ABSTRACT

Title of dissertation: BOUNDS ON THE SIZE OF CODES

Punarbasu Purkayastha
Doctor of Philosophy, 2010

Dissertation directed by: Professor Alexander Barg
Department of Electrical and Computer Engineering
and The Institute for Systems Research

In this dissertation we determine new bounds and properties of codes in three different finite metric spaces, namely the ordered Hamming space, the binary Hamming space, and the Johnson space.

The ordered Hamming space is a generalization of the Hamming space that arises in several different problems of coding theory and numerical integration. Structural properties of this space are well described in the framework of Delsarte's theory of association schemes. Relying on this theory, we perform a detailed study of polynomials related to the ordered Hamming space and derive new asymptotic bounds on the size of codes in this space which improve upon the estimates known earlier.

A related project concerns linear codes in the ordered Hamming space. We define and analyze a class of near-optimal codes, called near-Maximum Distance Separable codes. We determine the weight distribution and provide constructions of such codes. Codes in the ordered Hamming space are dual to a certain type of point distributions in the unit cube used in numerical integration. We show that near-Maximum Distance Separable codes are equivalently represented as certain near-optimal point distributions.

In the third part of our study we derive a new upper bound on the size of a family of subsets of a finite set with restricted pairwise intersections, which improves upon the well-known Frankl-Wilson upper bound. The new bound is obtained by analyzing a refinement of the association scheme of the Hamming space (the Terwilliger algebra) and intertwining functions of the symmetric group.

Finally, in the fourth set of problems we determine new estimates on the size of codes in the Johnson space. We also suggest a new approach to the derivation of the well-known Johnson bound for codes in this space. Our estimates are often valid in the region where the Johnson bound is vacuous. We show that these methods are also applicable to the case of multiple packings in the Hamming space (list-decodable codes). In this context we recover the best known estimate on the size of list-decodable codes in a new way.

BOUNDS ON THE SIZE OF CODES

by

Punarbasu Purkayastha

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2010

Advisory Committee:
Professor Alexander Barg, Chair/Advisor
Professor William Gasarch
Professor Prakash Narayan
Professor André L. Tits
Professor Lawrence C. Washington

Dedication

To my brother, mother and late father.

Acknowledgments

As I pen this acknowledgement memories of my early years as a graduate student fill my mind, when I was flitting between research topics. It was at this stage that my advisor, Professor Alexander Barg, channeled my attention to a rich and exciting area,- the study of extremal problems in coding theory. His individual lectures and precise guidance, frequently coupled with long hours of discussion, helped me comprehend this beautiful topic. He has been instrumental in often returning me to the correct point of view at various points during my research. His advice and guidance on matters both academic and non-academic, have proved invaluable time and again. My technical writing and presentation skills were essentially non-existent earlier, and I owe most of the improvement to the feedback and assistance provided by him. I am deeply grateful to Professor Barg for his help, patience, guidance and support throughout my years as a graduate student. He has been an inspiration for me and has left a lasting positive impression on my development as an individual and as a researcher.

I have gained much knowledge from the myriad courses offered at the Departments of ECE and Mathematics. The courses are very well organized and the teaching has been excellent. I have particularly enjoyed my interaction with Professor Prakash Narayan and Professor André Tits, both inside and outside the classroom. It has been a pleasure to attend their courses as their style of teaching and instruction shows a unique and distinct viewpoint on the subject. I am happy to have been a part of the Information and Coding Theory seminar series organized by Professor Barg and Professor Narayan. It has been an great experience and an ideal environment to learn in detail new and old topics of research.

I would like to thank my committee members,- Professors Alexander Barg, William Gasarch, Prakash Narayan, André Tits and Lawrence Washington, for their time and support. In particular, I thank Professor Narayan and Professor Tits for repeatedly stressing the need and importance of having a broader perspective, during both my Proposal and Defense examinations. I would like to specifically thank Professor Tits for providing me with crucial help and guidance that assisted me in solving the fractional programming in Chapter 7 of this thesis. I also thank Professor Gregory Kabatyanskii for providing me with a lot of useful feedback during my Defense examination.

I am fortunate to have gained several close and lasting friendships with Arun, Arya, Barna, Kaushik, Prasanth, Sirin and Vishwa. They have shared my happiness and sorrow, provided me with great advice, engaged me in many lively discussions (technical and otherwise), and in general have enriched my life at College Park. I am happy to have enjoyed the companionship of many good friends who have given me much joy and support over the years: Amit, Arijit, Biswadip, Dikpal, Himanshu, Kamlesh, Rajibul, Ravi, Shashi, Tania, Tushar, Vikas, and Woomyoung. I am grateful to my very close friend Joshi (at Purdue University) who has always been just a phone call away, and has always lent an ear to whatever I had to say, irrespective of however trite or insipid it might be!

This dissertation would not have been possible without the love, support and encouragement that I have always received from my family,- my elder brother, my mother and my late father. They had established an atmosphere conducive to academics during the early stages of my life. My brother has been most instrumental in providing me with the right books and guidance during my formative teenage years. Many of my subsequent achievements have been a direct or indirect result of his help and mentoring that I received during those years. As my roommate for most of my graduate life, he has also been a constant source of advice, love and support. My deepest gratitude goes to my mother who has endured long years of separation and has given me much love and affection through all these years.

I am thankful to the graduate staff at the Department of ECE, in particular Maria Hoo, Vivian Lu, and Tracy Chung for their help whenever I needed it. Handling any official matter has been a breeze because of their smooth and efficient functioning.

Finally, I thank the National Science Foundation for their grants CCF0515124, CCF0635271, DMS0807411, CCF0916919, CCF0830699, and CCF0635271, and to the National Security Agency for the grant H98230-06-1-0044, which supported my research and stay at the University of Maryland, College Park.

Table of Contents

List of Figures vii

Notation viii

1 Introduction 1
- 1.1 Research area . . . 1
- 1.2 Contributions . . . 3
- 1.3 Structure of the dissertation . . . 5

2 Algebraic combinatorics in coding theory 6
- 2.1 Association schemes . . . 7
- 2.2 Examples of association schemes . . . 11
- 2.3 The Terwilliger algebra of the binary Hamming space . . . 15
- 2.4 Functions on the binary Hamming space . . . 16
  - 2.4.1 The isometry group of the binary Hamming space . . . 16
  - 2.4.2 Decomposition of functions on the binary Hamming space . . 17
  - 2.4.3 Properties of the Krawtchouk polynomials . . . 19

3 Bounds on codes in the Hamming space 21
- 3.1 Sphere-covering and averaging bounds . . . 21
- 3.2 Krawtchouk polynomials and the LP bound . . . 24
  - 3.2.1 The method . . . 24
  - 3.2.2 The bound . . . 24

4 NMDS codes in ordered Hamming space 27
- 4.1 Definitions and basic properties . . . 28
  - 4.1.1 Poset metrics . . . 28
  - 4.1.2 NMDS poset codes . . . 29
- 4.2 NMDS codes and distributions . . . 32
- 4.3 Weight distribution of NMDS codes . . . 34
- 4.4 Constructions of NMDS codes . . . 38

5 Bounds on ordered codes and orthogonal arrays 40
5.1 Introduction . . . 41
5.2 Bounds on ordered codes and OOAs . . . 42
5.2.1 Existence bounds . . . 42
5.2.2 Nonexistence bounds . . . 42
5.2.3 Asymptotics . . . 46
5.2.3.1 Asymptotic bounds for digital (t, m, n)-nets . . . 46
5.3 Multivariate Krawtchouk polynomials . . . 47
5.3.1 Properties of the polynomials $K_f(e)$ . . . 48
5.4 An LP bound on ordered codes and OOAs . . . 54
5.4.1 The bound . . . 54
5.4.2 Spectral radius of $\mathbf{S}_\kappa$ . . . 56
5.4.3 Asymptotic estimate for codes and OOAs . . . 60
5.5 The case $r = 2$ . . . 60

6 Bounds on sets with few intersections 69
6.1 Intertwining functions of the symmetric group . . . 70
6.1.1 Decomposition of the Johnson space . . . 70
6.1.2 Decomposition of functions on the Johnson space . . . 70
6.1.3 Intertwining functions on the Hamming space . . . 71
6.2 New bound on sets with few intersections . . . 73
6.3 Concluding remarks . . . 79

7 Constant weight codes 80
7.1 Introduction . . . 80
7.2 Bound from weighted averages . . . 81
7.3 Bound obtained by using L^2 norm . . . 84
7.4 Numerical results . . . 85
7.5 Bounds on constant weight codes in $\mathcal{H}(q, n)$ . . . 86
7.6 Bounds for codes under list decoding . . . 90
7.7 Discussion . . . 93
7.8 Appendix . . . 94
7.8.1 Minimization on the RHS of (7.7) . . . 94
7.8.2 Minimization of (7.24) and (7.30) . . . 98
7.8.3 Minimization of (7.39) . . . 99

A Appendix 101

Index 102

Bibliography 104

List of Figures

4.1 To the proof of Corollary 4.10 36

5.1 The region L 63
5.2 Bounds on codes for $r = 2$. 67
5.3 Bounds on codes for $r = 3$. 68

7.1 Table of bounds on $A(n, d, w)$ 86
7.2 An example illustrating Lemma 7.9 95

Notation

Boldface low-case letters $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \ldots$ denote vectors. $\mathbf{0}$ denotes a zero vector or an all-zero matrix as appropriate. Working with the ordered Hamming space, we denote shape vectors by low-case letters $a, b, e, f, \ldots$ (and occasionally by F_i). Finite sets are denoted by capital letters A, B, F, etc. Ideals in a partial order are denoted I, $\overleftarrow{I}$, while $\mathcal{I}$ refers to a family of ideals.

Some commonly used notation and acronyms are summarized in the table below.

$\mathcal{C}$	code
$\mathcal{C}\ (n, M, d)$	a code $\mathcal{C}$ of block length n, size M and minimum distance d
$\mathcal{C}\ [n, k, d]$	a linear code $\mathcal{C}$ of length n, dimension k and minimum distance d
$\mathbb{C}$	the set of complex numbers
$d(\cdot, \cdot)$	distance between two vectors
$\mathcal{F}$	family of finite sets
$\mathbb{F}_q$	finite field of q elements
$h_q(\cdot)$	q-ary entropy function: $h_q(x) = -x \log_q \frac{x}{q-1} - (1-x)\log_q(1-x)$
$\mathcal{H}$	Hamming space
$\overleftarrow{\mathcal{H}}, \overrightarrow{\mathcal{H}}$	ordered Hamming space
$K_k(x; n)$	univariate Krawtchouk polynomial
$K_f(e)$	multivariate Krawtchouk polynomial
M	size of a code or a set
$[n]$	the set $\{1, 2, \ldots, n\}$
$\overleftarrow{\mathcal{P}}, \overrightarrow{\mathcal{P}}$	partially ordered sets (posets)
$p_{i,j}^k$	intersection number
$\mathcal{Q}$	additive group of order q
$\mathcal{R}$	set of relations in an association scheme
$\mathbb{R}$	the set of real numbers
$\operatorname{supp} \boldsymbol{x}$	the non-zero coordinates of the vector $\boldsymbol{x}$: $\operatorname{supp} \boldsymbol{x} = \{i : x_i \neq 0\}$
$\mathcal{S}_w$	a sphere of radius w
V_k	the space of univariate polynomials of degree up to k
V_κ	the space of multivariate polynomials of total degree up to κ
X	abstract finite metric space
l.a.	left-adjusted
LP	Linear Programming
MDS	Maximum Distance Separable
NMDS	Near Maximum Distance Separable
OOA	Ordered Orthogonal Array
r.a.	right-adjusted

CHAPTER 1

Introduction

1.1 Research area

This dissertation is devoted to algebraic and combinatorial properties of error-correcting codes. The theory of error-correcting codes began as an answer to the quest of reliable transmission of digital data over noisy channels. Its applications have since expanded into diverse areas of electrical engineering, algorithms and data structures, data security, and presently include copyright protection, property testing, computational biology, methods for reconstruction of under-sampled data, wireless transmission protocols, biometrics and much more. At the same time coding theory has given rise to advances in several mathematical disciplines including discrete and algebraic geometry, combinatorics, analysis, and computational algebra. This dissertation is focused on a range of problems in coding theory in the areas that originate in applied questions but rely on mathematical methods for their solution.

In the simplest setting, an error-correcting code is a set of binary words designed to transmit messages over a communication link ("channel") that introduces occasional errors in the transmission. The task of code design is to ensure recovery of the messages by the recipient under the reliability constraints specified by the system. The maximum number of messages that can be sent through the channel under these constraints has a direct impact on the efficiency of the overall system. Estimating the largest possible number of messages thus becomes one of the main problems encountered in system design.

The problem of estimating the largest size of an error-correcting code with a given recovery guarantee has led to the development of new methods in algebraic and enumerative combinatorics. One prominent example is P. Delsarte's theory of association schemes [26] that has shaped a new mathematical discipline and continues to be used in the analysis of extremal and structural properties of codes. Applications of this theory account for such spectacular discoveries as the best known estimate for the density of packing of spheres in the n-dimensional real space, deep structural results for binary error-correcting codes, extremal arrangements of spheres in the real space, and a universal combinatorial description of configurations in a large class of finite spaces.

To describe the main results of this dissertation, recall that the Hamming distance between two n-dimensional binary vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\{0,1\}^n$ equals the number of their distinct coordinates, i.e., $d_H(\boldsymbol{x},\boldsymbol{y}) = |\{i : x_i \neq y_i,\ i = 1,\dots,n\}|$. The distance d_H arises as an adequate performance measure for data transmission over a channel with independent equiprobable errors ("bit flips"), called the binary symmetric channel. A more complex transmission scenario occurs when the channel preferentially disrupts certain communication sublinks. This channel behavior can be modeled by a generalization of the distance d_H called the r-Hamming distance, defined on the binary space $\{0,1\}^{nr} = \{0,1\}^r \times \cdots \times \{0,1\}^r$ (here $r \geq 1$). For a vector $\boldsymbol{x} \in \{0,1\}^{nr}$ we can write $\boldsymbol{x} = (\boldsymbol{x}_1,\dots,\boldsymbol{x}_n)$, where each $\boldsymbol{x}_i,\ i = 1,\dots,n$ is a block of r bits. The distance between two vectors $\boldsymbol{x},\boldsymbol{y}$ is a sum of contributions of the blocks of r bits, where the distance between two blocks $\boldsymbol{x}_i,\boldsymbol{y}_i,\ i = 1,\dots,n$ is measured by the largest index of the nonzero bit in the difference $\boldsymbol{x}_i - \boldsymbol{y}_i$. This distance addresses the relative preference of the sub-channels, introducing an ordering on the set of coordinates of the vector and giving rise to the term "ordered Hamming weight" to describe the r-Hamming distance and the related norm[1].

