal{B}| \leq \binom{2\ell}{\ell} 2^{n-2\ell} . \tag{7.3}$$

Furthermore, if $|\mathcal{A}||\mathcal{B}| = \binom{2\ell}{\ell}2^{n-\ell}$, then there exists some choice of parameters κ, τ, n':

$$\begin{aligned} &\kappa \in \{2\ell-1, 2\ell\} \ , \ \tau \in \{0,\ldots,\kappa\} \ , \\ &\kappa + \tau \leq n' \leq n, \end{aligned} \tag{7.4}$$

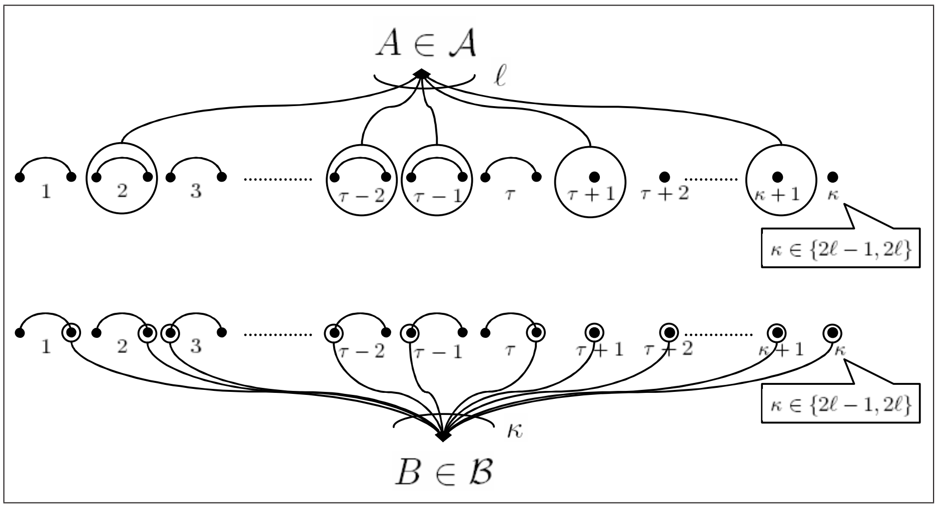

Figure 7.1: The extremal family (7.5) of ℓ-cross-intersecting pairs $\mathcal{A}, \mathcal{B}$ in case $n = \kappa + \tau$.

such that, up to a relabeling of the elements of $[n]$ *and swapping* $\mathcal{A}, \mathcal{B}$*, the following holds:*

$$\begin{aligned} \mathcal{A} &= \left\{ \bigcup_{T \in J} T \;:\; J \subset \left\{ \begin{array}{c} \{1, \kappa+1\}, \ldots, \{\tau, \kappa+\tau\}, \\ \{\tau+1\}, \ldots, \{\kappa\} \end{array} \right\}, \; |J| = \ell \right\} \times 2^X , \\ \mathcal{B} &= \left\{ L \cup \{\tau+1, \ldots, \kappa\} : \begin{array}{c} L \subset \{1, \ldots, \tau, \kappa+1, \ldots, \kappa+\tau\} \\ |L \cap \{i, \kappa+i\}| = 1 \textit{ for all } i \in [\tau] \end{array} \right\} \times 2^Y . \end{aligned} \tag{7.5}$$

where $X = \{\kappa+\tau+1, \ldots, n'\}$ *and* $Y = \{n'+1, \ldots, n\}$.

An illustration of the family of extremal pairs $\mathcal{A}, \mathcal{B}$ described in Theorem 7.1.1 appears in Figure 7.1. Indeed, this family satisfies:

$$|\mathcal{A}||\mathcal{B}| = \binom{\kappa}{\ell} \cdot 2^{|X|} \cdot 2^{\tau+|Y|} = \binom{\kappa}{\ell} 2^{n-\kappa} = \binom{2\ell}{\ell} 2^{n-2\ell} ,$$

where the last inequality is by the choice of $\kappa \in \{2\ell-1, 2\ell\}$. The construction of [2] fits the special case $\tau = 0$, $\kappa = 2\ell$.

The proof of Theorem 7.1.1 combines tools from linear algebra with techniques from extremal combinatorics, including the Littlewood-Offord Lemma, extensions of Sperner's Theorem and some large deviation estimates.

The rest of this chapter is organized as follows: Section 7.2 includes some of the ingredients needed for the proof of Theorem 7.1.1. In order to prove the main result, we first prove a weaker version of Theorem 7.1.1, which states that $P_\ell(n) \leq 2^{n+3}/\sqrt{\ell}$ for every sufficiently large ℓ (note that this result alone gives a positive answer to the above question of Sgall). This is shown in Section 7.3. In Section 7.4 we reduce the proof of Theorem 7.1.1 to two lemmas, Lemma 7.4.1 and Lemma 7.4.2. These lemmas are proved in Sections 7.5 and 7.6 respectively. Section 7.7 contains some concluding remarks and open problems.

Throughout the chapter, all logarithms are in base 2.

7.2 Preliminary Sperner-type Theorems

7.2.1 Sperner's Theorem and the Littlewood Offord Lemma

If P is a finite partially ordered set, an antichain of P is a set of pairwise incomparable elements. Sperner's Theorem [98] provides a tight upper bound on the maximal size of an antichain, when P is the collection of all subsets of an n-element set with the subset relation ($A \leq B$ iff $A \subset B$):

Theorem 7.2.1 ([98]). *If $\mathcal{A}$ is an antichain of an n-element set, then $|\mathcal{A}| \leq \binom{n}{\lfloor n/2 \rfloor}$.*

In [78], Littlewood and Offord studied a problem which has the following formulation in the 1-dimensional case: let $a_1, \ldots, a_n \in \mathbb{R}$ with $|a_i| > 1$ for all i. What is the maximal number of sub-sums $\sum_{i \in I} a_i$, $I \subset [n]$, which lie in an interval of length 1? An immediate lower bound is $\binom{n}{\lfloor n/2 \rfloor}$, when, for some $\alpha > 1$, half of the a_i-s is equal to α and the other half is equal to $-\alpha$.

Using Sperner's Theorem, Erdős [46] gave a tight upper bound of $\binom{n}{\lfloor n/2 \rfloor}$ for the 1-dimensional case of the so-called Littlewood-Offord Lemma. To see this, consider the maximal number of sub-sums of $a_1, \ldots, a_n$, which all belong to some unit interval. Without loss of generality, we may assume that all the a_i-s are positive (possibly shifting the target unit interval). Therefore, $a_i > 1$ for all i, implying that the desired family of subsets is an antichain.

The result now follows from Sperner's Theorem. Using a similar argument, Erdős proved the following stronger result:

Lemma 7.2.2 ([46]). *Let $a_1, \ldots, a_n \in \mathbb{R} \setminus \{0\}$, and let $\delta = \min\{|a_i|\}$. Let T be a union of m half-open intervals, each of width at most δ. Then the number of sub-sums $\sum_{i \in I} a_i$, $I \subset [n]$, which belong to T, is at most the sum of the m middle binomial coefficients in n.*

7.2.2 A bipartite extension of Sperner's Theorem

The following lemma gives an upper bound on the size of an antichain of $[n]$, which satisfies an additional requirement with respect to a pre-defined partition of $[n]$ into into two sets.

Lemma 7.2.3. *Let $U = [u]$ and $V = [n] \setminus U$, $u \leq n$. If $\mathcal{A}$ is an antichain of $[n]$, and in addition satisfies: $|A \cap V| = f(|A \cap U|)$, where $f : \mathbb{N} \to \mathbb{N}$ is some monotone increasing function, then $|\mathcal{A}| \leq \binom{u}{\lfloor u/2 \rfloor}\binom{n-u}{\lfloor (n-u)/2 \rfloor}$.*

The above lemma will follow from the next generalization of Sperner's Theorem:

Proposition 7.2.4. *Let $U = [u]$ and $V = [n] \setminus U$, $u \leq n$. If every two sets $A \neq B \in \mathcal{A}$ satisfy that either $A \cap U$, $B \cap U$ are incomparable or $A \cap V$, $B \cap V$ are incomparable, then $|\mathcal{A}| \leq \binom{u}{\lfloor u/2 \rfloor}\binom{n-u}{\lfloor (n-u)/2 \rfloor}$.*

Proof. Notice that the upper bound is tight, as it is achieved by a cartesian product of maximal antichains of U and V. The proof is based on Lubbell's proof [84] of Sperner's Theorem and the LYM inequality. For each $A \in \mathcal{A}$, let:

$$A_U = A \cap U \ , \ A_V = \{x - u : x \in A \cap V\} \ . \tag{7.6}$$

Let $\sigma \in S_u$ and $\pi \in S_{n-u}$ (where S_m is the symmetric group on m elements) denote two random permutations, chosen uniformly and independently . We define the event E_A for $A \in \mathcal{A}$ to be:

$$E_A = (\ A_U = \{\sigma(1), \ldots, \sigma(|A_U|)\} \ \wedge \ A_V = \{\pi(1), \ldots, \pi(|A_V|)\}\ ) \ ,$$

that is, the first entries of σ form A_U, and the first entries of π form A_V. The key observation is that the events E_A and E_B are disjoint for all $A \neq B \in \mathcal{A}$.

To see this, assume that $E_A \wedge E_B$ holds for some $A \neq B \in \mathcal{A}$. The fact that the first entries of σ form both A_U and B_U implies that either $A_U \subset B_U$ or $B_U \subset A_U$, and the same applies to A_V, B_V. Therefore, the assumption on $\mathcal{A}$ implies that the events E_A and E_B are indeed disjoint, and thus:

$$\sum_{A\in\mathcal{A}} \Pr[E_A] = \Pr[\bigcup_{A\in\mathcal{A}} E_A] \leq 1 \ .$$

Since:

$$\Pr[E_A] = \frac{1}{\binom{u}{|A_U|}\binom{n-u}{|A_V|}} \ ,$$

it follows that:

$$\sum_{A\in\mathcal{A}} \frac{1}{\binom{u}{|A_U|}\binom{n-u}{|A_V|}} \leq 1 \ . \tag{7.7}$$

Note that in the special case $u = n$ this is the LYM inequality. The left hand side of (7.7) is at most $\sum_{A\in\mathcal{A}} 1/\left(\binom{u}{\lfloor u/2\rfloor}\binom{n-u}{\lfloor (n-u)/2\rfloor}\right)$ and the desired result follows. ■

Proof of Lemma 7.2.3. Following the notation of Proposition 7.2.4, define A_U and A_V for each $A \in \mathcal{A}$ as in (7.6). By Proposition 7.2.4, it suffices to show that, for all $A \neq B \in \mathcal{A}$, either A_U, B_U are incomparable or A_V, B_V are incomparable. Assume the contrary, and let $A \neq B \in \mathcal{A}$ be a counterexample. Without loss of generality, assume that $A_U \subset B_U$. If $A_V \subset B_V$ then $A \subset B$, contradicting the fact that $\mathcal{A}$ is an antichain. It follows that $B_V \subsetneqq A_V$, and since f is monotone increasing, the following holds:

$$|A_V| > |B_V| = f(|B_U|) \geq f(|A_U|) \ ,$$

contradicting the assumption that $|A_V| = f(|A_U|)$. ■

7.3 An upper bound tight up to a constant

In this section we prove a weaker version of Theorem 7.1.1, whose arguments will be later extended to prove the precise lower bound.

Theorem 7.3.1. *For any sufficiently large $\ell \in \mathbb{N}$, every ℓ-cross-intersecting pair $\mathcal{A}, \mathcal{B} \subset 2^{[n]}$ satisfies:*

$$|\mathcal{A}||\mathcal{B}| \leq \frac{2^{n+3}}{\sqrt{\ell}} \ . \tag{7.8}$$

Proof. Let $\mathcal{A}$ and $\mathcal{B}$ be as above. A key observation is the following: it is sufficient to prove (7.8) for the case where both $\mathcal{A}$ and $\mathcal{B}$ are antichains. This follows from an induction on n, where in the case $n = \ell$, $|\mathcal{A}||\mathcal{B}| = 1$ and (7.8) clearly holds. Indeed, suppose that there exist $A_1, A_2 \in \mathcal{A}$ such that $A_1 \subset A_2$. As $(\mathcal{A}, \mathcal{B})$ are ℓ-cross-intersecting, this implies that:

$$B \cap (A_2 \setminus A_1) = \emptyset \text{ for all } B \in \mathcal{B} \ , \tag{7.9}$$

hence the restriction of the families $(\mathcal{A}, \mathcal{B})$ to $[n] \setminus (A_2 \setminus A_1)$, $(\mathcal{A}', \mathcal{B}')$, is an ℓ-cross-intersecting pair of an n'-element set, where $n' < n$. By (7.9), $|\mathcal{B}'| = |\mathcal{B}|$, and by the induction hypothesis:

$$|\mathcal{A}||\mathcal{B}| \leq 2^{n-n'}|\mathcal{A}'||\mathcal{B}'| \leq \frac{2^{n+3}}{\sqrt{\ell}} \ ,$$

as required.

For any subset $A \subset [n]$, let $\chi_A \in \{0,1\}^n$ denote its characteristic vector. Let $\mathcal{F}_\mathcal{A}$ and $\mathcal{F}_\mathcal{B}$ denote the linear subspaces of $\mathbb{R}^n$ formed by the characteristic vectors of $\mathcal{A}$ and $\mathcal{B}$ respectively:

$$\begin{array}{rcl} \mathcal{F}_\mathcal{A} & = & \mathrm{span}(\{\chi_A : A \in \mathcal{A}\}) \subset \mathbb{R}^n \ , \\ \mathcal{F}_\mathcal{B} & = & \mathrm{span}(\{\chi_B : B \in \mathcal{B}\}) \subset \mathbb{R}^n \ , \end{array} \tag{7.10}$$

and assume without loss of generality that $\dim(\mathcal{F}_\mathcal{A}) \geq \dim(\mathcal{F}_\mathcal{B})$. Choose an arbitrary set $B_1 \in \mathcal{B}$ and define:

$$\begin{array}{l} \mathcal{F}'_\mathcal{B} = \mathrm{span}(\{\chi_B - \chi_{B_1} : B \in \mathcal{B}\}) \ , \\ k = \dim(\mathcal{F}_\mathcal{A}) \ , \ h = \dim(\mathcal{F}'_\mathcal{B}) \leq \dim(\mathcal{F}_\mathcal{B}) \ . \end{array} \tag{7.11}$$

By the definition of ℓ-cross-intersection, it follows that $\mathcal{F}_\mathcal{A}, \mathcal{F}'_\mathcal{B}$ are two orthogonal linear subspaces of $\mathbb{R}^n$, and $k + h \leq n$. Note also that $k \geq h$ by the assumption on $\dim(\mathcal{F}_\mathcal{A})$.

Let $M_\mathcal{A}$ denote the $k \times n$ row-reduced echelon form matrix, which is the result of performing Gauss elimination on the row-vectors $\{\chi_A : A \in \mathcal{A}\}$

over $\mathbb{R}$, and let $M_{\mathcal{B}}$ denote the corresponding $h \times n$ matrix for the vectors $\{\chi_B - \chi_{B_1} : B \in \mathcal{B}\}$. As $\text{rank}M_{\mathcal{A}} = k$ and $\text{rank}M_{\mathcal{B}} = h$, without loss of generality we have:

$$M_{\mathcal{A}} = \left(\begin{array}{c|c} I_k & * \end{array} \right) \, , \, M_{\mathcal{B}} = \left(\begin{array}{c|c} I_h & * \end{array} \right) \, .$$

where I_r denotes the identity matrix of order r (and the order of the columns in $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ is not necessarily the same). This implies that any linear combination of the rows of $M_{\mathcal{A}}$ which belongs to $\{0,1\}^n$ has precisely two possible coefficients for each row: $\{0,1\}$, and in particular, $|\mathcal{A}| \leq 2^k$. Similarly, $|\mathcal{B}| \leq 2^h$ (the two possible coefficients in the affine combination are now determined by the vector χ_{B_1}), hence $|\mathcal{A}||\mathcal{B}| \leq 2^{k+h} \leq 2^n$, giving the known upper bound of [55]. Observe that if $k + h \leq n - \log n$, we get

$$|\mathcal{A}||\mathcal{B}| \leq \frac{2^n}{n} \, ,$$

and (7.8) clearly holds. Therefore, recalling that $k \geq h$, we may assume that:

$$\left\{ \begin{array}{rcl} \frac{n}{2} - \frac{1}{2}\log n < & k & \\ n - \log n < & k + h & \leq n \end{array} \right. . \tag{7.12}$$

We claim that the following statement, which clearly implies (7.8), holds:

$$|\mathcal{A}||\mathcal{B}| \leq \frac{2^{k+h+3}}{\sqrt{n}} . \tag{7.13}$$

To show this, we need the next lemma, which will be applied once on $M_{\mathcal{A}}, \mathcal{A}, k$ and once on $M_{\mathcal{B}}, \mathcal{B}, h$, to conclude that a constant fraction of the rows of $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ have precisely two non-zero entries, 1 and -1.

Lemma 7.3.2. *Let M denote a $d \times n$ matrix in row-reduced echelon form: $M = \left(\begin{array}{c|c} I_d & * \end{array} \right)$, and let $\mathcal{D}$ denote an antichain of subsets of $[n]$. Assume that:*

1. *The characteristic vectors of $\mathcal{D}$ belong to $w + \text{span}(M)$, the affine subspace formed by some fixed vector $w \in \{0,1\}^n$ and the span of the rows of M.*

2. *The antichain $\mathcal{D}$ satisfies $|\mathcal{D}| \geq 8 \cdot 2^d/\sqrt{n}$.*

Then there exists a subset of c rows of M, $C \subset [d]$, where $c \geq d - \frac{n}{20} - 10\log n$, such that:

1. *Every row i of C belongs to $\{0, \pm 1\}^n \setminus \{0,1\}^n$.*
2. *Every column of the $c \times n$ sub-matrix formed by C contains at most 1 non-zero entry.*

Proof. Our first step is to remove a small portion of the rows of M, such that the remaining rows will have at most one non-zero entry in each column.

Claim 7.3.3. *Let $M, \mathcal{D}, w$ satisfy the requirements of Lemma 7.3.2. There exists a set of rows $R \subset [d]$ such that $|R| \leq \frac{n}{25} + 10\log n$, and each column of M has at most one non-zero value in the remaining $d - |R|$ rows.*

Proof of Claim. Perform the following process of column-selection on M: first, set $M' = M$. If M' has no column with at least 2 non-zero entries, the process ends. Otherwise, perform the following step (step j, for $j \geq 1$):

- Let i_j denote the index of a column of M' with a maximal number of non-zero entries, r_j.
- Let R_j denote the set of rows where the column i_j is non-zero ($|R_j| = r_j$).
- Replace all these rows in M' by 0-rows, and continue the process.

The result is a sequence of indices, $i_1, \ldots, i_t$ ($t \geq 0$) and a sequence of sets of rows $R_1, \ldots, R_t$ of sizes $r_1 \geq r_2 \geq \ldots \geq r_t > 1$, such that the column i_j has r_j non-zero values in the rows R_j, and $R_j \cap R_{j'} = \emptyset$ for all $j \neq j'$. Finally, the sub-matrix formed by removing the rows $R = \cup_{j=1}^{t} R_j$ from M has at most 1 non-zero entry in every column.

Consider affine combinations (with the affine vector w) of the rows of M which produce a $\{0,1\}^n$-vector. As stated above, each row of M allows precisely two coefficients in such an affine combination, as the first d columns of M form the identity matrix. Clearly, the value of the affine combination at index i_1 depends precisely on the r_1 coefficients of the rows R_1. In general, if we already chose the coefficients for the rows $\cup_{j'<j} R_{j'}$, then the value of the

affine combination at index i_j depends only on the choice of the r_j coefficients for the rows R_j.

A simple argument will show that for $1 \leq j \leq t$, at most $\frac{3}{4}$ of the above 2^{r_j} combinations of coefficients for the rows R_j are indeed valid. To this end, recall the following simple fact, which corresponds to the Cauchy-Davenport Theorem when A, B are subsets of $\mathbb{Z}/p\mathbb{Z}$ instead of $\mathbb{R}$:

$$|A+B| \geq |A| + |B| - 1 \quad \text{for any two finite nonempty } A, B \subset \mathbb{R} \ , \tag{7.14}$$

where $A + B = \{a + b : a \in A,\ b \in B\}$. To see this, simply sort the values of A and B by order of magnitude, then produce distinct sums by iterating first on A, then on B.

Suppose we already chose coefficients for the rows $\cup_{j'<j} R_{j'}$, and consider the column i_j. Select 2 arbitrary rows $u, v \in R_j$, and fix the choice of coefficients for the remaining $r_j - 2$ rows. We are left with a choice between two coefficients for u, yielding two possible values a_1, a_2 contributed by u to the index i_j. Similarly, the row v contributes one of two possible values b_1, b_2 to the index i_j. Setting $A = \{a_1, a_2\}$ and $B = \{b_1, b_2\}$, the above fact implies that $|A + B| \geq 3$, hence at least one of the 4 possible combinations of u and v gives a non-$\{0,1\}$ value in index i_j of the resulting affine combination. Therefore, at most $\frac{3}{4}$ of the 2^{r_j} combinations for R_j result in a $\{0,1\}^n$ vector. We conclude that $|\mathcal{D}| \leq \left(\frac{3}{4}\right)^t 2^d$, and hence $t \leq 2 \log n$, otherwise we would get:

$$|\mathcal{D}| \leq \frac{2^d}{n^{2\log(4/3)}} < \frac{2^d}{\sqrt{n}} \ ,$$

contradicting the assumption on $|\mathcal{D}|$.

After providing an upper bound on t, we wish to bound the term $\sum_{i=1}^t r_i$. Let $0 \leq s \leq t$ denote the maximal index such that $r_s \geq 6$, that is:

$$r_1 \geq r_2 \geq \ldots \geq r_s \geq 6 \ ,$$
$$6 > r_{s+1} \geq r_{s+2} \geq \ldots \geq r_t > 1 \ .$$

As before, we consider the choice of coefficients for the rows R_j at step j, determining the i_j-th entry of the linear combination. By the Littlewood-Offord Lemma (Lemma 7.2.2), we conclude that there are at most $2\binom{r_j}{\lfloor r_j/2 \rfloor} <$

$\frac{2}{\sqrt{\frac{\pi}{2}r_j}}2^{r_j}$ possible combinations of the rows R_j which yield a $\{0,1\}$-value in the i_j column (note that the inequality $\binom{2x}{x} \leq 2^{2x}/\sqrt{\pi x}$ holds for every integer $x \geq 1$, by the improved approximation [92] of the error term in Stirling's formula). Applying this argument to $i_1, \ldots, i_s$, we obtain that:

$$|\mathcal{D}| \leq 2^d \prod_{i=1}^{s} \frac{2\sqrt{2/\pi}}{\sqrt{r_i}} \ . \tag{7.15}$$

Observe that every m reals $a_1, \ldots, a_m \geq 2$ satisfy:

$$\prod_{i=1}^{m} \frac{1}{a_i} \leq \frac{1}{\sum_{i=1}^{m} a_i}$$

(this follows by induction on m from the fact that $xy \geq x+y$ for $x, y \geq 2$). Therefore, as $r_i \geq 6 > 2 \cdot (2\sqrt{2/\pi})^2$ for $1 \leq i \leq s$, it follows that:

$$\prod_{i=1}^{s} \frac{2\sqrt{2/\pi}}{\sqrt{r_i}} \leq \frac{2\sqrt{2/\pi}}{\sqrt{\sum_{i=1}^{s} r_i}} \ .$$

Combining this with (7.15) we obtain that if $\sum_{i=1}^{s} r_i > n/25$, then $|\mathcal{D}| < 8 \cdot 2^d/\sqrt{n}$, contradicting the assumption on $|\mathcal{D}|$. Assume therefore that $\sum_{i=1}^{s} r_i \leq n/25$. Altogether, we obtain that $R = \cup_{j=1}^{t} R_j$ satisfies:

$$|R| = \sum_{i=1}^{t} r_i \leq (\sum_{i=1}^{s} r_i) + 5(t-s) \leq \frac{n}{25} + 10 \log n \ .$$

This completes the proof of the claim. ■

It remains to deal with rows which do not belong to $\{0,\pm 1\}^n \setminus \{0,1\}^n$. The next claim provides an upper bound on the number of such rows in M:

Claim 7.3.4. *Let $M, \mathcal{D}, w$ satisfy the requirements of Lemma 7.3.2, and let $R \subset [d]$ be a set of indices of rows of M as provided by Claim 7.3.3. Let S denote the set of indices in $[d] \setminus R$ of rows which do not belong to $\{0,\pm 1\}^n \setminus \{0,1\}^n$. Then $|S| < n/100$.*

Proof of Claim. To prove the claim, fix a linear combination u of the rows $[d] \setminus S$, and consider all the possible combinations of the rows of S which can be added to $w' = w + u$ to produce vectors of $\mathcal{D}$. We will show that the number of these combinations is at most $2^s/\sqrt{\pi s/2}$, where $s = |S|$, and the result will follow from the assumption on $|\mathcal{D}|$.

Put $S = S_{01} \cup S_{\overline{01}}$, where $S_{01} \subset S$ is the set of indices of rows in S which are $\{0,1\}^n$ vectors, and $S_{\overline{01}} = S \setminus S_{01}$. Recall that the first d columns of M form the identity matrix, and that $w \in \{0,1\}^n$, hence the only two coefficients which can be assigned to the row i to produce $\{0,1\}$ values in the i-th column are:

$$\left\{ \begin{array}{ll} \{0,1\} & \text{if } w_i = 0 \\ \{0,-1\} & \text{if } w_i = 1 \end{array} \right. . \tag{7.16}$$

It will be more convenient to have the coefficients $\{0,1\}$ for all rows of S: to obtain this, subtract each row $i \in S$, whose coefficients are $\{0,-1\}$, from w', and let w'' denote the resulting vector.

Let $i \in S_{\overline{01}}$ be an index of a row which does not belong to $\{0,\pm1\}^n$, and let j denote a column such that $M_{ij} = \lambda \notin \{0,\pm1\}$. Crucially, $S \cap R = \emptyset$, hence column j contains at most one non-zero entry in the rows of S. Therefore, the two possible values of the affine combination in index j are $\{w'_j, w'_j + \lambda\}$, and as $0 < |\lambda| \neq 1$ it follows that at least one of these values does not belong to $\{0,1\}$. We deduce that there is at most one valid choice of coefficients for all the rows $S_{\overline{01}}$. Denoting this unique combination of the rows of $S_{\overline{01}}$ by v, it follows that every linear combination of S which, when added to w', belongs to $\mathcal{D}$, is the sum of $z = w'' + v$ and a linear combination of S_{01}.

It remains to set the coefficients of the rows S_{01}, and since each row of S_{01} has $\{0,1\}$ as its coefficients, we are considering a sum of a subset of the rows of S_{01}. Each of these rows belongs to $\{0,1\}^n$, and in particular, is non-negative: we claim that the set of possible subsets of S_{01} is therefore an antichain. To see this, suppose that two distinct subsets $X, Y \subset S_{01}$, $X \subset Y$, produce (when added to z) two vectors $x, y \in \mathbb{R}^n$ which correspond to sets in $\mathcal{D}$. The values of x, y at the indices of S_{01} are determined by the sets X, Y (in fact, these values are equal to those of the corresponding characteristic vectors), hence $x \neq y$. Furthermore, as the rows of S_{01} are non-negative, and $X \subset Y$, we have $x_i \leq y_i$ for all $i \in [n]$. This contradicts the fact that $\mathcal{D}$ is an antichain. Let $s' = |S_{01}|$; Sperner's Theorem gives:

$$|\mathcal{D}| \leq 2^{d-s} \cdot \binom{s'}{\lfloor s'/2 \rfloor} \leq 2^{d-s} \cdot \binom{s}{\lfloor s/2 \rfloor} \leq \frac{2^d}{\sqrt{\pi s/2}} ,$$

and by the assumption on $|\mathcal{D}|$, we obtain that $s \leq n/100$, completing the proof of the claim. ■

Altogether, Claims 7.3.3 and 7.3.4 imply that we can delete at most

$$|R| + |S| \leq \frac{n}{20} + 10 \log n$$

rows of M, and obtain a subset of c rows, $d - \frac{n}{20} - 10 \log n \leq c \leq d$, satisfying the statements of the lemma. ■

Note that the requirements of Lemma 7.3.2 are satisfied both by $M_{\mathcal{A}}, \mathcal{A}$ and by $M_{\mathcal{B}}, \mathcal{B}$. Indeed, if either $|\mathcal{A}| < \frac{8}{\sqrt{n}} \cdot 2^k$ or $|\mathcal{B}| < \frac{8}{\sqrt{n}} \cdot 2^h$, then (7.13) holds and we are done. The remaining requirement on the characteristic vectors of $\mathcal{D}$ is satisfied by definition (for $\mathcal{A}$, w is the zero vector, whereas for $\mathcal{B}$, $w = \chi_{B_1}$).

Applying Lemma 7.3.2 to $M_{\mathcal{A}}, \mathcal{A}$, we obtain a set of at least $c_1 \geq k - \frac{n}{20} - 10 \log n$ rows, $C_1 \subset [k]$, such that each row has an entry of -1 at some index $j > k$, and each column has at most 1 non-zero entry in these rows. In particular, we get: $c_1 \leq n - k$, and thus:

$$k - \frac{n}{20} - 10 \log n \leq n - k ,$$

and by (7.12) we get:

$$\left\{ \begin{array}{rcl} \frac{n}{2} - \log n & \leq k \leq & \frac{21}{40}n + 5 \log n \\ \frac{19}{40}n - 6 \log n & \leq h \leq & \frac{n}{2} \end{array} \right. . \tag{7.17}$$

Next, let $C_1' \subset C_1$ denote the set of indices of rows of C_1 with precisely two non-zero entries. Notice that, as each of the columns $\{k+1, \ldots, n\}$ contains at most 1 non-zero entry in the rows C_1, and on the other hand, each of the rows C_1 contains a non-zero value in one of these columns, it follows that

$|C_1 \setminus C_1'| \le n - k - c_1$. The lower bound on c_1 and (7.17) give the following bound on $c_1' = |C_1'|$:

$$c_1' \ge c_1 - (n - k - c_1) \ge 3k - \frac{n}{10} - 20\log n - n \ge \frac{2}{5}n - 23\log n\ . \qquad (7.18)$$

Since each row $i \in C_1'$ has precisely 2 non-zero entries, it follows that it has the entry 1 at index i and the entry -1 at some index $j > k$.

Applying Lemma 7.3.2 to $M_{\mathcal{B}}$ and $\mathcal{B}$, we obtain a set of at least $c_2 \ge h - \frac{n}{20} - 10\log n$ rows, $C_2 \subset [h]$, and a similar argument to the one above implies that at most $n-h-c_2$ rows can contain more than 2 non-zero entries. Let $C_2' \subset C_2$ denote the set indices of rows of C_2 with precisely two non-zero entries, and let $c_2' = |C_2'|$. By the lower bound on c_2 and (7.17) we obtain:

$$c_2' \ge c_2 - (n - h - c_2) \ge 3h - \frac{n}{10} - 20\log n - n \ge \frac{13}{40}n - 38\log n\ . \qquad (7.19)$$

Note that each row $i \in C_2'$ has the entry 1 at the index i and the entry -1 at some index $j > h$.

Finally, notice that (7.18) and (7.19) imply that $c_1' + c_2' > n/2$ for a sufficiently large value of n. However, as the rows of $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ are orthogonal, the non-zero entries of each pair of rows $i \in C_1'$ and $j \in C_2'$ must be in pairwise disjoint columns. In particular, we obtain that $2c_1' + 2c_2' \le n$, yielding a contradiction. Thus, either $\mathcal{A}$ or $\mathcal{B}$ does not meet the requirements of Lemma 7.3.2, and we deduce that (7.13) holds. ■

7.4 Proof of Theorem 7.1.1 and two lemmas

Let $\mathcal{A}$ and $\mathcal{B}$ denote an ℓ-cross-intersection pair of families in $2^{[n]}$. Recall that in the proof of Theorem 7.3.1, we argued that if, for instance, $\mathcal{A}$ is not an antichain, then $\bigcup_{B\in\mathcal{B}} B \neq [n]$ (see (7.9)). In such a case, letting $i \in [n]$ be so that $i \notin B$ for all $B \in \mathcal{B}$, it follows that $\mathcal{A} = \mathcal{A}' \cup \{A \cup \{i\} : A \in \mathcal{A}'\}$ and $\mathcal{B} = \mathcal{B}'$, where $(\mathcal{A}', \mathcal{B}')$ is an optimal ℓ-cross-intersecting pair on $[n] \setminus \{i\}$. Therefore, by induction, the structure of $\mathcal{A}, \mathcal{B}$ is as specified in Theorem 7.1.1, where the parameter n' (determining the set X in (7.5)) accounts for the modification of $(\mathcal{A}', \mathcal{B}')$ to $(\mathcal{A}, \mathcal{B})$. The same consideration applies when $\bigcup_{A\in\mathcal{A}} A \neq [n]$, which follows when $\mathcal{B}$ is not an antichain (in this case, the set

Y in (7.5) treats the modification of $\mathcal{B}'$ to $\mathcal{B}$). Altogether, we may assume that $\mathcal{A},\mathcal{B}$ are both antichains, and furthermore:

$$\bigcup_{A\in\mathcal{A}} A = \bigcup_{B\in\mathcal{B}} B = [n]\ . \tag{7.20}$$

It remains to prove that in this case $|\mathcal{A}||\mathcal{B}| \leq \binom{2\ell}{\ell}2^{n-2\ell}$, and that equality holds iff for some

$$\begin{aligned} &\kappa \in \{2\ell-1, 2\ell\}\ ,\ \tau \in \{0,\ldots,\kappa\}\ ,\\ &\kappa+\tau = n, \end{aligned} \tag{7.21}$$

the following holds up to a relabeling of the elements of $[n]$ and swapping $\mathcal{A},\mathcal{B}$:

$$\begin{aligned} \mathcal{A} &= \left\{ \bigcup_{T\in J} T\ :\ J\subset \left\{ \begin{array}{c} \{1,\kappa+1\},\ldots,\{\tau,\kappa+\tau\}, \\ \{\tau+1\},\ldots,\{\kappa\} \end{array} \right\}\ ,\ |J| = \ell \right\}\ ,\\ \mathcal{B} &= \left\{ L\cup\{\tau+1,\ldots,\kappa\} : \begin{array}{l} L\subset\{1,\ldots,\tau,\kappa+1,\ldots,\kappa+\tau\} \\ |L\cap\{i,\kappa+i\}| = 1 \text{ for all } i\in[\tau] \end{array} \right\}\ . \end{aligned} \tag{7.22}$$

Following the notations of Theorem 7.3.1, define $\mathcal{F}_{\mathcal{A}},\mathcal{F}'_{\mathcal{B}},k,h$ as in (7.10) and (7.11), obtaining $k\geq h$. Recall that the proof of Theorem 7.3.1 implies that $|\mathcal{A}||\mathcal{B}| \leq 2^{k+h+3}/\sqrt{n}$ provided that ℓ is sufficiently large (equation (7.13)). This implies that if $k+h\leq n-4$ then:

$$|\mathcal{A}||\mathcal{B}| \leq \frac{1}{2}\cdot\frac{2^n}{\sqrt{n}}\ ,$$

and as $\frac{1}{2} < 1/\sqrt{\pi}$, the pair $\mathcal{A},\mathcal{B}$ is suboptimal. Assume therefore that $k+h\geq n-3$:

$$\left\{ \begin{array}{rcl} \frac{n-3}{2} \leq & k & \\ n-3 \leq & k+h & \leq n \end{array} \right. . \tag{7.23}$$

Observe that, as the rows of $M_{\mathcal{A}}$ are orthogonal to the rows of $M_{\mathcal{B}}$, we may assume without loss of generality that:

$$M_{\mathcal{A}} = \left(\ I_k\ \middle|\ *\ \right)\ ,\ M_{\mathcal{B}} = \left(\ *\ \middle|\ I_h\ \right)\ .$$

To see this, first perform Gauss elimination on a basis for $\mathcal{F}_{\mathcal{A}}$ to obtain $M_{\mathcal{A}}$. Next, perform Gauss elimination on a basis for $\mathcal{F}'_{\mathcal{B}}$, and notice that, as the rows of $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ are pairwise orthogonal, it is always possible to find a leading non-zero entry at some index $j > k$. Once $M_{\mathcal{B}}$ is in row-reduced echelon form, we may relabel the elements $k+1, \ldots, n$ to obtain the above structure.

We again apply the arguments of Lemma 7.3.2 on $\mathcal{A}, M_{\mathcal{A}}$ and on $\mathcal{B}, M_{\mathcal{B}}$, only this time we perform the calculations more carefully. Let $R_A \subset [k]$ denote the subset of the rows of $M_{\mathcal{A}}$ which are selected by the process described in Claim 7.3.3. That is, we repeatedly select an arbitrary column with at least 2 non-zero entries, while one exists, add the rows where it is non-zero to R_A, and delete them from $M_{\mathcal{A}}$. While in Claim 7.3.3 we repeatedly selected a column with a maximal number of non-zero entries, here we allow an arbitrary choice when selecting the next column with at least 2 non-zero entries. Let $r_A = |R_A|$, and define $R_B \subset [h]$ and $r_B = |R_B|$ similarly for $M_{\mathcal{B}}$.

Let $S_A \subset [k] \setminus R_A$ denote the indices of rows of $M_{\mathcal{A}}$, which belong neither to R_A nor to $\{0, \pm 1\}^n \setminus \{0,1\}^n$. That is, S_A denotes the rows which were treated by Claim 7.3.4. Let $s_A = |S_A|$, and define $S_B \subset [h] \setminus R_B$ and $s_B = |S_B|$ similarly for $M_{\mathcal{B}}$.

The following lemma, proved in Section 7.5, determines the optimal pairs $\mathcal{A}, \mathcal{B}$ when $r_A + s_A = o(n)$:

Lemma 7.4.1. *If there exists some order of column selection when producing the set R_A such that $r_A + s_A = o(n)$, then $|\mathcal{A}||\mathcal{B}| \le \binom{2\ell}{\ell} 2^{n-2\ell}$. Furthermore, equality holds iff either:*

$$M_{\mathcal{A}} = \left(\begin{array}{c|c|c||c} & 0 & & 0 \\ I_{k-1} & \vdots & -I_{k-1} & \vdots \\ & 0 & & 0 \\ \hline 0 & 1 & 1\ldots1 & 1 \end{array}\right) , \quad M_{\mathcal{B}} = \left(\begin{array}{c|c|c||c} & -1 & & 0 \\ I_{k-1} & \vdots & I_{k-1} & \vdots \\ & -1 & & 0 \\ \hline\hline 0 & -1 & 0\ldots0 & 1 \end{array}\right)$$
$$h \in \{2\ell-2, 2\ell-1\} \ , \ h+k=n \ , \ k \in \{\tfrac{n}{2}, \tfrac{n+1}{2}\} \ , \ B_1 = \cup_{i \in [\ell]} \{(i, k+i)\} \tag{7.24}$$

or :

$$M_{\mathcal{A}} = \left(\begin{array}{c|c|c|c||c|c} & 0 & 0 & & 0 & 0 \\ I_{k-2} & \vdots & \vdots & -I_{k-2} & \vdots & \vdots \\ & 0 & 0 & & 0 & 0 \\ \hline 0 & 1 & 0 & 1\dots1 & 1 & 1 \\ 0 & 0 & 1 & 1\dots1 & 1 & 1 \end{array}\right) ,$$

$$M_{\mathcal{B}} = \left(\begin{array}{c|c|c|c||c|c} & -1 & -1 & & 0 & 0 \\ I_{k-2} & \vdots & \vdots & I_{k-2} & \vdots & \vdots \\ & -1 & -1 & & 0 & 0 \\ \hline\hline 0 & -1 & -1 & 0\dots0 & 1 & 0 \\ 0 & -1 & -1 & 0\dots0 & 0 & 1 \end{array}\right) \tag{7.25}$$

$$h \in \{2\ell-2, 2\ell-1\} \ , \ h+k=n \ , \ k \in \{\tfrac{n}{2}, \tfrac{n+1}{2}, \tfrac{n}{2}+1\} \ ,$$
$$B_1 = \cup_{i\in[\ell]}\{(i,k+i)\}$$

up to a relabeling of the elements of $[n]$ *and the choice of* B_1*. In both cases above, the pair* $(\mathcal{A},\mathcal{B})$ *belongs to the family* (7.22) *with* $\kappa = h+1$, $\tau = k-1$ *and swapping* $\mathcal{A},\mathcal{B}$*.*

In the above figures (7.24) and (7.25), the columns to the right of the double-line-separators and the rows below the double-line-separators appear or not, depending on the value of k.

The remaining case is treated by the next lemma, which is proved in Section 7.6, and concludes the proof of the theorem:

Lemma 7.4.2. *If every order of column selection when producing the set* R_A *gives* $r_A + s_A = \Omega(n)$*, then* $|\mathcal{A}||\mathcal{B}| \leq \binom{2\ell}{\ell}2^{n-2\ell}$*. Furthermore, equality holds iff:*

$$M_{\mathcal{A}} = \left(\begin{array}{c|c|c} I_h & 0 & I_h \\ \hline 0 & I_{k-h} & 0 \end{array}\right) \ , \ M_{\mathcal{B}} = \left(\begin{array}{c|c|c} -I_h & 0 & I_h \end{array}\right) \tag{7.26}$$
$$k \in \{2\ell-1, 2\ell\} \ , \ h+k=n \ , \ B_1 = [k]$$

up to a relabeling of the elements of $[n]$*. In this case, the pair* $(\mathcal{A},\mathcal{B})$ *belongs to the family* (7.22) *with* $\kappa = k$ *and* $\tau = h$*.*

Remark 7.4.3: It is, in fact, not difficult to check that if, in one order of column selection we have $r_A + s_A = \Omega(n)$, so is the case in any order, but the above formulation suffices for our purpose.

7.5 Proof of Lemma 7.4.1

Let $C_1 = [k] \setminus (R_A \cup S_A)$. By the assumption on r_A, s_A and the fact that $k \geq \frac{n-3}{2}$ we deduce that $|C_1| = (1 - o(1))k$. Recall that each column of $M_{\mathcal{A}}$ contains at most one non-zero entry in the rows of C_1, and that each row of C_1 belongs to $\{0, \pm 1\}^n \setminus \{0, 1\}^n$. Hence, $n \geq k + |C_1| = (2 - o(1))k$. Altogether, we obtain that:

$$k = \left(\frac{1}{2} + o(1)\right) n \quad , \quad h = \left(\frac{1}{2} - o(1)\right) n \ . \tag{7.27}$$

The $\{1, -1\}$ entries in each row of C_1 account for $2|C_1| = (1 - o(1))n$ distinct columns, leaving at most $o(n)$ columns which may contribute additional values to rows of C_1. Again, as each column contains at most 1 non-zero entry in the rows of C_1, the set of all rows with non-zero entries either in these columns, or in columns $\{k + 1, \ldots, n - h\}$ (at most 3 columns), is of size $o(n)$. We obtain that, without loss of generality:

$$M_{\mathcal{A}} = \begin{array}{c} \begin{array}{ccccc} \multicolumn{2}{c}{\overleftrightarrow{\cdots\ k\ \cdots}} & \|\overleftrightarrow{\leq 3}\| & \multicolumn{2}{c}{\overleftrightarrow{\cdots\ h\ \cdots}} \end{array} \\ \left(\begin{array}{c|c|c|c|c} I_{k'} & 0 & 0 & -I_{k'} & 0 \\ 0 & I_{k-k'} & * & * & * \end{array}\right) \end{array} , \tag{7.28}$$

where $k' = (1 - o(1))k = (1 - o(1))h$. The above structure of $M_{\mathcal{A}}$ provides a quick bound on $|\mathcal{A}|$. Consider column $n - h + 1$; if this column contains at least 2 non-zero entries, then we gain a factor of $\frac{3}{4}$ by (7.14). Otherwise, the fact that $M_{n-h+1,1} = -1$ implies that the coefficient of row 1 is necessarily 0, giving a factor of $\frac{1}{2}$. Therefore:

$$|\mathcal{A}| \leq \frac{3}{4} \cdot 2^k \ . \tag{7.29}$$

For another corollary of (7.28), notice that for all $i \in [k']$, row i of $M_{\mathcal{A}}$ contains $1, -1$ in columns $i, n - h + i$ respectively (and 0 in the remaining columns), and is orthogonal to all rows of $M_{\mathcal{B}}$. It follows that columns $i, n - h + i$ are equal in $M_{\mathcal{B}}$ for all $i \in [k']$, and hence:

$$M_{\mathcal{B}} = \begin{array}{c} \begin{array}{ccccc} \multicolumn{2}{c}{\overleftrightarrow{\cdots\ k\ \cdots}} & \|\overleftrightarrow{\leq 3}\| & \multicolumn{2}{c}{\overleftrightarrow{\cdots\ h\ \cdots}} \end{array} \\ \left(\begin{array}{c|c|c|c|c} I_{k'} & * & * & I_{k'} & 0 \\ 0 & * & * & 0 & I_{h-k'} \end{array}\right) \end{array} . \tag{7.30}$$

We claim that the above structure of $M_{\mathcal{B}}$ implies that $r_B + s_B = (1-o(1))h$. Indeed, once we delete the rows $R_B \cup S_B$ from $M_{\mathcal{B}}$, each row must contain an entry of -1, which must reside in one of the columns $k'+1,\ldots,n-h$. As each column contains at most one non-zero entry in rows $[h] \setminus (R_B \cup S_B)$, we deduce that $n - h - k' \geq h - r_B - s_B$, and equivalently:

$$r_B + s_B \geq 2h + k' - n = (1-o(1))h = \left(\frac{1}{2} - o(1)\right) n \ ,$$

where the last two equalities are by (7.27) and the fact that $k' = (1-o(1))k$. Recall that the analysis of Claim 7.3.3 implies that, if R_B is nonempty, then at most $2^{r_B+1}/\sqrt{\frac{\pi}{2} r_B}$ linear combinations of the rows of R_B are valid in order to produce a $\{0,1\}^n$ vector from the rows of $M_{\mathcal{B}}$. Furthermore, if S_B is nonempty, then for each choice of coefficients for the rows $[h] \setminus S_B$, Claim 7.3.4 implies that at most $2^{s_B}/\sqrt{\frac{\pi}{2} s_B}$ combinations of the rows of S_B are valid in order to produce a $\{0,1\}^n$ antichain of vectors from the rows of $M_{\mathcal{B}}$. Since in our case we have $r_B + s_B = \Omega(n)$, at least one of r_B, s_B is $\Omega(n)$, and we deduce that:

$$|\mathcal{B}| = O(2^h/\sqrt{n}) \ . \tag{7.31}$$

Furthermore, if both $r_B = \omega(1)$ and $s_B = \omega(1)$ we get $|\mathcal{B}| = O(\frac{2^h}{\sqrt{r_B s_B}}) = o(2^h/\sqrt{n})$ and hence (regardless of the structure of $M_{\mathcal{A}}$)

$$|\mathcal{A}||\mathcal{B}| = o(2^{k+h}/\sqrt{n}) \leq o(2^n/\sqrt{\ell}) \ ,$$

showing this cannot be an optimal configuration, as required. The same consequence is obtained if either $r_A = \omega(1)$ or $s_A = \omega(1)$, as in this case $|\mathcal{A}| = o(2^k)$. Assume therefore that $r_A + s_A = O(1)$, and by the above arguments we obtain that:

$$k = \tfrac{n}{2} + O(1) \ , \ h = \tfrac{n}{2} - O(1) \ , \tag{7.32}$$

$$k' = k - O(1) \ , \tag{7.33}$$

$$r_B = O(1) \ , \ s_B = h - O(1) \quad \text{or} \quad r_B = h - O(1) \ , \ s_B = O(1) \ . \tag{7.34}$$

At this point, we claim that either $n = (4+o(1))\ell$, or the pair $\mathcal{A},\mathcal{B}$ is suboptimal:

Claim 7.5.1. *Let $\mathcal{A},\mathcal{B}$ be as above, then either $|\mathcal{A}||\mathcal{B}| = o(2^n/\sqrt{n})$ or $n = (4+o(1))\ell$.*

Proof. Fix a choice of coefficients for the last $k-k'$ rows of $M_{\mathcal{A}}$, yielding a linear combination w_A. By the structure of $M_{\mathcal{A}}$ specified in (7.28), if for some index $i \in [k']$, w_A does not equal 0 at index i or does not equal 1 at index $n-h+i$, then the i-th row of $M_{\mathcal{A}}$ has at most one valid coefficient. Thus, if there are $\omega(1)$ such indices, we deduce that there are at most $o(2^{k'})$ combinations of the rows $[k']$ of $M_{\mathcal{A}}$ which extend w_A to an element of $\mathcal{A}$. Therefore, by (7.31), this choice of w_A counts for at most $o(2^{k'+h}/\sqrt{n})$ pairs $(A,B) \in \mathcal{A}\times\mathcal{B}$. Summing over all $2^{k-k'}$ choices for w_A, this amounts to at most $o(2^n/\sqrt{n})$ pairs $(A,B) \in \mathcal{A}\times\mathcal{B}$, and we may thus assume that at least $k'-O(1)$ of the indices $j \in [k']$ satisfy

$$w_A^{(j)} = 0\ ,\ w_A^{(n-h+j)} = 1\ . \tag{7.35}$$

Next, fix a choice of coefficients for the last $h-k'$ rows of $M_{\mathcal{B}}$, yielding an affine combination (together with χ_{B_1}) w_B, and consider the structure of $M_{\mathcal{B}}$ specified in (7.30). Every index $j \in [k']$ for which $\chi_{B_1}^{(j)} \neq \chi_{B_1}^{(n-h+j)}$ implies that the row j has at most one valid coefficient. Thus, if there are $\omega(1)$ such indices, it follows that w_B can be extended to at most $o(2^h/\sqrt{n})$ elements of $\mathcal{B}$. To see this, take $m=\omega(1)$ and yet $m=o(n)$ such rows, arbitrarily; there is at most one legal combination for these rows. As $r_B+s_B=\Omega(n)$, the remaining rows have at most $O(2^{h-m}/\sqrt{n})$ combinations, and the result follows.

Altogether, we may assume that $k'-O(1)$ of the indices $j\in[k']$ satisfy:

$$\chi_{B_1}^{(j)} = \chi_{B_1}^{(n-h+j)}\ . \tag{7.36}$$

Let $L \subset [k']$ denote the indices of $[k']$ which satisfy both (7.35) and (7.36). It follows that $|L| = h - O(1)$, and for each $i \in L$, the choice of a coefficient for row i exclusively determines between the cases $i, n+h-i \in B$ and $i, n+h-i \notin B$.

Fix a choice of coefficients for the remaining rows of $M_{\mathcal{A}}$, and let A denote the resulting set, and fix a choice of coefficients for all rows of $M_{\mathcal{B}}$ except those whose indices are in L. For each $i \in L$, let X_i denote the variable whose value is 1 if we choose a coefficient for the row i such that $i, n+h-i \in B$ and 0 otherwise. Recall that A contains precisely one element from each pair $\{i, n+h-i \ : \ i \in L\}$. Therefore, any choice of coefficients of the rows L in $M_{\mathcal{B}}$ gives a set B which satisfies:

$$\ell = |A \cap B| = (\sum_{i \in L} X_i) + O(1) \ , \tag{7.37}$$

where the $O(1)$-term accounts for the intersection of A with at most $n - 2|L| = O(1)$ indices. Choose one of each pair of coefficients for each row of L uniformly at random and independently of the other rows, to obtain that $X = \sum_{i \in L} X_i$ has a binomial distribution $\mathrm{Bin}(\frac{n}{2} - O(1), \frac{1}{2})$. Fix some small $\varepsilon > 0$; by the Chernoff bound (see, e.g., [19], Chapter A.1):

$$\Pr[|X - \frac{n}{4}| > \varepsilon n] \leq O\left(\exp(-\Omega(n))\right) \ ,$$

thus if $|\ell - \frac{n}{4}| > \varepsilon n$ then at most $O(2^h / \exp(\Omega(n)))$ sets $B \in \mathcal{B}$ can be produced from w_B and we are done. We conclude that $\ell = (\frac{1}{4} + o(1))n$. ■

The last claim, along with (7.34), implies that the case $s_B = h - O(1)$ is suboptimal. Indeed, in this case:

$$|\mathcal{B}| \leq \frac{2^h}{\sqrt{\pi s_B/2}} = (1+o(1))\frac{2^h}{\sqrt{\pi h/2}} = (1+o(1))\frac{2^h}{\sqrt{\pi n/4}} = (1+o(1))\frac{2^h}{\sqrt{\pi \ell}} \ ,$$

where the last inequality is by Claim 7.5.1. Combining this with (7.29), we deduce that $|\mathcal{A}||\mathcal{B}|$ is at most $(\frac{3}{4} + o(1))2^n/\sqrt{\pi\ell}$, and that the pair $\mathcal{A}, \mathcal{B}$ is suboptimal.

It remains to deal with the case $r_B = h - O(1)$, in which case we have:

$$|\mathcal{B}| \leq \frac{2^{h+1}}{\sqrt{\pi r_B/2}} = (2+o(1))\frac{2^h}{\sqrt{\pi \ell}} \ , \tag{7.38}$$

and hence ($|\mathcal{A}| \leq \frac{3}{4} \cdot 2^k$), $|\mathcal{A}||\mathcal{B}| \leq (\frac{3}{2} + o(1))2^{k+h}/\sqrt{\pi\ell}$. If $k + h < n$, it follows that $|\mathcal{A}||\mathcal{B}|$ is at most $(\frac{3}{4} + o(1))2^n/\sqrt{\pi\ell}$, and again the pair $\mathcal{A}, \mathcal{B}$ is suboptimal. We may thus assume:

$$k + h = n \ , \ r_B = h - O(1) \ , \ s_B = O(1) \ .$$

To complete the proof of the lemma, we show that either $|\mathcal{A}||\mathcal{B}| \leq (\delta + o(1))2^n/\sqrt{\pi\ell}$ for some fixed $\delta < 1$, or all columns of $M_{\mathcal{B}}$ except either 1 or 2 have at most 1 non-zero entry, whereas the remaining columns are of the form $(-1, \ldots, -1)$. This will imply that either (7.24) holds or (7.25) holds. For this purpose, we must first concentrate on the $(k - k') \times k'$ sub-matrix of $M_{\mathcal{A}}$, on rows $\{k' + 1, \ldots, k\}$ and columns $\{k + 1, \ldots, k + k'\}$. This sub-matrix appears boxed in diagram (7.39), which reflects the form of $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ given the fact $k + h = n$:

$$\begin{array}{l}
M_{\mathcal{A}} = \begin{array}{c}
\begin{array}{cc} \longleftarrow \cdots\cdots k \cdots\cdots \longrightarrow & || \longleftarrow \cdots\cdots h \cdots\cdots \longrightarrow \end{array}\\
\left(\begin{array}{c|c|c|c} I_{k'} & 0 & -I_{k'} & 0 \\ 0 & I_{k-k'} & \boxed{*} & * \end{array}\right)
\end{array} \\
M_{\mathcal{B}} = \begin{array}{c}
\begin{array}{cc} \longleftarrow \cdots\cdots k \cdots\cdots \longrightarrow & || \longleftarrow \cdots\cdots h \cdots\cdots \longrightarrow \end{array}\\
\left(\begin{array}{c|c|c|c} I_{k'} & * & I_{k'} & 0 \\ 0 & * & 0 & I_{h-k'} \end{array}\right)
\end{array}
\end{array} . \tag{7.39}$$

Suppose the linear combination of rows $k'+1, \ldots, k$ of $M_{\mathcal{A}}$ is some vector $w_{\mathcal{A}}$. A key observation is the following: if $w_{\mathcal{A}}$ has $\omega(1)$ entries not equal to 1 in indices $\{k+1, \ldots, k+k'\}$, then at most $o(2^{k'})$ combinations of the remaining rows can be added to $w_{\mathcal{A}}$ to produce a vector in $\{0,1\}^n$. This follows directly from the structure of $M_{\mathcal{A}}$ in (7.28), as the fact that $w_{\mathcal{A}}^{(k+j)} \neq 1$ forces the coefficient of row j to be 0. Using the above observation, we will show that either $|\mathcal{A}| \leq (\frac{3}{8} + o(1))2^k$, or at most $O(1)$ columns of $M_{\mathcal{A}}$ with indices $\{k + 1, \ldots, k+k'\}$ are not of one of the forms $\{(-1, 1, 0, \ldots, 0), (-1, 1, 1, 0, \ldots, 0)\}$ (at some coordinate order). Consider the following three cases:

(I) $\omega(1)$ **columns of** $M_{\mathcal{A}}$ **contain at least** 3 **non-zero entries in rows** $\{k' + 1, \ldots, k\}$**:** Let S denote the indices of columns in $\{k + 1, \ldots, k + k'\}$ for which $M_{\mathcal{A}}$ has non-zero entries in rows $\{k' + 1, \ldots, k\}$. The Littlewood-Offord Lemma implies that, whenever there are t non-zero entries in a single column in these rows, then at most $m = 2^{k-k'-t}\binom{t}{\lfloor t/2 \rfloor}$ of the $2^{k-k'}$ possible linear combinations of these rows can produce a value of 1. Notice that for $t \geq 3$ we get $\binom{t}{\lfloor t/2 \rfloor}/2^t \leq \frac{3}{8}$, hence $m/2^{k-k'} \leq \frac{3}{8}$. Next, let each column which has at least 3 non-zero entries in rows $\{k' + 1, \ldots, k\}$ "rate" m linear combinations, including all those for which it gives a value of 1. It follows that choosing any combination for rows $\{k'+1,\ldots,k\}$ excluding the most popular set of m linear combinations, yields values not equal to 1 in at least $|S|/\binom{2^{k'-k}}{m} = \Omega(|S|) = \omega(1)$ columns, hence (by the above observation) such combinations contribute $o(2^k)$ vectors to $\mathcal{A}$. We deduce that $|\mathcal{A}| \leq (\frac{3}{8}+o(1))2^k$.

(II) **$\omega(1)$ columns of $M_{\mathcal{A}}$ contain 2 non-zero entries $\neq (1,1)$ in rows $\{k'+1,\ldots,k\}$:** The argument here is similar to the argument in the previous item. If a column has two non-zero entries $(x,y) \neq (1,1)$ in rows $k'+1,\ldots,k$, then the possible values of the linear combination at this column are $\{0,x,y,x+y\}$. At most 1 of these 4 values can be 1, hence at most $m = 2^{k-k'-2}$ of the combinations yield a value of 1 at this column. By the above argument, we deduce that $|\mathcal{A}| \leq (\frac{1}{4}+o(1))2^k$.

(III) **$\omega(1)$ columns of $M_{\mathcal{A}}$ contain at most 1 non-zero entry $\neq 1$ in rows $\{k'+1,\ldots,k\}$:** this case is the simplest, following directly from the observation. Indeed, every linear combination of the rows $k'+1,\ldots,k$ has $\omega(1)$ entries which do not equal 1 in columns $\{k+1,\ldots,k+k'\}$, hence $|\mathcal{A}| = o(2^k)$.