Another apparently rather different problem that gives rise to the ordered Hamming weight on binary vectors relates to numerical integration of continuous functions on the unit cube $U^n = [0,1)^n$ in $\mathbb{R}^n$. If f is such a function, then $\int_{U^n} f dx$ can be approximated by averaging f over a finite sample of points $\mathcal{M}$ in the cube, called a net. From the early results of E. Hlawka [45] it is known that the error $|\int_{U^n} f\ dx - \frac{1}{M}\sum_{x\in\mathcal{M}} f(x)|$, where $M = |\mathcal{M}|$, is bounded above by the deviation of the point set from a uniform distribution, called the star-discrepancy $D^*(\mathcal{M})$ of the net. It is a long standing conjecture, proved for $n = 1, 2$, that for any net $\mathcal{M}$, the smallest possible star-discrepancy has the order of magnitude $O(M^{-1}(\log M)^{n-1})$. This led H. Niederreiter [67] to introduce a special class of nets, called (t,m,n)-nets, whose star-discrepancy has the conjectured optimal scaling order.

As a result of the works of Lawrence [51], Mullen and Schmid [63], and Martin and Stinson [60] it became clear that codes in the ordered Hamming space and (t,m,n)-nets form classes of dual objects. It therefore became possible to give a unified treatment of these seemingly unrelated notions. For a given error of integration it is desired that the size of the net be as small as possible. As a result of the duality, upper estimates on the size of codes in the ordered Hamming space result in lower estimates on the size of nets in the unit cube. The ordered Hamming metric has since arisen in a surprisingly large number of disparate applications: it turns out to be an adequate quality measure for communication over slowly fading channels [82, 40], appears in a recent list decoding algorithm of Reed-Solomon codes [65], and in the study of linear complexity of sequences [61].

In this dissertation, we consider the following two problems related to the ordered Hamming space. The first one is concerned with the study of optimal linear codes. A class of optimal linear codes in the ordered Hamming space is given by Maximum Distance Separable (MDS) codes[2]. These codes have the largest size possible for a given value of the minimum separation between distinct code points

[1] These somewhat informal definitions will be made precise in Section 2.2.

[2] A well-known class of MDS codes in the usual Hamming space is the family of Reed-Solomon codes (see, e.g., Roth [73]). The construction of Reed-Solomon codes has also been extended to the ordered Hamming space [72].

and give rise to optimal point distributions in the unit cube. Properties of MDS codes in the ordered Hamming space were studied in a number of papers [72, 78, 30]. Addressing this area, we study properties of near-optimal, or near-MDS (NMDS) codes, find their parameters and establish their relation to point distributions.

The second problem area is related to the study of the combinatorial structure of the ordered Hamming space (its association scheme). We establish properties of the eigenvalues of the scheme and show that they give rise to a family of orthogonal polynomials of several discrete variables (the multivariate Krawtchouk polynomials). This enables us to address the problem of estimates of the size of optimal codes in the ordered Hamming space, studied previously in [72, 59, 19], both in the asymptotic and finite-length setting. By the duality between nets and codes, these estimates also yield lower bounds on the size of (t, m, n)-nets.

A well-studied class of problems in extremal combinatorics deals with estimates of the maximum size of a family of subsets of the n-set that afford a restricted number of pairwise intersections, or distances, or satisfy other conditions of this kind. The study of problems of this nature dates back to the work of Fisher on experimental designs [36]. A classical example in this area is the Erdös-Ko-Rado theorem [34] about the maximum number of subsets such that no two of them are disjoint. Another problem deals with the maximum size of a binary code in which the distance between every two (distinct) vectors takes one of l given values [26, 70, 13]. Extensions of these results were obtained in [5, 38, 39] and many other works. We study the problem of Frankl and Wilson [37] that deals with bounds on the number of subsets with l intersections, employing algebraic ideas brought forth recently in A. Schrijver's study of bounds on codes in the Hamming space [74].

The final group of problems deals with estimates of the size of codes in a subset of the Hamming space formed of vectors of a fixed Hamming weight, or constant weight codes. Bounds on the size of constant weight codes form the contents of an extensive survey [4]. One of the most well-known results is the Johnson bound which has recently gained prominence because of its relation to Sudan-type list decoding of Reed-Solomon and algebraic-geometric codes [73]. Averaging arguments involved in its proof gave rise to some of the classical inequalities in coding theory, e.g., [76]. Following this line of research, we consider embeddings of codes in the real space and bounds on constant weight and list-decodable codes.

1.2 Contributions

This dissertation makes the following contributions to the problems discussed above.

NMDS codes in the ordered Hamming space: Maximum distance separable (MDS) codes in the ordered Hamming space were studied in [72, 78, 47]. They are known to correspond to certain *optimal distributions* of points in the unit cube. In this context, we consider near-maximum distance separable (NMDS) codes, i.e., codes whose minimum distance is just one less than that of MDS codes. We develop elements of the theory of linear codes in the ordered Hamming space and use it to establish structural properties of ordered NMDS codes. We also determine the weight distribution of such codes and show that they are equivalently represented as distributions of points in the unit cube which have properties close to optimal distributions. We also give some constructions of ordered NMDS codes. Most of the results in this context can be readily extended to a generalization of the ordered Hamming metric called the *poset metric space*. For this reason our derivations are phrased in the language of arbitrary poset metrics while results for the ordered case are derived as their specializations. This group of problems is studied in Chapter 4 of the thesis (see also [15]).

Bounds on the size of codes in the ordered Hamming space: We derive new bounds on the maximum size of codes in the ordered Hamming space for a given minimum separation between two distinct code points (minimum distance). Our results improve upon earlier bounds of Rosenbloom and Tsfasman [72], Martin [57, 59], and Bierbrauer [19]. In particular, we obtain the best known asymptotic upper bounds on the size of ordered codes. Owing to the duality between codes and nets, the bounds on codes also result in lower bounds on the size of nets. These results form the contents of Chapter 5 (see also [14, 11]).

Bounds on sets with few intersections: Let $\mathcal{F} = \{U_1, U_2, \dots\}$ be a family of subsets of $[n] = \{1, \dots, n\}$ such that any two distinct subsets satisfy the condition $|U_j \cap U_k| \in \{r_1, \dots, r_l\},\ 0 \le r_i \le n-1,\ i = 1, \dots, l.$ By a well-known theorem of Frankl and Wilson [37], the size of such a family satisfies $|\mathcal{F}| \le \sum_{i=0}^{l} \binom{n}{i}$. This bound is attained if the intersections are taken to be $\{r_1, \dots, r_l\} = \{0, 1, \dots, l-1\}$ [37, 69]. We use coding-theoretic methods to improve the Frankl-Wilson bound for the case when the intersections are $\{r_1, \dots, r_l\} = \{n-1, \dots, n-l\}$ for $l \le \lfloor n/2 \rfloor$. The proof relies on an approach based on the properties of a certain matrix algebra derived from allocations of triples of points in the Hamming space (the Terwilliger algebra of the Hamming space [83, 74]) and linear-algebraic considerations.

Bounds on constant weight codes in the Hamming space: The Johnson bound on the size of constant weight codes is proved by estimating the average distance between distinct vectors in the code. Refining this technique, we use "weighted averages" and some classical inequalities to prove new bounds on the size of such codes. The values of these bounds are at times exact and meet the table of bounds on constant weight codes of small length [4]. We also show that this technique can be adopted to provide new bounds on non-binary constant weight codes. A further generalization of the Johnson bound is related to the concept of multiple packings of the Hamming space (also termed list-decodable codes). The corresponding result in this case was derived by Blinovskii [20]. We show that our methods yield a new proof of his result, and identify obstacles in the way of its improvement.

1.3 Structure of the dissertation

- In Chapter 2 we overview the basic notions of algebraic combinatorics as they apply to coding theory in finite metric spaces. We begin with a discussion of association schemes and the related duality notion for codes and orthogonal arrays. Most of Section 2.2 is devoted to the association scheme that describes the ordered Hamming space. We elaborate on the duality between codes and nets in this context and state a linear programming bound on codes and designs.

- Chapter 3 contains a synopsis of the techniques that are used to derive bounds on codes. These techniques are used in Chapters 5 and 7 of the thesis.

- Chapter 4 is devoted to MDS and NMDS codes in the ordered Hamming space. The new results obtained in this chapter include: computation of the weight distribution of ordered NMDS codes (Theorem 4.9), a relation of NMDS codes to (t, m, n)-nets (Theorem 4.7), and constructions of NMDS codes (Section 4.4).

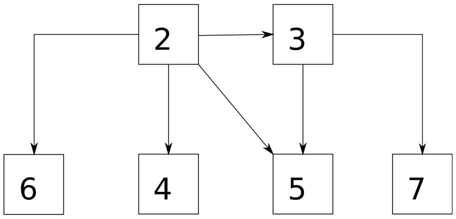
- Chapter 5 is devoted to properties of multivariate Krawtchouk polynomials and estimates of the size of codes and orthogonal arrays. The new results here are: most of Section 5.3, and several new bounds for codes and orthogonal arrays (Theorems 5.6, 5.12, 5.17).

- Chapter 6 begins with a brief introduction to polynomials related to the Terwilliger algebra of the Hamming space. This enables us to introduce the problem of bounds on families of sets with few intersections by showing a connection to these polynomials. A new bound on the size of such families is given in Theorem 6.3.

- Chapter 7 is concerned with bounds on constant weight codes in the Hamming space. We first provide a bound on constant weight codes in the binary Hamming space and then show how this technique can be extended to the non-binary case. We also make a connection to bounds on list-decodable binary codes. The new results in this chapter are related to the proof method and bounds on constant weight codes of Theorems 7.2, 7.4, 7.5 and 7.6.

- The appendix (Chapter A) contains some theorems from linear algebra which are used in other parts of this dissertation.

Chapter Dependencies:



CHAPTER 2

Algebraic combinatorics in coding theory

In this chapter we give a brief overview of methods of algebraic combinatorics and harmonic analysis on groups that are used in later parts of the thesis to derive bounds on the size of codes.

Section 2.1 is devoted to the theory of association schemes. The notions of dual codes, designs and orthogonal arrays arise naturally as a part of this theory. It also leads to a formulation of general linear programming problems whose solutions give bounds on codes in various finite and infinite spaces. For the so-called polynomial association schemes, these bounds are expressed as solutions of optimization problems for certain classes of orthogonal polynomials supported by the scheme.

Examples of association schemes that are used in later chapters of the thesis are given in Section 2.2. In particular, we discuss the Hamming scheme, the Johnson scheme, and the ordered Hamming scheme.

In Section 2.3 we discuss an extension of association schemes given by the Terwilliger algebra. The polynomials supported by this algebra are used in Chapter 6 to provide new bounds on sets with few intersections.

Another facet of the theory of bounds on codes arises when we consider the action of the isometry group on the underlying metric space. These considerations tie bounds on codes to group representations and special functions. In this context, in Section 2.4 we give a brief overview of the decomposition of the space of functions on the Hamming space under the action of its isometry group. The relevant family of orthogonal polynomials that arise under this action are the well-known Krawtchouk polynomials. In Section 2.4.3 we state some useful properties of Krawtchouk polynomials.

The ideas and methods discussed in this chapter form a classic part of combinatorics. They are covered in a vast body of literature. A good source for the theory of association schemes is Delsarte's thesis [26]. The group-theoretic approach to bounds on codes was pioneered by Kabatyanskii and Levenshtein [52] following the foundational works of Gelfand [41], Krein [53], Schoenberg [75], Bochner [21] and others.

2.1 Association schemes

The concept of an association scheme is one of the most important in algebraic combinatorics. Its use in coding theory was initiated by Delsarte in his groundbreaking work [26] and has since resulted in a unifying approach to structural and extremal properties of codes and combinatorial designs. Most results in this section are due to Delsarte [26].

Definition 2.1 *Let X be a finite set and let $\mathcal{R} = \{R_0, \dots, R_N\}$ be a set of $N+1$ binary relations on X. $\mathcal{A} = \{X, \mathcal{R}\}$ is called an N-class* association scheme *if*

1. *$R_0 = \{(\boldsymbol{x}, \boldsymbol{x}) : \boldsymbol{x} \in X\}$ is the identity relation.*

2. *For every $\boldsymbol{x}, \boldsymbol{y} \in X$, $(\boldsymbol{x}, \boldsymbol{y}) \in R_i$ for exactly one i. Thus, the set $\mathcal{R}$ forms a partition of $X \times X$.*

3. *For each i in $\{0, \dots, N\}$, R_i is symmetric: $(\boldsymbol{x}, \boldsymbol{y}) \in R_i \Leftrightarrow (\boldsymbol{y}, \boldsymbol{x}) \in R_i$ for any $\boldsymbol{x}, \boldsymbol{y} \in X$.*

4. *There exist non-negative numbers p_{ij}^k called* intersection numbers*, defined as*

$$p_{ij}^k = |\{\boldsymbol{z} \in X : (\boldsymbol{x}, \boldsymbol{z}) \in R_i, (\boldsymbol{y}, \boldsymbol{z}) \in R_j\}|, \quad \text{for any } (\boldsymbol{x}, \boldsymbol{y}) \in R_k, \quad i, j, k \in \{0, \dots, N\}.$$

The numbers p_{ij}^k depend only on i, j and k, and satisfy the condition $p_{ij}^k = p_{ji}^k$.

An immediate consequence of the definition is that for each $i = 0, \dots, N$, the numbers $v_i = p_{i,i}^0 = |\{\boldsymbol{y} \in X : (\boldsymbol{x}, \boldsymbol{y}) \in R_i\}|$ are independent of $\boldsymbol{x}$. These numbers are called the *valencies* of the scheme.

An equivalent definition of an association scheme is given in terms of a particular matrix algebra $\mathbb{A}$ called the *Bose-Mesner algebra*. For $i = 0, \dots, N$ let D_i be the adjacency matrix of R_i, i.e., an $|X| \times |X|$ matrix such that $(D_i)_{\boldsymbol{x},\boldsymbol{y}} = 1((\boldsymbol{x}, \boldsymbol{y}) \in R_i)$, where $1(\cdot)$ is the indicator function. Consider the complex vector space $\mathbb{A}$ generated by the adjacency matrices D_i, i.e. $\mathbb{A} \triangleq \{c_0 D_0 + \dots + c_N D_N : c_i \in \mathbb{C}\}$. The vector space $\mathbb{A}$ has the following properties.

1. The all 1 matrix of size $|X|$, $J = D_0 + \dots + D_N$, belongs to $\mathbb{A}$.

2. The adjacency matrix is real and symmetric: $D_i = D_i^T,\ i = 0, \dots, N$. Hence $\mathbb{A}$ is closed under conjugate transposition.

3. For any $i, j \in \{0, \dots, N\},\ D_i D_j = \sum_{k=0}^N p_{ij}^k D_k$.

By property *4* of the association scheme $\mathcal{A}$ we obtain the relation $D_i D_j = D_j D_i$. Thus, $\mathbb{A}$ is commutative and closed under matrix multiplication, and is hence a commutative algebra over $\mathbb{C}$.

The Bose-Mesner algebra also has a basis of minimal idempotents which we proceed to describe. Consider the complex vector space $U = \langle \boldsymbol{e}_{\boldsymbol{x}}, \boldsymbol{x} \in X \rangle$, where $\boldsymbol{e}_{\boldsymbol{x}}$

is the indicator vector of $\boldsymbol{x}$ in X. Since the Bose-Mesner algebra $\mathbb{A}$ is Hermitian and commutative, by the spectral decomposition theorem there exists a unitary matrix that simultaneously diagonalizes all matrices in $\mathbb{A}$. Therefore, the space U splits into an orthogonal direct sum of the common eigenspaces $\{U_i, i = 0, 1, \dots, N\}$ of the adjacency matrices. Let $E_i : U \to U_i$ denote the orthogonal projection of U onto the i-th eigenspace U_i.

Then we obtain the following properties of the matrices E_i.

1. The matrices E_i, $i = 0, \dots, N$ are idempotent and positive semidefinite. The set $\{E_0, \dots, E_N\}$ forms a basis for $\mathbb{A}$.

2. $\sum_{j=0}^N E_j = I$, where I is an identity matrix of order $|X|$.

3. $|X|E_0 = J$.

4. The adjacency matrices can be written in the basis of the idempotent matrices as $D_i = \sum_{j=0}^N p_i(j)E_j$ and thus, for any $i, j \in \{0, \dots, N\}$,
$$D_iE_j = p_i(j)E_j.$$

The numbers $p_i(j)$ are called the *first eigenvalues* of $\mathcal{A}$.

Define the Schur (elementwise) product of two matrices A and B by $(A \circ B)_{ij} = A_{ij}B_{ij}$. Note that
$$D_i \circ D_j = \delta_{i,j}D_i,$$
i.e., the Bose-Mesner algebra is also closed under the Schur product. The basis of idempotents $\{E_0, \dots, E_N\}$ satisfies a set of properties under the Schur product that are similar to the properties of the adjacency matrices under the matrix multiplication.

1. For any $i, j \in \{0, \dots, N\}$ the idempotent matrices satisfy $|X|(E_i \circ E_j) = \sum_{k=0}^N q_{ij}^k E_k$. The numbers q_{ij}^k are non-negative and are called the *Krein numbers* of the scheme.

2. Since $\{D_0, \dots, D_N\}$ is also a basis of $\mathbb{A}$, for each $j = 0, \dots, N$ we can expand E_j in terms of $\{D_0, \dots, D_N\}$ as
$$|X|E_j = \sum_{i=0}^N q_j(i)D_i.$$
Therefore, $|X|E_j \circ D_i = q_j(i)D_i$. The coefficients $q_j(i)$ of this expansion are called the *second eigenvalues* of the scheme.

3. $\mathrm{Tr}(E_j) = \mathrm{rank}(E_j) = m_j$, where m_j are called the *multiplicities* of the scheme, and correspond to the multiplicities of the eigenvalues $p_i(j)$ of D_i.

4. $|X|(E_k)_{\boldsymbol{x},\boldsymbol{y}} = q_k(i)$, if $(\boldsymbol{x}, \boldsymbol{y}) \in R_i$.

The first and the second eigenvalues of the scheme $\mathcal{A}$ satisfy the following important relations.

$$\text{Orthogonality:}\quad \sum_{i=0}^{N} m_i p_j(i) p_k(i) = |X| v_j \delta_{j,k}, \quad \sum_{i=0}^{N} v_i q_j(i) q_k(i) = |X| m_j \delta_{j,k}. \tag{2.1}$$

$$m_i p_j(i) = v_j q_i(j). \tag{2.2}$$

$$p_j(i) p_k(i) = \sum_{l=0}^{N} p_{j,k}^{l} p_l(i), \ 0 \le i \le N. \tag{2.3}$$

Duality in association schemes: Two N-class association schemes are called *Delsarte duals* of each other if the adjacency matrices D_i, the first eigenvalues $p_i(j)$ and the intersection numbers p_{ij}^k of one scheme are respectively the idempotents E_i, the second eigenvalues $q_i(j)$, and the Krein numbers q_{ij}^k of the other. The duality also exchanges the roles of the matrix and Schur multiplication. A scheme is called *self dual* if it is equal (isomorphic) to its dual. A scheme is called *formally self dual* if there exists an ordering of the idempotents E_i under which $p_i(j) = q_i(j)$ for each $i, j \in \{0, \dots, N\}$. In a formally self dual scheme $v_i = m_i$ and $p_{ij}^k = q_{ij}^k$ also hold for all $i, j, k \in \{0, \dots, N\}$. A formally self-dual scheme may or may not have a dual scheme.

P- and Q-polynomial association schemes: Suppose that for every $i = 0, \dots, N$ there exist polynomials $P_i(x)$ (resp. $Q_i(x)$) of degree i such that $P_i(j) = p_i(j)$ (resp. $Q_i(j) = q_i(j)$). In this case the association scheme is called *P-polynomial* (resp. *Q-polynomial*). From (2.1) it follows that the polynomials $P_i(x)$ and $Q_i(x)$ are orthogonal on $\{0, \dots, N\}$ with weights m_i and v_i, respectively. Polynomial association schemes are particularly important in the derivation of bounds on the size of codes and designs.

Codes in association schemes: Let $\mathcal{A}(X, \mathcal{R})$ be an association scheme with N classes and let $\mathcal{C} \subset X$ be a code. Suppose that X is a metric space with distance function $d(\cdot,\cdot)$ and $R_i \triangleq \{(\boldsymbol{x}, \boldsymbol{y}) : d(\boldsymbol{x}, \boldsymbol{y}) = i\}$. Let $\chi = (1(\boldsymbol{x} \in \mathcal{C}), \boldsymbol{x} \in X)$ be the characteristic (column) vector of $\mathcal{C}$ in X. Define the *inner distribution* of the code $\mathcal{C}$ as $(B_0, B_1, \dots, B_N)$, where B_i is the average number of ordered pairs of code points that fall in R_i (are distance i apart):

$$\begin{aligned} B_i &\triangleq \frac{1}{|\mathcal{C}|} \chi^T D_i \chi \qquad\qquad (2.4) \\ &= \frac{1}{|\mathcal{C}|} |\mathcal{C}^2 \cap R_i|, \quad i = 0, \dots, N. \end{aligned}$$

It is readily seen that $B_0 = 1$ and $\sum_{k=0}^{N} B_k = |\mathcal{C}|$. The *minimum distance* of $\mathcal{C}$ is the smallest non-zero index i for which $B_i \neq 0$.

The *dual distribution* of $\mathcal{C}$ is defined as

$$B_k^{\perp} = \frac{1}{|\mathcal{C}|} \sum_{i=0}^{N} B_i q_k(i), \quad k = 0, \ldots, N. \tag{2.5}$$

An important observation, due to Delsarte, is that the quantities $B_k^{\perp}$ are nonnegative:

$$\begin{aligned} B_k^{\perp} &= \frac{1}{|\mathcal{C}|} \sum_{i=0}^{N} B_i q_k(i) \\ &= \frac{1}{|\mathcal{C}|^2} \chi^T \left(\sum_{i=0}^{N} q_k(i) D_i \right) \chi \\ &= \frac{|X|}{|\mathcal{C}|^2} \chi^T E_k \chi \geq 0, \end{aligned}$$

where the last step follows because E_k is positive semidefinite. This implies that the size of the code can be bounded above as follows.

Theorem 2.1 *Let $\mathcal{C} \subset X$ be a code of size M and minimum distance d. Then $M \leq 1 + LP$, where LP is the optimal value of the following linear programming problem:*

$$\begin{aligned} \max\ & B_d + \cdots + B_N \\ s.\ t.\ & B_i \geq 0,\ i = d, \ldots, N \\ & \sum_{i=d}^{N} B_i q_k(i) \geq -q_k(0),\ k = 1, \ldots, N. \end{aligned}$$

Equivalently, LP is the optimal value of the following dual linear program:

$$\begin{aligned} \min\ & \beta_1 q_1(0) + \cdots + \beta_N q_N(0) \\ s.\ t.\ & \beta_k \geq 0,\ k = 1, \ldots, N \\ & \sum_{k=1}^{N} \beta_k q_k(i) \leq -1,\ i = d, \ldots, N. \end{aligned}$$

Dual codes and designs in association schemes: A subset $\mathcal{C}$ of X is called a *t-design* if $B_k^{\perp} = 0,\ k = 1, \ldots, t$. Designs have been the subject of a large number of studies in combinatorics and applied statistics. Examples of designs include such well-studied combinatorial objects as orthogonal arrays, balanced incomplete block designs or BIBDs, and Steiner systems [55]. The notion of an orthogonal array is also a special case of weakly-biased random variables [6].

In the particular case that X is a vector space over a finite field $\mathbb{F}_q$, we call $\mathcal{C}$ a *linear code* if it is a subspace of X. We define its *dual code* $\mathcal{C}^{\perp}$ as

$$\mathcal{C}^{\perp} \triangleq \{\boldsymbol{y} \in X : \langle \boldsymbol{c}, \boldsymbol{y} \rangle = 0\ \forall \boldsymbol{c} \in \mathcal{C}\},$$

where the inner product is computed over $\mathbb{F}_q$. In this case, by the well-known MacWilliams theorem [26, Theorem 6.3] the inner distribution of $\mathcal{C}^\perp$ coincides with the dual distribution of $\mathcal{C}$. Thus if $\mathcal{C}$ is linear, the dual code $\mathcal{C}^\perp$ forms a t-design with $t = d - 1$. If the dimension of $\mathcal{C}$ is k then the dual code $\mathcal{C}^\perp$ has dimension $n - k$.

The duality between codes and designs is of fundamental nature and is manifested in many important results of combinatorial coding theory [26]. One of them is a simultaneous linear programming bound on the size of codes and designs. To develop it, suppose that the association scheme $\mathcal{A}$ is Q-polynomial. Then we can substitute the second eigenvalues $q_k(i)$ by the evaluations of a polynomial $Q_k(x)$. Since the polynomials $Q_k(x)$ are orthogonal, any polynomial of degree at most N can be written in the basis of the Q-polynomials. Finally, we note that any feasible solution to the dual program in Theorem 2.1 is an upper bound on the value of LP. These observations lead to the following linear programming bounds on codes and designs.

Theorem 2.2 *Consider a polynomial $f(x) = f_0 + \sum_{i=1}^N f_i Q_i(x)$ such that $f_0 > 0$, $f_i \geq 0,\ \forall i = 1, \ldots, N$, and $f(i) \leq 0, \forall i = d, \ldots, N$. If $\mathcal{C} \subset X$ is a code of size M and minimum distance d, then*

$$M \leq f(0)/f_0. \tag{2.6}$$

A design $\mathcal{C} \subset X$ of size M' and strength $t = d - 1$ satisfies

$$M' \geq |X| f_0 / f(0).$$

Remarks:

1. Note that we can identify $f_i, i = 1, \ldots, N$ with $\beta_i f_0$ of Theorem 2.1.
2. Bounds on codes obtained by the application of this theorem are called *linear programming (LP) bounds.* To derive a bound, rather than attempting numerical linear programming, we rely on analytical methods to find a suitable polynomial $f(x)$.
3. The bounds obtained from this theorem lead to some of the best asymptotic upper bounds on the size of the code. However, the application of this method faces non-trivial analytical challenges and asymptotic computations.

2.2 Examples of association schemes

Hamming scheme: Let $\mathcal{Q} = \{0, \ldots, q-1\}$ be a finite alphabet of size q viewed as an additive group mod q. The Hamming space $\mathcal{H}(q, n)$ is the set of n-strings over $\mathcal{Q}$ equipped with the *Hamming distance* $d_H(\boldsymbol{x}, \boldsymbol{y}) = |\{x_i \neq y_i,\ i = 1, \ldots, n\}|$. The *Hamming weight* $\mathrm{w}_H(\cdot)$ of a vector $\boldsymbol{x}$ is the number of non-zero coordinates in it: $\mathrm{w}_H(\boldsymbol{x}) = |\{i : x_i \neq 0,\ i = 1, \ldots, n\}|$.

Let $R_i = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{H}^2(q,n) : d_H(\boldsymbol{x}, \boldsymbol{y}) = i\},\ i = 0, 1, \ldots, n$. The space $\mathcal{H}(q,n)$ together with the relations R_i forms a self-dual association scheme. The valencies (and multiplicities) of the scheme are $v_i = m_i = (q-1)^i \binom{n}{i},\ i = 0, \ldots, n$. Its Q-polynomials (and P-polynomials) are given by a particular family of polynomials of a discrete variable $K_k(x; n),\ k = 0, \ldots, n$ called the Krawtchouk polynomials. By (2.1) the Krawtchouk polynomials are orthogonal on the set $\{0, \ldots, n\}$ with weights $m(i) = (q-1)^i \binom{n}{i},\ i = 0, \ldots, n$. Designs in the Hamming scheme are the classical orthogonal arrays, defined below.

Definition 2.2 *An $M \times n$ matrix O with entries from the alphabet $\mathcal{Q}$ is called an* orthogonal array *(OA) of size M, n constraints, q levels, strength t and index θ if any set of t columns of O contains all the q^t possible row vectors, each exactly θ times.*

Binary Johnson scheme: The subset of the binary Hamming space $\mathcal{H}(2,n)$ that consists of all vectors of a given weight w (with $w \leq n/2$) is called the Johnson space. We denote the space by $\mathcal{S}_w = \{\boldsymbol{x} \in \mathcal{H}(2,n) : \mathrm{w}_H(\boldsymbol{x}) = w\}$. The distance between two vectors $\boldsymbol{x}, \boldsymbol{y}$ in $\mathcal{S}_w$ is defined as $d_J(\boldsymbol{x}, \boldsymbol{y}) = 1/2\ d_H(\boldsymbol{x}, \boldsymbol{y})$. The association scheme on $\mathcal{S}_w$ is given by the relations $R_i = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{S}_w^2 : d_J(\boldsymbol{x}, \boldsymbol{y}) = i\},\ i = 0, \ldots, w$. The Q-polynomials of the scheme are given by a certain family of polynomials $Q_k(x),\ k = 0, \ldots, w$ orthogonal on $\{0, \ldots, w\}$, called the Hahn polynomials. The Johnson scheme has no dual scheme.

Ordered Hamming space: We now describe the ordered Hamming space, on which the results in Chapters 4 and 5 are based. We will follow this with a description of the association scheme on this space. Let $\mathcal{Q} = \{0, \ldots, q-1\}$ be as above. Consider the set $\mathcal{Q}^{n,r}$ of vectors of length nr over $\mathcal{Q}$. A vector $\boldsymbol{x} \in \mathcal{Q}^{n,r}$ is written as a concatenation of n blocks of length r each, $\boldsymbol{x} = (x_{11}, \ldots, x_{1r}; \ldots; x_{n1}, \ldots, x_{nr})$.