Note that if $|\mathcal{A}| \leq (\frac{3}{8}+o(1))2^k$, then $|\mathcal{A}||\mathcal{B}| \leq (\frac{3}{4}+o(1))2^n/\sqrt{\pi\ell}$ by (7.38), as required. Assume therefore that $M_{\mathcal{A}}$ has at most $O(1)$ columns among $\{k+1,\ldots,k+k'\}$, whose set of non-zero entries in rows $\{k'+1,\ldots,k\}$ is neither $\{1\}$ nor $\{1,1\}$. We use the abbreviation $\{1\}$-columns and $\{1,1\}$-columns for the $k'-O(1)$ remaining columns whose non-zero entries in rows $\{k'+1,\ldots,k\}$ of $M_{\mathcal{A}}$ are $\{1\}$ and $\{1,1\}$ respectively; according to this formulation:

$$\begin{aligned} &k'-O(1) \text{ of columns } \{k+1,\ldots,k'\} \text{ of } M_{\mathcal{A}} \text{ are} \\ &\qquad \{1\}\text{-columns or } \{1,1\}\text{-columns}\,. \end{aligned} \tag{7.40}$$

The two cases of whether there are $\omega(1)$ or $O(1)$ $\{1\}$-columns, are treated by Claims 7.5.2 and 7.5.3 respectively, and determine which of the two optimal families, stated in (7.24),(7.25), is obtained. These two claims are stated and proved in Subsections 7.5.1 and 7.5.2.

7.5.1 The optimal family (7.24)

Claim 7.5.2. *If $\omega(1)$ of columns $\{k+1,\dots,k+k'\}$ of $M_{\mathcal{A}}$ are $\{1\}$-columns, then (7.24) holds.*

Proof. It follows that some row of $\{k'+1,\dots,k\}$ contains a value of 1, which is the single non-zero entry of this column in these rows, in $\omega(1)$ columns of $\{k+1,\dots,k+k'\}$ (take the most popular row of $\{k'+1,\dots,k\}$). Without loss of generality, assume that this row is row k, the last row of $M_{\mathcal{A}}$. By the observation above, the coefficient for row k of $M_{\mathcal{A}}$ must be 1, otherwise only $o(2^k)$ combinations of the remaining rows produce vectors in $\{0,1\}^n$. This has several consequences:

(1) Row k contains the value 1 in columns $\{k+1,\dots,k+k'\}$. To see this, notice that if $(M_{\mathcal{A}})_{k,k+j} \neq 1$ for some $j \in [k']$, then $|\mathcal{A}| \leq \left(\frac{1}{4}+o(1)\right)2^k$: either the coefficient for row k is 0, contributing $o(2^k)$ vectors to $|\mathcal{A}|$, or it is 1, forcing the coefficient of row j to be 0.

(2) Row k contains $\{0,1\}$ values in columns $\{k+k'+1,\dots,n\}$. Indeed, if $(M_{\mathcal{A}})_{k,k+j} \notin \{0,1\}$ for some $j \in \{k'+1,\dots,n-k\}$, then the all-zero choice of coefficients for rows $\{k'+1,\dots,k-1\}$ becomes illegal when giving row k the coefficient 1, implying that $|\mathcal{A}| \leq \left(\frac{\delta}{2}+o(1)\right)2^k$, where $\delta = 1-2^{-(k-k')}$.

(3) If $M'_{\mathcal{A}}$ is the $(k-1)\times n$ sub-matrix of rows $\{1,\dots,k-1\}$ of $M_{\mathcal{A}}$ (that is, the matrix obtained by erasing the last row of $M_{\mathcal{A}}$), then every column of $M'_{\mathcal{A}}$ contains at most 1 non-zero entry, and every row of $M'_{\mathcal{A}}$ belongs to $\{0,\pm1\}^n \setminus \{0,1\}^n$. To see this, notice that the coefficient of row k is set to 1, otherwise we obtain at most $o(2^k)$ vectors. We can thus regard this row as an affine vector in $\{0,1\}^n$, and consider the 2^{k-1} combinations for the remaining rows. Now, a column of $M'_{\mathcal{A}}$ with at least 2 non-zero entries implies that the number of such legal combinations (resulting in a vector in $\{0,1\}^n$) is at most $\frac{3}{4}\cdot 2^{k-1}$, and a row which does not belong to $\{0,\pm1\}^n \setminus \{0,1\}^n$ implies that this number is at most 2^{k-2}. In both cases, we get $|\mathcal{A}| \leq (\frac{3}{8}+o(1))2^k$.

(4) Every row of $M'_{\mathcal{A}}$ has at most 2 non-zero values: assume that the converse holds, that is, that row $m \in [k-1]$ contains at least 2 non-zero entries in indices $\{k+1,\ldots,n\}$. Since each of the $k-1$ rows of $M'_{\mathcal{A}}$ must contain a -1 value in an exclusive column, it leaves at most $n-k-(k-1) = n-2k+1 \leq 1$ column (recall that $k \geq \frac{n}{2}$), which can contribute 1 additional non-zero value to row m. We deduce that row m has precisely two non-zero entries at columns $\{k+1,\ldots,n\}$. However, in this case column m of $M_{\mathcal{B}}$ has precisely two non-zero entries , since (7.39) and the orthogonality of $M_{\mathcal{A}}, M_{\mathcal{B}}$ imply that:

$$(M_{\mathcal{A}})_{i,k+j} = -(M_{\mathcal{B}})_{j,i} \;\text{ for all } i \in [k] \text{ and } j \in [h] \tag{7.41}$$

(the inner product of row i of $M_{\mathcal{A}}$ and row j of $M_{\mathcal{B}}$ is $(M_{\mathcal{A}})_{i,k+j}+(M_{\mathcal{B}})_{j,i} = 0$). From the same reason, column k of $M_{\mathcal{B}}$ has at least k' non-zero entries (as row k of $M_{\mathcal{A}}$ has the value 1 in columns $\{k+1,\ldots,k+k'\}$). Therefore, performing the process of Claim 7.3.3 first on column m and then on column k of $M_{\mathcal{B}}$ gives $|\mathcal{B}| \leq \frac{3}{4} \cdot \frac{2+o(1)}{\sqrt{\pi\ell}}$, hence the pair $\mathcal{A},\mathcal{B}$ is suboptimal.

Items (3) and (4) imply that, if the pair $\mathcal{A},\mathcal{B}$ is optimal, then without loss of generality, $M'_{\mathcal{A}}$ is of the form $\left(I_{k-1}|0|-I_{k-1}|0 \right)$, as each row has $1,-1$ in exclusive columns and 0 everywhere else. In particular, $k' = k-1$, and since $k \geq n/2$ and $k+k' \leq n$, we get:

$$k = h = \frac{n}{2} \quad \text{or} \quad (k = \frac{n+1}{2}\ ,\ h = \frac{n-1}{2})\ , \tag{7.42}$$

and without loss of generality (using the orthogonality of $M_{\mathcal{A}}, M_{\mathcal{B}}$):

$$M_{\mathcal{A}} = \left(\begin{array}{c|c|c||c} & 0 & & 0 \\ I_{k-1} & \vdots & -I_{k-1} & \vdots \\ & 0 & & 0 \\ \hline 0 & 1 & 1\ldots1 & 0/1 \end{array}\right)\ ,\ M_{\mathcal{B}} = \left(\begin{array}{c|c|c||c} & -1 & & 0 \\ I_{k-1} & \vdots & I_{k-1} & \vdots \\ & -1 & & 0 \\ \hline\hline 0 & 0/-1 & 0\ldots0 & 1 \end{array}\right)\ , \tag{7.43}$$

where the last column of $M_{\mathcal{A}}$ and the last row and column of $M_{\mathcal{B}}$ do not exist in case $k = (n+1)/2$. If $h = n/2$ and $(M_{\mathcal{B}})_{h,k} = 0$ (as opposed to -1), then $|\mathcal{B}| \leq (1+o(1))2^h/\sqrt{\pi\ell}$: the first $h-1$ rows have at most

$(2+o(1))2^{h-1}/\sqrt{\pi\ell}$ combinations by the usual Littlewood-Offord argument on column k, and when adding row h we must form an antichain. It follows that if $k=h=n/2$, then $(M_{\mathcal{B}})_{h,k}=-1$ and, by orthogonality, $(M_{\mathcal{A}})_{k,n}=1$:

$$M_{\mathcal{A}}=\left(\begin{array}{c|c|c||c} \multicolumn{3}{c||}{\ddots} & \vdots \\ \multicolumn{3}{c||}{} & 0 \\ \hline 0 & 1 & 1\ldots1 & 1 \end{array}\right)\ ,\ M_{\mathcal{B}}=\left(\begin{array}{c|c|c||c} \multicolumn{3}{c||}{\ddots} & \vdots \\ \multicolumn{3}{c||}{} & 0 \\ \hline\hline 0 & -1 & 0\ldots0 & 1 \end{array}\right)\ .$$

Finally, notice that the above structure of $M_{\mathcal{A}}$ implies that the coefficient for row k is always 1: a coefficient of 0 necessarily results in the all-zero vector, which is forbidden in $\mathcal{A}$ (for instance, since $|\mathcal{A}|$ is an antichain, or since $\ell>0$). Therefore:

$$|\mathcal{A}|\leq 2^{k-1}\ .$$

If $\chi_{B_1}^{(j)}\neq\chi_{B_1}^{(k+j)}$ for some $j\in[k-1]$, we must assign the coefficient 0 to row j of $M_{\mathcal{B}}$, and we are done, as in this case $|\mathcal{B}|\leq(1+o(1))2^h/\sqrt{\pi\ell}$. Assume therefore that $\chi_{B_1}^{(j)}=\chi_{B_1}^{(k+j)}$ for all $j\in[k-1]$, and define:

$$P=\{i\in[h]:k+i\notin B_1\}=\{i\in[h]:\chi_{B_1}^{(k+i)}=0\}\ ,\ Q=[h]\setminus P\ .$$

Every row $i\in P$ of $M_{\mathcal{B}}$ has $\{0,1\}$ as the set of possible coefficients, and every row $i\in Q$ has $\{0,-1\}$ as the possible coefficients. Take $B\in\mathcal{B}$, and suppose that the affine combination which produces B assigns the coefficient 1 to p rows of P ($0\leq p\leq|P|$), and assigns the coefficient -1 to q rows of Q ($0\leq q\leq|Q|$). It follows from (7.43) that for all $A\in\mathcal{A}$:

$$\ell=|A\cap B|=p+(|Q|-q)+\chi_B^{(k)}\ . \tag{7.44}$$

Let $\mathcal{B}_0$ denote the sets $\{B\in\mathcal{B}:k\notin B\}$, and let $\mathcal{B}_1=\mathcal{B}\setminus\mathcal{B}_0$. By (7.44), we obtain that $q=p+|Q|-\ell$ if $k\notin B$, hence:

$$|\mathcal{B}_0|\leq\sum_{p=0}^{|P|}\binom{|P|}{p}\binom{|Q|}{p+|Q|-\ell}=\sum_{p=0}^{|P|}\binom{|P|}{p}\binom{|Q|}{\ell-p}=\binom{h}{\ell}\ .$$

Similarly, if $k\in B$ then $q=p+|Q|-\ell+1$, and it follows that: $|\mathcal{B}_1|\leq\binom{h}{\ell-1}$. Altogether:

$$|\mathcal{B}|=|\mathcal{B}_0|+|\mathcal{B}_1|\leq\binom{h}{\ell}+\binom{h}{\ell-1}=\binom{h+1}{\ell}\ ,$$

and as $|\mathcal{A}| \leq 2^{k-1}$:

$$|\mathcal{A}||\mathcal{B}| \leq \binom{h+1}{\ell} 2^{n-h-1} \ . \tag{7.45}$$

As the maxima of the function $f(x) = \binom{x}{\ell}2^{-x}$ on the domain $\mathbb{N}$ are achieved at $x \in \{2\ell-1, 2\ell\}$, we conclude that $h \in \{2\ell-2, 2\ell-1\}$ (otherwise $|\mathcal{A}||\mathcal{B}| < \binom{2\ell}{\ell}2^{n-2\ell}$). Finally, recalling that:

$$\chi_B^{(k)} = q - p + \chi_{B_1}^{(k)} \ , \tag{7.46}$$

and combining (7.44) and (7.46) we get:

$$\ell = |Q| + \chi_{B_1}^{(k)} \ .$$

Therefore, whenever $\chi_{B_1}^{(k)} = 0$ we get $|Q| = \ell$, hence $B = \cup_{i \in [\ell]}\{(i, k+i)\}$ for some $B \in \mathcal{B}$. Letting B_1 denote this set B without loss of generality, we obtain the statement of (7.24).

Finally, let us link the above to the optimal family (7.22). Define:

$$X = \begin{cases} \{k, n\} & \text{if } k = \frac{n}{2} \\ \{k\} & \text{if } k = \frac{n+1}{2} \end{cases} \ .$$

Each set $A \in \mathcal{A}$ is obtained by choosing one out of each pair of elements $\big\{\{i, k+i\} : i \in [k-1]\big\}$, then adding these $k-1$ chosen elements to the elements of X. Define:

$$Y = \begin{cases} \big\{\{i, k+i\} : i \in [k-1]\big\} \cup \big\{\{n\}\big\} & \text{if } k = \frac{n}{2} \\ \big\{\{i, k+i\} : i \in [k-1]\big\} & \text{if } k = \frac{n+1}{2} \end{cases} \ .$$

Each set $B \in \mathcal{B}_1$ (that is, those sets which contain k) has, in addition to k, $\ell - 1$ objects of Y. Each set $B \in \mathcal{B}_0$ is the union of ℓ objects of Y, and altogether, all sets $B \in \mathcal{B}$ are the union of ℓ objects of $Y \cup \big\{\{k\}\big\}$. As the last set holds the $k-1$ pairs $\{i, k+i\}$ for $i \in [k-1]$ and the single elements corresponding to X, this fits the description of (7.22) for $\kappa = h+1$, $\tau = k-1$ and swapping $\mathcal{A}, \mathcal{B}$. ■

7.5.2 The optimal family (7.25)

Claim 7.5.3. *If $O(1)$ of columns $\{k+1,\ldots,k+k'\}$ of $M_{\mathcal{A}}$ are $\{1\}$-columns, then (7.25) holds.*

Proof. By the assumption and by (7.40), we obtain that $k'-O(1)$ of the columns $\{k+1,\ldots,k+k'\}$ are $\{1,1\}$-columns, that is, there are $k'-O(1)$ columns $j \in \{k+1,\ldots,k+k'\}$ where there are precisely two non-zero entries in rows $\{k'+1,\ldots,k\}$, and both entries are equal to 1. For each such column j, let $i_1(j), i_2(j) \in \{k'+1,\ldots,k\}$ denote the rows where these two entries are located. Assume that, without loss of generality, the pair of rows $k-1, k$ is the most popular pair among the above pairs of rows $\{(i_1(j), i_2(j)) :$ j is a $\{1,1\}$-column$\}$; it follows that there are $\omega(1)$ columns (and in fact, $\Omega(k')$ columns) $j \in \{k+1,\ldots,k+k'\}$ such that:

$$\left\{\begin{array}{l}(M_{\mathcal{A}})_{k-1,j} = (M_{\mathcal{A}})_{k,j} = 1 \ , \\ (M_{\mathcal{A}})_{i,j} = 0 \text{ for all } i \in \{k'+1,\ldots,k-2\} \ .\end{array}\right.$$

Hence, if we assign the same coefficient to rows $k-1, k$ then we obtain $\omega(1)$ values which differ from 1 in columns $\{k+1,\ldots,k'\}$, and contribute $o(2^k)$ vectors to $\mathcal{A}$. We must therefore assign the coefficient 1 to precisely one of the rows $k-1, k$ (and assign the coefficient 0 to the other).

The arguments given in the proof of Claim 7.5.2 regarding row k readily imply the following analogous results on rows $k-1, k$:

(1) Rows $k-1, k$ contain the value 1 in columns $\{k+1,\ldots,k\}$.

(2) Rows $k-1, k$ belong to $\{0,1\}^n$.

(3) If $M'_{\mathcal{A}}$ is the $(k-2) \times n$ sub-matrix of rows $\{1,\ldots,k-2\}$ of $M_{\mathcal{A}}$, then every column of $M'_{\mathcal{A}}$ contains at most 1 non-zero entry, and every row of $M'_{\mathcal{A}}$ belongs to $\{0,\pm 1\}^n \setminus \{0,1\}^n$.

(4) Every row of $M'_{\mathcal{A}}$ contains at most 2 non-zero entries.

By the last two items, we deduce that if $\mathcal{A}, \mathcal{B}$ is an optimal pair, then without loss of generality, $M'_{\mathcal{A}} = \left(I_{k-2} | 0 | -I_{k-2} | 0 \right)$, and in particular, $k' = k-2$. The constraints $k \geq n/2$ and $k + k' \leq n$ now imply:

$$k = h = \frac{n}{2} \quad \text{or} \quad (k = \frac{n+1}{2}\ ,\ h = \frac{n-1}{2}) \quad \text{or} \quad (k = \frac{n}{2} + 1\ ,\ h = \frac{n}{2} - 1)\ , \tag{7.47}$$

and by orthogonality:

$$M_{\mathcal{A}} = \left(\begin{array}{c|c|c|c||c|c} & 0 & 0 & & 0 & 0 \\ I_{k-2} & \vdots & \vdots & -I_{k-2} & \vdots & \vdots \\ & 0 & 0 & & 0 & 0 \\ \hline 0 & 1 & 0 & 1\ldots1 & 0/1 & 0/1 \\ \hline 0 & 0 & 1 & 1\ldots1 & 0/1 & 0/1 \end{array}\right)\ ,$$

$$M_{\mathcal{B}} = \left(\begin{array}{c|c|c|c||c|c} & -1 & -1 & & 0 & 0 \\ I_{k-2} & \vdots & \vdots & I_{k-2} & \vdots & \vdots \\ & -1 & -1 & & 0 & 0 \\ \hline\hline 0 & 0/-1 & 0/-1 & 0\ldots0 & 1 & 0 \\ \hline 0 & 0/-1 & 0/-1 & 0\ldots0 & 0 & 1 \end{array}\right)\ , \tag{7.48}$$

where the last two columns of $M_{\mathcal{A}}$ and the last two rows and columns of $M_{\mathcal{B}}$ are optional, depending on whether $k = \frac{n}{2} + 1$, $k = \frac{n+1}{2}$ or $k = \frac{n}{2}$ (where we have 0, 1 or 2 of the last columns of $M_{\mathcal{A}}$ and the last rows and columns of $M_{\mathcal{B}}$ respectively).

By (7.48), it now follows that choosing the same coefficient for both rows $k-1, k$ does not produce sets in $\mathcal{A}$ (so far we only showed that it produces $o(2^k)$ sets in $\mathcal{A}$). Indeed, assigning the coefficient 0 to both these rows can only yield the all-zero vector, forbidden in $\mathcal{A}$ (for instance, as $\ell > 0$). Assigning the coefficient 1 to rows $k-1, k$ can only yield a vector which is 1 in every coordinate $j \in [2k-2]$, and is the sum of the two rows $k-1, k$ in columns $2k-1, 2k$ if these columns exist. Hence, if this vector belongs to $\{0,1\}^n$, then it contains any set which can be produced from $M_{\mathcal{A}}$, and we have $|\mathcal{A}| = 1$, and a suboptimal pair $\mathcal{A}, \mathcal{B}$. It follows that:

$$|\mathcal{A}| \leq 2^{k-1}\ .$$

Our next goal is to show that if row $q \in \{k-1, k\}$ of $M_{\mathcal{B}}$ exists, then its entries in columns $k-1, k$ (marked by $0/-1$ in (7.48)) are both -1. Let

$q \in \{k-1,k\}$ denote a row of $M_{\mathcal{B}}$, let $m \in \{1,2\}$ denote the number of rows of $\{k-1,k\}$ in $M_{\mathcal{B}}$, and let $q' \neq q$ denote the additional row of $\{k-1,k\}$ in $M_{\mathcal{B}}$ if $m=2$. Since $m=1$ iff $k=\frac{n+1}{2}$ and $m=2$ iff $k=\frac{n}{2}$, it follows that $m = 2-(k-h)$.

First, assume that $(\mathbf{M}_{\mathcal{A}})_{\mathbf{q},\mathbf{k-1}} = (\mathbf{M}_{\mathcal{A}})_{\mathbf{q},\mathbf{k}} = \mathbf{0}$. It follows that row q is in $\{0,1\}^n$, and since $\mathcal{B}$ is an antichain, we get an additional factor of $\frac{1}{2}$ on $|\mathcal{B}|$ (first apply the Littlewood-Offord Lemma on the remaining rows with respect to column k, then consider the coefficient for row q). It follows that $|\mathcal{B}| \leq (1+o(1))\frac{2^h}{\sqrt{\pi\ell}}$, and that $|\mathcal{A}||\mathcal{B}| \leq (\frac{1}{2}+o(1))2^n/\sqrt{\pi\ell}$.

Second, assume that $(\mathbf{M}_{\mathcal{A}})_{\mathbf{q},\mathbf{k-1}} \neq (\mathbf{M}_{\mathcal{A}})_{\mathbf{q},\mathbf{k}}$. Let t_1 denote the number of sets $B \in \mathcal{B}$ produced from $M_{\mathcal{B}}$ by assigning the coefficient $\alpha \neq 0$ to row q, and the coefficient 0 to row q' (if this row exists), and let $t_2 = |\mathcal{B}| - t_1$. Consider a set B counted by t_2: since row q' does not take part in the affine combinations, the combination of rows $[k-2]$ together with χ_{B_1} sums up to the same value, some λ, in the two columns $k-1,k$ (these two columns are identical in rows $[k-2]$). The fact that indices $k-1,k$ of the resulting vector, χ_B, are $\{\lambda, \lambda-\alpha\}$, forces λ to be equal to α. We can thus apply the Littlewood-Offord Lemma on rows $[k-2]$ (with respect to column k, which has 1 target value), and deduce that:

$$t_1 \leq (1+o(1))\frac{2^{k-2}}{\sqrt{\pi\ell}} \ .$$

To obtain an upper bound on t_2, for each of the remaining $2^m - 1$ combinations of rows $\{k-1,k\}$ in $M_{\mathcal{B}}$, column k has at most 2 target values (in order to give a $\{0,1\}$ final value), hence, by the Littlewood-Offord Lemma:

$$t_2 \leq (2^m-1)(2+o(1))\frac{2^{k-2}}{\sqrt{\pi\ell}} \ .$$

It follows that:

$$|\mathcal{B}| = t_1 + t_2 \leq (2 - 2^{-m} + o(1))\frac{2^{m+k-2}}{\sqrt{\pi\ell}} = (2-2^{-m}+o(1))\frac{2^h}{\sqrt{\pi\ell}} \ ,$$

where in the last equality we used the fact that $m = 2-(k-h)$. The fact that $|\mathcal{A}| \leq 2^{k-1}$ now implies that the pair $\mathcal{A},\mathcal{B}$ is suboptimal.

Having ruled out the cases $(M_{\mathcal{A}})_{q,k-1} = (M_{\mathcal{A}})_{q,k} = 0$ and $(M_{\mathcal{A}})_{q,k-1} \neq (M_{\mathcal{A}})_{q,k}$, we deduce that:

$$(M_{\mathcal{A}})_{q,k-1} = (M_{\mathcal{A}})_{q,k} = -1 \ ,$$

hence the structure of $M_{\mathcal{A}}, M_{\mathcal{B}}$ is:

$$M_{\mathcal{A}} = \left(\begin{array}{c|c|c|c||c|c} \multicolumn{4}{c||}{\ddots} & \vdots & \vdots \\ \multicolumn{4}{c||}{} & 0 & 0 \\ \hline 0 & 1 & 0 & 1\ldots1 & 1 & 1 \\ \hline 0 & 0 & 1 & 1\ldots1 & 1 & 1 \end{array}\right), M_{\mathcal{B}} = \left(\begin{array}{c|c|c|c||c|c} \multicolumn{4}{c||}{\ddots} & \vdots & \vdots \\ \multicolumn{4}{c||}{} & 0 & 0 \\ \hline\hline 0 & -1 & -1 & 0 & 1 & 0 \\ \hline 0 & -1 & -1 & 0 & 0 & 1 \end{array}\right),$$

as specified in (7.25). To conclude the proof of the claim, recall that every $A \in \mathcal{A}$ has precisely one of the elements $k-1, k$, hence the analysis of $|A \cap B|$ for all $B \in \mathcal{B}$ is exactly the same as in Claim 7.5.2 (precisely one of the columns $k-1, k$ of $M_{\mathcal{B}}$ effects the intersection). It follows that $|\mathcal{A}||\mathcal{B}| \leq \left(\binom{h}{\ell} + \binom{h}{\ell-1}\right) 2^{n-h-1} = \binom{h+1}{\ell} 2^{n-h-1}$, and hence $h \in \{2\ell-2, 2\ell-1\}$, otherwise $\mathcal{A}, \mathcal{B}$ is a suboptimal pair. Similarly, the arguments of Claim 7.5.2 imply that $|Q| = \ell$, where Q is the set of indices $\{i \in [h] : k+i \in B_1\}$, and without loss of generality, we can take B_1 to be $\cup_{i\in[\ell]}\{i, k+i\}$. Altogether, (7.25) holds.

It remains to link the above to the optimal family (7.22). Define:

$$X = \left\{ \begin{array}{ll} \{n-1, n\} & \text{if } k = \frac{n}{2} \\ \{n\} & \text{if } k = \frac{n+1}{2} \\ \emptyset & \text{if } k = \frac{n}{2}+1 \end{array} \right. .$$

Recall that precisely one of the rows $k-1, k$ receives the coefficient 1 in a linear combination which produces some $A \in \mathcal{A}$ from $M_{\mathcal{A}}$. It follows that each set $A \in \mathcal{A}$ is obtained by choosing one out of each pair of elements $\big\{\{i, k+i\} : i \in [k-2]\big\} \cup \big\{\{k-1, k\}\big\}$, then adding these $k-1$ chosen elements to the elements of X. Define:

$$Y = \left\{ \begin{array}{ll} \big\{\{i, k+i\} : i \in [k-2]\big\} \cup \big\{\{n-1\}, \{n\}\big\} & \text{if } k = \frac{n}{2} \\ \big\{\{i, k+i\} : i \in [k-2]\big\} \cup \big\{\{n\}\big\} & \text{if } k = \frac{n+1}{2} \\ \big\{\{i, k+i\} : i \in [k-2]\big\} & \text{if } k = \frac{n}{2}+1 \end{array} \right. .$$

Recall that, for all $B \in \mathcal{B}$, the elements $k-1, k$ are either both in B or both not in B. If $k-1, k \notin B$, then B is the union of ℓ elements of Y.

Otherwise, B contains, in addition to $\{k-1,k\}$, the union of $\ell-1$ elements of Y. Altogether, all sets $B\in\mathcal{B}$ are the union of ℓ objects of $Y\cup\big\{\{k-1,k\}\big\}$. As the last set holds the $k-2$ pairs $\{i,k+i\}$ for $i\in[k-2]$, the pair $\{k-1,k\}$ and the single elements corresponding to X, this fits the description of (7.22) for $\kappa=h+1$, $\tau=k-1$ and swapping $\mathcal{A},\mathcal{B}$.

This completes the proof of Claim 7.5.3 and of Lemma 7.4.1. ■

7.6 Proof of Lemma 7.4.2

The assumption that $r_A+s_A=\Omega(n)$ implies that $|\mathcal{A}|=O(2^k/\sqrt{n})$. Thus, if $r_B+s_B=\omega(1)$ we deduce that $|\mathcal{A}||\mathcal{B}|=o(2^n/\sqrt{n})$ and we are done. Assume therefore that $r_B+s_B=O(1)$, and let $C_2=[h]\setminus(R_B\cup S_B)$. By definition of R_B and S_B, the following holds:

- Every column of $M_\mathcal{B}$ contains at most 1 non-zero value in the rows of C_2.
- Every row of C_2 belongs to $\{0,\pm1\}^n\setminus\{0,1\}^n$.

We wish to show that $M_\mathcal{B}$ is roughly of the form $(-I_h\mid 0\mid I_h)$, although so far we did not obtain any restriction on the number of rows in C_2 with more than 2 non-zero entries in $M_\mathcal{B}$. In contrast to the analysis of $M_\mathcal{A}$ in Lemma 7.4.1, this does not follow directly from the fact that $r_B+s_B=O(1)$, as h might be substantially smaller than $n/2$ (as opposed to k).

We therefore return to $M_\mathcal{A}$ and claim that at most $O(1)$ columns of $M_\mathcal{A}$ contain at least 2 non-zero entries in a **cascading** manner. In other words, the process where we repeatedly select an arbitrary column of $M_\mathcal{A}$ with at least two non-zero entries, and remove the rows where it is non-zero from the matrix, ends after at most $O(1)$ steps. To see this, assume that $\omega(1)$ such columns exist: $j_1,\ldots,j_m$. Perform the process of creating R_A, beginning with the above columns: choose column j_i at step i for $i\le m$, and complete the process in an arbitrary order of column selection, $j_{m+1},\ldots,j_t$. By the assumption of the lemma, $r_A+s_A=\Omega(n)$, hence two cases are possible:

- $r_A=o(n)$: in this case $s_A=\Omega(n)$. Clearly, $r_A\ge 2m=\omega(1)$ by the assumption, and the additional $O(1/\sqrt{n})$ factor resulting from the rows S_A implies that $|\mathcal{A}|=o(2^k/\sqrt{n})$.

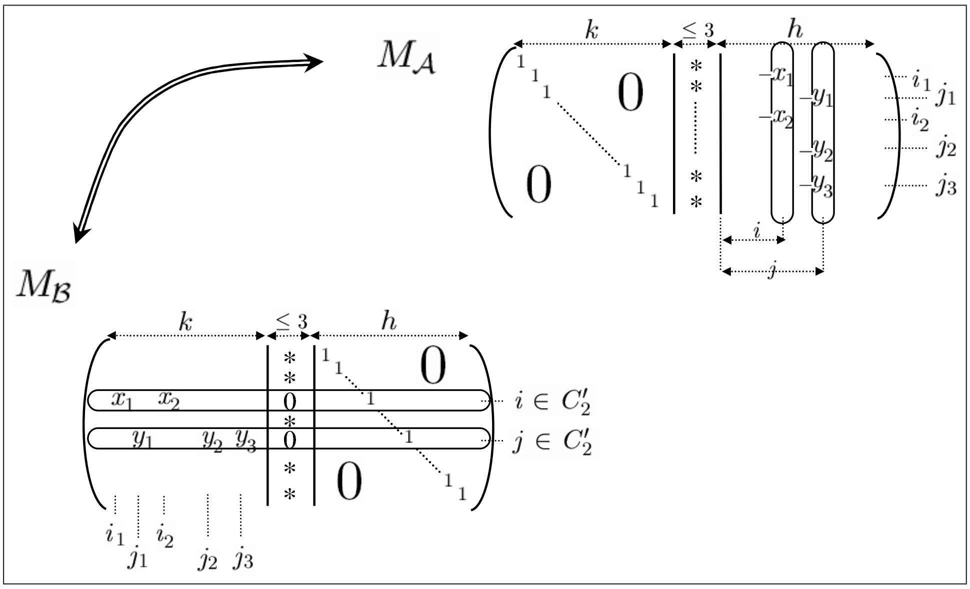

Figure 7.2: The duality between $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ when selected rows of $M_{\mathcal{B}}$ have 0 entries in columns $\{k+1,\ldots,n-h\}$.

- $r_A = \Omega(n)$: by definition, $r_A = \sum_{i=1}^{t} r_i$. If for some $i, j \leq t$ we have $r_i, r_j = \omega(\sqrt{n})$ then $|\mathcal{A}| = o(2^k/\sqrt{n})$. Recall that if $t \geq 2\log n$, then $|\mathcal{A}| \leq 2^k/(\frac{3}{4})^t \leq 2^k/n$. These two facts imply that precisely one i satisfies $r_i = \Omega(n)$. Therefore, column i gives a factor of $O(1/\sqrt{n})$, and the remaining $t-1$ columns give a factor of $o(1)$ as $t \geq m = \omega(1)$ and each such column contributes a factor of at most $\frac{3}{4}$. Altogether, we deduce that $|\mathcal{A}| = o(2^k/\sqrt{n})$.

Assume therefore that $M_{\mathcal{A}}$ contains at most $O(1)$ columns which contain at least 2 non-zero entries in a cascading manner. As we next show, returning to $M_{\mathcal{B}}$, this implies that at most $O(1)$ rows of C_2 contain more than 2 non-zero entries. First, recall that $k + h \geq n - 3$ and that each column contains at most one non-zero value in the rows of C_2. Thus, we can remove at most 3 rows from C_2 and obtain a set C_2', each remaining row of which does not contain non-zero entries in indices $k+1,\ldots,n-h$. Second, suppose rows $i, j \in C_2'$ each contains more than 2 non-zero entries. Let $i_1,\ldots,i_r \in [k]$, $r \geq 2$, denote the indices of the non-zero entries of row i excluding its value of 1 at index $n-h+i$ (recall that columns $n-h+1,\dots,n$ of $M_{\mathcal{B}}$ form the identity matrix of order h). Similarly, let $j_1,\dots,j_m \in [k]$, $m \geq 2$, denote the corresponding indices of row j :

$$(M_{\mathcal{B}})_{i,i_t} \neq 0 \text{ for } 1 \leq t \leq r \ , \ (M_{\mathcal{B}})_{i,n-h+i} = 1 \ ,$$

$$(M_{\mathcal{B}})_{j,j_t} \neq 0 \text{ for } 1 \leq t \leq m \ , \ (M_{\mathcal{B}})_{j,n-h+j} = 1 \ .$$

Since the rows of $M_{\mathcal{A}}$ are orthogonal to the rows of $M_{\mathcal{B}}$, and columns $1,\dots,k$ of $M_{\mathcal{A}}$ form the identity matrix of order k, we deduce that:

$$(M_{\mathcal{A}})_{n-h+i_t,i} \neq 0 \text{ for } 1 \leq t \leq r \ ,$$

$$(M_{\mathcal{A}})_{n-h+j_t,j} \neq 0 \text{ for } 1 \leq t \leq m \ .$$

See Figure 7.2 for an illustration of the above relation between $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$. As the sets $\{i_1,\dots,i_r\}$ and $\{j_1,\dots,j_m\}$ are disjoint, columns $n-h+i$ and $n-h+j$ of $M_{\mathcal{A}}$ each contains at least 2 non-zero entries in pairwise distinct indices. In general, if m rows in C'_2 contain more than 2 non-zero entries, we deduce that m columns in $M_{\mathcal{A}}$ contain at least 2 non-zero entries in a cascading manner. As argued above, there are at most $O(1)$ such columns in $M_{\mathcal{A}}$, hence $m = O(1)$: let C''_2 denote the set C'_2 after removing these m rows, and let $h' = |C''_2| = h - O(1)$. Each row of C''_2 is in $\{0,\pm 1\}^n \setminus \{0,1\}^n$ and contains at most 2 non-zero values, and we deduce that without loss of generality:

$$M_{\mathcal{B}} = \begin{array}{c} \begin{array}{ccccc} \multicolumn{2}{c}{\leftarrow\!\!-\,-\cdots\; k \;\cdots-\,-\!\!\rightarrow} & \multicolumn{1}{||c||}{\leftarrow\!\!-\,-\leq 3\,-\,-\!\!\rightarrow} & \multicolumn{2}{c}{\leftarrow\!\!-\,-\cdots\; h \;\cdots-\,-\!\!\rightarrow} \end{array} \\ \left(\begin{array}{c|c|c|c|c} -I_{h'} & 0 & 0 & I_{h'} & 0 \\ * & * & * & 0 & I_{h-h'} \end{array}\right) \end{array} . \tag{7.49}$$

Since the rows of $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ are orthogonal, it follows that:

$$M_{\mathcal{A}} = \begin{array}{c} \begin{array}{ccccc} \multicolumn{2}{c}{\leftarrow\!\!-\,-\cdots\; k \;\cdots-\,-\!\!\rightarrow} & \multicolumn{1}{||c||}{\leftarrow\!\!-\,-\leq 3\,-\,-\!\!\rightarrow} & \multicolumn{2}{c}{\leftarrow\!\!-\,-\cdots\; h \;\cdots-\,-\!\!\rightarrow} \end{array} \\ \left(\begin{array}{c|c|c|c|c} I_{h'} & 0 & * & I_{h'} & * \\ 0 & I_{k-h'} & * & 0 & * \end{array}\right) \end{array} . \tag{7.50}$$

The above structure of $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ provides an upper bound on ℓ in terms of k, which we prove in Subsection 7.6.1:

Claim 7.6.1. *Let $\mathcal{A}$ and $\mathcal{B}$ be as above. If $|\mathcal{A}||\mathcal{B}| = \Omega(2^n/\sqrt{n})$, then:*

$$\ell \leq \left(\frac{1}{2} + o(1)\right) k \ . \tag{7.51}$$

The proof of the lemma is completed by the next two claims, which are proved in Subsections 7.6.2 and 7.6.3:

Claim 7.6.2. *Let $\mathcal{A}$ and $\mathcal{B}$ be as above. If $r_A = o(n)$ then $|\mathcal{A}||\mathcal{B}| \leq \binom{2\ell}{\ell}2^{n-2\ell}$. Furthermore, equality holds iff (7.26) holds.*

Claim 7.6.3. *Let $\mathcal{A}$ and $\mathcal{B}$ be as above. If $r_A = \Omega(n)$ then the pair $\mathcal{A}, \mathcal{B}$ is suboptimal.*

7.6.1 Proof of Claim 7.6.1

Fix a choice of coefficients for the rows $h'+1, \ldots, h$ of $M_{\mathcal{B}}$, and let w_B denote the result of adding this combination to χ_{B_1}. As argued in the proof of Claim 7.5.1, the structure of $M_{\mathcal{B}}$ in (7.49) implies that each index $j \in [h']$ such that

$$w_B^{(j)} \neq 1 - w_B^{(n-h+j)} \tag{7.52}$$

eliminates at least one of the two possible coefficients for the row j of $M_{\mathcal{B}}$ (compare this to the treatment of the vector w_A in (7.35)). Thus, if there are $\omega(1)$ such coefficients, then w_B allows at most $o(2^{h'})$ combinations of the remaining rows of $M_{\mathcal{B}}$ to produce sets in $\mathcal{B}$. Since $|\mathcal{A}| = O(2^k/\sqrt{n})$ (recall that $r_A + s_A = \Omega(n)$), summing over at most $2^{h-h'}$ combinations for such vectors w_B gives $o(2^{k+h}/\sqrt{n})$ pairs $(A, B) \in \mathcal{A} \times \mathcal{B}$.

It remains to treat vectors w_B in which at most $O(1)$ indices $j \in [h']$ satisfy (7.52). Note that each $B \in \mathcal{B}$ produced from w_B and a combination of rows $1, \ldots, h'$ of $M_{\mathcal{B}}$ satisfies:

$$|B \cap \{j, n-h+j\}| = 1 \text{ for all but at most } O(1) \text{ indices } j \in [h'] \ . \tag{7.53}$$

Let $A \in \mathcal{A}$, and let $X_i \in \{0,1\}$ denotes the coefficient of the row i of $M_{\mathcal{A}}$ in the linear combination which produces A. By (7.53) and the structure of $M_{\mathcal{A}}$ in (7.50), we obtain that:

$$\big|A \cap B \cap \big([h'] \cup \{n-h+1, \ldots, n-h+h'\}\big)\big| = (\sum_{i=1}^{h'} X_i) + O(1) \ . \tag{7.54}$$

Furthermore, the structure of $M_{\mathcal{A}}$ in (7.50) gives:

$$|A \cap B \cap \{h'+1,\ldots,k\}| \leq |A \cap \{h'+1,\ldots,k\}| = \sum_{i=h'+1}^{k} X_i \ . \tag{7.55}$$

Combining (7.54) and (7.55) with the fact that $k + h' = n - O(1)$, we obtain that:

$$\ell = |A \cap B| \leq \left(\sum_{i=1}^{k} X_i \right) + O(1) \ .$$

Let $\varepsilon > 0$, and assume that $\ell > (1+\varepsilon)\frac{k}{2}$. By the Chernoff bound, the number of assignments of $\{0,1\}$ to the variables $X_1,\ldots,X_k$, which satisfy $\sum_{i=1}^{k} X_i > (1+\varepsilon)\frac{k}{2}$, is at most $2^k/\exp(\Omega(k)) = 2^k/\exp(\Omega(n))$. Therefore, the assumption on ℓ implies that at most $O(2^k/\exp(\Omega(n))$ sets $A \in \mathcal{A}$ satisfy $|A \cap B| = \ell$, and summing over all sets B whose vector w_B is as above gives at most $2^{k+h}/\exp(\Omega(n))$ pairs $(A,B) \in \mathcal{A} \times \mathcal{B}$. This contradicts the assumption that $|\mathcal{A}||\mathcal{B}| = \Omega(2^n/\sqrt{n})$, and we conclude that $\ell \leq (\frac{1}{2} + o(1))k$, as required. ■

7.6.2 Proof of Claim 7.6.2

The assumptions $r_A + s_A = \Omega(n)$ and $r_A = o(n)$ imply that $s_A = \Omega(n)$, and, as before, we may assume that $r_A = O(1)$, otherwise we get $|\mathcal{A}| = o(2^k/\sqrt{n})$, leading to a suboptimal pair $\mathcal{A},\mathcal{B}$. Thus, each column of $M_{\mathcal{A}}$ has at most $O(1)$ non-zero entries. Since $n - (k+h) \leq 3$ and $h - h' = O(1)$, it follows that at most $O(1)$ rows of $M_{\mathcal{A}}$ have non-zero entries in columns $\{k+1,\ldots,n-h\} \cup \{n-h+h'+1,\ldots,n\}$. Without loss of generality, reorder the indices of these rows to $k'+1,\ldots,k$ (where $k' = k - O(1)$), and let $h'' = h' - O(1)$ reflect the reordering of rows whose original indices belonged to $[h']$. We obtain that:

$$M_{\mathcal{A}} = \begin{array}{c} \begin{array}{cccccc} \multicolumn{3}{c}{\longleftarrow \cdots\cdots\cdots\ k\ \cdots\cdots\cdots \longrightarrow} & || \longleftarrow \leq 3 \longrightarrow || & \multicolumn{2}{c}{\longleftarrow \cdots\ h\ \cdots \longrightarrow} \end{array} \\ \left(\begin{array}{c|c|c|c|c|c} I_{h''} & 0 & 0 & 0 & I_{h''} & 0 \\ 0 & I_{k'-h''} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{k-k'} & * & 0 & * \end{array} \right) \end{array} , \tag{7.56}$$

and by the orthogonality of $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$:

$$M_{\mathcal{B}} = \begin{array}{c} \begin{array}{ccc} \leftarrow\!-\,\cdots\ k'\ \cdots-\!\rightarrow & ||\leftarrow\!-\ O(1)\ -\!\rightarrow|| & \leftarrow\!-\cdots\quad h\quad\cdots-\!\rightarrow \end{array} \\ \left(\begin{array}{c|c|c|c|c} -I_{h''} & 0 & 0 & I_{h''} & 0 \\ 0 & 0 & * & 0 & I_{h-h''} \end{array} \right) \end{array} . \tag{7.57}$$

Notice that the first k' rows of $M_{\mathcal{A}}$ form an antichain on the first k' elements, hence:

$$|\mathcal{A}| \leq (1+o(1))\frac{2^k}{\sqrt{\pi k'/2}} \leq (1+o(1))\frac{2^k}{\sqrt{\pi \ell}} \ ,$$

where the last inequality is by (7.51). This yields an upper bound on $|\mathcal{A}||\mathcal{B}|$ which is asymptotically tight, hence any additional constant factor bounded away from 1 which multiplies either $|\mathcal{A}|$ or $|\mathcal{B}|$ implies that the pair $(\mathcal{A},\mathcal{B})$ is suboptimal. In particular:

(i) If $k+h<n$, we have a suboptimal pair: $|\mathcal{A}||\mathcal{B}| \leq \left(\frac{1}{2}+o(1)\right) 2^n/\sqrt{\pi\ell}$. Assume therefore that $k+h=n$.

(ii) If $M_{\mathcal{B}}$ has a column with more than 1 non-zero entry, we gain a multiplicative factor of at most $\frac{3}{4}$ and we are done. The same applies to $M_{\mathcal{A}}$: such a column has $O(1)$ non-zero entries, as $r_A = O(1)$, and once we set the combination of these rows (gaining a factor of at most $\frac{3}{4}$) as well as of rows $k'+1,\ldots,k$, the remaining $k'-O(1)$ rows out of $[k']$ must still form an antichain.

(iii) If $M_{\mathcal{A}}$ has a row with more than 2 non-zero entries, by Item (i) it corresponds to a column with more than 1 non-zero entry in $M_{\mathcal{B}}$ (since statement (7.41) holds), which does not exist according to Item (ii). The same applies to the rows of $M_{\mathcal{B}}$.

(iv) Each row of $M_{\mathcal{B}}$ must belong to $\{0,\pm 1\}^n \setminus \{0,1\}^n$, otherwise the arguments of Claim 7.3.4 imply a constant multiplicative factor of at most $\frac{1}{2}$.

Items (iii) and (iv) imply that every row of $M_{\mathcal{B}}$ has precisely two non-zero entries: $\{1,-1\}$, and without loss of generality, $h''=h$. Recalling (7.56) and

(7.57), $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$ take the following form:

$$\begin{aligned} M_{\mathcal{A}} &= \left(\begin{array}{c|c|c} I_h & 0 & I_h \\ 0 & I_{k-h} & 0 \end{array} \right) , \\ M_{\mathcal{B}} &= \left(\begin{array}{c|c|c} -\ I_h & 0 & I_h \end{array} \right) . \end{aligned} \tag{7.58}$$

Notice that the above structure of $M_{\mathcal{B}}$ implies that $\chi_B^{(j)} = \chi_{B_1}^{(j)}$ for all $j \in \{h+1, \ldots, k\}$ and $B \in \mathcal{B}$. As we assumed in (7.20) that $\bigcup_{B \in \mathcal{B}} \mathcal{B} = [n]$, it follows that $\{h+1, \ldots, k\} \in B_1$.

Consider the rows of $M_{\mathcal{B}}$, let w_B take the initial value of the vector χ_{B_1}, then subtract from w_B each row i of $M_{\mathcal{B}}$ for which $k+i \in B_1$. This translates the possible coefficients for each row i of $M_{\mathcal{B}}$ to $\{0,1\}$; hence, the characteristic vector of every element of $\mathcal{B}$ is a sum of w_B with a sub-sum of the rows of $M_{\mathcal{B}}$. First, $w_B^{(j)} = \chi_{B_1}^{(j)} = 1$ for all $j \in \{h+1, \ldots, k\}$. Second, the structure of $M_{\mathcal{B}}$ (7.58) implies that, if $w_B^{(j)} \neq 1$ for some $j \in [h]$, then row j cannot be added to w_B to yield a vector in $\{0,1\}^n$. Since this leads to a suboptimal pair $(\mathcal{A}, \mathcal{B})$ (of size at most $(\frac{1}{2} + o(1))2^n/\sqrt{\pi \ell}$), we deduce that:

$$w_B = \big(\overbrace{1 \ldots 1}^{k} \ \overbrace{0 \ldots 0}^{h} \big) \ .$$

The structure of $M_{\mathcal{B}}$ (7.58) implies that for every $B \in \mathcal{B}$, χ_B is of the form:

$$\chi_B = \big(\overbrace{0/1 \ldots 0/1}^{h} \ \overbrace{1 \ldots 1}^{k-h} \ \overbrace{1/0 \ldots 1/0}^{h} \big) \ ,$$

where precisely one index in each of the pairs $\{(1, k+1), \ldots, (h, k+h)\}$ is equal to 1 in χ_B. If $X_i \in \{0,1\}$ denotes the coefficient of row i of $M_{\mathcal{A}}$ in a combination that produces some $A \in M_{\mathcal{A}}$, it follows from (7.58) that $\ell = |A \cap B| = \sum_{i=1}^{k} X_i$ for all $B \in \mathcal{B}$. By the properties of the binomial distribution, we deduce that $|\mathcal{A}| \leq \binom{k}{\ell}$, and altogether:

$$|\mathcal{A}||\mathcal{B}| \leq 2^{n-k} \binom{k}{\ell} \ .$$

The expression above realizes the bound (7.3) iff either $k = 2\ell$ or $k = 2\ell - 1$, hence the final structure of the optimal pair $(\mathcal{A}, \mathcal{B})$ is as described in Lemma 7.4.2. ■

7.6.3 Proof of Claim 7.6.3

The assumption $r_A = \Omega(n)$ implies that, unless $s_A = O(1)$, we get $|\mathcal{A}| = o(2^k/\sqrt{k}) = o(2^k/\sqrt{n})$ as required. However, if we remove the rows R_A from $[k]$, (7.50) implies that only the columns $\{k+1, \ldots, n-h\} \cup \{n-h+h', \ldots, n\}$ can contribute -1 entries to the remaining rows, and each column has at most 1 non-zero entry in each of these rows. Since $n-(k+h) \leq 3$ and $h-h' = O(1)$, we deduce that $[k] - r_A - s_A = O(1)$, and altogether:

$$r_A = k - O(1) \ .$$

Definition 7.1. *A column of $M_{\mathcal{A}}$ is called "heavy" if it contains $k - O(1)$ non-zero entries.*

The next argument shows that there exists a heavy column in $M_{\mathcal{A}}$. There are at most $O(1)$ columns which may contain more than 1 non-zero entry in $M_{\mathcal{A}}$ (as columns $[k]$ and $\{n-h+1, \ldots, n-h+h'\}$ contain a single non-zero entry of 1). Therefore, there exists some column $q \in [n]$ of $M_{\mathcal{A}}$ with $\Omega(r_A) = \Omega(k)$ non zero entries. If some other column has $\omega(1)$ non-zero entries in a cascading manner, we obtain $|\mathcal{A}| = o(2^k/\sqrt{n})$, and we are done. We deduce the column q has $r_A - O(1) = k - O(1)$ non-zero entries, therefore column q is heavy. Applying the Littlewood-Offord Lemma to the $k - O(1)$ rows where column q is non-zero at, we obtain that:

$$|\mathcal{A}| \leq (2+o(1))\frac{2^k}{\sqrt{\pi k/2}} \leq (2+o(1))\frac{2^k}{\sqrt{\pi \ell}} \ , \tag{7.59}$$

where the last inequality is by (7.51).

Let q denote a heavy column of $M_{\mathcal{A}}$. Lemma 7.2.3 enables us to eliminate the case where all non-zero entries of q are ± 1. To see this, assume the converse, and let:

$$U = \{i \in [k] : (M_{\mathcal{A}})_{i,q} = 1\} \ , \ V = \{i \in [k] : (M_{\mathcal{A}})_{i,q} = -1\} \ .$$

Recall that $|U| + |V| = k - O(1)$, and take $\varepsilon > 0$. If $|U| \geq (\frac{1}{2} + \varepsilon)k$, then Chernoff's bound implies that the number of sub-sums of the rows $U \cup V$ which give a value of $\{0, 1\}$ in this column is at most $2^k/\exp(\Omega(k))$. We deduce $|U| = (\frac{1}{2} + o(1))k$ and that $|V| = (\frac{1}{2} + o(1))k$.

Set $m = n - (k+h) + (h - h') = O(1)$. For each possible set of values $\underline{x} \in \{0,1\}^m$ for columns $\{k+1,\ldots,n-h\} \cup \{n-h+h',\ldots,n\}$, the family of all sets $A \in \mathcal{A}$ which matches the pattern $\underline{x}$ in the above set of columns is an antichain, and either $|A \cap V| = |A \cap U|$ or $|A \cap V| = |A \cap U| - 1$. Therefore, Lemma 7.2.3 implies that $|\mathcal{A}| = O(2^k/k) = O(2^k/n)$. We may therefore assume that:

$$\text{Every heavy column } q \text{ of } M_{\mathcal{A}} \text{ satisfies } (M_{\mathcal{A}})_{i,q} \notin \{0,\pm 1\} \text{ for some } i \in [k] \ . \tag{7.60}$$

This provides an upper bound on $|\mathcal{B}|$:

$$|\mathcal{B}| \leq 2^{n-k-1} \ . \tag{7.61}$$

The above bound follows immediately if $h < n-k$, so consider the case $k+h=n$, and let q denote a heavy column of $M_{\mathcal{A}}$. By the orthogonality of $M_{\mathcal{A}}, M_{\mathcal{B}}$, (7.41) holds, and (7.60) now implies that $(M_{\mathcal{B}})_{q-k,i} \notin \{0,\pm1\}$ for some $i \in [k]$. In particular, row $q-k$ of $M_{\mathcal{B}}$ does not belong to $\{0,\pm1\}^n$, and hence $|\mathcal{B}| \leq 2^{h-1}$ (as enumerating on the coefficients for rows $[h] \setminus \{q-k\}$ of $M_{\mathcal{B}}$ leaves at most one legal coefficient for row $q-k$).

Combining (7.61) with (7.59) yields an asymptotically tight upper bound on $|\mathcal{A}||\mathcal{B}|$:

$$|\mathcal{A}||\mathcal{B}| \leq (1+o(1))\frac{2^n}{\sqrt{\pi k/2}} \leq (1+o(1))\frac{2^n}{\sqrt{\pi \ell}} \ .$$

Let $\varepsilon > 0$; if $k \geq (2+\varepsilon)\ell$, then the first inequality of the bound above implies that the pair $\mathcal{A}, \mathcal{B}$ is suboptimal. Therefore, adding this to (7.51), we may assume that:

$$k = (2+o(1))\ell \ . \tag{7.62}$$

Next, we wish to eliminate the case where some column q has $k - O(1)$ non-zero entries, all of which have the same sign. In this case, let $Q = \{i : (M_{\mathcal{A}})_{i,q} \neq 0\}$. As all the entries in rows Q and column q of $M_{\mathcal{A}}$ have the same sign, only the all-zero linear combination of these rows can produce the value 0 at index q. Applying the Littlewood-Offord Lemma to the rows Q, we obtain an upper bound on the number of combinations which produce the value 1, and altogether:

$$|\mathcal{A}| \leq 2^{k-|Q|}(\binom{|Q|}{\lfloor |Q|/2 \rfloor} + 1) = (1+o(1))\frac{2^k}{\sqrt{\pi\ell}} \ ,$$

where in the last inequality we used the fact that $|Q| \geq (2+o(1))\ell$, as $|Q| = k - O(1)$. By (7.61), this implies that $|\mathcal{A}||\mathcal{B}| \leq (\frac{1}{2}+o(1))2^n/\sqrt{\pi\ell}$, implying the statement of the claim. We thus assume that:

$$\text{Every heavy column } q \text{ of } M_{\mathcal{A}} \text{ contains both positive and negative entries .} \tag{7.63}$$

Using the last statement, we prove the next claim:

Claim 7.6.4. *Let $\lambda \in \{0,1\}$, $L \subset [k]$ and $d > 0$, and let q denote a heavy column of $M_{\mathcal{A}}$. Define:*

$$\mathcal{A}^{(q)}_{L,d,\lambda} = \{A \in \mathcal{A} : |A \cap L| = d \;, \chi^{(q)}_A = \lambda\} \;. \tag{7.64}$$

If $d = (1+o(1))\ell$ and $|L| \geq (1+o(1))\ell$ then:

$$|\mathcal{A}^{(q)}_{L,d,\lambda}| \leq \left(\frac{3}{4}+o(1)\right)\frac{2^k}{\sqrt{\pi\ell}} \;. \tag{7.65}$$

Proof. Let Q denote the indices of the rows in which column q of $|\mathcal{A}|$ has a non-zero entry. Observe that if $Q \nsubseteq L$, then the rows of L have at most $\binom{|L|}{d}$ legal combinations, and the remaining rows $[k] \setminus L$ have at most $2^{k-|L|-1}$ legal combinations, as these rows contain non-zero entries in column q, which must combine to a final value of λ. Hence, in this case:

$$\begin{aligned}|\mathcal{A}^{(q)}_{L,d,\lambda}| &\leq \frac{1}{2}\cdot 2^{k-|L|}\binom{|L|}{d} \leq \frac{1}{2}2^{k-|L|}\binom{|L|}{\lfloor |L|/2 \rfloor} \\ &= \frac{1+o(1)}{2}\cdot\frac{2^k}{\sqrt{\pi|L|/2}} \leq \left(\frac{1}{\sqrt{2}}+o(1)\right)\frac{2^k}{\sqrt{\pi\ell}} \;,\end{aligned}$$

where the last inequality is by the fact that $|L| \geq (1+o(1))\ell$. Assume therefore that $Q \subset L$, and notice that, as $|Q| = k - O(1)$ and $L \subset [k]$, then $|L| = k - O(1)$, and by (7.62):

$$|L| = (2+o(1))\ell = (2+o(1))d \;.$$

Fix an enumeration on the coefficients of the rows $[k] \setminus L$, and let $\mathcal{S} \subset 2^L$ denote the d-element subsets of the rows of L which extend this enumeration to elements of $\mathcal{A}^{(q)}_{L,d,\lambda}$. Let $j_1, j_2 \in L$ be two indices such that $(M_{\mathcal{A}})_{j_1,q} \neq (M_{\mathcal{A}})_{j_2,q}$ (such indices exist by (7.63) and since $Q \subset L$), and define:

$$\mathcal{S}_0 = \{S \subset [L] : |S| = d\ ,\ |S \cap \{j_1, j_2\}| = 1\}\ .$$

Notice that, as $j_1 \neq j_2$, the function $f : \mathcal{S}_0 \to \mathcal{S}_0$ which swaps j_1, j_2 is a bijection, which satisfies the following property for all $S \in \mathcal{S}_0$: at most one of the subsets $\{S, f(S)\}$ can belong to $\mathcal{S}$. Furthermore, if S is a random d-element set of L, then:

$$\Pr[S \in \mathcal{S}_0] = \frac{2\binom{|L|-2}{d-1}}{\binom{|L|}{d}} = \frac{2d(|L|-d)}{|L|(|L|-1)} = \frac{1}{2} + o(1)\ ,$$

and thus $|\mathcal{S}_0| = (\frac{1}{2} + o(1))\binom{|L|}{d}$, and we deduce that:

$$|\mathcal{S}| \leq \binom{|L|}{d} - \frac{|\mathcal{S}_0|}{2} = \left(\frac{3}{4} + o(1)\right)\binom{|L|}{d}\ .$$

Therefore:

$$|\mathcal{A}^{(q)}_{L,d,\lambda}| \leq 2^{k-|L|}|\mathcal{S}| \leq \left(\frac{3}{4} + o(1)\right)\frac{2^k}{\sqrt{\pi|L|/2}} = \left(\frac{3}{4} + o(1)\right)\frac{2^k}{\sqrt{\pi\ell}}\ ,$$

as required. ■

In order to deduce the claim from (7.65), we treat the two cases $k+h < n$ and $k + h = n$ in Claims 7.6.5 and 7.6.6 below.