Definition 2.3 *For a given vector $\boldsymbol{x} \in \mathcal{Q}^{n,r}$ let*

$$\mathrm{w}_r(\boldsymbol{x}) = \sum_{i=1}^{n} \max\{j : x_{ij} \neq 0\}$$

be its ordered Hamming weight*, where the maximum is taken to be* 0 *if the set $\{j : x_{ij} \neq 0\}$ is empty. The* ordered Hamming distance *between vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{Q}^{n,r}$ is equal to $d_r(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{w}_r(\boldsymbol{x} - \boldsymbol{y})$. Define the* dual weight *as*

$$\overline{\mathrm{w}}_r(\boldsymbol{x}) = \sum_{i=1}^{n} \max\{j : x_{i,r-j+1} \neq 0\}.$$

Note that in the case $r = 1$ both w_r and $\overline{\mathrm{w}}_r$ correspond to the usual Hamming distance on $\mathcal{Q}^n$.

Let $e_i,\ i = 1, \ldots, r$ be the number of r-blocks of $\boldsymbol{x}$ whose rightmost nonzero entry is in the i-th position counting from the beginning of the block, i.e., $e_i = |\{j :$

$\max\{k : x_{jk} \neq 0\} = i\}|$. The r-vector $e = (e_1, \dots, e_r)$ will be called the *shape* of $\boldsymbol{x}$, denoted $\mathrm{shape}(\boldsymbol{x})$. For brevity we write

$$|e| = \sum_{i=1}^{n} e_i, \quad |e|' = \sum_{i=1}^{n} i e_i, \quad e_0 = n - |e|.$$

In particular, $\mathrm{w}_r(\boldsymbol{x}) = |e|'$. A shape vector $e = (e_1, \dots, e_r)$ defines a partition of a number $N \leq n$ into a sum of r parts. Let $\mathbb{N}$ denote the set of all positive integers and let

$$\Delta_{n,r} = \Big\{ e \in (\mathbb{N} \cup \{0\})^r : \sum_{i=1}^{n} e_i \leq n \Big\}$$

be the set of all such partitions.

The ordered weight was first introduced by Niederreiter [67] in his study of low-discrepancy point sets for numerical integration of functions on the unit cube in $\mathbb{R}^n$. Later, Rosenbloom and Tsfasman [72] independently defined the weight $\overline{\mathrm{w}}_r(\boldsymbol{x})$ and studied codes in $\mathcal{Q}^{n,r}$ with respect to it. The set $\mathcal{Q}^{n,r}$ together with the distance function $d_r(\cdot, \cdot)$ will be called the *ordered Hamming space* (the *NRT space*) and denoted by $\overrightarrow{\mathcal{H}} = \overrightarrow{\mathcal{H}}(q, n, r)$. Unless specified otherwise below, in this section by distance (weight) we mean the ordered distance (weight) for some fixed value of r.

The ordered Hamming scheme: The association scheme for the ordered Hamming space was constructed and studied by Martin and Stinson [60].

To define the ordered Hamming scheme, we need the notion of extensions of association schemes [26, p. 17]. A scheme $(X^N, \mathcal{R})$ is called an *N-th degree extension* of an r-class scheme $\mathcal{K} = \big(X, \mathcal{D} = (D_0, D_1, \dots, D_r)\big)$ if its point set is the N-fold Cartesian product of X and the relations $R_e, e \in \Delta_{N,r}$ are given by

$$R_e = \big\{ \big((\boldsymbol{x}_1, \dots, \boldsymbol{x}_N), (\boldsymbol{y}_1, \dots, \boldsymbol{y}_N)\big) \in X^N \times X^N : |\{j : (\boldsymbol{x}_j, \boldsymbol{y}_j) \in D_i\}| = e_i, \\ i = 0, 1, \dots, r \big\}.$$

We begin with an auxiliary notion. An r-class *kernel scheme* $\overrightarrow{\mathcal{K}}$ on the space $\overrightarrow{\mathcal{H}}(q, 1, r)$ has relations

$$R_i = \{(\boldsymbol{x}, \boldsymbol{y}) \in \overrightarrow{\mathcal{H}}(q, 1, r) \times \overrightarrow{\mathcal{H}}(q, 1, r) : d_r(\boldsymbol{x}, \boldsymbol{y}) = i\}, \quad i = 0, 1, \dots, r.$$

The scheme $\overrightarrow{\mathcal{K}}$ is formally self-dual and has a dual scheme $\overleftarrow{\mathcal{K}}$ that is defined on the space $\overleftarrow{\mathcal{H}}(q, 1, r)$ which consists of the same set of points $\mathcal{Q}^{1,r}$, but with distance defined as $\bar{d}_r(\boldsymbol{x}, \boldsymbol{y}) = \overline{\mathrm{w}}_r(\boldsymbol{x} - \boldsymbol{y})$.

The association scheme $\overrightarrow{\mathcal{A}}$ of the ordered Hamming space is defined as the n-th degree extension of $\overrightarrow{\mathcal{K}}$. In other words, the set X that affords $\overrightarrow{\mathcal{A}}$ is $\mathcal{Q}^{n,r}$. The relations in the scheme $\overrightarrow{\mathcal{A}}$ are naturally indexed by shape vectors $e \in \Delta_{n,r}$. In particular,

$$(\boldsymbol{x}, \boldsymbol{y}) \in R_e \Leftrightarrow \mathrm{shape}(\boldsymbol{x} - \boldsymbol{y}) = e.$$

For a vector $\boldsymbol{x} \in \mathcal{Q}^{n,r}$ define $\overline{\mathrm{shape}}(\boldsymbol{x}) = (e_1, \dots, e_r)$, where $e_j = |\{i : 1 \leq i \leq n, \overline{\mathrm{w}}_r(x_{i1}, \dots, x_{ir}) = j\}|$.

Theorem 2.3 (Martin and Stinson [60]) *The space $X = \mathcal{Q}^{n,r}$ together with the relations $\mathcal{R} = \{R_e : e \in \Delta_{n,r}\}$, where*

$$R_e = \{(\boldsymbol{x}, \boldsymbol{y}) \in X \times X : \ \mathrm{shape}(\boldsymbol{x} - \boldsymbol{y}) = e\}, \quad e \in \Delta_{n,r}$$

forms a formally self-dual association scheme called the ordered Hamming scheme $\overrightarrow{\mathcal{A}}$. It can be constructed as an n-fold Delsarte extension of $\overrightarrow{\mathcal{K}}$.

The dual scheme of $\overrightarrow{\mathcal{A}}$ is $\overleftarrow{\mathcal{A}}$ whose point set is $\mathcal{Q}^{n,r}$ and the set of relations is given by

$$R_e = \{(\boldsymbol{x}, \boldsymbol{y}) \in X \times X : \overline{\mathrm{shape}}(\boldsymbol{x} - \boldsymbol{y}) = e\}, \quad e \in \Delta_{n,r}.$$

The ordered Hamming space corresponding to the scheme $\overleftarrow{\mathcal{A}}$ is denoted by $\overleftarrow{\mathcal{H}}(q, n, r)$. The intersection numbers of the ordered Hamming scheme $\overrightarrow{\mathcal{A}}$ are defined by the shapes:

$$p_{ef}^{g} = |\{\boldsymbol{z} \in \overrightarrow{\mathcal{H}}(q, n, r) : \mathrm{shape}(\boldsymbol{x} - \boldsymbol{z}) = e, \ \mathrm{shape}(\boldsymbol{y} - \boldsymbol{z}) = f, \ \mathrm{shape}(\boldsymbol{x} - \boldsymbol{y}) = g\}|, \tag{2.7}$$

and p_{ef}^{g} does not depend on the vectors $\boldsymbol{x}, \boldsymbol{y} \in \overrightarrow{\mathcal{H}}(q, n, r)$. The valencies (and multiplicities) of the scheme are given by

$$v_e = \binom{n}{e_0, \dots, e_r} (q-1)^{|e|} q^{|e|' - |e|}, \quad e \in \Delta_{n,r}. \tag{2.8}$$

The eigenvalues of the scheme are evaluations on the set $\Delta_{n,r}$ of certain multivariate orthogonal polynomials, called the (multivariate) Krawtchouk polynomials . Since the relations of $\overrightarrow{\mathcal{A}}$ are indexed by the shape vectors, this induces the numbering of the polynomials. Therefore, we denote the multivariate Krawtchouk polynomial by $K_f(e)$ where e, f are shapes. The orthogonality relation (2.1) of the eigenvalues takes the following form:

$$\sum_{e \in \Delta_{n,r}} v_e K_f(e) K_g(e) = q^{nr} v_f \delta_{f,g}. \tag{2.9}$$

We refer to Sec 5.3, Chapter 5 for a more explicit description and properties of the polynomials $K_f(e)$.

Codes and orthogonal arrays in the ordered Hamming space: An arbitrary subset $\mathcal{C}$ of $\overrightarrow{\mathcal{H}}(q, n, r)$ is called an *ordered code*. If the ordered code has minimum distance d and size M, we will call it an (nr, M, d) code. If $\mathcal{C}$ is a linear ordered code of dimension k over a finite field $\mathbb{F}_q$, we will use the notation $[nr, k, d]$ instead. If $\mathcal{C}$ is a linear code in $\overrightarrow{\mathcal{H}}(q, n, r)$ the dual code is given by $\mathcal{C}^{\perp} = \{\boldsymbol{y} \in \mathbb{F}_q^{n,r} : \sum_{i=1}^{n} \sum_{j=1}^{r} c_{ij} y_{ij} = 0 \ \forall \boldsymbol{c} \in \mathcal{C}\}$. Because of the ordering imposed by the duality of the association schemes, the dual code is a subset of the space $\overleftarrow{\mathcal{H}}(q, n, r)$.

In analogy with the definition of an orthogonal array in the Hamming space, one can define an ordered orthogonal array (OOA) in the ordered Hamming space. Let us call a subset of coordinates $I \subset \{1, \dots, rn\}$ *left-adjusted* if with any coordinate $ir + j, \ 0 \le i \le n - 1, \ 1 \le j \le r$ in the i-th block it also contains all the coordinates $(ir + 1, \dots, ir + j - 1)$ of the same block.

Definition 2.4 *A subset $O \subset \overrightarrow{\mathcal{H}}(q,n,r)$ of size $|O| = M$ is called a (t,n,r,q)* OOA *of strength t if its projection on any left-adjusted set of t coordinates contains all the q^t rows an equal number, say θ, of times. The parameter θ is called the* index *of O.*

It follows that $M = \theta q^t$. If $\mathcal{Q}$ is equipped with the structure of an additive group, then one can construct *additive* OOAs. OOAs (also called hypercubic designs) were introduced in Lawrence [51] and Mullen and Schmid [63] as a combinatorial equivalent of point sets suitable for numerical integration over the unit cube.

Remark: Carrying over the theory of linear codes from the classical (Hamming) case to the context considered encounters a number of obstacles. In particular, the relation between a code and its dual code becomes far less straightforward than in the standard situation. Most importantly, in the present situation, the combinatorial structure of the linear space for a linear code is not identical to the structure for its dual code. This leads to a number of subtle changes in the standard facts about linear codes and related combinatorial configurations.

Let $\mathcal{C}$ be a linear $[nr,k,d]$ code in $\overrightarrow{\mathcal{H}}(q,n,r)$ and let $\mathcal{C}^\perp$ be the set of vectors orthogonal to $\mathcal{C}$ with respect to the usual inner product.

1. Considered as an OOA, $\mathcal{C}^\perp$ has parameters $(d-1,n,r,q)$ with index $\theta = q^{nr-k-d+1}$, and is a subset of $\overrightarrow{\mathcal{H}}(q,n,r)$.
2. Considered as a code with minimum distance $d^\perp$, the same set of vectors $\mathcal{C}^\perp$ is a linear subspace of dimension $n-k$ in $\overleftarrow{\mathcal{H}}(q,n,r)$.

In the ordered Hamming space the LP bound of Theorem 2.2 on the size of codes and designs takes the following form.

Theorem 2.4 [26, 60] *Let $F(f) = F_0 + \sum_{e\neq 0} F_e K_e(f)$ be a polynomial that satisfies*

$$F_0 > 0; \quad F_e \ge 0 \text{ for } e \neq 0; \quad F(e) \le 0 \text{ for all } e \text{ such that } \textstyle\sum_{i=1}^r ie_i \ge d. \tag{2.10}$$

Then the size of any (nr,M,d) code satisfies

$$M \le F(0)/F_0, \tag{2.11}$$

and the size M' of any OOA of strength $t = d-1$ satisfies

$$M' \ge q^{nr} F_0/F(0). \tag{2.12}$$

2.3 The Terwilliger algebra of the binary Hamming space

In this section we give a brief description of a refinement of the Bose-Mesner algebra, called the Terwilliger algebra, focusing on the binary Hamming space $\mathcal{H} = \mathcal{H}(2,n)$. Polynomials related to this algebra are used in Chapter 6 to derive bounds on sets with few intersections.

For a vector $\boldsymbol{u} \in \mathcal{H}$ define the *support* of $\boldsymbol{u}$ as the set of non-zero coordinates in $\boldsymbol{u}$: $\operatorname{supp} \boldsymbol{u} = \{i : u_i \neq 0,\ i = 1, \dots, n\}$. For vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{H}$ we write $U = \operatorname{supp} \boldsymbol{u},\ V = \operatorname{supp} \boldsymbol{v}$. For nonnegative integers s, t, r, consider the $|\mathcal{H}| \times |\mathcal{H}|$ matrix $M_{s,t}^r$ with entries

$$(M_{s,t}^r)_{\boldsymbol{u},\boldsymbol{v} \in \mathcal{H}} = 1(|U| = s, |V| = t, |U \cap V| = r),$$

where $1(\cdot)$ is the indicator function. Since $d_H(\boldsymbol{u}, \boldsymbol{v}) = s + t - 2r$, it is readily seen that the matrices can also be indexed by the distances in the set $\{\boldsymbol{u}, \boldsymbol{v}, \mathbf{0}\}$, where $\mathbf{0}$ is the origin.

The $\mathbb{C}*$ algebra

$$\mathbb{T} = \left\{ \sum_{s,t,r=0}^{n} c_{s,t}^r M_{s,t}^r : c_{s,t}^r \in \mathbb{C} \right\}$$

is called the *Terwilliger algebra* of the Hamming space with respect to $\mathbf{0}$ [83, 74].

The algebra $\mathbb{T}$ contains the identity matrix and is closed under addition, matrix multiplication, and taking the adjoint. It follows from the Artin-Wedderburn theory [50] that there exists a unitary matrix that simultaneously reduces all the matrices in $\mathbb{T}$ to a block-diagonal form. The entries of the block-diagonal matrices are, up to scaling factors, the evaluations of the Hahn polynomials [74, 85].

The Bose-Mesner algebra of $\mathcal{H}$ is a commutative algebra defined by the relations between pairs of points in the Hamming space. Using a diagonalization of the Bose-Mesner algebra, bounds on codes are obtained upon setting a linear programming problem. The Terwilliger algebra of the Hamming space is a non-commutative algebra obtained by considering relations between triplets of points one of which is fixed to $\mathbf{0}$. Improved bounds on codes in the Hamming space have been obtained by block-diagonalizing the Terwilliger algebra and using semidefinite programming [74, 43]. The constraints in the semidefinite programs subsume the constraints in the linear programs obtained from the Bose-Mesner algebra. We refer to [74, 43, 85] for a detailed discussion of these results.

2.4 Functions on the binary Hamming space

In this section we develop an alternative view of Q-polynomials of the association scheme and their generalizations. Our focus will be on the binary Hamming space because this is the case for which these ideas will be employed in Chapter 6. The results developed in this section are primarily due to Kabatyanskii and Levenshtein [52] (the general case) and Dunkl [32] (the Hamming case). They are rooted in the theory of special functions and harmonic analysis of non-commutative groups.

2.4.1 The isometry group of the binary Hamming space

A typical isometry of the binary Hamming space $\mathcal{H} = \mathcal{H}(2, n)$ acts by a permutation σ of the coordinates followed by a translation by a fixed vector $\boldsymbol{x}$, i.e., for $\boldsymbol{c} \in \mathcal{H}$, $(\boldsymbol{x}, \sigma) \cdot \boldsymbol{c} = \boldsymbol{x} + \sigma(\boldsymbol{c})$. Let us compute a composition of two such isometries.

Suppose that

$$(\boldsymbol{z}, \pi) = (\boldsymbol{x}, \sigma)(\boldsymbol{y}, \tau)$$

is applied to a vector $\boldsymbol{c} \in \mathcal{H}$. Considering the action $(\boldsymbol{z}, \pi) \cdot \boldsymbol{c}$ it is easy to see that $\pi = \sigma\tau$ and $\boldsymbol{z} = \boldsymbol{x} + \sigma(\boldsymbol{y})$. In particular,

$$(\tau^{-1}(\boldsymbol{y}), \tau^{-1})(\boldsymbol{y}, \tau) = (\boldsymbol{0}, e),$$

where e is the identity permutation. Thus, the pairs $(\boldsymbol{x}, \sigma), \boldsymbol{x} \in \mathcal{H}, \sigma \in S_n$ form a group, called the *semidirect product* of the additive group $\mathcal{H}$ and the symmetric group S_n, and denoted $G = \mathcal{H} \rtimes S_n$. It is well known that G is the full isometry group of the Hamming space. We note that G is not commutative (if it were, a direct product would suffice to combine isometries from $\mathcal{H}$ and S_n).

Note that the action of G on $\mathcal{H}$ is *transitive*, i.e, for any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{H}$ there exists $g \in G$ such that $\boldsymbol{y} = g \cdot \boldsymbol{x}$. Thus, $\mathcal{H}$ is a homogeneous space of the group G. Let H be the subgroup of G that fixes the point $\boldsymbol{0}$, i.e., $H = \{(\boldsymbol{0}, \sigma), \sigma \in S_n\}$, $H \cong S_n$. We can realize $\mathcal{H}$ as the set of (left) cosets G/H by identifying a vector $\boldsymbol{x}$ with the coset $(\boldsymbol{x}, e)H$ and writing $\mathcal{H} = G/H$. This identification is consistent with the group action. Indeed, if $h = (\boldsymbol{0}, \sigma) \in H$ and $gH,\ g = (\boldsymbol{x}, e) \in G$ is the coset that corresponds to a vector $\boldsymbol{x} \in \mathcal{H}$, then

$$h(gH) = (\sigma(\boldsymbol{x}), e)H.$$

An orbit of this action contains all the vectors obtained from $\boldsymbol{x}$ after permuting its coordinates, i.e., $H(gH) = \{\boldsymbol{x}' \in \mathcal{H} : \mathrm{w}_H(\boldsymbol{x}') = \mathrm{w}_H(\boldsymbol{x})\}$. In other words, the set of double cosets $H\backslash G/H$ represents the set of spheres $\mathcal{S}_k \subset \mathcal{H}, k = 0, 1, \ldots, n$, where $\mathcal{S}_k = \{\boldsymbol{x} \in \mathcal{H}(2, n) : \mathrm{w}_H(\boldsymbol{x}) = k\}$. The spheres form a partition of $\mathcal{H}$.

2.4.2 Decomposition of functions on the binary Hamming space

Consider the space $L^2(\mathcal{H})$ of real "square-integrable" functions on $\mathcal{H}$, i.e. functions $f : \mathcal{H} \to \mathbb{R}$ with the inner product

$$\langle f_1, f_2 \rangle = \frac{1}{|\mathcal{H}|} \sum_{\boldsymbol{x} \in \mathcal{H}} f_1(\boldsymbol{x}) f_2(\boldsymbol{x})$$

and the norm $\|f\| = \langle f, f \rangle^{1/2}$. The group G acts on $L^2(\mathcal{H})$ by $g \cdot f(\boldsymbol{x}) = f(g^{-1} \cdot \boldsymbol{x})$. Under this action the space $L^2(\mathcal{H})$ decomposes into a direct sum of pairwise orthogonal irreducible subspaces

$$L^2(\mathcal{H}) = H_0 \perp \cdots \perp H_n, \tag{2.13}$$

where each H_k is a subspace of functions invariant under the action of S_n. This is a classical result in harmonic analysis that for the Hamming case is discussed in detail in [32, 85]. The dimension of the space H_k is $\binom{n}{k}$. An orthogonal basis of the space H_k is given by the functions $e_{k,j}(\boldsymbol{x}) = (-1)^{x_{i_1} + \cdots + x_{i_k}}$, where $1 \le i_1 < \cdots < i_k \le n$ is the j-th k-subset of $[n]$ (under some numbering of such subsets).

The analysis of the space $L^2(\mathcal{H})$ is accomplished by studying functions on the group G [32]. As above, let H be the stationary subgroup of the point $\boldsymbol{x}_0 = \mathbf{0}$. A real function $\phi(g), g \in G$ is called spherical with respect to H if it is constant on a (left) coset of H, i.e., $\phi(gh) = \phi(g)$ for all $h \in H$. A spherical function can be thought of as an element of the space $L^2(\mathcal{H})$. Under this approach the basis $\{e_{k,j} : 0 \le k \le n, 1 \le j \le \binom{n}{k}\}$ is identified as the basis of *spherical harmonics*[1]. Of particular interest are *zonal spherical functions* ϕ_k, i.e., functions constant on double cosets $H\backslash G/H$, and the associated zonal spherical kernels. It is known that there exists a unique, up to a constant factor, zonal spherical kernel of degree k, given by the expression

$$K_k(\boldsymbol{x}, \boldsymbol{y}) = \sum_{j=1}^{\binom{n}{k}} e_{k,j}(\boldsymbol{x}) e_{k,j}(\boldsymbol{y}). \tag{2.14}$$

It is readily seen that $K_k(\boldsymbol{x}, \boldsymbol{y})$ is invariant under the action of H: for any permutation σ we have $K_k(\sigma \cdot \boldsymbol{x}, \sigma \cdot \boldsymbol{y}) = K_k(\boldsymbol{x}, \boldsymbol{y})$. The corresponding zonal spherical harmonic ϕ_k satisfies

$$\phi_k(h_1 g h_2) = \phi_k(g), \quad h_1, h_2 \in H,\ g \in G.$$

Moreover, if $g_{\boldsymbol{x}}$ and $g_{\boldsymbol{y}}$ are such that $g_{\boldsymbol{x}} \cdot \mathbf{0} = \boldsymbol{x}$ and $g_{\boldsymbol{y}} \cdot \mathbf{0} = \boldsymbol{y}$. then

$$K_k(\boldsymbol{x}, \boldsymbol{y}) = \phi_k(g_{\boldsymbol{y}}^{-1} g_{\boldsymbol{x}}).$$

Furthermore, the function $K_k(\boldsymbol{x}, \boldsymbol{y})$ can be easily seen to depend only on the Hamming distance $d(\boldsymbol{x}, \boldsymbol{y})$ and thus becomes a univariate polynomial. It has the following explicit expression:

$$K_k(x; n) = \sum_{i=0}^{k} (-1)^i \binom{x}{i} \binom{n-x}{k-i}, \tag{2.15}$$

where x is a "discrete" real variable. The polynomials $K_k(x; n), k = 0, 1, \dots, n$ form an orthogonal basis in the space of univariate polynomials of x with respect to the weights $w(i) = \binom{n}{i} 2^{-n},\ i = 0, \dots, n$. They form a particular case of Krawtchouk polynomials well known in the theory of classical orthogonal polynomials [81]. Note that this weight function is inherited from the inner product on $L^2(\mathcal{H})$: the uniform probability distribution becomes binomial once we pass from functions of $\boldsymbol{x}$ to functions of x. Consequently, the inner product on $L^2(\mathcal{H})$ changes into

$$\langle f_1, f_2 \rangle = \sum_{i=0}^{n} w(i) f_1(i) f_2(i). \tag{2.16}$$

Decomposition (2.13) implies that any polynomial $f(x)$ of degree at most n can be written as a Fourier series

$$f(x) = \sum_{i=0}^{n} f_i K_i(x; n), \quad f_i \in \mathbb{R},$$

[1] The term is inherited from functions on the sphere in $\mathbb{R}^n$.

where $f_i = \langle f, K_i\rangle / \langle K_i, K_i\rangle$.

Remark: The group view in the last two sections applies to all association schemes including the Johnson scheme and the ordered Hamming scheme, as well as to infinite homogeneous spaces (see, for instance, Bachoc [8]).

2.4.3 Properties of the Krawtchouk polynomials

The binary case: Consider the space V_n of all real polynomials in one variable x defined on $\{0, \ldots, n\}$, and define the inner product between two polynomials $f(x), g(x)$ on $\{0, \ldots, n\}$ as in (2.16). Let $K_k(x; n)$ be the Krawtchouk polynomial of degree k. It satisfies the orthogonality relation

$$\langle K_k, K_j\rangle = \binom{n}{k}\delta_{kj}.$$

The normalized Krawtchouk polynomial is given by

$$\tilde{K}_k(x; n) = \frac{K_k(x; n)}{\sqrt{\binom{n}{k}}},$$

and thus

$$\langle \tilde{K}_k, \tilde{K}_j\rangle = \delta_{kj}.$$

The Krawtchouk polynomials satisfy the following three-term recurrence relation

$$(n - 2x)\tilde{K}_k(x; n) = a_k\tilde{K}_{k+1}(x; n) + a_{k-1}\tilde{K}_{k-1}(x; n). \tag{2.17}$$

where $a_k = \sqrt{(n-k)(k+1)}$. We also have from (2.3)

$$\tilde{K}_i(x; n)\tilde{K}_j(x; n) = \sum_{k=0}^{n} \tilde{p}^k_{i,j}\tilde{K}_k(x; n), \quad x = 0, \ldots, n, \tag{2.18}$$

where the numbers $\tilde{p}^k_{i,j} \geq 0$ are related in an obvious way to the intersection numbers of the Hamming scheme (Section 2.1). The k-th polynomial kernel (the Christoffel-Darboux kernel) $U_k(x, a)$ is defined as

$$U_k(x, a) = \sum_{i=0}^{k} \tilde{K}_i(x; n)\tilde{K}_i(a; n). \tag{2.19}$$

It has the following reproducing property: $\langle U_k(\cdot, a), f(\cdot)\rangle = f(a)$ for any polynomial $f, \deg(f) \leq k$.

Finally,

$$\tilde{K}_k(0; n) = \sqrt{\binom{n}{k}}. \tag{2.20}$$

The q-ary case, $q \geq 2$: It is possible to extend the considerations in this section to the case of a q-ary Hamming space with arbitrary q. The corresponding Krawtchouk polynomials have the following form:

$$K_k(x;n) = \sum_{i=0}^{k} (-1)^i (q-1)^{k-i} \binom{x}{i} \binom{n-x}{k-i}, \quad k = 0, 1, \dots, n. \tag{2.21}$$

The value of q will be clear from the context, so we keep the same notation. They are orthogonal on $\{0, 1, \dots, n\}$ with weights $\binom{n}{i}(q-1)^i q^{-n}$, $i = 0, \dots, n$. We will use the following properties of the polynomials $K_k(x;n)$ whose proofs are found for instance in [54]. Let $x_i(n,k)$, $i = 1, \dots, k$ be the roots of $K_k(x;n)$ in the ascending order. Then

$$0 < x_i(n-1,k) < x_i(n,k) < x_i(n-1,k-1) < x_i(n,k-1) < x_{i+1}(n,k) < n, \\ 1 < k < n, \ i = 1, \dots, k-1. \tag{2.22}$$

Let $n \to \infty, k/n \to y$. Then

$$\lim_{n \to \infty} \frac{x_1(n,k)}{n} = \gamma(y) \triangleq \frac{q-1}{q} - \frac{q-2}{q} y - \frac{2}{q} \sqrt{(q-1)y(1-y)}. \tag{2.23}$$

The polynomials $K_k(x;n)$ satisfy the recurrence

$$K_k(x;n) = K_k(x;n-1) + (q-1)K_{k-1}(x;n-1) \tag{2.24}$$

and the Christoffel-Darboux formula

$$q(x-y) \sum_{k=0}^{l} \frac{K_k(x;n)K_k(y;n)}{K_k(0;n)} = \frac{l+1}{K_l(0;n)} \big(K_{l+1}(y;n)K_l(x;n) - K_{l+1}(x;n)K_l(y;n) \big). \tag{2.25}$$

Notes: The theory on association schemes is presented in detail in the works of Delsarte [26], Bannai and Ito [9], and Brouwer, Cohen and Neumaier [23]. The Hamming and the Johnson schemes were identified in [26] and were extensively studied thereafter (see [9] or MacWilliams and Sloane [55]). The material on the ordered Hamming scheme is a summary of the theory presented in Martin and Stinson [60]. The discussion on the Terwilliger algebra is an outline of the basic concepts from Schrijver [74] and Vallentin [85]. The discussion on Krawtchouk polynomials in Section 2.4 is a short summary of the general theory in Vilenkin [86] and in Kabatyanskii and Levenshtein [52]. Dunkl's paper [32] treats the case of the Hamming space in detail (see also Chapter 6). The properties of the univariate Krawtchouk polynomials are present in MacWilliams and Sloane [55] and in Levenshtein [54].

CHAPTER 3

Bounds on codes in the Hamming space

This chapter presents an overview of the methods used to derive the bounds on the size of the codes in the binary Hamming space $\mathcal{H} = \mathcal{H}(2, n)$. Even though this setting is restrictive, the techniques are universal and have been used to address similar problems in a vast range of metric spaces. The methods presented here are classical in nature; however, their adaptation to different special cases requires nontrivial extensions and computations. The study of various special cases continues to this day [7, 10, 59], representing an established branch of coding theory.