Claim 7.6.5. *Let $\mathcal{A}, \mathcal{B}$ be as above. If $k + h < n$, then the pair $\mathcal{A}, \mathcal{B}$ is suboptimal.*

Proof. In this case, we may assume that $k + h = n - 1$, otherwise (7.59) implies that $|\mathcal{A}||\mathcal{B}| \leq (\frac{1}{2} + o(1))2^n/\sqrt{\pi\ell}$. Recalling (7.49) and (7.50), we have:

$$\begin{array}{rc}
 & \begin{array}{ccccc} \multicolumn{2}{c}{\longleftarrow \cdots\cdots k \cdots\cdots \longrightarrow} & \|\longleftarrow 1 \longrightarrow\| & \multicolumn{2}{c}{\longleftarrow \cdots h \cdots \longrightarrow} \end{array} \\
M_{\mathcal{A}} = & \left(\begin{array}{c|c|c|c|c} I_{h'} & 0 & * & I_{h'} & * \\ 0 & I_{k-h'} & * & 0 & * \end{array}\right) \\
 & \begin{array}{ccccc} \multicolumn{2}{c}{\longleftarrow \cdots\cdots k \cdots\cdots \longrightarrow} & \|\longleftarrow 1 \longrightarrow\| & \multicolumn{2}{c}{\longleftarrow \cdots h \cdots \longrightarrow} \end{array} \\
M_{\mathcal{B}} = & \left(\begin{array}{c|c|c|c|c} -I_{h'} & 0 & 0 & I_{h'} & 0 \\ * & * & * & 0 & I_{h-h'} \end{array}\right)
\end{array}\ . \qquad (7.66)$$

Let $m = h - h' = O(1)$, and consider a choice of coefficients for rows $h' + 1, \ldots, h$ of $M_{\mathcal{B}}$, yielding (together with χ_{B_1}) a vector w_B. First, by (7.59), each of the $2^m - 1$ choices of coefficients such that $(w_B^{(n-m+1)} \ldots w_B^{(n)}) \neq 0$ can each be completed to a pair $(A, B) \in \mathcal{A} \times \mathcal{B}$, in at most

$$2^{h-m} \cdot (2 + o(1)) \frac{2^k}{\sqrt{\pi \ell}} = (1 + o(1)) \frac{2^{n-m}}{\sqrt{\pi \ell}}$$

ways. Let $\mathcal{B}_0$ denote the sets $B \in \mathcal{B}$ which can be produced from the remaining combination for w_B (the one for which $w_B^{(n-m+1)} = \ldots = w_B^{(n)} = 0$). In order to show that $\mathcal{A}, \mathcal{B}$ is suboptimal, it is enough to show that:

$$|\mathcal{A}||\mathcal{B}_0| \leq (\alpha + o(1)) \frac{2^{n-m}}{\sqrt{\pi \ell}} \text{ for some } \alpha < 1 \ , \tag{7.67}$$

since this would imply:

$$\begin{aligned} |\mathcal{A}||\mathcal{B}| &\leq (2^m - 1)(1 + o(1)) \frac{2^{n-m}}{\sqrt{\pi \ell}} + (\alpha + o(1)) \frac{2^{n-m}}{\sqrt{\pi \ell}} \\ &= \left(1 - \frac{1-\alpha}{2^m} + o(1)\right) \frac{2^n}{\sqrt{\pi \ell}} \ . \end{aligned} \tag{7.68}$$

If for some index $j \in [h']$ we have $w_B^{(n-h+j)} \neq 1 - w_B^{(j)}$, then row j of $M_{\mathcal{B}}$ has at most one legal coefficient, hence $|\mathcal{B}_0| \leq 2^{h-m-1}$, and the same holds in case $w_B \notin \{0,1\}^n$ (if $j \in \{h'+1, \ldots, n-h\}$ is such that $w_B^{(j)} \notin \{0,1\}$, then $\mathcal{B}_0 = \emptyset$). As $|\mathcal{A}| \leq (2 + o(1)) \frac{2^k}{\sqrt{\pi \ell}}$ and $k + h < n$, it follows that in the above two cases $|\mathcal{A}||\mathcal{B}_0| \leq \left(\frac{1}{2} + o(1)\right) \frac{2^{n-m}}{\sqrt{\pi \ell}}$, satisfying (7.67) for $\alpha = \frac{1}{2}$.

Assume therefore that $w_B^{(n-h+j)} = 1 - w_B^{(j)}$ for all $j \in [h']$, and that $w_B \in \{0,1\}^n$, and define:

$$L = [h'] \cup \{h' + 1 \leq i \leq k : w_B^{(i)} = 1\} \ .$$

Recalling that $w_B^{(n-h+h'+1)} = \ldots = w_B^{(n)} = 0$, (7.66) implies that every B produced from w_B satisfies:

$$\ell = |A \cap B| = \mathbf{1}_{\{k+1 \in A \cap B\}} + \sum_{i \in L} X_i \ , \tag{7.69}$$

for all $A \in \mathcal{A}$, where $X_i \in \{0,1\}$ denotes the coefficient for row i in a combination which produces A from $M_{\mathcal{A}}$. We may assume that $\mathcal{B}_0 \neq \emptyset$ (otherwise (7.67) immediately holds), and by (7.69) we obtain that $|L| \geq \ell - 1$, and in particular, $|L| \geq (1+o(1))\ell$.

If column $k+1$ of $M_{\mathcal{A}}$ has $o(k) = o(|L|)$ non-zero entries in some rows U, fix an enumeration on the coefficients of these rows, and let $L' = L \setminus U$, noting that $|L'| = (1-o(1))|L| \geq (1-o(1))\ell$. The enumeration on the coefficients for the rows U determines whether or not $k+1 \in A \cap B$, and by (7.69), this determines the value of $\sum_{i \in L'} X_i$. Therefore, by the properties of the binomial distribution, there are at most $\binom{|L'|}{\lfloor |L'|/2 \rfloor} \leq 2^{|L'|}/\sqrt{\pi|L'|/2}$ combinations for the coefficients of the rows L'. We conclude that:

- In case $|L| \geq (1-o(1))k$, recalling (7.62), we get $|\mathcal{A}| \leq (1+o(1))\frac{2^k}{\sqrt{\pi\ell}}$.
- Otherwise, $k - |L| = \Omega(k)$, and after choosing a combination for the rows L', we are left with rows $[k] \setminus (L \cup U)$ which contain $\Omega(k)$ non-zero entries in some heavy column q of $M_{\mathcal{A}}$ (recall that each heavy column has $k - O(1)$ non-zero entries). The Littlewood-Offord Lemma gives a factor of $O(1/\sqrt{k})$ on the number of combinations for the remaining rows, which, when multiplied by the previous factor of $O(1/\sqrt{|L'|}) = O(1/\sqrt{k})$ gives $|\mathcal{A}| \leq O(2^k/k) = O(2^k/\ell)$. In particular, we have $|\mathcal{A}| \leq (1+o(1))\frac{2^k}{\sqrt{\pi\ell}}$ (with room to spare).

Altogether, as $|\mathcal{B}_0| \leq 2^{h-m} \leq 2^{n-m-k-1}$, in both cases we obtain that (7.67) holds for $\alpha = \frac{1}{2}$.

It remains to treat the case where column $k+1$ of $M_{\mathcal{A}}$ has $\Omega(k)$ non-zero entries; by the arguments in the beginning of the proof of Claim 7.6.3, it follows that column $k+1$ is heavy. Therefore, recalling that $\mathcal{B}_0 \neq \emptyset$ and using the definition (7.64), it follows that:

$$|\mathcal{A}| = \begin{cases} |\mathcal{A}^{(k+1)}_{L,\ell,0}| + |\mathcal{A}^{(k+1)}_{L,\ell,1}| & \text{if } w_B^{(k+1)} = 0 \\ \\ |\mathcal{A}^{(k+1)}_{L,\ell,0}| + |\mathcal{A}^{(k+1)}_{L,\ell-1,1}| & \text{if } w_B^{(k+1)} = 1 \end{cases} .$$

Applying Claim 7.6.4 (recall that $|L| \geq \ell - 1$) gives:

$$|\mathcal{A}| \leq 2 \cdot \left(\frac{3}{4} + o(1)\right) \frac{2^k}{\sqrt{\pi\ell}} = \left(\frac{3}{2} + o(1)\right) \frac{2^k}{\sqrt{\pi\ell}} ,$$

and as $|\mathcal{B}_0| \leq 2^{h-m} \leq 2^{n-m-k-1}$, (7.67) holds for $\alpha = \frac{3}{4}$, as required. ■

Claim 7.6.6. *Let $\mathcal{A}, \mathcal{B}$ be as above. If $k + h = n$, then the pair $\mathcal{A}, \mathcal{B}$ is suboptimal.*

Proof. The proof will follow from arguments similar to those in the proof of Claim 7.6.5; the factor of $\frac{1}{2}$ which followed from the case $k + h < n$ is replaced by the duality between $M_{\mathcal{A}}, M_{\mathcal{B}}$ (7.41) when $k + h = n$. The assumption $k + h = n$ gives (7.49) and (7.50) the following form:

$$\begin{array}{rcccc}
 & \multicolumn{2}{c}{\longleftarrow - \cdots\cdots k \cdots\cdots - \longrightarrow} & \multicolumn{2}{c}{|| \longleftarrow - \cdots h \cdots - \longrightarrow} \\
M_{\mathcal{A}} = & \multicolumn{4}{c}{\left(\begin{array}{c|c|c|c} I_{h'} & 0 & I_{h'} & * \\ 0 & I_{k-h'} & 0 & * \end{array}\right)} \\
 & \multicolumn{2}{c}{\longleftarrow - \cdots\cdots k \cdots\cdots - \longrightarrow} & \multicolumn{2}{c}{|| \longleftarrow - \cdots\cdots h \cdots\cdots - \longrightarrow} \\
M_{\mathcal{B}} = & \multicolumn{4}{c}{\left(\begin{array}{c|c|c|c} -I_{h'} & 0 & I_{h'} & 0 \\ * & * & 0 & I_{h-h'} \end{array}\right)}
\end{array} .$$

Let $q \in [n]$ denote a heavy column of $M_{\mathcal{A}}$; by the above structure of $M_{\mathcal{A}}$, we can assume without loss of generality that $q = n$. Let $p \in [k]$ be such that $(M_{\mathcal{A}})_{p,n} \notin \{0, \pm 1\}$ (such a p exists by (7.60)). Recall that, as $k + h = n$, the orthogonality of $M_{\mathcal{A}}, M_{\mathcal{B}}$ implies that (7.41) holds, and thus $(M_{\mathcal{B}})_{h,p} = -(M_{\mathcal{B}})_{p,n} \notin \{0, \pm 1\}$.

Consider the following set of rows of $M_{\mathcal{B}}$:

$$W = \begin{cases} \{p\} \cup \{h' + 1, \ldots, h - 1\} & \text{if } p \in [h'] \ , \\ \{h' + 1, \ldots, h - 1\} & \text{otherwise} \ . \end{cases}$$

Let $m = |W|$, and consider one of the $2^m - 1$ choices of coefficients for the rows W of $M_{\mathcal{B}}$, such that the sum of χ_{B_1} and the resulting combination of these rows, satisfies $w_B^{(k+j)} \neq 0$ for some $j \in W$. Observe that w_B allows at most one coefficient for row h of $M_{\mathcal{B}}$, since all the remaining rows $[h-1] \setminus W$ have 0 entries at column p, whereas $(M_{\mathcal{B}})_{h,p} \notin \{0, \pm 1\}$. Therefore, by (7.59), each of the $2^m - 1$ possibilities for such vectors w_B can produce at most:

$$2^{h-m-1} \cdot (2 + o(1)) \frac{2^k}{\sqrt{\pi \ell}} = (1 + o(1)) \frac{2^{n-m}}{\sqrt{\pi \ell}}$$

pairs $(A, B) \in \mathcal{A} \times \mathcal{B}$. Consider the remaining combination of the rows W, satisfying $w_B^{(k+j)} = 0$ for all $j \in W$, and let $\mathcal{B}_0$ denote the sets $B \in \mathcal{B}$ which can be produced from w_B. Using this notation, it is enough to show that (7.67) holds, and the claim will follow from the resulting calculation (7.68).

As before, the fact that $(M_\mathcal{B})_{h,p} \notin \{0, \pm 1\}$ and that the remaining rows $[h-1] \setminus W$ have 0 entries in column p, implies that there is at most one coefficient possible for row h. If no coefficient for row h is legal, we get $\mathcal{B}_0 = \emptyset$ and (7.67) holds, otherwise let $\tilde{w}_B$ denote the sum of w_B with the appropriate multiple of row h of $M_\mathcal{B}$. We are left with $h - m - 1$ rows of $M_\mathcal{B}$ whose coefficients were not yet determined: rows $[h-1] \setminus W = [h'] \setminus \{p\}$.

If $\tilde{w}_B^{(j)} \neq 1 - \tilde{w}_B^{(k+j)}$ for some $j \in [h'] \setminus \{p\}$ or $\tilde{w}_B \neq \{0,1\}^n$, we obtain an additional factor of at most $\frac{1}{2}$ from one of the remaining rows of $M_\mathcal{B}$, and $|\mathcal{B}_0| \leq 2^{h-m-2}$. Combining this with (7.59) implies that (7.67) holds for $\alpha = \frac{1}{2}$. Assume therefore that $\tilde{w}_B^{(j)} = 1 - \tilde{w}_B^{(k+j)}$ for all $j \in [h'] \setminus \{p\}$ and that $\tilde{w}_B \in \{0,1\}^n$, and define:

$$L = [h'] \setminus \{p\} \cup \left\{ i \in \{h'+1, \ldots, k\} \cup \{p\} : \tilde{w}_B^{(i)} = 1 \right\} .$$

Since every set B produced from $\tilde{w}_B$ satisfies $|B \cap \{j, k+j\}| = 1$ for all $j \in [h'] \setminus \{p\}$ and $k + j \notin B$ for all $j \in W$, we deduce that, if $p \notin [h']$ (in which case $W = \{h'+1, \ldots, h-1\}$):

$$\ell = |A \cap B| = \mathbf{1}_{\{n \in A \cap B\}} + \sum_{i \in L} X_i \ , \tag{7.70}$$

for all $A \in \mathcal{A}$, where $X_i \in \{0,1\}$ denotes the coefficient for row i in a combination which produces A from $M_\mathcal{A}$. On the other hand, if $p \in [h']$, then $p \in W$ and it follows that $\tilde{w}_B^{(k+p)} = 0$, and:

- If $\tilde{w}_B^{(p)} = 0$, then $p \notin L$, and indeed, X_p does not contribute to $|A \cap B|$ for all $A \in \mathcal{A}$ and B produced by $\tilde{w}_B$, as neither p nor $k + p$ belong to B.

- If $\tilde{w}_B^{(p)} = 1$, then $p \in L$, and indeed X_p contributes 1 to $|A \cap B|$ for all $A \in \mathcal{A}$ and B produced by $\tilde{w}_B$, as $p \in B$ and $k + p \notin B$.

We deduce that (7.70) holds for $p \in [h']$ as-well. Recalling that $\mathcal{B}_0 \neq \emptyset$ (otherwise (7.67) immediately holds) (7.70) gives $|L| \geq \ell - 1$, and in particular, $|L| \geq (1 + o(1))\ell$. Using the definition (7.64), it follows that:

$$|\mathcal{A}| = \begin{cases} |\mathcal{A}^{(n)}_{L,\ell,0}| + |\mathcal{A}^{(n)}_{L,\ell,1}| & \text{if } \tilde{w}^{(n)}_B = 0 \\ \\ |\mathcal{A}^{(n)}_{L,\ell,0}| + |\mathcal{A}^{(n)}_{L,\ell-1,1}| & \text{if } \tilde{w}^{(n)}_B = 1 \end{cases} .$$

Applying Claim 7.6.4 (recall that $|L| \geq \ell - 1$) gives:

$$|\mathcal{A}| \leq 2 \cdot \left(\frac{3}{4} + o(1)\right) \frac{2^k}{\sqrt{\pi \ell}} = \left(\frac{3}{2} + o(1)\right) \frac{2^k}{\sqrt{\pi \ell}} ,$$

and as $|\mathcal{B}_0| \leq 2^{h-m-1}$, (7.67) holds for $\alpha = \frac{3}{4}$, as required. ■

This completes the proof of Claim 7.6.3 and of Lemma 7.4.2.

7.7 Concluding remarks and open problems

- We have shown that if two families of subsets of an n-element set, $\mathcal{A}, \mathcal{B}$, are ℓ-cross-intersecting, and ℓ is sufficiently large, then $|\mathcal{A}||\mathcal{B}| \leq \binom{2\ell}{\ell} 2^{n-2\ell}$, and in addition, we have given a complete characterization of all the extremal pairs $\mathcal{A}, \mathcal{B}$ for which equality is achieved.

- It would be interesting to prove that the above result holds for all values of ℓ (instead of all $\ell \geq \ell_0$ for some ℓ_0). Perhaps knowing the precise structure of the extremal pairs $\mathcal{A}, \mathcal{B}$, as described in Theorem 7.1.1 (assuming that this holds for all ℓ), will assist in proving this result.

- Finally, one may consider the corresponding problem where the pair $\mathcal{A}, \mathcal{B}$ does not have one possible cross-intersection, but rather a set L of legal cross-intersections. Such notions have been studied in [2], [95], [71], with different restrictions on L, and it would be interesting to derive tight bounds on $|\mathcal{A}||\mathcal{B}|$, and possibly describe the structure of all the extremal pairs, when in addition, each member of L is larger than some predefined integer ℓ.

Part IV

Tensor graph powers and graph isoperimetric inequalities

Chapter 8

Independent sets in tensor graph powers

The results of this chapter appear in [12]

The tensor product of two graphs, G and H, has a vertex set $V(G)\times V(H)$ and an edge between (u,v) and (u',v') iff both $uu' \in E(G)$ and $vv' \in E(H)$. Let $A(G)$ denote the limit of the independence ratios of tensor powers of G, $\lim \alpha(G^n)/|V(G^n)|$. This parameter was introduced in [37], where it was shown that $A(G)$ is lower bounded by the vertex expansion ratio of independent sets of G. In this chapter we study the relation between these parameters further, and ask whether they are in fact equal. We present several families of graphs where equality holds, and discuss the effect the above question has on various open problems related to tensor graph products.

8.1 Introduction

The *tensor* product (also dubbed as categorical or weak product) of two graphs, $G \times H$, is the graph whose vertex set is $V(G) \times V(H)$, where two vertices (u,v),(u',v') are adjacent iff both $uu' \in E(G)$ and $vv' \in E(H)$, i.e., the vertices are adjacent in each of their coordinates. Clearly, this product is associative and commutative, thus G^n is well defined to be the tensor product of n copies of G.

The tensor product has attracted a considerable amount of attention ever since Hedetniemi conjectured in 1966 ([66]) that $\chi(G \times H)$ is equal to $\min\{\chi(G), \chi(H)\}$ (where $\chi(G)$ denotes the chromatic number of G), a problem which remains open (see [109] for an extensive survey of this problem). For further work on colorings of tensor products of graphs, see [9], [61], [74], [100], [101], [108], [110].

It is easy to verify that Hedetniemi's conjecture is true when there is a homomorphism from G to H, and in particular when $G = H$, by examining a copy of G in $G \times H$, and it follows that $\chi(G^n) = \chi(G)$ for every integer n. Furthermore, a similar argument shows that $\omega(G \times H)$ (the clique number of $G \times H$) equals $\min\{\omega(G), \omega(H)\}$ for every two graphs G and H, and in particular, $\omega(G^n) = \omega(G)$ for every integer n. However, the behavior of the independence ratios of the graphs G^n is far more interesting. Let $i(G) = \alpha(G)/|V(G)|$ denote the independence ratio of G. Notice that for every two graphs G and H, if I is an independent set of G, then the cartesian product $I \times V(H)$ is independent in $G \times H$, hence every two graphs G and H satisfy:

$$i(G \times H) \geq \max\{i(G), i(H)\} \ . \tag{8.1}$$

Therefore, the series $i(G^n)$ is monotone non-decreasing and bounded, hence its limit exists; we denote this limit, introduced in [37], where it is called the Ultimate Categorical Independence Ratio of G, by $A(G)$. In contrast to the clique numbers and chromatic numbers, $A(G)$ may indeed exceed its value at the first power of G, $i(G)$. The authors of [37] proved the following simple lower bound for $A(G)$: if I is an independent set of G, then $A(G) \geq \frac{|I|}{|I|+|N(I)|}$, where $N(I)$ denotes the vertex neighborhood of I. We thus have the following lower bound on $A(G)$: $A(G) \geq a(G)$, where

$$a(G) = \max_{I \text{ ind. set}} \frac{|I|}{|I| + |N(I)|} \ .$$

It easy to see that $a(G)$ resembles $i(G)$ in the sense that $a(G \times H) \geq \max\{a(G), a(H)\}$ (to see this, consider the cartesian product $I \times V(H)$, where I is an independent set of G which attains the ratio $a(G)$). However, as opposed to $i(G)$, it is not clear if there are any graphs G, H such that $a(G \times H) > \max\{a(G), a(H)\}$ and yet $a(G), a(H) \leq \frac{1}{2}$. This is further discussed later.

It is not difficult to see that if $A(G) > \frac{1}{2}$ then $A(G) = 1$, thus $A(G) \in (0, \frac{1}{2}] \cup \{1\}$, as proved in [37] (for the sake of completeness, we will provide short proofs for this fact and for the fact that $A(G) \geq a(G)$ in Section 8.2). Hence, we introduce the following variant of $a(G)$:

$$a^*(G) = \begin{cases} a(G) & \text{if } a(G) \leq \frac{1}{2} \\ 1 & \text{if } a(G) > \frac{1}{2} \end{cases},$$

and obtain that $A(G) \geq a^*(G)$ for every graph G. The following question seems crucial to the understanding of the behavior of independence ratios in tensor graph powers:

Question 8.1.1. *Does every graph G satisfy $A(G) = a^*(G)$?*

In other words, are non-expanding independent sets of G the only reason for an increase in the independence ratio of larger powers? If so, this would immediately settle several open problems related to $A(G)$ and to fractional colorings of tensor graph products. Otherwise, an example of a graph G satisfying $A(G) > a^*(G)$ would demonstrate a thus-far unknown way to increase $A(G)$. While it may seem unreasonable that the complicated parameter $A(G)$ translates into a relatively easy property of G, so far the intermediate results on several conjectures regarding $A(G)$ are consistent with the consequences of an equality between $A(G)$ and $a^*(G)$.

As we show later, Question 8.1.1 has the following simple equivalent form:

Question 8.1.1'. *Does every graph G satisfy $a^*(G^2) = a^*(G)$?*

Conversely, is there a graph G which satisfies the following two properties:

1. Every independent set I of G has at least $|I|$ neighbors (or equivalently, $a(G) \leq \frac{1}{2}$).

2. There exists an independent set J of G^2 whose vertex-expansion ratio, $\frac{|N(J)|}{|J|}$, is strictly *smaller* than $\frac{|N(I)|}{|I|}$ for every independent set I of G.

In this chapter, we study the relation between $A(G)$ and $a^*(G)$, show families of graphs where equality holds, and discuss the effects of Question 8.1.1 on several conjectures regarding $A(G)$ and fractional colorings of tensor graph products. The rest of the chapter is organized as follows:

In Section 8.2 we present several families of graphs where equality holds between $A(G)$ and $a^*(G)$. First, we extend some of the ideas of [37] and obtain a characterization of all graphs G which satisfy the property $A(G) = 1$, showing that for these graphs $a^*(G)$ and $A(G)$ coincide. In the process, we obtain a polynomial time algorithm for determining whether a graph G satisfies $A(G) = 1$. We conclude the section by observing that $A(G) = a(G)$ whenever G is vertex transitive, and when it is the disjoint union of certain vertex transitive graphs.

Section 8.3 discusses the parameters $i(G)$ and $a(G)$ when G is a tensor product of two graphs, G_1 and G_2. Taking $G_1 = G_2$, we show the equivalence between Questions 8.1.1 and 8.1.1'. Next, when G_1 and G_2 are both vertex transitive, the relation between $i(G)$ and $a(G)$ is related to a fractional version of Hedetniemi's conjecture, raised by Zhu in [108]. We show that for every two graphs G and H, $A(G+H) = A(G\times H)$, where $G+H$ is the disjoint union of G and H. This property links the above problems, along with Question 8.1.1, to the problem of determining $A(G + H)$, raised in [37] (where it is conjectured to be equal to $\max\{A(G), A(H)\}$). Namely, the equality $A(G+H) = A(G \times H)$ implies that if $A(H) = a^*(H)$ for $H = G_1+G_2$, then:

$$i(G_1 \times G_2) \leq a^*(G_1 + G_2) = \max\{a^*(G_1), a^*(G_2)\} \ .$$

This raises the following question, which is a weaker form of Question 8.1.1:

Question 8.1.2. *Does the inequality $i(G \times H) \leq \max\{a^*(G), a^*(H)\}$ hold for every two graphs G and H?*

We proceed to demonstrate that several families of graphs satisfy this inequality, and in the process, obtain several additional families of graphs G which satisfy $A(G) = a(G) = a^*(G)$.

Section 8.4 is devoted to concluding remarks and open problems. We list several additional interesting questions which are related to $a(G)$, as well as summarize the main problems which were discussed in the previous sections. Among the new mentioned problems are those of determining or estimating the value of $A(G)$ for the random graph models $\mathcal{G}_{n,d}$, $\mathcal{G}_{n,\frac{1}{2}}$ and for the random graph process.

8.2 Equality between $A(G)$ and $a^*(G)$

8.2.1 Graphs G which satisfy $A(G) = 1$

In this section we prove a characterization of graphs G satisfying $A(G) = 1$, showing that this is equivalent to the non-existence of a fractional perfect matching in G. A *fractional matching* in a graph $G = (V, E)$ is a function $f : E \to \mathbb{R}^+$ such that for every $v \in V$, $\sum_{v \in e} f(e) \leq 1$ (a matching is the special case of restricting the values of f to $\{0, 1\}$). The value of the fractional matching is defined as $f(E) = \sum_{e \in E} f(e)$ $(\leq \frac{|V|}{2})$. A *fractional perfect matching* is a fractional matching which achieves this maximum: $f(E) = \frac{|V|}{2}$.

Theorem 8.2.1. *For every graph G, $A(G) = 1$ iff $a^*(G) = 1$ iff G does not contain a fractional perfect matching.*

The proof of Theorem 8.2.1 relies on the results of [37] mentioned in the introduction; we recall these results and provide short proofs for them.

Claim 8.2.2 ([37])**.** *For every graph G, $A(G) \geq a(G)$.*

Proof. Let I be an independent set which attains the maximum of $a(G)$. Clearly, for every $k \in \mathbb{N}$, all vertices in G^k, which contain a member of $I \cup N(I)$ in one of their coordinates, and in addition, whose first coordinate out of $I \cup N(I)$ belongs to I, form an independent set. As k tends to infinity, almost every vertex has a member of $I \cup N(I)$ in at least one of its coordinates, and the second restriction implies that the fractional size of the set above tends to $\frac{|I|}{|I|+|N(I)|} = a(G)$. ■

Claim 8.2.3 ([37])**.** *If $A(G) > \frac{1}{2}$ then $A(G) = 1$.*

Proof. Assume, without loss of generality, that $i(G) > \frac{1}{2}$, and let I be a maximum independent set of G. For every power k, the set of all vertices of G^k, in which strictly more than $\frac{k}{2}$ of the coordinates belong to I, is independent. Clearly, since $\frac{|I|}{|G|} > \frac{1}{2}$, the size of this set tends to $|V(G)|^k$ as k tends to infinity (as the probability of more Heads than Tails in a sufficiently long sequence of tosses of a coin biased towards Heads is nearly 1), hence $A(G) = 1$. ■

Proof of Theorem 8.2.1. By Claims 8.2.2 and 8.2.3, if $a(G) > \frac{1}{2}$ (or equivalently, $a^*(G) = 1$) then $A(G) = 1$. Conversely, assuming that $a^*(G) = a(G) \leq \frac{1}{2}$, we must show that $A(G) < 1$. This will follow from the following simple lemma, proved by Tutte in 1953 (cf., e.g., [82] p. 216):

Lemma 8.2.4. *For a given set $S \subset V(G)$, let $N(S)$ denote that set of all vertices of G which have a neighbor in S; then every set $S \subset V(G)$ satisfies $|N(S)| \geq |S|$ iff every independent set $I \subset V(G)$ satisfies $|N(I)| \geq |I|$.*

Proof of lemma. One direction is obvious; for the other direction, take a subset S with $|N(S)| < |S|$. Define S' to be $\{v \in S \mid N(v) \cap S \neq \emptyset\}$, and examine $I = S \setminus S'$. Since $S' \subset N(S)$ and $|N(S)| < |S|$, I is nonempty, and is obviously independent. Therefore $|N(I)| \geq |I|$, however $|N(I)| \leq |N(S)| - |S'| < |S| - |S'| = |I|$, yielding a contradiction. ■

Returning to the proof of the theorem, observe that by our assumption that $a(G) \leq \frac{1}{2}$ and the lemma, Hall's criterion for a perfect matching applies to the bipartite graph $G \times K_2$ (where K_2 is the complete graph on two vertices). Therefore, G contains a factor $H \subset G$ of vertex disjoint cycles and edges (to see this, as long as the matching is nonempty, repeatedly traverse it until closing a cycle and omit these edges). Since removing edges from G may only increase $A(G)$, it is enough to show that $A(H) < 1$.

We claim that the subgraph H satisfies $A(H) \leq \frac{1}{2}$. To see this, argue as follows: direct H according to its cycles and edges (arbitrarily choosing clockwise or counter-clockwise orientations), and examine the mapping from each vertex to the following vertex in its cycle. This mapping is an invertible function $f : V \to V$, such that for all $v \in V$, $vf(v) \in E(H)$. Now let I be an independent set of H^k. Pick a random vertex $\underline{u} \in V(H^k)$, uniformly over all the vertices, and consider the pair $\{\underline{u}, \underline{v}\}$, where $\underline{v} = f(\underline{u})$ is the result of applying f on each coordinate of $\underline{u}$. Obviously $\underline{v}$ is uniformly distributed over H^k as-well, thus:

$$\mathbb{E}\,|I \cap \{\underline{u}, \underline{v}\}| \geq \frac{2}{|H^k|}|I| \; .$$

Choosing a vertex $\underline{u}$ for which $|I \cap \{\underline{u}, \underline{v}\}|$ is at least its expected value, and recalling that $\underline{u}$ and $\underline{v}$ are adjacent in H^k, we get:

$$\frac{2}{|H^k|}|I| \leq |I \cap \{\underline{u}, \underline{v}\}|) \leq 1 \ .$$

Hence, $i(H^k) \leq \frac{1}{2}$, and thus $A(H) \leq \frac{1}{2}$.

An immediate corollary from the above proof that $A(G) = 1$ iff $a^*(G) = 1$ is the equivalence between the property $A(G) \leq \frac{1}{2}$ and the existence of a fractional perfect matching in the graph G. It is well known (see for instance [82]) that for every graph G, the maximal fractional matching of G can be achieved using only the weights $\{0, \frac{1}{2}, 1\}$. Therefore, a fractional perfect matching is precisely a factor $H \subset G$, comprised of vertex disjoint cycles and edges, and we obtain another format for the condition $a(G) \leq \frac{1}{2}$: $A(G) \leq \frac{1}{2}$ iff G has a fractional perfect matching; otherwise, $A(G) = 1$.

Notice that a fractional perfect matching f of G immediately induces a fractional perfect matching on G^k for every k (assign an edge of G^k a weight equaling the product of the weights of each of the edges in the corresponding coordinates). As it is easy to see that a fractional perfect matching implies that $i(G) \leq \frac{1}{2}$, this provides an alternative proof that if $a(G) \leq \frac{1}{2}$ then $A(G) \leq \frac{1}{2}$. ■

Since Lemma 8.2.4 also provides us with a polynomial algorithm for determining whether $a(G) > \frac{1}{2}$ (determine whether Hall's criterion applies to $G \times K_2$, using network flows), we obtain the following corollary:

Corollary 8.2.5. *Given an input graph G, determining whether $A(G) = 1$ or $A(G) \leq \frac{1}{2}$ can be done in polynomial time.*

8.2.2 Vertex transitive graphs

The observation that $A(G) = a(G)$ whenever G is vertex transitive (notice that $A(G) \leq \frac{1}{2}$ for every nontrivial regular graph G) is a direct corollary of the following result of [9] (the proof of this fact is by covering G^k uniformly by copies of G):

Proposition 8.2.6 ([9]). *If G is vertex transitive, then $A(G) = i(G)$.*

Clearly, for every graph G, $i(G) \leq a(G)$. Hence, for every vertex transitive graph G the following holds:

$$A(G) = i(G) \leq a(G) \leq A(G) \ ,$$

proving the following corollary:

Observation 8.2.7. *For every vertex transitive graph G, $A(G) = a^*(G) = a(G)$.*

We conclude this section by mentioning several families of vertex transitive graphs G and H whose disjoint union $G + H$ satisfies $A(G + H) = a(G + H) = \max\{A(G), A(H)\}$. These examples satisfy both the property of Question 8.1.1 and the disjoint union conjecture of [37].

The next two claims follow from the results of Section 8.3, as we later show. For the first claim, recall that a circular complete graph (defined in [108]), $K_{n/d}$, where $n \geq 2d$, has a vertex set $\{0, \ldots, n-1\}$ and an edge between i, j whenever $d \leq |i - j| \leq n - d$. A Kneser graph, $KN_{n,k}$, where $k \leq n$, has $\binom{n}{k}$ vertices corresponding to k-element subsets of $\{1, \ldots, n\}$, and two vertices are adjacent iff their corresponding subsets are disjoint.

Claim 8.2.8. *Let G and H be two vertex transitive graphs, where H is one of the following: a Kneser graph, a circular complete graph, a cycle or a complete bipartite graph. Then $G + H$ satisfies $A(G + H) = a(G + H) = \max\{A(G), A(H)\}$.*

Claim 8.2.9. *Let G and H be two vertex transitive graphs satisfying $\chi(G) = \omega(G) \leq \omega(H)$. Then $A(G + H) = a(G + H) = \max\{A(G), A(H)\}$.*

8.3 The tensor product of two graphs

8.3.1 The expansion properties of G^2

Question 8.1.1, which discusses the relation between the expansion of independent sets of G, and the limit of independence ratios of tensor powers of G, can be translated into a seemingly simpler question (stated as Question 8.1.1') comparing the vertex expansions of a graph and its square: can the minimal expansion ratio $|N(I)|/|I|$ of independent sets I decrease in the second power of G?

To see the equivalence between Questions 8.1.1 and 8.1.1', argue as follows: assuming the answer to Question 8.1.1 is positive, every graph G satisfies:

$$a^*(G) = A(G) = A(G^2) = a^*(G^2) \ ,$$

and hence $a^*(G^2) = a^*(G)$ (recall that every graph H satisfies $a(H^2) \geq a(H)$). Conversely, suppose that there exists a graph G such that $A(G) > a^*(G)$. By the simple fact that every graph H satisfies $i(H) \leq a^*(H)$ we conclude that there exists an integer k such that $a^*(G^{2^k}) \geq i(G^{2^k}) > a^*(G)$, and therefore there exists some integer $\ell \leq k$ for which $a(G^{2^\ell}) > a(G^{2^{\ell-1}})$.

8.3.2 The relation between the tensor product and disjoint unions

In this section we prove the following theorem, which links between the quantities $i(G_1 \times G_2)$, $a(G_1 \times G_2)$, $\chi_f(G_1 \times G_2)$ and $A(G_1 + G_2)$, where $\chi_f(G)$ denotes the fractional chromatic number of G:

Theorem 8.3.1. *For every two vertex transitive graphs G_1 and G_2, the following statements are equivalent:*

$$\begin{aligned} i(G_1 \times G_2) &\leq \max\{a^*(G_1), a^*(G_2)\} && (8.2)\\ a^*(G_1 \times G_2) &\leq \max\{a^*(G_1), a^*(G_2)\} && (8.3)\\ \chi_f(G_1 \times G_2) &= \min\{\chi_f(G_1), \chi_f(G_2)\} && (8.4)\\ A(G_1 + G_2) &= \max\{A(G_1), A(G_2)\} && (8.5) \end{aligned}$$

Proof. The proof of Theorem 8.3.1 relies on the following proposition:

Proposition 8.3.2. *For every two graphs G and H, $A(G+H) = A(G \times H)$.*
We note that this generalizes a result of [37], which states that $A(G+H)$ is at least $\max\{A(G), A(H)\}$. Indeed, that result immediately follows from the fact that $A(G \times H)$ is always at least the maximum of $A(G)$ and $A(H)$ (by (8.1)).

Proof of Proposition 8.3.2. Examine $(G+H)^n$, and observe that a vertex whose i-th coordinate is taken from G is disconnected from all vertices whose i-th coordinate is taken from H. Hence, we can break down the n-th power of the disjoint union $G+H$ to 2^n disjoint graphs, and obtain:

$$\alpha\left((G+H)^n\right) = \sum_{k=0}^{n} \binom{n}{k} \alpha\left(G^k H^{n-k}\right) \ . \tag{8.6}$$

To prove that $A(G+H) \geq A(G\times H)$, fix $\varepsilon > 0$, and let N denote a sufficiently large integer such that $i\left((G\times H)^N\right) \geq (1-\varepsilon)A(G\times H)$. The following is true for every $n > 2N$ and $N \leq k \leq n-N$:

$$\begin{aligned} i(G^k H^{n-k}) &= i\left((G\times H)^N G^{k-N} H^{n-k-N}\right) \\ &\geq i\left((G\times H)^N\right) \geq (1-\varepsilon)A(G\times H) \ , \end{aligned}$$

where the first inequality is by (8.1). Using this inequality together with (8.6) yields:

$$\begin{aligned} & i\left((G+H)^n\right) \geq \frac{1}{|(G+H)^n|} \sum_{k=N}^{n-N} \binom{n}{k} \alpha(G^k H^{n-k}) \\ & \geq \frac{\sum_{k=N}^{n-N} \binom{n}{k} |G|^k |H|^{n-k}}{(|G|+|H|)^n} (1-\varepsilon) A(G\times H) \\ & \xrightarrow[n\to\infty]{} (1-\varepsilon)A(G\times H) \ . \end{aligned}$$

Therefore $A(G+H) \geq (1-\varepsilon)A(G\times H)$ for any $\varepsilon > 0$, as required.

It remains to show that $A(G+H) \leq A(G\times H)$. First observe that (8.1) gives the following relation:

$$\forall\ k,l \geq 1\ , \quad i(G^k H^l) \leq i(G^k H^l \times G^l H^k) = i(G^{k+l} H^{k+l}) \leq A(G\times H)\ . \tag{8.7}$$

Using (8.6) again, we obtain:

$$\begin{aligned} & i\left((G+H)^n\right) = \sum_{k=0}^{n} \binom{n}{k} \frac{\alpha(G^k H^{n-k})}{(|G|+|H|)^n} = \sum_{k=0}^{n} \binom{n}{k} i(G^k H^{n-k}) \cdot \frac{|G|^k |H|^{n-k}}{(|G|+|H|)^n} \\ & \leq \frac{|G|^n}{(|G|+|H|)^n} i(G^n) + \frac{|H|^n}{(|G|+|H|)^n} i(H^n) + \left(1 - \frac{|G|^n + |H|^n}{(|G|+|H|)^n}\right) A(G\times H) \\ & \leq \frac{|G|^n}{(|G|+|H|)^n} A(G) + \frac{|H|^n}{(|G|+|H|)^n} A(H) + \left(1 - \frac{|G|^n + |H|^n}{(|G|+|H|)^n}\right) A(G\times H) \\ & \longrightarrow A(G\times H)\ , \end{aligned}$$

where the first inequality is by (8.7), and the second is by definition of $A(G)$. ■

Equipped with the last proposition, we can now prove that Question 8.1.2 is indeed a weaker form of Question 8.1.1, namely that if $A(G) = a^*(G)$ for every G, then $i(G_1 \times G_2) \leq \max\{a^*(G_1), a^*(G_2)\}$ for every two graphs G_1, G_2. Indeed, if $A(G_1 + G_2) = a^*(G_1 + G_2)$ then inequality (8.2) holds, as well as the stronger inequality (8.3):

$$\begin{aligned} a^*(G_1 \times G_2) &\leq A(G_1 \times G_2) = A(G_1 + G_2) \\ &= a^*(G_1 + G_2) = \max\{a^*(G_1), a^*(G_2)\} \ , \end{aligned}$$

as required.

Having shown that a positive answer to Question 8.1.1 implies inequality (8.3) (and hence inequality (8.2) as well), we show the implications of inequality (8.2) when the two graphs are vertex transitive.

Recall that for every two graphs G_1 and G_2,

$$i(G_1 \times G_2) \geq \max\{i(G_1), i(G_2)\} \ ,$$

and consider the case when G_1, G_2 are both vertex transitive and have edges. In this case, $i(G_i) = a(G_i) = a^*(G_i)$ $(i = 1, 2)$, hence inequalities (8.2) and (8.3) are equivalent, and are both translated into the form

$$i(G_1 \times G_2) = \max\{i(G_1), i(G_2)\} \ .$$

Next, recall that for every vertex transitive graph G, $i(G) = 1/\chi_f(G)$. Hence, inequality (8.2) (corresponding to Question 8.1.2), when restricted to vertex transitive graphs, coincides with (8.4). Furthermore, by Observation 8.2.7 and Proposition 8.3.2, for vertex transitive G_1 and G_2 we have:

$$\begin{aligned} i(G_1 \times G_2) &= A(G_1 \times G_2) = A(G_1 + G_2) \\ &\geq \max\{A(G_1), A(G_2)\} = \max\{i(G_1), i(G_2)\} \ , \end{aligned}$$

hence in this case (8.4) also coincides with (8.5). Thus, all four statements are equivalent for vertex transitive graphs. ■

By the last theorem, the following two conjectures, raised in [37] and [108], coincide for vertex transitive graphs:

Conjecture 8.3.3 ([37]). *For every two graphs G and H, the following holds: $A(G+H) = \max\{A(G), A(H)\}$.*

Conjecture 8.3.4 ([108]). *For every two graphs G and H, the following holds: $\chi_f(G \times H) = \min\{\chi_f(G), \chi_f(H)\}$.*

The study of Conjecture 8.3.4 is somewhat related to the famous and long studied Hedetniemi conjecture (stating that $\chi(G \times H) = \min\{\chi(G), \chi(H)\}$), as for every two graphs G and H, $\omega(G \times H) = \min\{\omega(G), \omega(H)\}$, and furthermore $\omega(G) \leq \chi_f(G) \leq \chi(G)$.

It is easy to see that the inequality

$$\chi_f(G \times H) \leq \min\{\chi_f(G), \chi_f(H)\}$$

is always true. It is shown in [100] that Conjecture 8.3.4 is not far from being true, by proving that for every graphs G and H,

$$\chi_f(G \times H) \geq \frac{1}{4} \min\{\chi_f(G), \chi_f(H)\} \ .$$

So far, Conjecture 8.3.4 was verified (in [108]) for the cases in which one of the two graphs is either a Kneser graph or a circular-complete graph. This implies the cases of H belonging to these two families of graphs in Claim 8.2.8. Claim 8.2.9 is derived from the the following remark, which provides another family of graphs for which Conjecture 8.3.4 holds.

Remark 8.3.5: Let G and H be graphs such that $\chi(G) = \omega(G) \leq \omega(H)$. It follows that $\omega(G \times H) = \omega(G) = \chi(G \times H)$, and thus $\chi_f(G \times H) = \min\{\chi_f(G), \chi_f(H)\}$, and $\chi(G \times H) = \min\{\chi(G), \chi(H)\}$. In particular, this is true when G and H are perfect graphs.

8.3.3 Graphs satisfying the property of Question 8.1.2

In this subsection we note that several families of graphs satisfy inequality (8.2) (and the property of Question 8.1.2). This appears in the following propositions:

Proposition 8.3.6. *For every graph G and integer ℓ, the following holds: $i(G \times C_\ell) \leq \max\{a(G), a(C_\ell\}$ (and hence, $i(G \times C_\ell) \leq \max\{a^*(G), a^*(C_\ell)\}$). This result can be extended to $G \times H$, where H is a disjoint union of cycles.*

Proof. We need the following lemma:

Lemma 8.3.7. *Let G and H be two graphs which satisfy at least one of the following conditions:*

1. *$a(G) \geq \frac{1}{2}$, and every $S \subsetneqq V(H)$ satisfies $|N(S)| > |S|$.*
2. *$a(G) > \frac{1}{2}$ and every $S \subset V(H)$ satisfies $|N(S)| \geq |S|$.*

Then every maximum independent set $I \subset V(G \times H)$ contains at least one "full copy" of H, i.e., for each such I there is a vertex $v \in V(G)$, such that $\{(v, w) : w \in H\} \subset I$.

Proof of lemma. We begin with the case $a(G) \geq \frac{1}{2}$ and $|N(S)| > |S|$ for every $S \subsetneqq V(H)$. Let J be a smallest (with respect to either size or containment) nonempty independent set in G such that $\frac{|J|}{|J|+|N(J)|} \geq \frac{1}{2}$. Clearly, $|N(J)| \leq |J|$. We claim that this inequality proves the existence of a one-one function $f : N(J) \to J$, such that $vf(v) \in E(G)$ (that is, there is a matching between $N(J)$ and J which saturates $N(J)$). To prove this fact, take any set $S \subset N(J)$ and assume $|N(S) \cap J| < |S|$; it is thus possible to delete $N(S) \cap J$ from J (and at least $|S|$ vertices from $N(J)$) and since $|N(S) \cap J| < |S| \leq |N(J)| \leq |J|$ we are left with a nonempty $J' \subsetneqq J$ satisfying $|N(J')| \leq |J'|$. This contradicts the minimality of J. Now we can apply Hall's Theorem to match a unique vertex in J for each vertex in $N(J)$.

Assume the lemma is false, and let I be a counterexample. Examine the intersection of I with a pair of copies of H, which are matched in the matching above between $N(J)$ and J. As we assumed that there are no full H copies in I, each set S of vertices in a copy of H has at least $|S|+1$ neighbors in an adjacent copy of H. Thus, each of the matched pairs of $N(J) \to J$ contains at most $|H| - 1$ vertices of I. Define I' as the result of adding all missing vertices from the H copies of J to I, and removing all existing vertices from the copies of $N(J)$ (all other vertices remain unchanged). Then I' is independent, and we obtain a contradiction to the maximality of I.

The case of $a(G) > \frac{1}{2}$ and $|N(S)| \geq |S|$ for every $S \subset V(H)$ is essentially the same. The set J is now the smallest independent set of G for which

$|N(J)| < |J|$, and again, this implies the existence of a matching from $N(J)$ to J, which saturates $N(J)$. By our assumption on H, each pair of copies of H in the matching contributes at most $|H|$ vertices to a maximum independent set I of $G \times H$, and by the assumption on I, the unmatched copies of H (recall $|J| > |N(J)|$) are incomplete. Therefore, we have strictly less than $|H||J|$ vertices of I in $(N(J) \cup J) \times H$, contradicting the maximality of I. ■

Returning to the proof of the proposition, let I be a maximum independent set of $G \times C_\ell$. Remove all vertices, which belong to full copies of C_ℓ in I, if there are any, along with all their neighbors (note that these neighbors are also complete copies of C_ℓ, but this time empty ones). These vertices contribute a ratio of at most $a(G)$, since their copies form an independent set in G. Let G' denote the induced graph of G on all remaining copies. The set I', defined to be $I \cap (G' \times C_\ell)$, is a maximum independent set of $G' \times C_\ell$, because for any member of I we removed, we also removed all of its neighbors from the graph.

Notice that C_ℓ satisfies the expansion property required from H in Lemma 8.3.7: for every k, every set $S \subsetneqq V(C_{2k+1})$ satisfies $|N(S)| > |S|$, and every set $S \subset V(C_{2k})$ satisfies $|N(S)| \geq |S|$. We note that, in fact, by the method used in the proof of Lemma 8.2.4 it is easy to show that every regular graph H satisfies $|N(S)| \geq |S|$ for every set $S \subset V(H)$, and if in addition H is non-bipartite and connected, then every $S \subsetneqq V(H)$ satisfies $|N(S)| > |S|$.

We can therefore apply the lemma on $G' \times C_\ell$. By definition, there are no full copies of C_ℓ in I', hence, by the lemma, we obtain that $a(G') \leq \frac{1}{2}$ (and even $a(G') < \frac{1}{2}$ in case ℓ is odd). In particular, we can apply Hall's Theorem and obtain a factor of edges and cycles in G'. Each connected pair of non-full copies has an independence ratio of at most $i(C_\ell) = a(C_\ell)$ (by a similar argument to the one stated in the proof of the lemma), and double counting the contribution of the copies in the cycles we conclude that $\frac{|I'|}{|G'||C_\ell|} \leq a(C_\ell)$. Therefore $i(G \times C_\ell)$ is an average between values, each of which is at most $\max\{a(G), a(C_\ell)\}$, completing the proof. ■

Proposition 8.3.8. *For every graph G and integer k, the following holds: $i(G \times K_k) \leq \max\{a(G), a(K_k)\}$ (and hence, such graphs satisfy the inequality of Question 8.1.2). This result can be extended to $G \times H$, where H is a disjoint* *union of complete graphs.*

Proof. Let I denote a maximum independent set of $G \times K_k$, and examine all copies of K_k which contain at least two vertices of I. Such a copy of K_k in $G \times K_k$ forces its neighbor copies to be empty (since two vertices of K_k have the entire graph K_k as their neighborhood). Therefore, by the maximality of I, such copies must contain *all* vertices of K_k. Denote the vertices of G which represent these copies by $S \subset V(G)$; then S is an independent set of G, and the copies represented by $S \cup N(S)$ contribute an independence ratio of at most $a(G)$. Each of the remaining copies contains at most one vertex, giving an independence ratio of at most $\frac{1}{k} = a(K_k)$. Therefore, $i(G \times K_k)$ is an average between values which are at most $\max\{a(G), a(K_k)\}$, and the result follows. ∎

Corollary 8.3.9. *Let G be a graph satisfying $a(G) = \frac{1}{2}$; then for every graph H the following inequality holds: $i(G \times H) \leq \max\{a(G), a(H)\}$.*

Proof. By Theorem 8.2.1 we deduce that G contains a fractional perfect matching; let G' be a factor of G consisting of vertex disjoint cycles and edges. Since $a(G') \leq \frac{1}{2}$ as-well, it is enough to show that $i(G' \times H) \leq \max\{a(G'), a(H)\}$. Indeed, since G' is a disjoint union of the form $C_{\ell_1} + \ldots + C_{\ell_k} + K_2 + \ldots + K_2$, the result follows from Proposition 8.3.6 and Proposition 8.3.8. ∎

8.4 Concluding remarks and open problems

We have seen that answering Question 8.1.1 is imperative to the understanding of the behavior of independent sets in tensor graph powers. While it is relatively simple to show that $A(G)$ equals $a(G)$ whenever G is vertex transitive, proving this equality for $G = G_1 + G_2$, the disjoint union of two vertex transitive graphs G_1 and G_2, seems difficult; it is equivalent to Conjecture 8.3.4, the fractional version of Hedetniemi's conjecture, for vertex transitive graphs. These two conjectures are consistent with a positive answer to Question 8.1.1, and are in fact direct corollaries in such a case.

The assertion of Conjecture 8.3.3 for several cases can be deduced from the spectral bound for $A(G)$ proved in [9]. For a regular graph G with n vertices and eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, denote $\Lambda(G) = \frac{-\lambda_n}{\lambda_1 - \lambda_n}$. As observed in [9], the usual known spectral upper bound for the independence number of a graph implies that for each regular G, $A(G) \leq \Lambda(G)$. It is not difficult to check that for regular G and H, $\Lambda(G \times H) = \max\{\Lambda(G), \Lambda(H)\}$. Therefore, by Proposition 8.3.2, if G and H are regular and satisfy $\Lambda(G) \leq \Lambda(H) = A(H)$, then the assertion of Conjecture 8.3.3 holds for G and H. Several examples of graphs H satisfying $\Lambda(H) = A(H)$ are mentioned in [9].

It is interesting to inspect the expected values of $A(G)$ for random graph models. First, consider $G^t \sim \mathcal{G}_{n,t}$, the random graph process on n vertices after t steps, where there are t edges chosen uniformly out of all possible edges (for more information on the random graph process, see [34]). It is not difficult to show, as mentioned in [24], that $a(G^t)$ is given by the minimal degree of G^t, $\delta(G^t)$, as long as $\delta(G^t)$ is fixed and $|G|$ is sufficiently large. When considering $A(G)$, the following is a direct corollary of the fractional perfect matching characterization for $A(G) = 1$ (Theorem 8.1.1), along with the fact that the property "G contains a fractional perfect matching" almost surely has the same hitting time as the property "$\delta(G) \geq 1$":

Remark 8.4.1: With high probability, the hitting time of the property $A(G) < 1$ equals the hitting time of $\delta(G) \geq 1$. Furthermore, almost every graph process at that time satisfies $A(G) = \frac{1}{2}$.

Question 8.4.2. *Does almost every graph process satisfy $A(G) = \frac{1}{\delta(G)+1}$ as long as $\delta(G)$ is fixed?*

Second, the expected value of $a(G)$ for a random regular graph $G \sim \mathcal{G}_{n,d}$ is easily shown to be $\Theta(\frac{\log d}{d})$, as the independence ratio of $\mathcal{G}_{n,d}$ is almost surely between $\frac{\log d}{d}$ and $2\frac{\log d}{d}$ as $n \to \infty$ (see [32], [107]). As for $A(G)$, the following is easy to prove, by the spectral upper bound $\Lambda(G)$ mentioned above, and by the eigenvalue estimations of [57]:

Remark 8.4.3: Let G denote the random regular graph $\mathcal{G}_{n,d}$; almost surely: $\Omega(\frac{\log d}{d}) \leq A(G) \leq O(\frac{1}{\sqrt{d}})$ as $d \to \infty$.

Question 8.4.4. *Is the expected value of $A(G)$ for the random regular graph $G \sim \mathcal{G}_{n,d}$ equal to $\Theta(\frac{\log d}{d})$?*

The last approach can be applied to the random graph $G \sim \mathcal{G}_{n,\frac{1}{2}}$ as well. To see this, consider a large regular factor (see [96]), and use the eigenvalue estimations of [59] to obtain that almost surely $\Omega(\frac{\log n}{n}) \leq A(G) \leq O(\sqrt{\frac{\log n}{n}})$, whereas $a(G)$ is almost surely $(2+o(1))\frac{\log_2 n}{n}$.

Question 8.4.5. *Is the expected value of $A(G)$ for the random graph $G \sim \mathcal{G}_{n,\frac{1}{2}}$ equal to $\Theta(\frac{\log n}{n})$?*

We conclude with the question of the decidability of $A(G)$. Clearly, deciding if $a(G) > \beta$ for a given value β is in NP, and we can show that it is in fact NP-complete. It seems plausible that $A(G)$ can be calculated (though not necessarily by an efficient algorithm) up to an arbitrary precision:

Question 8.4.6. *Is the problem of deciding whether $A(G) > \beta$, for a given graph G and a given value β, decidable?*

Chapter 9

The isoperimetric constant of the random graph process

The results of this chapter appear in [24]

The isoperimetric constant of a graph G on n vertices, $i(G)$, is the minimum of $\frac{|\partial S|}{|S|}$, taken over all nonempty subsets $S \subset V(G)$ of size at most $n/2$, where ∂S denotes the set of edges with precisely one end in S. A random graph process on n vertices, $\widetilde{G}(t)$, is a sequence of $\binom{n}{2}$ graphs, where $\widetilde{G}(0)$ is the edgeless graph on n vertices, and $\widetilde{G}(t)$ is the result of adding an edge to $\widetilde{G}(t-1)$, uniformly distributed over all the missing edges. We show that in almost every graph process $i(\widetilde{G}(t))$ equals the minimal degree of $\widetilde{G}(t)$ as long as the minimal degree is $o(\log n)$. Furthermore, we show that this result is essentially best possible, by demonstrating that along the period in which the minimum degree is typically $\Theta(\log n)$, the ratio between the isoperimetric constant and the minimum degree falls from 1 to $\frac{1}{2}$, its final value.

9.1 Introduction

Let $G = (V, E)$ be a graph. For each subset of its vertices, $S \subseteq V$, we define its edge boundary, ∂S, as the set of all edges with exactly one endpoint in S:

$$\partial S = \{(u, v) \in E : u \in S, v \notin S\} .$$

The isoperimetric constant, or isoperimetric number, of $G = (V, E)$, $i(G)$, is defined to be:

$$i(G) = \min_{\emptyset \neq S \subset V} \frac{|\partial S|}{\min\{|S|, |V \setminus S|\}} = \min_{\substack{\emptyset \neq S \subset V \\ |S| \leq \frac{1}{2}|V|}} \frac{|\partial S|}{|S|} \ .$$

It is well known that this parameter, which measures edge expansion properties of a graph G, is strongly related to the spectral properties of G. Indeed:

$$\frac{\lambda}{2} \leq i(G) \leq \sqrt{\lambda(2\Delta(G) - \lambda)} \ , \tag{9.1}$$

where $\Delta(G)$ denotes the maximal degree of G, and λ denotes the second smallest eigenvalue of the Laplacian matrix of G (defined as $D(G) - A(G)$, where $D(G)$ is the diagonal matrix of degrees of vertices of G, and $A(G)$ is the adjacency matrix of G): for proofs of these facts, see [17] and [87]. The upper bound in (9.1) can be viewed as a discrete version of the Cheeger inequality bounding the first eigenvalue of a Riemannian manifold, and indeed, there is a natural relation between the study of isoperimetric inequalities of graphs and the study of Cheeger constants in spectral geometry. For instance, see [38], where the author relates isoperimetric constants and spectral properties of graphs with those of certain Riemann surfaces. The eigenvalue bounds in (9.1) also relate $i(G)$ (as well as a variation of it, the conductance of G) to the mixing time of a random walk in G, defined to be the minimal time it takes a random walk on G to approach the stationary distribution within a variation distance of $1/2$.

A closely related variant of the isoperimetric constant is the Cheeger constant of a graph, where the edge boundary of S is divided by its volume (defined to be the sum of its degrees) instead of by its size. For further information on this parameter, its relation to the isoperimetric constant, and its corresponding eigenvalue bounds (analogous to (9.1)), see [40], as well as [41], Chapter 2.