Below, we provide only the bounds and techniques which are used in the later chapters. In Section 3.1 we present the Gilbert-Varshamov bound, the Plotkin bound, the Johnson bound, and the Bassalygo-Elias bound on the size of codes. In Section 3.2 we discuss the linear programming (LP) bound.

None of the results in this chapter are new. The material in Section 3.1 can be found in MacWilliams and Sloane [55]. The material in Section 3.2 draws on the ideas of Bachoc [7] and can be found in Barg and Nogin [10].

3.1 Sphere-covering and averaging bounds

The main argument used in proving existence of large-size codes is given in the following theorem.

Theorem 3.1 (Gilbert-Varshamov bound[1]) *There exists a code $\mathcal{C}$ of size M and minimum distance d in $\mathcal{H}$ provided*

$$M \geq \frac{2^n}{\sum_{i=0}^{d-1} \binom{n}{i}}.$$

Proof: The proof proceeds by a greedy sphere covering argument. Pick $\boldsymbol{c}_1 \in \mathcal{H}$ and let $\mathcal{C} = \{\boldsymbol{c}_1\}$.

1. In the i-th step pick a point $\boldsymbol{c}_{i+1} \in \mathcal{H}$ outside $\cup_{j=1}^{i} B_{d-1}(\boldsymbol{c}_j)$, where $B_{d-1}(\boldsymbol{c}_j)$ is a ball of radius $d-1$ around $\boldsymbol{c}_j$. Let $\mathcal{C} = \mathcal{C} \cup \{\boldsymbol{c}_{i+1}\}$.
2. Increment i by 1 and repeat the previous step till no such point $\boldsymbol{c}_{i+1} \in \mathcal{H}$ exists. ■

[1] Although not quite accurate, this name is often associated with the technique presented here.

Below we state some upper bounds (non-existence results) on the size of the code $\mathcal{C}$. Call the matrix whose rows are the codewords of $\mathcal{C}$ the *codematrix* of $\mathcal{C}$.

Theorem 3.2 (Plotkin bound) *Let $\mathcal{C} \subset \mathcal{H}$ be an (n, M, d) code of size M and minimum distance d. Then*

$$M \leq \frac{2d}{2d-n}, \quad 2d > n.$$

Proof: The proof proceeds by an averaging argument on all pairs of distances in the code. The sum $\sum_{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}} d_H(\boldsymbol{u},\boldsymbol{v})$ is bounded in two ways, as shown below. First

$$M(M-1)d \leq \sum_{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}} d_H(\boldsymbol{u},\boldsymbol{v}).$$

Next, consider the $M \times n$ codematrix of $\mathcal{C}$. Let ν_1^l be the number of ones in the l-th column of the codematrix. Then

$$\begin{aligned}\sum_{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}} d_H(\boldsymbol{u},\boldsymbol{v}) &= \sum_{l=1}^{n} 2\nu_1^l(M-\nu_1^l) \\ &\leq \frac{nM^2}{2},\end{aligned}$$

where the inequality in the last step is obtained by maximizing over all $0 \leq \nu_1^l \leq M,\ l = 1, \ldots, n$. ■

The Johnson bound is an upper bound on the size of codes in the sphere $\mathcal{S}_w$ of radius w in $\mathcal{H}$. This bound is also proved by estimating the average distance in the code. It is an analog of the Plotkin bound for the constant weight space $\mathcal{S}_w$.

Theorem 3.3 (Johnson bound) *Let $\mathcal{C}$ be an (n, M, d) code in $\mathcal{S}_w$. For any $w < \frac{n}{2}\left(1 - \sqrt{1 - \frac{2d}{n}}\right)$,*

$$M \leq \left\lfloor \frac{dn}{dn - 2wn + 2w^2} \right\rfloor. \tag{3.1}$$

Proof: Let ν_1^l be the number of ones in the l-th column of the $M \times n$ codematrix of $\mathcal{C}$. The total number of ones in the codematrix is $Mw = \sum_{l=1}^{n} \nu_1^l$. Then

$$\begin{aligned}M(M-1)d &\leq \sum_{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}} d_H(\boldsymbol{u},\boldsymbol{v}) \\ &= \sum_{l=1}^{n} 2\nu_1^l(M-\nu_1^l) \\ &\leq \frac{M^2}{n}(2wn - 2w^2),\end{aligned}$$

where in the last step we maximize under the constraints $0 \leq \nu_1^l \leq M,\ l = 1, \ldots, n$, and $\sum_{l=1}^{n} \nu_1^l = Mw$. The condition on w in the theorem arises from the requirement that the denominator be positive. ■

Bassalygo-Elias bound: It is obtained by a conjunction of the "Bassalygo-Elias inequality" and the Johnson bound. The inequality is stated in the lemma below.

Lemma 3.4 *Let $\mathcal{C}$ be an (n, M, d) code in the Hamming space $\mathcal{H}$, $\mathcal{S}_w$ be a sphere in $\mathcal{H}$, and let $A(n, d, w)$ denote the maximum size of a code of minimum distance d in $\mathcal{S}_w$. Then*

$$M|\mathcal{S}_w| \le A(n, d, w)2^n,$$

Proof: We have

$$\sum_{\boldsymbol{x}\in\mathcal{H}} |(\mathcal{C} - \boldsymbol{x}) \cap \mathcal{S}_w| = \sum_{\boldsymbol{x}\in\mathcal{H}} \sum_{\boldsymbol{c}\in\mathcal{C}} 1\big((\boldsymbol{c} - \boldsymbol{x}) \in \mathcal{S}_w\big) = \sum_{\boldsymbol{c}\in\mathcal{C}} \sum_{\boldsymbol{x}\in\mathcal{H}} 1\big((\boldsymbol{c} - \boldsymbol{x}) \in \mathcal{S}_w\big) = |\mathcal{C}||\mathcal{S}_w|.$$

Since for every $\boldsymbol{x}$, $(\mathcal{C} - \boldsymbol{x}) \cap \mathcal{S}_w$ is a constant weight code, we have

$$\sum_{\boldsymbol{x}\in\mathcal{H}} |(\mathcal{C} - \boldsymbol{x}) \cap \mathcal{S}_w| \le 2^n A(n, d, w).$$

■

Combining the Johnson bound with this lemma, we obtain

Theorem 3.5 (Bassalygo-Elias bound) *Let $\mathcal{C} \subset \mathcal{H}$ be an (n, M, d) code. For any $w < \frac{n}{2}\left(1 - \sqrt{1 - \frac{2d}{n}}\right)$*

$$M \le \frac{2^n dn}{(dn - 2wn + 2w^2)|\mathcal{S}_w|}.$$

Let us find the asymptotic version of the Bassalygo-Elias bound. Let $A(n, d)$ be the maximum size of a code $\mathcal{C} \in \mathcal{H}$ of distance d. Let $\delta = d/n$ be the *relative distance* of the code and let

$$R(\delta) = \limsup_{n\to\infty} \frac{1}{n} \log_2 A(n, \lfloor \delta n \rfloor).$$

be the *asymptotic rate* of the code. Let $w = \frac{n}{2}\Big(1 - \sqrt{1 - \frac{2d}{n}}\Big) - 1$. From the Theorem 3.5 we obtain the following asymptotic upper bound:

$$R(\delta) \le 1 - h_2\Big(\frac{1}{2} - \frac{1}{2}\sqrt{1 - 2\delta}\Big),$$

where $h_2(\cdot)$ is the binary entropy function. It is known that $R(\delta) = 0$ for $\delta \ge 1/2$ and $R(0) = 1$; otherwise the exact behavior of $R(\delta)$ is unknown.

Remark: In Lemma 3.4 the set $\mathcal{S}_w$ can be replaced by any subset $B \subset \mathcal{H}$. Then on the right-hand side the quantity $A(n, d, w)$ is replaced by the maximum size of a code with distance d in the set B. Generally estimating this size is difficult, which is why we state this lemma for $B = \mathcal{S}_w$. In [77] $\mathcal{S}_w$ is replaced with another subset, resulting in an improvement to the Bassalygo-Elias bound on codes in the non-binary Hamming space.

3.2 Krawtchouk polynomials and the LP bound

3.2.1 The method

In the next section we apply the Delsarte method (Theorem 2.2) to derive an improvement of the Bassalygo-Elias bound (Theorem 3.5). Our aim is to derive the result of McEliece et al. [62] which uses the Delsarte method to prove the best known asymptotic bounds on the size of a code. The proof method employed below is different from the one used in [62]. It will be generalized for the case of the ordered Hamming space in Chapter 5.

To derive a bound on the size M of the code we need to choose a polynomial that satisfies the conditions of Theorem 2.2. Without loss of generality assume that $f(0) = 1$. According to (2.6), we need to find a polynomial with the largest possible value of the expectation $Ef = 2^{-n}\sum_{i=0}^{n} f(i)\binom{n}{i} = f_0$. Computing the stationary point of the functional Ef in some class of polynomials, one finds that on the space V_k of polynomials of degree $\le k$ it is approximated by the Christoffel-Darboux kernel $U_k(x, a)$ (2.19) (see [12]). This approach is taken in [62] and many papers after it. The derivation of the bound relies on the behavior of extremal roots of the Krawtchouk polynomials.

If the polynomials involved in the LP bound are multivariate, the above approach becomes infeasible because we do not have control over the roots. For instance, this is the case for the ordered Hamming space. The spectral method [7], explained below, is a clever argument designed to circumvent this difficulty. It proceeds by finding a linear operator on V_k for which U_k is an eigenfunction, and then performing optimization on k within the limits imposed by the conditions in Theorem 2.2.

These ideas are illustrated in the next section to derive the bound on the code rate [62] for $\mathcal{H}$. The purpose of this derivation, taken from [10], is to clarify the ideas which may be obscured by substantial technical difficulties encountered for the ordered Hamming space.

3.2.2 The bound

Let V_k be the space of univariate polynomials of degree $\le k$ considered as a subspace of the space V_n of polynomials on $\{0, 1, \dots, n\}$ with the inner product (2.16). Below we use regular letters to denote operators acting on V_n and bold letters to denote their matrices in the basis $\{\tilde{K}_i\}$. Let E_k be the orthogonal projection from V_n to V_k. Consider the operator

$$S_k = E_k \circ (n - 2x) : \; V_k \to V_k,$$

i.e., multiplication by $(n - 2x)$ followed by projection on V_k.

The argument that follows relies on the fact that the operator S_k is self-adjoint with respect to the bilinear form $\langle \cdot, \cdot \rangle$. Indeed, both multiplication by a function and the orthogonal projection are self-adjoint operators. Therefore, the matrix $\mathbf{S}_k$

is symmetric. Its explicit form is as follows:

$$\mathbf{S}_k = \begin{bmatrix} 0 & a_0 & 0 & \dots & 0 \\ a_0 & 0 & a_1 & \dots & 0 \\ 0 & a_1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & a_{k-1} \\ 0 & 0 & \dots & a_{k-1} & 0 \end{bmatrix},$$

where the coefficients $a_i, i = 0, 1, \dots, k-1$ are given by (2.17). The matrix $\mathbf{S}_k$ is irreducible (see Definition A.1 in the Appendix). Hence by the Perron-Frobenius theorem (see Theorem A.4) the matrix $\mathbf{S}_k$ has a unique positive maximum eigenvalue $\lambda_{\max}(\mathbf{S}_k)$.

Theorem 3.6 [10] *Let $\mathcal{C} \subset \mathcal{H}$ be an (n, M, d) code. Then*

$$M \le \frac{4(n-k)}{n - \lambda_{\max}(\mathbf{S}_k)} \binom{n}{k}$$

for all k such that $\lambda_{\max}(\mathbf{S}_{k-1}) \ge n - 2d$.

PROOF: Let $g = \sum_{i=1}^{k} g_i \tilde{K}_i \in V_k$. Consider the operator $T_k : V_k \to V_k$ defined by

$$T_k g = S_k g - (n-k) g_k \tilde{K}_k \tag{3.2}$$

and let θ_k be its largest eigenvalue. Recall that $\mathbf{T}_k$ is the matrix of this operator in the basis $\{\tilde{K}_i\}$. ($\mathbf{T}_k$ is the same as $\mathbf{S}_k$ except that $(\mathbf{T}_k)_{k+1,k+1} = -(n-k)$.) Let us "shift" the matrix $\mathbf{T}_k$ by a multiple of the identity matrix I to make all of its elements nonnegative. For instance, we have $\mathbf{T}_k + (n-k)I \ge 0$. By Lemma A.3 and the Perron-Frobenius theorem,

$$\lambda_{\max}(\mathbf{S}_{k-1}) < \theta_k < \lambda_{\max}(\mathbf{S}_k),$$

because the same inequalities hold for the largest eigenvalues of the shifted matrices. Moreover, the eigenvalue θ_k is of multiplicity one. Denote by $f \in V_k$ the eigenvector that corresponds to it. By (2.17) and (3.2) we have

$$(n-2x)f = S_k f + f_k a_k \tilde{K}_{k+1} = \theta_k f + (n-k) f_k \tilde{K}_k + f_k a_k \tilde{K}_{k+1},$$

so

$$f = \frac{(n-k)\tilde{K}_k + a_k \tilde{K}_{k+1}}{n - 2x - \theta_k} f_k.$$

Consider the polynomial $F = (n - 2x - \theta_k) f^2 = \big((n-k)\tilde{K}_k + a_k \tilde{K}_{k+1}\big) f_k f$. By Theorem A.4, f can be chosen to have positive coordinates. Therefore by (2.18), the coefficients of the expansion of F into the basis $\{\tilde{K}_i\}$ are nonnegative. Next, if $n - 2d \le \lambda_{\max}(\mathbf{S}_{k-1})$ then $F(x) \le 0$ for $x \ge d$.

Since multiplication by f is a self-adjoint operator, we compute

$$\begin{aligned} F_0 &= \Big\langle \big((n-k)\tilde{K}_k + a_k\tilde{K}_{k+1}\big) f_k f, 1 \Big\rangle \\ &= f_k \Big\langle \big((n-k)\tilde{K}_k + a_k\tilde{K}_{k+1}\big), f \Big\rangle \\ &= (n-k) f_k^2 > 0, \end{aligned}$$

and

$$F(0) = \frac{\left((n-k)\sqrt{\binom{n}{k}} + a_k\sqrt{\binom{n}{n-k}}\right)^2}{n-\theta_k} f_k^2.$$

Substituting $a_k = \sqrt{(k+1)(n-k)}$ we find

$$F(0) = \frac{4(n-k)^2 f_k^2}{n-\theta_k}\binom{n}{k} < \frac{4(n-k)^2 f_k^2}{n-\lambda_{\max}(\mathbf{S}_k)}\binom{n}{k}$$

provided that $\lambda_{\max}(\mathbf{S}_k) < n$. The claimed estimate is obtained by using the polynomial F in Theorem 2.2. ■

Next we compute the asymptotic behavior of the largest eigenvalue.