There has been much study of the isoperimetric constants of various graphs, such as grid graphs, torus graphs, the n-cube, and more generally, cartesian products of graphs. See, for instance, [36],[35],[42],[68], [87]. In [33], Bollobás studied the isoperimetric constant of random d-regular graphs, and used probabilistic arguments to prove that, for each d, infinitely many d-regular graphs G satisfy $i(G) \geq \frac{d}{2} - O(\sqrt{d})$. Alon proved in [5] that this inequality is in fact tight, by providing an upper bound of $i(G) \leq \frac{d}{2} - c\sqrt{d}$, where $c > 0$ is some absolute constant, for any d-regular graph G on a sufficiently large number of vertices.

In this chapter, we study the isoperimetric constant of general random graphs $\mathcal{G}(n,p)$, $\mathcal{G}(n,M)$, and the random graph process, and show that in these graphs, the ratio between the isoperimetric constant and the minimal degree exhibits an interesting behavior.

We briefly recall several elementary details on these models (for further information, c.f., e.g., [34], Chapter 2). The random graph $\mathcal{G}(n,p)$ is a graph on n vertices, where each pair of distinct vertices is adjacent with probability p, and independently of all other pairs of vertices. The distribution of $\mathcal{G}(n,p)$ is closely related with that of $\mathcal{G}(n,M)$, the uniform distribution on all graphs on n vertices with precisely M edges, if we choose $p = M/\binom{n}{2}$. The random graph process on n vertices, $\widetilde{G}(t)$, is a sequence of $\binom{n}{2}$ graphs, where $\widetilde{G}(0)$ is the edgeless graph on n vertices, and $\widetilde{G}(t)$ is the result of adding an edge to $\widetilde{G}(t-1)$, uniformly distributed over all the missing edges. Notice that at a given time $0 \leq t \leq \binom{n}{2}$, $\widetilde{G}(t)$ is distributed as $\mathcal{G}(n,M)$ with $M = t$.

For a given graph process $\widetilde{G}$ on n vertices, we define the hitting time of a monotone graph property $\mathcal{A}$ (a family of graphs closed under isomorphism and the addition of edges) as:

$$\tau(\mathcal{A}) = \min\ \left\{0 \leq t \leq \binom{n}{2} : \widetilde{G}(t) \in \mathcal{A}\right\} .$$

We use the abbreviation $\tau(\delta = d)$ for the hitting time $\tau(\{G : \delta(G) \geq d\})$ of a given graph process, where $\delta(G)$ denotes the minimal degree of G. Finally, we say that a random graph G satisfies some property *with high probability*, or *almost surely*, or that *almost every* graph process satisfies a property, if the probability for the corresponding event tends to 1 as the number of vertices tends to infinity.

Consider the beginning of the random graph process. It is easy to see that for every graph G, $i(G)$ is at most $\delta(G)$, the minimal degree of G (choose a set S consisting of a single vertex of degree $\delta(G)$). Hence, at the beginning of the graph process, $i(\widetilde{G}(0)) = 0 = \delta(\widetilde{G}(0))$, and this remains the case as long as there exists an isolated vertex in $\widetilde{G}(t)$. Next, consider the time where the minimal degree and maximal degree of the random graph process become more or less equal. At this point, we can examine random δ-regular graphs for intuition as to the behavior of the isoperimetric constant, in which case the results of [5] and [33] suggest that $i(\widetilde{G}(t))$ is roughly $\delta/2$. Hence, at some point along the random graph process, the behavior of the isoperimetric constant changes, and instead of being equal to δ it drifts towards $\delta/2$ (it is easy to confirm that the isoperimetric constant of the complete graph is $\lceil \frac{n}{2} \rceil$). The following results summarize the behavior of the isoperimetric constant of the random graph process (and, resulting from which, of the appropriate random graphs models):

In Section 9.2 we prove that, for almost every graph process, there is equality between the isoperimetric constant and the minimal degree, as long as the minimal degree is $o(\log n)$. In other words, we prove a hitting time result: the minimal degree increases by 1 exactly when the isoperimetric constant increases by 1 throughout the entire period in which $\delta = o(\log n)$.

Theorem 9.1.1. *Let $\ell = \ell(n)$ denote a function satisfying $\ell(n) = o(\log n)$. Almost every graph process $\widetilde{G}$ on n vertices satisfies $i(\widetilde{G}(t)) = \delta(\widetilde{G}(t))$ for every $t \in [0, \tau(\delta = \ell)]$. Furthermore, with high probability, for every such t, every set S which attains the minimum of $i(\widetilde{G}(t))$ is an independent set of vertices of degree $\delta(\widetilde{G}(t))$.*

In Section 9.3 we show that the $o(\log n)$ bound in Theorem 9.1.1 is essentially best possible. Indeed, during the period in which the minimal degree is $\Theta(\log n)$, $i(G)$ drifts towards $\frac{1}{2}\delta(G)$, as the next theorem demonstrates:

Theorem 9.1.2. *For every $0 < \varepsilon < \frac{1}{2}$ there exists a constant $C = C(\varepsilon) > 0$, such that the random graph $G \sim \mathcal{G}(n,p)$, where $p = C\frac{\log n}{n}$, almost surely satisfies:*

$$i(G) \leq \left(\frac{1}{2} + \varepsilon\right) \delta(G) = \Theta(\log n) \ .$$

Furthermore, with high probability, every set S of size $\lfloor \frac{n}{2} \rfloor$ satisfies: $\frac{|\partial S|}{|S|} < \left(\frac{1}{2} + \varepsilon\right) \delta(G)$.

An analogous statement holds for $\mathcal{G}(n, M)$ as well, where $M = Cn \log n$ for a sufficiently large $C = C(\varepsilon)$.

We note that throughout the chapter, all logarithms are natural.

9.2 The behavior of $i(G)$ when $\delta = o(\log n)$

9.2.1 Proof of Theorem 9.1.1

Since every graph G satisfies $i(G) \leq \delta(G)$, proving that, for every $d \leq \ell$, with high probability, at time $\tau(\delta = d)$ the isoperimetric constant of G is at least d, will prove the theorem. We show that for every $d = d(n) = o(\log n)$, the probability for this event is at least $1 - o(\frac{1}{\log n})$, and the theorem follows from the union bound on the events corresponding to all possible values of $d \leq \ell$, as we explain in the end of this section.

Recall that almost every graph process $\widetilde{G}$ satisfies $\delta(\widetilde{G}) \leq d - 1$ at time

$$m_d = \binom{n}{2} \frac{\log n + (d-1)\log\log n - \omega(n)}{n} \ ,$$

and $\delta(\widetilde{G}) \geq d$ at time

$$M_d = \binom{n}{2} \frac{\log n + (d-1)\log\log n + \omega(n)}{n} \ ,$$

where $d \geq 1$ is some fixed integer, the $\omega(n)$-term represents a function growing to infinity arbitrarily slowly while satisfying $\omega(n) \leq \log\log\log n$ (see, e.g., [34], Chapter 3). Hence, $\tau(\delta = d)$ is between m_d and M_d. Using the same methods described in [34], it is easy to extend this statement to every $d = d(n) = o(\log n)$, as the next proposition summarizes:

Proposition 9.2.1. *Let $\ell = \ell(n) = o(\log n)$. For every $1 \leq d \leq \ell$ define:*

$$r = r(n) = \frac{\log n}{d} \ .$$

Next, define the following threshold functions:

$$m_d = \binom{n}{2} \frac{\log n + (d-1)\log r - (2d + \omega(n))}{n} \ , \tag{9.2}$$

and:

$$M_d = \binom{n}{2}\frac{\log n + (d-1)\log r + (2d+\omega(n))}{n} \,, \tag{9.3}$$

where $\omega(n) \leq \log\log r$ *and* $\lim_{n\to\infty}\omega(n) = \infty$. *Then, almost every graph process* $\widetilde{G}$ *satisfies* $\delta(\widetilde{G}(m_d)) \leq d-1$ *and* $\delta(\widetilde{G}(M_d)) \geq d$ *for every* $1 \leq d \leq \ell$.

Notice that $r \leq \log n$, and that r tends to infinity as $n \to \infty$, hence these definitions coincide with the previous definitions of m_d and M_d for a fixed d, and it is left to verify them for $1 \ll d \ll \log n$. Proposition 9.2.1 follows from standard first moment and second moment considerations, and we postpone its proof to Section 9.2.2. Assume therefore, throughout the proof of Theorem 9.1.1, that the hitting time $\tau(\delta = d)$ is almost surely in the interval $(m_d, M_d]$ for every $1 \leq d \leq \ell$.

Consider a set $S \subset V$ of size $|S| \leq n/2$; we need to show that, with high probability, every such set satisfies $|\partial S| \geq \delta(\widetilde{G}(t))|S|$ at every time $t \leq \tau(\delta = \ell)$ in the random graph process. Clearly, at a given time $t = M$, the random variable $|\partial S|$ has a binomial distribution with parameters $\mathcal{B}\left(|S|(n-|S|), p\right)$, where $p = M/\binom{n}{2}$. When $|S|$ is sufficiently large (namely, larger than $n^{1/4}$), the result follows from standard large deviation bounds and bounds on the tail of the binomial distribution. However, these bounds are not tight enough for small values of $|S|$, which require a separate and more delicate treatment.

Throughout the rest of this section, fix $d = d(n) = o(\log n)$, and define m_d, M_d and r according to Proposition 9.2.1.

The following lemma shows that every small set S has a boundary of size at least $\delta(G)|S|$ almost surely:

Lemma 9.2.2. *With probability at least* $1 - o(n^{-1/5})$, *the random graph process* $\widetilde{G}$ *satisfies that every* $G \in \{\widetilde{G}(t) \;:\; m_d \leq t \leq \tau(\delta = d)\}$ *has the property* $|\partial S| \geq \delta(G)|S|$ *for every set* S *of size* $|S| \leq n^{1/4}$. *Furthermore, if such a set* S *satisfies* $|\partial S| = \delta(G)|S|$, *it is necessarily an independent set of vertices whose degrees are* $\delta(G)$.

Proof. Given a graph $G = (V, E)$, we call a set $S \subset V$ *bad* if it satisfies $|\partial S| < \delta(G)|S|$. The idea of the proof is as follows: we show that, with high probability, every induced subgraph on $k \leq n^{1/4}$ vertices has a low average degree. Since bad sets have a boundary of at most $\delta(G)|S|$, this implies that bad sets, as well as sets which are "almost" bad, must contain many vertices whose degrees are low in G. The result is derived from several properties of the set of all vertices of low degrees. We begin with defining this set of vertices and examining its properties:

Definition 9.1. *Let $G = (V, E)$. The set of vertices* $\textsc{Small}(G)$ *is defined to be:*

$$\textsc{Small} = \textsc{Small}(G) = \{v \in V \ : \ d(v) < 4(d+6)\} \ .$$

Claim 9.2.3. *With probability at least $1 - o(n^{-1/5})$, the random graph process $\widetilde{G}$ has the following property: for every $m_d \leq t \leq M_d$,* $\textsc{Small}$ *is an independent set, and every two vertices of* $\textsc{Small}$ *have no common neighbors in V.*

Proof. Notice that the set $\textsc{Small}$ changes along the random graph process, as vertices are removed from it once they reach a degree of $4(d+6)$. We show a slightly stronger result: if S_0 denotes $\textsc{Small}(\widetilde{G}(m_d))$, then S_0 satisfies the above properties almost surely for every $m_d \leq t \leq M_d$. Since $\textsc{Small}(\widetilde{G}(t)) \subseteq S_0$ for every $t \geq m_d$, this will imply the claim. In order to prove this result, we show that, with high probability, S_0 satisfies the above properties at time $t = m_d$, and that the addition of $M_d - m_d$ edges almost surely does not harm these properties of S_0.

Let $p = m_d/\binom{n}{2}$, and let $G_0 \sim \mathcal{G}(n, p)$. The same consideration will show that $\textsc{Small}$ satisfies the properties of the claim with the mentioned probability, both in $\mathcal{G}(n, p)$ and in $\mathcal{G}(n, m_d)$; for the sake of simplicity, we perform the calculations in the $\mathcal{G}(n, p)$ model, and note that they hold for the $\mathcal{G}(n, m_d)$ model as well. Indeed, the main tool in the proof is an upper bound on the probability that a given vertex would have a low degree (a degree of $L = o(n)$ when the edge probability is p), and the probabilities of the relevant events in $\mathcal{G}(n, m_d)$ are already upper bounded by the corresponding probabilities in $\mathcal{G}(n, p)$.

Both of the properties mentioned in the claim are immediate consequences of the next upper bound for the probability of the event $\{\mathcal{B}(n - L, p) \leq D\}$, where $4d \leq D \leq 30d$ and $L = o(n)$. We use the fact that, by this choice of parameters, $D = o\left((n - L)p\right)$, implying the following monotonicity of the binomial distribution:

$$\begin{aligned}\Pr[\mathcal{B}(n-L,p)\leq D] &\leq (D+1)\binom{n-L}{D}p^D(1-p)^{n-L-D}\\ &\leq (D+1)\left(\frac{\mathrm{e}pn}{D}\right)^D \mathrm{e}^{-p(1-o(1))n}\\ &\leq (30d+1)\left(\frac{(\mathrm{e}+o(1))\log n}{4d}\right)^{30d}\mathrm{e}^{-(1-o(1))\log n}\\ &\leq (30d+1)r^{30d}\mathrm{e}^{-(1-o(1))\log n}\\ &= \exp\left(O(1)+\log d+30d\log r-(1-o(1))\log n\right)\\ &= \exp\left(O(1)+\log d+30\log n\frac{\log r}{r}-(1-o(1))\log n\right)\\ &= \exp\left(-(1-o(1))\log n\right)=o(n^{-0.9})\ .\end{aligned}$$

Set $D=4(d+6)$, and let $A_{u,v}$ denote the event that the edge (u,v) belongs to the induced graph on SMALL, for a given pair of vertices $u,v\in V$. The following holds:

$$\Pr[A_{u,v}]=p\Pr[\mathcal{B}(n-2,p)<D-1]^2\leq\frac{(1+o(1))\log n}{n^{2.8}}=o(n^{-2.5})\ .$$

Thus, the probability that there exists such a pair of vertices is at most $\binom{n}{2}\Pr[A_{u,v}]=o(n^{-1/2})$, and SMALL$(G_0)$ is an independent set with probability $1-o(n^{-1/2})$. Next, let $A_{u,v,w}$ denote the event that $u,v\in$ SMALL(G_0) and w is a common neighbor of u and v, for some $u,v,w\in V$. Again, we get:

$$\begin{aligned}\Pr[A_{u,v,w}]=p^2\Big(&p\Pr[\mathcal{B}(n-3,p)<D-2]^2\\ &+(1-p)\Pr[\mathcal{B}(n-3,p)<D-1]^2\Big)\leq p^2n^{-1.8}=o(n^{-3.5})\ ,\end{aligned}$$

and therefore $\binom{n}{3}\Pr[A_{u,v,w}]=o(n^{-1/2})$.

We have shown that with probability at least $1-o(n^{-1/2})$, SMALL(G_0) satisfies the two properties of the claim, and by the same argument, $S_0=$ SMALL$(\widetilde{G}(m_d))$ satisfies the two properties of the claim with probability at least $1-o(n^{-1/2})$. We now give a rough upper bound on the size of S_0 using the above upper bound on $\mathcal{B}(n,p)$:

$$\mathbb{E}|S_0|\leq n\Pr[\mathcal{B}(n-1,p)<D]=o(n^{0.1})\ .$$

Hence, by Markov's inequality, $\Pr[|S_0| \geq n^{0.3}] \leq n^{-1/5}$. Altogether, we have shown that, with probability $1 - o(n^{-1/5})$, the set SMALL at time $t = m_d$ satisfies the requirements of the claim, and is of size at most $n^{0.3}$.

Assume that indeed $|S_0| \leq n^{0.3}$ and that the distance between every pair of vertices of S_0 is at least 3 at time m_d. We wish to show that this property is maintained throughout the period $t \in (m_d, M_d]$. Notice that the probability that an edge will be added between a given pair of vertices u, v in this period is

$$\hat{p} = (1 + o(1))\,(M_d - m_d)\,/\binom{n}{2} = (2 + o(1))\frac{2d + \omega(n)}{n}\ .$$

Hence, the probability that an internal edge is added to S_0 is at most:

$$\binom{|S_0|}{2}\hat{p} \leq \frac{n^{0.6}(1 + o(1))(2d + \omega(n))}{n} = o(n^{-1/5})\ .$$

Since the set of neighbors of S_0, $N(S_0)$, consists of at most $4(d + 6)|S_0|$ vertices, the probability that an edge is added between $N(S_0)$ and a vertex of S_0 is at most:

$$|N(S_0)||S_0|\hat{p} \leq \frac{n^{0.6}(2 + o(1))4(d + 6)(2d + \omega(n))}{n} = o(n^{-1/5})\ .$$

Finally, the probability that two edges are added between one vertex of $V \setminus S_0$ and two vertices of S_0 is at most:

$$n\binom{|S_0|}{2}\hat{p}^2 \leq \frac{n^{1.6}(2 + o(1))(2d + \omega(n))^2}{n^2} = o(n^{-1/5})\ .$$

Altogether, with probability $1 - o(n^{-1/5})$ the set S_0 maintains the property that the distance between each pair of its vertices is at least 3 in the period $m_d \leq t \leq M_d$. This completes the proof of the claim. ■

The following claim is crucial to the handling of small sets in G, showing that the average degree of the subgraph induced by a small set is small:

Claim 9.2.4. *With probability at least $1 - o(n^{-1/5})$, the random graph process $\widetilde{G}$ has the following property: for every $t \leq M_d$, every induced subgraph of $\widetilde{G}(t)$ on $k \leq n^{1/4}$ vertices contains at most $2k$ edges.*

Proof. Since this property is monotone with respect to the removal of edges, it is enough to prove the claim for $t = M_d$. Let $p = M_d/\binom{n}{2}$ and $G \sim \mathcal{G}(n,p)$. Fix $1 \leq k \leq n^{1/4}$; the probability that an induced subgraph H on k vertices has at least $2k$ edges is:

$$\begin{aligned}\Pr[|E(H)| \geq 2k] = \Pr[\mathcal{B}(\binom{k}{2}, p) \geq 2k] &\leq \binom{\binom{k}{2}}{2k} p^{2k} \\ &\leq (kp)^{2k} \leq \left(\frac{(1+o(1))\log n}{n^{3/4}}\right)^{2k} .\end{aligned}$$

Summing over all the subgraphs of size at most k, we obtain that the probability that such a subgraph exists is at most:

$$\begin{aligned}\sum_{k \leq n^{1/4}} \sum_{|H|=k} \Pr[|E(H)| \geq 2k] &\leq \sum_{k \leq n^{1/4}} \binom{n}{k} \left(\frac{(1+o(1)\log n}{n^{3/4}}\right)^{2k} \\ &\leq \sum_{k \leq n^{1/4}} \left(n^{-\frac{1}{2}+o(1)}\right)^{k} = o(n^{-1/5}) .\end{aligned}$$

Again, performing the same calculation in $\mathcal{G}(n, M_d)$ gives the same result: the probability that a specific set of $2k$ edges belongs to $\mathcal{G}(n, M_d)$ is $\binom{N-2k}{M_d-2k}/\binom{N}{M_d}$ (where $N = \binom{n}{2}$), which equals $((1+o(1))M_d/N)^{2k} = ((1+o(1))p)^{2k}$. ■

Equipped with Claim 9.2.3 and Claim 9.2.4, we are ready to prove Lemma 9.2.2.

Recall that a set S is bad iff $|\partial S| < \delta(G)|S|$. We call a bad set S *elementary* if it does not contain a smaller bad set, i.e., every $T \subset S$, $T \neq S$ is not bad. Clearly, in order to show that there are no bad sets of size at most $n^{1/4}$, it is enough to show that there are no elementary bad sets of such size. With high probability, every $G \in \{\widetilde{G}(t) : m_d \leq t \leq M_d\}$ satisfies both Claim 9.2.3 and Claim 9.2.4. Since $m_d < \tau(\delta = d) \leq M_d$, every graph $G = \widetilde{G}(t)$ in the interval $m_d \leq t \leq \tau(\delta = d)$ satisfies both claims, as well as $\delta(G) \leq d$. We claim that this implies the required result; to see this, consider a graph G which satisfies the above properties, and let $\delta = \delta(G)$. We first prove that there are no elementary bad sets of size at most $n^{1/4}$ in G:

Let S denote an elementary bad set S of size $k \leq n^{1/4}$. Notice that necessarily $k \geq 2$, since a single vertex has at least δ edges and hence cannot account for a bad set. By Claim 9.2.4, the induced graph H on S contains at most $2k$ edges. Since the boundary of S contains at most $\delta k - 1 \leq dk$ edges, this implies that $|S \cap \textsc{Small}| \geq \frac{3}{4}k$, otherwise the number of edges in H would satisfy:

$$|E(H)| = \frac{1}{2}\sum_{v \in S} d_H(v) \geq \frac{1}{2}\left(\frac{k}{4}4(d+6) - dk\right) \geq 3k \ ,$$

leading to a contradiction. Assume therefore that at most $k/4$ vertices in S do not belong to $\textsc{Small}$. We define $A = S \cap \textsc{Small}$, and $B = S \setminus A$. By Claim 9.2.3, A is an independent set, and furthermore, no two vertices of A have a common neighbor in B. Hence, each vertex of B is adjacent to at most one vertex of A, and if we denote by $A' \subseteq A$ the vertices of A, which are not adjacent to any vertex of S, the following holds:

$$|A'| \geq |A| - |B| \geq (\frac{3}{4} - \frac{1}{4})k = \frac{1}{2}k \ .$$

In particular, A' is nonempty; we claim that this contradicts the fact that S is elementary. Indeed, each vertex $v \in A'$ is not adjacent to any vertex in S, hence it contributes $d(v)$ edges to ∂S. Removing the vertex v would result in a nonempty ($k \geq 2$) strictly smaller subset T of S which satisfies:

$$|\partial T| = |\partial S| - d(v) \leq |\partial S| - \delta < \delta(|S| - 1) = \delta|T| \ ,$$

establishing a contradiction. We conclude that G does not contain bad sets of size at most $n^{1/4}$.

Next, consider a set S of size $|S| \leq n^{1/4}$ which satisfies $|\partial S| = \delta|S|$. If $|S| = 1$, obviously S consists of a single vertex of degree δ and we are done. Otherwise, repeating the above arguments for bad sets, we deduce that $|S \cap \textsc{Small}| \geq \frac{3}{4}|S|$ (this argument merely required that $|\partial S| \leq \delta|S|$) and that S contains a nonempty set A', whose vertices are not adjacent to any vertex of S. Consider a vertex $v \in A'$; this vertex contributes $d(v) \geq \delta$ edges to ∂S. However, $d(v)$ cannot be greater than δ, otherwise the set $S' = S \setminus \{v\}$ would satisfy $|\partial S'| < \delta|S'|$, contradicting the fact that there are no bad sets of size at most $n^{1/4}$ in G. Therefore, all the vertices of A' are of degree δ, and are not adjacent to any of the vertices of S. If we denote the remaining vertices by $S' = S \setminus A'$, S' satisfies $|\partial S'| = \delta|S| - \delta|A'| = \delta|S'|$, and, by induction, the result follows.

This completes the proof of Lemma 9.2.2. ■

The large sets are handled by the following lemma, which shows that even at time m_d (when the minimal degree is still at most $d-1$) these sets already have boundaries of size at least $d|S|+1$.

Lemma 9.2.5. *With probability at least $1-o(n^{-1/5})$, the graph $\widetilde{G}(m_d)$ satisfies $|\partial S| > d|S|$ for every set S of size $n^{1/4} \leq |S| \leq n/2$ (and hence $\widetilde{G}(t)$ has this property for every $t \geq m_d$ with probability at least $1-o(n^{-1/5})$).*

Proof. Define $p = m_d/\binom{n}{2}$. For the sake of simplicity, the calculations are performed in the $\mathcal{G}(n,p)$ model and we note that by the same considerations the results apply for the corresponding $\mathcal{G}(n,m_d)$ model as well. To show that, with probability $1-o(n^{-1/5})$, the random graph $G \sim \mathcal{G}(n,p)$ satisfies $|\partial S| > d|S|$ for sets S of the given size, argue as follows:

Fix a set $S \subset V$ of size k, $\frac{n}{\log n} \leq k \leq n/2$, and let A_S denote the event $\{|\partial S| \leq dk\}$. Let μ denote $\mathbb{E}|\partial S| = k(n-k)p$. By the Chernoff bound, $\Pr[|\partial S| < \mu - t] \leq \exp\left(-\frac{1}{2\mu}t^2\right)$. Therefore, setting $t = \mu - (dk+1)$, we get:

$$\begin{aligned}
&\Pr[A_S] = \Pr[|\partial S| < dk+1] \leq \exp\left(-\frac{1}{2}\left(1-\frac{d+\frac{1}{k}}{(n-k)p}\right)^2 k(n-k)p\right)\\
&\leq \exp\left(-\frac{1}{2}\left(1-\frac{(2+o(1))d}{\log n}\right)^2 k\left(\frac{1}{2}-o(1)\right)\log n\right)\\
&= \exp\left(-\frac{1-o(1)}{4}k\log n\right) .
\end{aligned}$$

Hence, the probability that there exists such a set S is at most:

$$\sum_{k=\frac{n}{\log n}}^{n/2} \binom{n}{k} \exp\left(-\frac{1-o(1)}{4} k \log n\right) \leq \sum_{k=\frac{n}{\log n}}^{n/2} \left(\mathrm{e}\frac{n}{k}\right)^k \exp\left(-\frac{1-o(1)}{4} k \log n\right)$$
$$\leq \sum_{k=\frac{n}{\log n}}^{n/2} \exp\left(k(\log\log n + 1) - \frac{1-o(1)}{4} k \log n\right)$$
$$\leq \sum_{k=\frac{n}{\log n}}^{n/2} \left(n^{-\frac{1}{4}+o(1)}\right)^k = o(n^{-1/5}) \ .$$

Let $S \subset V$ be a set of size $n^{1/4} \leq k \leq \frac{n}{\log n}$. Notice that:

$$(n-k)p = (1+o(1)) \log n \ , \qquad (9.4)$$

and hence, $dk < \mu$, and we can give the following upper bound on the probability that $|\partial S| \leq dk$:

$$\begin{aligned}
\Pr[|\partial S| \leq dk] &\leq (dk+1) \Pr[|\partial S| = dk] \\
&= (dk+1) \binom{k(n-k)}{dk} p^{dk} (1-p)^{k(n-k)-dk} \\
&\leq (dk+1) \left(\frac{\mathrm{e}k(n-k)p}{dk}\right)^{dk} \mathrm{e}^{-pk(n-k-d)} \\
&= (dk+1)(\mathrm{e}/d)^{dk} \left(p(n-k)\right)^{dk} \mathrm{e}^{-kpn+pk^2+pkd} \ .
\end{aligned}$$

We now use (9.4) and the facts that $pk \leq 1 + o(1)$ and $d = o(k)$, and obtain:

$$\begin{aligned}
\Pr[|\partial S| \leq dk] &\leq O(1) dk (\mathrm{e}/d)^{dk} \frac{(\log n)^{dk} \mathrm{e}^{(2d+\omega(n)+1+o(1))k+d}}{n^k r^{k(d-1)}} \\
&\leq \left(\frac{\mathrm{e}^{\omega(n)+2d+O(1)} \log n}{n}\right)^k \ .
\end{aligned}$$

Summing over all sets S of size k, we get:

$$\sum_{|S|=k} \Pr[|\partial S| \leq dk] \leq \left(\frac{\mathrm{e}n}{k}\right)^k \left(\frac{\mathrm{e}^{\omega(n)+2d+O(1)} \log n}{n}\right)^k = \left(\frac{\mathrm{e}^{\omega(n)+2d+O(1)} \log n}{k}\right)^k$$
$$\leq \left(\frac{O(1) n^{2/r} \log r \log n}{n^{1/4}}\right)^k = \left(n^{-\frac{1}{4}+o(1)}\right)^k \ .$$

Thus:

$$\sum_{n^{1/4}\leq|S|\leq\frac{n}{\log n}} \Pr[|\partial S|\leq d|S|]\leq \sum_{k\geq n^{1/4}} \left(n^{-\frac{1}{4}+o(1)}\right)^k = o(n^{-1/5})\ .$$

■

Since $\widetilde{G}$ satisfies the properties of both Lemma 9.2.2 and Lemma 9.2.5 for a given $d \leq \ell = o(\log n)$ with probability at least $1 - o(n^{-1/5})$, the union bound over all possible values of d implies that these properties are satisfied almost surely for every $d \leq \ell$. Theorem 9.1.1 follows directly: to see this, assume that indeed a random graph process $\widetilde{G}$ satisfies the mentioned properties for every $d \leq \ell$, and consider some $d \leq \ell$. By the properties of Lemma 9.2.2, in the period $t \in [m_d, \tau(\delta = d)]$ every set of size $k \leq n^{1/4}$ has at least δk edges in its corresponding cut, and if there are precisely δk edges in the cut, then S is an independent set of vertices of degree δ. In particular, at time $t = \tau(\delta = d)$, every set S of at most $n^{1/4}$ vertices has a ratio $\frac{|\partial S|}{|S|}$ of at least d, and a ratio of precisely d implies that S is an independent set of vertices of degree d. By monotonicity, this is true for every $t \in [\tau(\delta = d), \tau(\delta = d+1))$. Next, by the properties of Lemma 9.2.5, every set of size $k \geq n^{1/4}$ has at least $dk + 1$ edges in its corresponding cut at time $t = m_d$. In particular, for every $t \in [\tau(\delta = d), \tau(\delta = d+1))$, every set S, larger than $n^{1/4}$ vertices, has a ratio $\frac{|\partial S|}{|S|}$ strictly larger than d. These two facts imply the theorem. ■

9.2.2 Proof of Proposition 9.2.1

A standard first moment consideration shows that indeed, with high probability, $\delta(\mathcal{G}(n, M_d)) \geq d$ for every $d \leq \ell$. We perform the calculations in the $\mathcal{G}(n,p)$ model and note that the same applies to $\mathcal{G}(n, M_d)$.

For each $v \in V(G)$, let A_v and B_v denote the events $\{d(v) = d-1\}$ and $\{d(v) \leq d-1\}$ respectively, and set $Y_d = |\{v : d(v) = d-1\}|$ and $Z_d = |\{v : d(v) \leq d-1\}|$. Recall that $d = o(\log n)$, and furthermore, we may assume that d tends to infinity as $n \to \infty$, since m_d and M_d coincide with the well known threshold functions for constant values of d. Choosing

$p = M_d/\binom{n}{2}$, the following holds:

$$\begin{aligned}\Pr[A_v] &= \binom{n-1}{d-1}p^{d-1}(1-p)^{n-d}\\ &\leq \left(\frac{(1+o(1))\mathrm{e}\log n}{d}\right)^{d-1}\mathrm{e}^{-(1-\frac{d}{n})(\log n+(d-1)\log r+2d+\omega(n))}\\ &\leq \frac{1}{n^{1-d/n}}\left(\frac{(1+o(1))\mathrm{e}r}{r^{1-d/n}}\right)^{d-1}\mathrm{e}^{-(1-o(1))(2d+\omega(n))}\\ &= \frac{n^{d/n}}{n}\left((1+o(1))\mathrm{e}r^{\frac{\log n}{rn}}\right)^{d-1}\mathrm{e}^{-(1-o(1))(2d+\omega(n))} \leq \frac{1}{n}\mathrm{e}^{-(1-o(1))(d+\omega(n))} \; .\end{aligned} \tag{9.5}$$

Since $d \leq (n-1)p$, we have:

$$\Pr[B_v] \leq d\Pr[A_v] \leq \frac{1}{n}\mathrm{e}^{-(1-o(1))(d+\omega(n))} \; .$$

Hence,

$$\mathbb{E}Z_d \leq \mathrm{e}^{-(1-o(1))(d+\omega(n))} \; ,$$

and summing over every $d \leq \ell$ we obtain:

$$\sum_{d\leq\ell}\Pr[Z_d > 0] \leq \mathrm{e}^{-(1-o(1))\omega(n)}\sum_{d\leq\ell}\mathrm{e}^{-(1-o(1))d} = o(1) \; .$$

A second moment argument proves that almost surely $\delta(\mathcal{G}(n,p)) \leq d-1$ for every $d \leq \ell$. To see this, argue as follows (again, calculations are performed in the $\mathcal{G}(n,p)$ model): following the same definitions, only this time with $p = m_d/\binom{n}{2}$, apply the bound $\binom{a}{b} \geq \left(\frac{a}{b}\right)^b$ and the well known bound $1-x \geq \mathrm{e}^{-x/(1-x)}$ for $0 \leq x < 1$, to obtain:

$$\Pr[A_v] = \binom{n-1}{d-1}p^{d-1}(1-p)^{n-d} \geq$$

$$\geq \left(\frac{(1+o(1))\log n}{d}\right)^{d-1}\mathrm{e}^{(-\log n-(d-1)\log r+2d+\omega(n))/(1-p)} \geq \frac{1}{n}\Omega(\mathrm{e}^{d+\omega(n)}) \; ,$$

where in the last inequality we omitted the the $1/(1-p)$ factor in the exponent, since, for instance, $n^{1-\frac{1}{1-p}} = n^{\frac{-p}{1-p}} \geq n^{-O(1)\frac{\log n}{n}} = \mathrm{e}^{o(1)}$. Therefore:

$$\mathbb{E}Y_d = \Omega(\mathrm{e}^{d+\omega(n)}) \; .$$

Take $u \in V(G)$ with $u \neq v$; denoting $\mathcal{P}^K_L = \Pr[\mathcal{B}(K,p) = L]$, the following holds:

$$\begin{aligned}\mathrm{Cov}(A_u, A_v) &= \Pr[A_u \wedge A_v] - \Pr[A_u]\Pr[A_v] \\ &= p(\mathcal{P}^{n-2}_{d-2})^2 + (1-p)(\mathcal{P}^{n-2}_{d-1})^2 - (\mathcal{P}^{n-1}_{d-1})^2 \ .\end{aligned}$$

Since $\mathcal{P}^{n-1}_{d-1} = p\mathcal{P}^{n-2}_{d-2} + (1-p)\mathcal{P}^{n-2}_{d-1}$, we get:

$$\begin{aligned}\mathrm{Cov}(A_u, A_v) &= p(1-p)(\mathcal{P}^{n-2}_{d-2})^2 + (1-p)p(\mathcal{P}^{n-2}_{d-1})^2 - 2p(1-p)\mathcal{P}^{n-2}_{d-2}\mathcal{P}^{n-2}_{d-2} \\ &= p(1-p)(\mathcal{P}^{n-2}_{d-1} - \mathcal{P}^{n-2}_{d-2})^2 \leq p(\mathcal{P}^{n-2}_{d-1})^2 \ .\end{aligned}$$

Notice that $\mathcal{P}^{n-2}_{d-1}$ corresponds to the event A_v for a graph on $n-1$ vertices, and hence a similar calculation to the one in (9.5) shows that $\mathcal{P}^{n-2}_{d-1} = O(\exp(3d+\omega(n))/n)$. Altogether we get:

$$\mathrm{Cov}(A_u, A_v) \leq O(1)p\frac{\mathrm{e}^{6d+2\omega(n)}}{n^2} \leq O(1)\mathbb{E}Y_d\frac{\mathrm{e}^{5d+\omega(n)}\log n}{n^3} = o(n^{-2})\mathbb{E}Y_d \ ,$$

which gives the following upper bound on the variance of Y_d:

$$\mathrm{Var}(Y_d) \leq \mathbb{E}Y_d + \sum_{u \neq v}\mathrm{Cov}(A_u, A_v) \leq \mathbb{E}Y_d + n^2 o(n^{-2})\mathbb{E}Y_d = (1+o(1))\mathbb{E}Y_d \ .$$

Applying Chebyshev's inequality gives:

$$\Pr[Y_d = 0] \leq \frac{\mathrm{Var}(Y_d)}{(\mathbb{E}Y_d)^2} \leq \frac{1+o(1)}{\mathbb{E}Y_d} \leq O(\mathrm{e}^{-d-\omega(n)}) \ ,$$

and summing over every $d \leq \ell$ we obtain:

$$\sum_{d \leq \ell}\Pr[Y_d = 0] \leq O(1)\mathrm{e}^{-\omega(n)}\sum_{d \leq \ell}\mathrm{e}^{-d} = o(1) \ ,$$

as required. ■

9.3 The behavior of $i(G)$ when $\delta = \Omega(\log n)$

Proof of Theorem 9.1.2: A bisection of a graph G on n vertices is a partition of the vertices into two disjoint sets (S,T), where $|S| = \lfloor \frac{n}{2} \rfloor$ and

$T = V \setminus S$. Fix $\varepsilon_1 > 0$; we first prove that, with high probability, every bisection $(S, V \setminus S)$ of $\mathcal{G}(n,p)$ has strictly less than $\left(\frac{1}{2} + \varepsilon_1\right) np|S|$ edges in the cut it defines, provided that $\lim_{n\to\infty} np = \infty$. We omit the floor and ceiling signs to simplify the presentation of the proof.

Let S be an arbitrary set of $n/2$ vertices. The number of edges in the boundary of S has a binomial distribution with parameters $\mathcal{B}(n^2/4, p)$, hence (by our assumption on p) its expected value μ tends to infinity faster than n. By the Chernoff bound, $\Pr[|\partial S| \geq (1+t)\mu] \leq \exp(-\mu t^2/4)$ provided that $t < 2\mathrm{e} - 1$, thus we get:

$$\Pr[|\partial S| \geq \left(\frac{1}{2} + \varepsilon_1\right) np|S|] = \Pr[|\partial S| \geq (1 + 2\varepsilon_1)\,\mu] \leq \exp(-\Omega(\mu)) \ .$$

Since this probability is $o(2^{-n})$, the expected number of bisections, in which the corresponding cuts contain at least $\left(\frac{1}{2} + \varepsilon_1\right) np|S|$ edges, is $o(1)$.

Next, fix $\varepsilon_2 > 0$. We claim that the minimal degree of $\mathcal{G}(n,p)$, where $p = C\frac{\log n}{n}$ and $C = C(\varepsilon_2)$ is sufficiently large, is at least $(1-\varepsilon_2)np$. Applying the Chernoff bound on the binomial distribution representing the degree of a given vertex v gives:

$$\begin{aligned}\Pr[d(v) \leq (1-\varepsilon_2)np] &= \Pr[d(v) \leq (1 - \varepsilon_2 + o(1))\mathbb{E}d(v)] \\ &\exp\left(-C\frac{\varepsilon_2^2}{2}(1-o(1))\log n\right) \ ,\end{aligned}$$

and for $C > \frac{2}{\varepsilon_2^2}$ this probability is smaller than $\frac{1}{n}$.

Altogether, for a sufficiently large C, the following holds with high probability: every bisection $(S, V \setminus S)$ satisfies:

$$\frac{|\partial S|}{|S|} < \frac{\frac{1}{2} + \varepsilon_1}{1 - \varepsilon_2}\delta(G) = \left(\frac{1}{2} + \varepsilon\right)\delta(G) \ ,$$

where $\varepsilon = \frac{\varepsilon_1 + \varepsilon_2/2}{1-\varepsilon_2}$. ■

Remark 9.3.1: We note that the above argument gives a crude estimate on the value of $C = C(\varepsilon)$. Since the first claim, concerning the behavior of bisections, holds for every value of C, we are left with determining when typically the minimal degree of G becomes sufficiently close to the average degree. This threshold can be easily computed, following arguments similar to the ones in the proof of Proposition 9.2.1; the following value of $C(\varepsilon)$ is sufficient for the properties of the theorem to hold with high probability:

$$C > \frac{1+2\varepsilon}{2\varepsilon - \log(1+2\varepsilon)} \ .$$

Remark 9.3.2: Theorem 9.1.2 provides an upper bound on $i(G)$, which is almost surely arbitrarily close to $\frac{\delta}{2}$ while the graph satisfies $\delta = \Theta(\log n)$. We note that the arguments of Theorem 9.1.1 can be repeated (in a simpler manner) to show that with high probability $i(\widetilde{G}(t)) \geq \delta/2$ for every t, and hence the bound in Theorem 9.1.2 is tight.

9.4 Concluding remarks

We have shown that there is a phase transition when the minimal degree changes from $o(\log n)$ to $\Omega(\log n)$; it would be interesting to give a more accurate description of this phase transition. Theorem 9.1.1 treats $\delta(G) = o(\log n)$, and Theorem 9.1.2 shows that, almost surely, $i(G) < \delta(G)$ once $p = C \log n/n$, where $C > 2/(1-\log 2) \approx 6.52$, in which case $\delta(G) > (C/2)\log n$. Hence we are left with the range in which $\delta(G) = c\log n$, where

$$0 < c \leq 1/(1-\log 2) \approx 3.26 \ .$$

It seems plausible that in this range $i(G) = \delta(G)$, i.e., that the isoperimetric constant is determined either by the typical minimal degree, or by the typical size of a bisection.

The vertex version of the isoperimetric constant (minimizing the ratio $|\delta S|/|S|$, where $\delta S \subset V \setminus S$ is the vertex neighborhood of S) is less natural, since the minimum has to be defined on all nonempty sets of size at most $n/(K+\varepsilon)$ if we wish to allow the constant to reach the value K. Nevertheless, the methods used to prove Theorem 9.1.1 can prove similar results for the vertex case, at least as long as the minimum degree is constant. Indeed, in that case, the probability for two vertices to have a common neighbor is small enough not to have an effect on the results.

Finally, it is interesting to consider the isoperimetric constant of certain subgraphs along the random graph process. To demonstrate this, we consider the period of $\widetilde{G}$ in which the minimal degree is 0, i.e., $t \leq \tau(\delta = 1)$. The existence of isolated vertices in $\widetilde{G}(t)$ implies that $i(\widetilde{G}(t)) = 0$, however we may consider a connected component of $G(t)$ and analyze its isoperimetric constant. For instance, the largest component at time t, $\mathcal{C}_1(t)$, after a short while (say, at $t = cn$ for some $c > 0$) satisfies $i(\mathcal{C}_1(t)) < \varepsilon$ for every $\varepsilon > 0$. In general, an easy calculation shows that for such values of t, with high probability there exist small sets which have an edge boundary smaller than their size. For instance, when $p = c/n$ for some $c < 1$, $\mathcal{G}(n,p)$ almost surely satisfies that all connected components are of size $O(\log n)$, hence each component $\mathcal{C}$ has a ratio $\frac{|\partial\mathcal{C}|}{|\mathcal{C}|}$ of 0. Furthermore, if we take $p = C/n$ for some $C > 1$, and consider the giant component H (recall that for this value of p, almost surely there is a single component of size $\Theta(n)$, and all other components are of size $O(\log n)$), $i(H) < \varepsilon$ for every $\varepsilon > 0$. One way to see this, is to consider a collection of arbitrarily long paths, each of which connects to the giant component at precisely one end.

Bibliography

[1] H.L. Abbott, A note on Ramsey's theorem, Canad. Math. Bull. 15 (1972), pp. 9-10.

[2] R. Ahlswede, N. Cai, and Z. Zhang, A general 4-words inequality with consequences for 2-way communication complexity, Adv. in Appl. Math. 10 (1989), 75-94.

[3] N. Alon, Explicit Ramsey graphs and orthonormal labelings, The Electronic Journal of Combinatorics, 1 (1994), R12, 8pp.

[4] N. Alon, Graph Powers, in: Contemporary Combinatorics, (B. Bolloba's, ed.), Bolyai Society Mathematical Studies, Springer 2002, pp. 11-28.

[5] N. Alon, On the edge-expansion of graphs, Combinatorics, Probability and Computing 6 (1997), 145-152.

[6] N. Alon, Problems and results in extremal combinatorics- I, Discrete Math. 273 (2003), 31-53.

[7] N. Alon, Probabilistic methods in extremal finite set theory, Extremal Problems for Finite Sets, (P. Frankl, Z. Füredi, G.O.H. Katona and D. Miklós Eds.), Bolyai Society Mathematical Studies, 3, Visegrád, Hungary, 1991, 39-57.

[8] N. Alon, The Shannon capacity of a union, Combinatorica 18 (1998), 301-310.

[9] N. Alon, I. Dinur, E. Friedgut and B. Sudakov, Graph products, fourier analysis and spectral techniques, Geometric and Functional Analysis 14 (2004), 913-940.

[10] N. Alon and E. Lubetzky, Codes and Xor graph products, Combinatorica 27 (2007), 13-33.

[11] N. Alon and E. Lubetzky, Graph powers, Delsarte, Hoffman, Ramsey and Shannon, SIAM J. Discrete Math 21 (2007), 329-348.

[12] N. Alon and E. Lubetzky, Independent sets in tensor graph powers, J. Graph Theory 54 (2007), 73-87.

[13] N. Alon and E. Lubetzky, Privileged users in zero-error transmission over a noisy channel, Combinatorica, to appear.

[14] N. Alon and E. Lubetzky, The Shannon capacity of a graph and the independence numbers of its powers, IEEE Transactions on Information Theory 52 (2006), 2172-2176.

[15] N. Alon and E. Lubetzky, Uniformly cross intersecting families, to appear.

[16] N. Alon, E. Lubetzky and U. Stav, The broadcast rate of a graph, to appear.

[17] N. Alon and V.D. Milman, λ_1, isoperimetric inequalities for graphs and superconcentrators, J. Combinatorial Theory, Ser. B 38 (1985), 73-88.

[18] N. Alon and A. Orlitsky, Repeated communication and Ramsey graphs, IEEE Transactions on Information Theory 41 (1995), 1276-1289.

[19] N. Alon and J.H. Spencer, *The Probabilistic Method*, Second Edition, Wiley, New York, 2000.

[20] L. Babai and P. Frankl, *Linear Algebra Methods in Combinatorics*, Preliminary Version 2. Dept. of Computer Science, The Univesity of Chicago, 1992.

[21] B. Barak, A. Rao, R. Shaltiel and A. Wigderson, 2-source dispersers for sub-polynomial entropy and ramsey graphs beating the Frankl-Wilson construction, Proc. of 38th STOC (2006), pp. 671-680.

[22] D.A.M. Barrington, R. Beigel and S. Rudich, Representing boolean functions as polynomials modulo composite numbers, Computational Complexity 4 (1994), 367-382. An extended abstract appeared in Proc. of the 24th Annual ACM symposium on Theory of computing (STOC 1992), pp. 462-467.

[23] Z. Bar-Yossef, Y. Birk, T.S. Jayram and T. Kol, Index coding with side information, Proc. of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), pp. 197-206.

[24] I. Benjamini, S. Haber, M. Krivelevich and E. Lubetzky, The isoperimetric constant of the random graph process, Random Structures and Algorithms 32 (2008), 101-114.

[25] C. Berge and M. Simonovits, The coloring numbers of direct product of two hypergraphs, In C. Berge and D. Ray-Chaudhuri, editors, Hypergraph Seminar, Lecture Notes on Mathematics, # 411. Springer Verlag, 1974.

[26] C. Berge and J.L. Ramírez Alfonsín, Origins and genesis, in: J.L. Rami'rez-Alfonsi'n and B.A. Reed (Eds.), *Perfect Graphs*, Wiley, 2001, pp. 1–12.

[27] Y. Birk and T. Kol, Coding-on-demand by an informed source (ISCOD) for efficient broadcast of different supplemental data to caching clients, IEEE Transactions on Information Theory 52 (2006), 2825-2830. An earlier version appeared in INFOCOM '98.

[28] T. Bohman, A limit theorem for the Shannon capacities of odd cycles. I., Proc. Amer. Math. Soc. 131 (2003), no. 11, 3559–3569 (electronic).

[29] T. Bohman, Private communication.

[30] T. Bohman and R. Holzman, A nontrivial lower bound on the Shannon capacities of the complements of odd cycles. IEEE Trans. Inform. Theory 49 (2003), no. 3, 721–722.

[31] R.C. Bollinger and C.L. Burchard, Lucas's Theorem and some related results for extended Pascal triangles, The American Mathematical Monthly 97 (1990), pp. 198-204.

[32] B. Bollobás, The independence ratio of regular graphs, Proc. Amer. Math. Soc. 83 (1981), 433-436.

[33] B. Bollobás, The isoperimetric number of random regular graphs, Europ. J. Combinatorics 9 (1988), 241-244.

[34] B. Bollobás, *Random Graphs*, Second Edition, Cambridge University Press, Cambridge, 2001.

[35] B. Bollobás and I. Leader, An isoperimetric inequality on the discrete torus, SIAM J. Disc. Math. 3 (1990) 32-37.

[36] B. Bollobás and I. Leader, Edge-isoperimetric inequalities in the grid, Combinatorica 11(1991) 299-314.

[37] J.I. Brown, R.J. Nowakowski and D. Rall, The ultimate categorical independence ratio of a graph, SIAM J. Discrete Math. 9 (1996), 290-300.

[38] P. Buser, Cubic graphs and the first eigenvalue of a Riemann surface, Mathematische Zeitschrift, 162 (1978), 87-99.

[39] M. Chudnovsky, N. Robertson, P.D. Seymour, and R. Thomas, The strong perfect graph theorem, Ann. Math. 164 (2006), 51-229.

[40] F.R.K. Chung, Laplacians of graphs and Cheeger's inequalities, in: Proc. Int. Conf. "Combinatorics, Paul Erdős is Eighty", Keszthely (Hungary), 1993, 2, 116.

[41] F.R.K. Chung, *Spectral Graph Theory*, American Mathematical Society, no. 92 in the Regional Conference Series in Mathematics, Providence, RI, 1997.

[42] F.R.K. Chung and P. Tetali, Isoperimetric inequalities for cartesian products of graphs, Combinatorics, Probability and Computing 7 (1998), 141-148.

[43] P. Delsarte, Bounds for unrestricted codes by linear programming, Philips Res. Rep. 27 (1972), 272-289.

[44] P. Delsarte, An algebraic approach to the association schemes of coding theory, Philips Res. Rep. Suppl. 10 (1973), 1-97.

[45] I. Dinur, E. Friedgut and O. Regev, Independent sets in graph powers are almost contained in juntas, GAFA, to appear.

[46] P. Erdős, On a lemma of Littlewood and Offord, Bull. Amer. Math. Soc. (2nd ser.) 51 (1945), 898-902.

[47] P. Erdős, On the number of complete subgraphs contained in certain graphs, Publ. Math. Inst. Hung. Acad. Sci., VII, ser. A 3 (1962), 459-464.

[48] P. Erdős, Problems and results in graph theory and combinatorial analysis, Proc. of the Fifth British Comb. Conf. 1975 Aberdeen, 169-192. Congressus Numerantium, No. XV, Utilitas Math., Winnipeg, Man., 1976.

[49] P. Erdős, Some remarks on the theory of graphs, Bull. AMS 53 (1947), 292–294.

[50] P. Erdős, C. Ko and R. Rado, Intersection theorems for systems of finite sets, Quart. J. Math. Oxford Ser. 2, 12 (1961) 313-320.

[51] U. Feige, Randomized graph products, chromatic numbers, and the Lovász ϑ-function, Proc. of the 27th ACM Symposium on Theory of Computing (STOC 1995), pp. 635-640.

[52] F. Franek and V. Rödl, 2-colorings of complete graphs with a small number of monochromatic K_4 subgraphs, Discrete Math. 114 (1993), 199-203.

[53] P. Frankl, Extremal set systems, in: R.L. Graham, M. Grötschel, L. Lovász (Eds.), *Handbook of Combinatorics*, Vol. 1, 2, 1293-1329, Elsevier, Amsterdam, 1995.

[54] P. Frankl and Z. Füredi, Forbidding just one intersection, J. Combin. Theory Ser. A 39 (1985), no. 2, 160-176.

[55] P. Frankl and V. Rödl, Forbidden intersections, Trans. Amer. Math. Soc. 300 (1987), 259-286.

[56] P. Frankl and R. Wilson, Intersection theorems with geometric consequences, Combinatorica 1 (1981), 357-368.

[57] J. Friedman, On the second eigenvalue and random walks in random d-regular graphs, Combinatorica 11 (1991), 331-362.

[58] D.R. Fulkerson, Blocking and Anti-Blocking Pairs of Polyhedra, Math. Program. 1 (1971), 168-194.

[59] Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices, Combinatorica 1 (1981), 233-241.

[60] C. Godsil and G. Royle, Algebraic Graph Theory, volume 207 of Graduate Text in Mathematics, Springer, New York, 2001.

[61] D. Greenwell and L. Lovász, Applications of product colorings, Acta Math. Acad. Sci. Hungar. 25 (3-4) (1974) pages 335-340.

[62] V. Grolmusz, Low rank co-diagonal matrices and ramsey graphs, Electr. J. Comb 7 (2000).

[63] W. Haemers, An upper bound for the Shannon capacity of a graph, Colloquia Mathematica Societatis János Bolyai, 25: Algebraic Methods in Graph Theory, Szeged (Hungary), 1978, 267-272.

[64] W. Haemers, On some problems of Lovász concerning the Shannon capacity of a graph, IEEE Transactions on Information Theory, 25(2) (1979), 231–232.

[65] M. Hall, *Combinatorial Theory*, Second Edition, Wiley, New York, 1986.

[66] S.H. Hedetniemi, Homomorphisms of graphs and automata, University of Michigan Technical Report 03105-44-T, 1966.

[67] A.J. Hoffman, On eigenvalues and colorings of graphs, B. Harris Ed., Graph Theory and its Applications, Academic, New York and London, 1970, 79-91.

[68] C. Houdré and P. Tetali, Isoperimetric constants for product Markov chains and graph products, Combinatorica Vol. 24 (2004), 359-388.

[69] H. Iwaniec and J. Pintz, Primes in short intervals, Monatsh. Math. 98 (1984), 115-143.

[70] G. Katona, Intersection theorems for systems of finite sets, Acta Math. Acad. Sci. Hungar 15 (1964), 329-337.

[71] P. Keevash and B. Sudakov, On a restricted cross-intersection problem, J. Combinatorial Theory Ser. A, 113 (2006), 1536-1542.

[72] D.E. Knuth, The Sandwich Theorem, Electronic J. Combinatorics 1 (1994), 1-48.

[73] M. Krivelevich, B. Sudakov, V. Vu and N. Wormald, Random regular graphs of high degree, Random Structures and Algorithms 18 (2001), 346-363.

[74] B. Larose and C. Tardif, Hedetniemi's conjecture and the retracts of a product of graphs, Combinatorica 20 (2000), 531-544.

[75] N. Linial and U. Vazirani, Graph products and chromatic numbers, Proc. of the 30th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1989), pp. 124-128.

[76] J.H. van Lint and R.M. Wilson, *A Course in Combinatorics*, Second Edition, Cambridge University Press, Cambridge, 2001.

[77] S. Litsyn, New upper bounds on error exponents, IEEE Transactions on Information Theory 45 (1999), no. 2, 385-398.

[78] J. Littlewood and C. Offord, On the number of real roots of a random algebraic equation III, Mat. Sbornik 12 (1943), 277-285.

[79] L. Lovász, Normal hypergraphs and the perfect graph conjecture, Discrete Math. 2 (1972), 253-267.

[80] L. Lovász, On the ratio of optimal integral and fractional covers, Discrete Math. 13 (1975), no. 4, 383-390.

[81] L. Lovász, On the Shannon capacity of a graph, IEEE Transactions on Information Theory 25 (1979), 1-7.

[82] L. Lovász and M.D. Plummer, *Matching Theory*, North-Holland, Amsterdam, 1986.

[83] E. Lubetzky and U. Stav, Non-linear index coding outperforming the linear optimum, Proc. of the 48th IEEE FOCS (2007), 161-167.

[84] D. Lubell, A short proof of Sperner's theorem, Journal of Combinatorial Theory 1 (1966), 299.

[85] F.J. MacWilliams and N.J.A. Sloane, The Theory of Error-Correcting Codes, North-Holland, Amsterdam, 1977.

[86] R.J. McEliece and E.C. Posner, Hide and seek, data storage, and entropy, The Annals of Mathematical Statistics, 42(5):1706-1716, 1971.

[87] B. Mohar, Isoperimetric numbers of graphs, Journal of Combinatorial Theory, Series B, 47 (1989), 274-291.

[88] E. Mossel, R. O'Donnell, and K. Oleszkiewicz, Noise stability of functions with low influences: invariance and optimality. In Proc. 46th FOCS (2005).

[89] M. Naor, Constructing Ramsey graphs from small probability spaces, IBM Research Report RJ 8810 (1992).

[90] O. Ore, *Theory of graphs*, AMS Colloquium Publications, Vol. 38, Providence, RI. 1962.

[91] D.K. Ray-Chaudhuri and R.M. Wilson. On t-designs, Osaka J. Math., 12 (1975), 737-744.

[92] H. Robbins, A remark on Stirling's formula, Amer. Math. Monthly 62, (1955), 26-29.

[93] M. Rosenfeld, On a Problem of C. E. Shannon in Graph Theory, Proceedings of the American Mathematical Society 18(2) (1967), 315-319.

[94] A. Schrijver, A comparison of the Delsarte and Lovász bounds, IEEE Trans. Inform. Theory, 25(4):425-429, 1979.

[95] J. Sgall, Bounds on pairs of families with restricted intersections, Combinatorica 19 (1999), 555-566.

[96] E. Shamir and E. Upfal, On factors in random graphs, Israel J. Math. 39 (1981), 296-302.

[97] C.E. Shannon, The zero-error capacity of a noisy channel, IRE Transactions on Information Theory, 2(3):8-19, 1956.

[98] E. Sperner, Ein Satz über Untermengen einer endlichen Menge, Math. Z. 27, 544-548, 1928.

[99] G. Szegő, *Orthogonal Polynomials*, 4th Edition, AMS Colloquium Publications, Vol. 23, Providence, RI, 1975.

[100] C. Tardif, The fractional chromatic number of the categorical product of graphs, Combinatorica 25 (2005), 625-632.

[101] C. Tardif, The chromatic number of the product of two graphs is at least half the minimum of the fractional chromatic numbers of the factors, Comment. Math. Univ. Carolin. 42 (2001) 353-355.

[102] A. Thomason, Graph products and monochromatic multiplicities, Combinatorica 17 (1997), 125-134.

[103] A. Thomason, A disproof of a conjecture of Erdős in Ramsey theory, J. London Math. Soc. 39 (1989), 246-255.

[104] H.N. Ward, Divisible codes, Arch. Math. (Basel) 36 (1981), no. 6, 485-494.

[105] H.N. Ward, Divisible Codes - A Survey, Serdica Math. J. 27 (2001), 263-278.

[106] H.S. Witsenhausen, The zero-error side information problem and chromatic numbers, IEEE Transactions on Information Theory, 22(5) (1976), 592-593.

[107] N.C. Wormald, Differential equations for random processes and random graphs, Annals of Applied Probability 5 (1995), 1217-1235.

[108] X. Zhu, The fractional chromatic number of the direct product of graphs, Glasgow Mathematical J. 44 (2002), 103-115.

[109] X. Zhu, A survey on Hedetniemi's conjecture, Taiwanese Journal of Mathematics 2 (1998), 1-24.

[110] X. Zhu, Star chromatic number and products of graphs, J. Graph Theory 16 (1992), 557-569.