Lemma 3.7 *Let $k < n/2$. For all $s = 2, \dots, k+1$,*

$$\lambda_{\max}(\mathbf{S}_k) \ge \frac{2(s-1)}{s}\sqrt{(k-s+2)(n-k+s-1)},$$

$$\lambda_{\max}(\mathbf{S}_k) \le 2\sqrt{k(n-k+1)}.$$

In particular, for $n \to \infty$, $k/n \to \tau$, and $s = o(n)$,

$$\lim_{n\to\infty} \frac{\lambda_{\max}(\mathbf{S}_k)}{n} = 2\sqrt{\tau(1-\tau)}.$$

The first inequality is obtained from Rayleigh-Ritz inequality (Theorem A.1) and the second inequality results from Lemma A.2.

Theorem 3.6 and Lemma 3.7 together lead to the following asymptotic result (the asymptotic MRRW bound for binary codes [62]):

$$R \le h_2\big(1/2 - \sqrt{\delta(1-\delta)}\big).$$

Indeed, let $\limsup_{n\to\infty} \frac{1}{n}\log_2 M = R$, $\lim_{n\to\infty}\frac{d}{n} = \delta$ and assume that $\delta \le 1/2$. We need to choose k so that $n^{-1}\lambda_{\max}(\mathbf{S}_{k-1}) \ge (1-2\delta)(1+o(1))$ as $n \to \infty$. In the limit, this amounts to taking τ that satisfies $2\sqrt{\tau(1-\tau)} \ge 1-2\delta$, or $\tau \ge 1/2 - \sqrt{\delta(1-\delta)}$. The result now follows by the Stirling approximation.

CHAPTER 4

NMDS codes in ordered Hamming space

In this chapter we study a class of near-optimal codes in the ordered Hamming space, called near-Maximum Distance Separable (NMDS) codes. A Maximum Distance Separable (MDS) code in the Hamming space is a set of vectors $\mathcal{C} = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_M\}$ in $\mathcal{H}(q, n)$ such that the minimum distance is d and the size of the code is $M = q^{n-d+1}$. By the well-known Singleton bound of coding theory, this is the maximum possible number of points with the given separation d. If $\mathcal{C}$ is an MDS code that forms an $\mathbb{F}_q$-linear space of dimension k, then the parameters n, k, d, of the code satisfy the relation $d = n - k + 1$. MDS codes are known to be linked to classical old problems in finite geometry and to a number of other combinatorial questions related to the Hamming space [73]. At the same time, the length of MDS codes cannot be very large; in particular, in all the known cases, $n \leq q + 2$. This restriction has led to the study of classes of codes with distance properties close to MDS codes, such as t-th rank MDS codes [87], NMDS codes [28] and almost-MDS codes [22]. The distance of these codes is only slightly less than $n-k+1$, and at the same time they still have many of the structural properties associated with MDS codes.

MDS codes have been extended to the ordered Hamming space, and their properties are well understood [72, 47, 78]. In this chapter we extend the study of NMDS codes to the ordered Hamming space. As observed by Skriganov [78], MDS codes correspond to distributions of points in the unit cube. In particular, distributions that arise from MDS codes are optimal in some well-defined sense. In the same way, NMDS codes correspond to distributions that are not far from optimal (they are characterized exactly in Section 4.2 below).

The ordered Hamming metric is an example of a wide class of distance functions on strings called the poset metrics [24]. We present our results for this general case because this requires only a small additional effort.

We begin with a review of the definition and properties of the poset metric in the next section. This section also includes elements of the theory of linear codes in the poset space that we derive following the approach in the usual Hamming space. After this we define linear NMDS codes over a finite field alphabet $\mathbb{F}_q$ in the poset metric and derive some of their properties. In Section 4.2 we establish a relation between NDMS codes and distributions in the unit cube. Section 4.3 is concerned with the derivation of the weight distribution of NMDS codes. Finally, in Section 4.4 we provide some constructions of NMDS codes in the ordered Hamming space.

Publications: The results of this chapter appear in [15, 16].

4.1 Definitions and basic properties

4.1.1 Poset metrics

We begin with defining poset metrics on q-ary strings of a fixed length and introduce the ordered Hamming metric as a special case of the general definition. Entries of a string $\boldsymbol{x} = (x_1, x_2, \dots)$ are indexed by a finite set $[n] = \{1, \dots, n\}$ which we call the set of coordinates. Let $\overrightarrow{\mathcal{P}}$ be an arbitrary partial order ($\le$) on $[n]$. Together $[n]$ and $\overrightarrow{\mathcal{P}}$ form a *poset*. An *ideal* of the poset is a subset $I \subset [n]$ that is "downward closed" under the $\le$ relation, which means that the conditions $i, j \in [n]$, $j \in I$ and $i \le j$ imply that $i \in I$. For the reasons that will become clear below, such ideals will be called *left-adjusted* (l.a.).

A *chain* is any linearly ordered subset of the poset. The *dual poset* $\overleftarrow{\mathcal{P}}$ is the set $[n]$ with the same set of chains as $\overrightarrow{\mathcal{P}}$, but the order within each of them reversed. In other words $j \le i$ in $\overleftarrow{\mathcal{P}}$ if and only if $i \le j$ in $\overrightarrow{\mathcal{P}}$. An ideal in the dual poset will be termed *right-adjusted* (r.a.). For a subset $S \subseteq \overrightarrow{\mathcal{P}}$ we denote by $\langle S \rangle = \langle S \rangle_{\overrightarrow{\mathcal{P}}}$ the smallest $\overrightarrow{\mathcal{P}}$-ideal containing the set S (we write $S \subseteq \overrightarrow{\mathcal{P}}$ to refer to a subset $S \subseteq [n]$ whose elements are ordered according to $\overrightarrow{\mathcal{P}}$). The support of a sequence $\boldsymbol{x}$ is the subset $\operatorname{supp} \boldsymbol{x} \subseteq [n]$ formed by the indices of all the nonzero entries of $\boldsymbol{x}$. The set $\langle \operatorname{supp} \boldsymbol{x} \rangle \subseteq \overrightarrow{\mathcal{P}}$ will be called the l.a. support of $\boldsymbol{x}$. The r.a. support is defined analogously.

Definition 4.1 (Brualdi et al. [24]) *Let $\overrightarrow{\mathcal{P}}$ be a poset defined on $[n]$ and let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_q^n$ be two strings. Define the weight of $\boldsymbol{x}$ with respect to $\overrightarrow{\mathcal{P}}$ as $\mathrm{w}(\boldsymbol{x}) = |\langle \operatorname{supp} \boldsymbol{x} \rangle|$, i.e., the size of the smallest $\overrightarrow{\mathcal{P}}$-ideal that contains the support of $\boldsymbol{x}$. The distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as $d_{\overrightarrow{\mathcal{P}}}(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{w}(\boldsymbol{x} - \boldsymbol{y}) = |\langle \operatorname{supp}(\boldsymbol{x} - \boldsymbol{y}) \rangle|$.*

A code $\mathcal{C}$ of minimum distance d is a subset of $\mathbb{F}_q^n$ such that any two distinct vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of $\mathcal{C}$ satisfy $d_{\overrightarrow{\mathcal{P}}}(\boldsymbol{x}, \boldsymbol{y}) \ge d$. It is similarly possible to consider codes whose distance is measured relative to $\overleftarrow{\mathcal{P}}$. Given a linear code $\mathcal{C} \subset \mathbb{F}_q^n$ its *dual code* $\mathcal{C}^\perp$ is the set of vectors $\{\boldsymbol{y} \in \mathbb{F}_q^n : \forall\, \boldsymbol{c} \in \mathcal{C}\ \sum_{i=1}^n c_i y_i = 0\}$. The weight of the dual code $\mathcal{C}^\perp$ is considered with respect to the dual poset $\overleftarrow{\mathcal{P}}$. If $\mathcal{C}$ has dimension k then $\mathcal{C}^\perp$ has dimension $n - k$.

A subset of $\mathbb{F}_q^n$ is called an *orthogonal array* of strength t and index θ with respect to $\overrightarrow{\mathcal{P}}$ if any t l.a. columns contain any vector $\boldsymbol{z} \in \mathbb{F}_q^t$ exactly θ times. In particular, the dual of a *linear* poset code is also a *linear* orthogonal array.

For instance, the Hamming metric is defined by the partial order $\overrightarrow{\mathcal{P}}$ which is a single antichain of length n (no two elements are comparable). Accordingly, the distance between two sequences is given by the number of coordinates in which they differ. In this case, $\overrightarrow{\mathcal{P}} = \overleftarrow{\mathcal{P}}$.

The *ordered Hamming metric*, which is introduced in Chapter 2, page 12, can be defined by a poset $\overrightarrow{\mathcal{P}}$ which is a disjoint union of n chains of equal length r. We recall that a vector (sequence) is written as $\boldsymbol{x} = (x_{11}, \dots, x_{1r}; \dots; x_{n1}, \dots, x_{nr}) \in \mathbb{F}_q^{n,r}$. According to Definition 4.1, the weight of $\boldsymbol{x}$ is given by (see also Definition 2.3)

$$\mathrm{w}_r(\boldsymbol{x}) = \sum_{i=1}^{n} \max\{j : x_{ij} \neq 0\}.$$

If e is the shape of $\boldsymbol{x}$ then the weight is equivalently written as $\mathrm{w}_r(\boldsymbol{x}) = |e|'$. For $I = \langle \operatorname{supp} \boldsymbol{x} \rangle$ we will denote the shape of the ideal I as shape(I)= e. By analogy with the properties of ideals in the ordered Hamming space, we use the term "left-adjusted" for ideals in general posets $\overrightarrow{\mathcal{P}}$.

Recall that a linear ordered code $\mathcal{C} \subset \mathbb{F}_q^{n,r}$ with parameters $[nr, k, d]$ is a linear subspace of dimension k and minimum ordered distance d. The dual $\mathcal{C}^{\perp}$ of a linear code has its distance derived from the dual order $\overleftarrow{\mathcal{P}}$, i.e., from r.a. ideals in $\overleftarrow{\mathcal{H}}(q, n, r)$.

The notion of orthogonal arrays in the ordered Hamming space is derived from the general definition given above. As mentioned in Section 2.2 they are called *ordered orthogonal arrays* (OOAs). We write (t, n, r, q) OOA for an orthogonal array of strength t in $\mathbb{F}_q^{n,r}$.

4.1.2 NMDS poset codes

We begin our study of NMDS codes in the poset space with several definitions that are generalized directly from the corresponding definitions in the Hamming space [87, 28]. The t-th *generalized poset weight* of a linear $[n, k]$ code $\mathcal{C}$ is defined as

$$d_t(\mathcal{C}) \triangleq \min\{|\langle \operatorname{supp} \mathcal{D} \rangle| : \mathcal{D} \text{ is an } [n, t] \text{ subcode of } \mathcal{C}\},$$

where $\operatorname{supp} \mathcal{D}$ is the union of the supports of all the vectors in $\mathcal{D}$. Note that $d_1(\mathcal{C}) = d$, the minimum distance of the code $\mathcal{C}$. Generalized poset weights have properties analogous to the well-known set of properties of generalized Hamming weights. Below we denote the $(n-k) \times k$ parity check matrix (whose rows are the $n-k$ basis vectors of $\mathcal{C}^{\perp}$) of $\mathcal{C}$ by H, and we denote the generator matrix (whose rows are the k basis vectors of $\mathcal{C}$) of $\mathcal{C}$ by G.

Lemma 4.1 *Let $\mathcal{C}$ be a linear $[n, k]$ poset code in $\mathbb{F}_q^n$. Then*

1. $0 < d_1(\mathcal{C}) < d_2(\mathcal{C}) < \dots < d_k(\mathcal{C}) \leq n$.

2. *Generalized Singleton bound:* $d_t(\mathcal{C}) \leq n - \dim(\mathcal{C}) + t, \quad \forall t \geq 1$.

3. *If $\mathcal{C}^{\perp}$ is the dual code of $\mathcal{C}$ then*

$$\{d_1(\mathcal{C}), d_2(\mathcal{C}), \dots, d_k(\mathcal{C})\} \cup \big(n + 1 - \{d_1(\mathcal{C}^{\perp}), d_2(\mathcal{C}^{\perp}), \dots, d_{n-k}(\mathcal{C}^{\perp})\}\big) = [n].$$

4. *H is the parity check matrix of $\mathcal{C}$ with $d_t(\mathcal{C}) = \delta$ if and only if*

(a) Every $\delta - 1$ l.a. columns of H have rank at least $\delta - t$.

(b) There exist δ l.a. columns of H with rank exactly $\delta - t$.

PROOF:

1. Let $\mathcal{D}_t \subseteq \mathcal{C}$ be a linear subspace such that $|\langle \operatorname{supp} \mathcal{D}_t \rangle| = d_t(\mathcal{C})$ and $\operatorname{rank}(\mathcal{D}_t) = t,\ t \geq 1$. Let $\Omega(\mathcal{D}_t)$ denote the maximal elements of the ideal $\langle \operatorname{supp} \mathcal{D}_t \rangle$. For each coordinate in $\Omega(\mathcal{D}_t)$, $\mathcal{D}_t$ has at least one vector with a nonzero component in that coordinate. We pick $i \in \Omega(\mathcal{D}_t)$ and let $\mathcal{D}_t^i$ be obtained by retaining only those vectors $\boldsymbol{v}$ in $\mathcal{D}_t$ which have $v_i = 0$. Then
$$d_{t-1}(\mathcal{C}) \leq |\langle \operatorname{supp} \mathcal{D}_t^i \rangle| \leq d_t(\mathcal{C}) - 1.$$

2. This is a consequence of the fact that $d_{t+1} \geq d_t + 1$ and $d_k \leq n$.

3. This proof is analogous to [87]. The reason for giving it here is to assure oneself that no complications arise from the fact that the weights in $\mathcal{C}^\perp$ are measured with respect to the dual poset.

   We show that for any $1 \leq s \leq n - k - 1$,
$$n + 1 - d_s(\mathcal{C}^\perp) \notin \{d_r(\mathcal{C}) : 1 \leq r \leq k\}.$$

   Let $t = k + s - d_s(\mathcal{C}^\perp)$. We consider two cases (one of which can be void), namely, $r \leq t$ and $r \geq t + 1$ and show that for each of them, $n + 1 - d_s(\mathcal{C}^\perp) \neq d_r(\mathcal{C})$.

   Take an s-dimensional subcode $\mathcal{D}_s \subseteq \mathcal{C}^\perp$ such that $|\langle \operatorname{supp} \mathcal{D}_s \rangle_{\overleftarrow{\mathcal{P}}}| = d_s(\mathcal{C}^\perp)$. Form a parity-check matrix of the code $\mathcal{C}$ whose first rows are some s linearly independent vectors from $\mathcal{D}_s$. Let D be the complement of $\langle \operatorname{supp} \mathcal{D}_s \rangle$ in the set of coordinates. Let the submatrix of H formed of all the columns in D be denoted by $H[D]$. The rank of $H[D]$ is at most $n - k - s$ and its corank (dimension of the null space) is at least
$$|D| - (n - k - s) = n - d_s(\mathcal{C}^\perp) - n + k + s = k + s - d_s(\mathcal{C}^\perp).$$

   Then $d_t(\mathcal{C}) \leq |D| = n - d_s(\mathcal{C}^\perp)$ and so $d_r(\mathcal{C}) \leq n - d_s(\mathcal{C}^\perp)$, $1 \leq r \leq t$.

   Now let us show that $d_{t+i}(\mathcal{C}) \neq n + 1 - d_s(\mathcal{C}^\perp)$ for all $1 \leq i \leq k - t$. Assume the contrary and consider a generator matrix G of $\mathcal{C}$ with the first $t + i$ rows corresponding to the subcode $\mathcal{D}_{t+i} \subseteq \mathcal{C}$ with $|\langle \operatorname{supp} \mathcal{D}_{t+i} \rangle_{\overrightarrow{\mathcal{P}}}| = d_{t+i}(\mathcal{C})$. Let D be the complement of $\langle \operatorname{supp} \mathcal{D}_{t+i} \rangle$ in the set of coordinates. Then $G[D]$ is a $k \times (n - d_{t+i}(\mathcal{C}))$ matrix of rank $k - t - i$. By part 2 of the lemma, $n - d_{t+i}(\mathcal{C}) \geq k - t - i$, so
$$\begin{aligned} \dim\ker(G[D]) &\geq n - d_{t+i}(\mathcal{C}) - k + t + i \\ &= s + i - (d_s(\mathcal{C}^\perp) + n - d_{t+i}(\mathcal{C})) \\ &= s + i - 1, \end{aligned}$$

   where the first equality follows on substituting the value of k and the second one by using the assumption. Hence $d_{s+i-1}(\mathcal{C}^\perp) \leq |D| = d_s(\mathcal{C}^\perp) - 1$, which contradicts part 1 of the lemma.

4. Follows by standard linear-algebraic arguments. ■

Definition 4.2 *A linear code $\mathcal{C}[n,k,d]$ is called NMDS if $d(\mathcal{C}) = n-k$ and $d_2(\mathcal{C}) = n-k+2$.*

Closely related is the notion of *almost-MDS code* where we have only the constraint that $d(\mathcal{C}) = n-k$ and there is no constraint on $d_2(\mathcal{C})$. In this chapter, we focus only on NMDS codes. The next set of properties of NMDS codes can be readily obtained as generalizations of the corresponding properties of NMDS codes in the Hamming space [28].

Lemma 4.2 *Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a linear $[n,k,d]$ code in the poset $\overrightarrow{\mathcal{P}}$.*

1. *$\mathcal{C}$ is NMDS if and only if all the following conditions hold*
   (a) *Every $n-k-1$ l.a. columns of the parity check matrix H are linearly independent.*
   (b) *There exist $n-k$ l.a. linearly dependent columns of H.*
   (c) *Every l.a. $n-k+1$ columns of H are full ranked.*
2. *If $\mathcal{C}$ is NMDS, so is its dual $\mathcal{C}^\perp$.*
3. *$\mathcal{C}$ is NMDS if and only if $d(\mathcal{C}) + d(\mathcal{C}^\perp) = n$.*
4. *If $\mathcal{C}$ is NMDS then there exists an NMDS code with parameters $[n-1,k-1,d]$ and an NMDS code with parameters $[n-1,k,d]$.*

PROOF:

1. Any linear code $\mathcal{C}$ has minimum distance $d = n-k$ if and only if parts (a) and (b) hold. Part (c) follows from Lemma 4.1.

2. From Lemma 4.1 we obtain
$$\{n+1-d_t(\mathcal{C}^\perp), 1 \le t \le n-k\} = \{1,\dots,n-k-1,n-k+1\}.$$
Hence $d(\mathcal{C}^\perp) = k$ and $d_2(\mathcal{C}^\perp) = k+2$.

3. Let $d(\mathcal{C}) + d(\mathcal{C}^\perp) = n$. Then
$$d_2(\mathcal{C}^\perp) \ge d(\mathcal{C}^\perp) + 1 = n - d(\mathcal{C}) + 1,$$
but then by Lemma 4.1, part 3, $d_2(\mathcal{C}^\perp) \ge n - d(\mathcal{C}) + 2$. Next,
$$n \ge d_{n-k}(\mathcal{C}^\perp) \ge d_2(\mathcal{C}^\perp) + n - k - 2 \ge 2n - k - d,$$
which implies that $d \ge n-k$. This leaves us with the possibilities of $d = n-k$ or $n-k+1$, but the latter would imply that $d(\mathcal{C}) + d(\mathcal{C}^\perp) = n+2$, so $d = n-k$. Further, $d_2(\mathcal{C}) \ge n - d(\mathcal{C}^\perp) + 2 = n-k+2$, as required. The converse is immediate.

4. To get a $[n-1, k-1, d]$ NMDS code, delete a column of the parity check matrix H of $\mathcal{C}$ preserving a set of $n-k$ l.a. linearly dependent columns. To get a $[n-1, k, d]$ NMDS code, delete a column of the generator matrix G of $\mathcal{C}$ preserving a set of $k+1$ r.a. columns which contains k r.a. linearly dependent columns. ∎

Lemma 4.3 *Let $\mathcal{C}$ be a linear poset code in $\overrightarrow{\mathcal{P}}$ with distance d and let $\mathcal{C}^{\perp}$ be its dual code. Then the matrix M whose rows are the codewords of $\mathcal{C}^{\perp}$ forms an orthogonal array of strength $d-1$ with respect to $\overleftarrow{\mathcal{P}}$.*

Proof: Follows because (1), $\mathcal{C}^{\perp}$ is the linear span of the parity-check matrix H of $\mathcal{C}$; and (2), any $d-1$ l.a. columns of H are linearly independent. ∎

4.2 NMDS codes and distributions

In this section we prove a characterization of NMDS poset codes and then use this result to establish a relationship between NMDS codes in the ordered Hamming space $\overrightarrow{\mathcal{H}}(q, n, r)$ and uniform distributions of points in the unit cube $U^n = [0,1)^n$. In our study of NMDS codes in the following sections, we analyze the properties of the code simultaneously as a linear code and as a linear orthogonal array.

Define the I-neighborhood of a poset code $\mathcal{C}$ with respect to an ideal I as

$$B_I(\mathcal{C}) = \bigcup_{\boldsymbol{c}\in\mathcal{C}} B_I(\boldsymbol{c}),$$

where $B_I(\boldsymbol{x}) = \{\boldsymbol{v} \in \mathbb{F}_q^n : \operatorname{supp}(\boldsymbol{v}-\boldsymbol{x}) \subseteq I\}$. We will say that a linear k-dimensional code $\mathcal{C}$ forms an *I-tiling* if there exists a partition $\mathcal{C} = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_{q^{k-1}}$ into equal parts such that the I-neighborhoods of its parts are disjoint. If in addition the I-neighborhoods form a partition of $\mathbb{F}_q^n$, we say $\mathcal{C}$ forms a *perfect I-tiling.*

Theorem 4.4 *Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be an $[n,k,d]$ linear code in the poset $\overrightarrow{\mathcal{P}}$. $\mathcal{C}$ is NMDS if and only if*

1. *For any $I \subset \overrightarrow{\mathcal{P}}, |I| = n-k+1$, the code $\mathcal{C}$ forms a perfect I-tiling.*

2. *There exists an ideal $I \subset \overrightarrow{\mathcal{P}}, |I| = n-k$ with respect to which $\mathcal{C}$ forms an I-tiling. No smaller-sized ideals with this property exist.*

Proof: Let $\mathcal{C}$ be NMDS and let I be an ideal of size $n-k+1$. Let $H[I]$ be the submatrix of the parity-check matrix H of $\mathcal{C}$ obtained from H by deleting all the columns not in I. Since $\operatorname{rank}(H[I]) = n-k$, the space $\ker(H[I])$ is one-dimensional. Let $\mathcal{C}_1 = \ker(H[I])$ and let $\mathcal{C}_j$ be the jth coset of $\mathcal{C}_1$ in $\mathcal{C}$, $j = 2, \dots, q^{k-1}$. By Lemma 4.3 the code $\mathcal{C}$ forms an orthogonal array of strength $k-1$ and index q in $\overleftarrow{\mathcal{P}}$. Therefore, every vector $\boldsymbol{z} \in \mathbb{F}_q^{k-1}$ appears exactly q times in the restrictions of the codevectors $\boldsymbol{c} \in \mathcal{C}$ to the coordinates of $J = I^c$. Thus, $\boldsymbol{c}'[J] = \boldsymbol{c}''[J]$ for any two vectors $\boldsymbol{c}', \boldsymbol{c}'' \in \mathcal{C}_i, i = 1, \dots, q^{k-1}$ and $\boldsymbol{c}'[J] \neq \boldsymbol{c}''[J]$ $\boldsymbol{c}' \in \mathcal{C}_i, \boldsymbol{c}'' \in \mathcal{C}_j, 1 \le i < j \le q^{k-1}$. This implies that $\mathcal{C}$ forms a perfect I-tiling, which proves assumption 1 of the theorem. To prove assumption 2, let $\boldsymbol{c}$ be a minimum-weight codeword of $\mathcal{C}$ and let $I = \langle \text{supp}(\boldsymbol{c}) \rangle, |I| = n - k$. Let $\mathcal{C}_1 = \{\alpha \boldsymbol{c} : \alpha \in \mathbb{F}_q\}$ and let $\mathcal{C}_2, \dots, \mathcal{C}_{q^{k-1}}$ be the cosets of $\mathcal{C}_1$ in $\mathcal{C}$. Then $\mathcal{C} = \cup_i \mathcal{C}_i$ forms an I-tiling of $\mathbb{F}_q^n$.

To prove the converse, let $I \subseteq \overrightarrow{\mathcal{P}}, |I| = n-k+1$ be an ideal and let $\mathcal{C}_1, \dots, \mathcal{C}_{q^{k-1}}$ be a partition of $\mathcal{C}$ with $|\mathcal{C}_i| = q$ for all i, that forms a perfect $I-$tiling. This implies that $\boldsymbol{c}'[I^c] \neq \boldsymbol{c}''[I^c]$, $\boldsymbol{c}' \in \mathcal{C}_i, \boldsymbol{c}'' \in \mathcal{C}_j, 1 \le i < j \le q^{k-1}$. In other words, $\mathcal{C}$ forms an orthogonal array with respect to $\overleftarrow{\mathcal{P}}$ of index q and strength $k-1$. We conclude that $d(\mathcal{C}^\perp) = k$ or $k+1$. If it is the latter, then $\mathcal{C}^\perp$ is MDS with respect to $\overleftarrow{\mathcal{P}}$ and so is $\mathcal{C}$ with respect to $\overrightarrow{\mathcal{P}}$, in violation of assumption 2. So $d(\mathcal{C}^\perp) = k$ and $d(\mathcal{C}) \le n - k$. If the inequality is strict, there exists an ideal I of size $< n - k$ that supports a one-dimensional subcode of $\mathcal{C}$. Then $\mathcal{C}$ forms an I-tiling which contradicts assumption 2.

It remains to prove that $d_2(\mathcal{C}) = n-k+2$. Assume the contrary, i.e., that there exists a 2-dimensional subcode $\mathcal{B} \subset \mathcal{C}$ whose l.a. support forms an ideal $I \subset \overrightarrow{\mathcal{P}}$ of size $n - k + 1$. The q^2 vectors of $\mathcal{B}$ all have zeros in I^c which contradicts the fact that $\mathcal{C}$ forms an orthogonal array of index q. ■

Next, we use this characterization to relate codes in the ordered Hamming space $\overrightarrow{\mathcal{H}}(q, n, r)$ to distributions. An idealized uniformly distributed point set $\mathcal{C}$ would satisfy the property that for any measurable subset $A \subset U^n$,

$$\frac{1}{|\mathcal{C}|} \sum_{\boldsymbol{c} \in \mathcal{C}} 1(\boldsymbol{c} \in A) = \text{vol}(A).$$

Distributions that we consider, and in particular (t, m, n)-nets, approximate this property by restricting the subsets A to be boxes with sides parallel to the coordinate axes.

Let

$$\mathcal{E} \triangleq \left\{ \prod_{i=1}^{n} \left[\frac{a_i}{q^{l_i}}, \frac{a_i + 1}{q^{l_i}} \right) : 0 \le a_i < q^{l_i}, 0 \le l_i \le r, 1 \le i \le n \right\}$$

be a collection of *elementary intervals* in the unit cube $U^n = [0, 1)^n$. An arbitrary collection of q^k points in U^n is called an $[nr, k]$ *distribution* in the base q (with respect to $\mathcal{E}$). A distribution is called *optimal* if every elementary interval of volume q^{-k} contains exactly one point [78]. The related notion of (t, m, n)-*nets*, introduced by Niederreiter [67], is obtained if we remove the upper bound on l_i (i.e., allow that $0 \le l_i < \infty$) and require that every elementary interval of volume q^{t-m} contain exactly q^t points.

An ordered code gives rise to a distribution of points in the unit cube via the following procedure. A codevector $\boldsymbol{c} = (c_{11}, \dots, c_{1r}; \dots; c_{n1}, \dots, c_{nr}) \in \overrightarrow{\mathcal{H}}(q, n, r)$ is mapped to $\boldsymbol{x} = (x_1, \dots, x_n) \in U^n$ by letting

$$x_i = \sum_{j=1}^{r} c_{ij} q^{j-r-1}, 1 \le i \le n. \tag{4.1}$$

In particular, an $(m-t,n,r,q)$ OOA of index q^t and size q^m corresponds to a distribution in which every elementary interval of volume q^{t-m} contains exactly q^t points. OOAs are related to (t,m,n)-nets by the following theorem.

Theorem 4.5 (Lawrence [51], Mullen/Schmid [63]) *There exists a (t,m,n)-net if and only if there exists an $(m-t,n,m-t,q)$ OOA of index q^t and size $M=q^m$.*

The relation between ordered MDS codes and optimal distributions was established by Skriganov, as shown in the following theorem.

Proposition 4.6 (Skriganov [78]) *An $[nr,k,d]$ MDS code in the ordered metric exists if and only if there exists an optimal $[nr,k]$ distribution.*

Skriganov [79] also considers the concept of *nearly-MDS* codes whose distance asymptotically tends to the distance of MDS codes, and shows how these codes can give rise to distributions.

The next theorem whose proof is immediate from Theorem 4.4 relates ordered NMDS codes and distributions.

Theorem 4.7 *Let $\mathcal{C}$ be a linear $[nr,k,d]$ code in $\overrightarrow{\mathcal{H}}(q,n,r)$ and let $P(\mathcal{C})$ be the corresponding set of points in U^n. Then $\mathcal{C}$ is NMDS if and only if*

1. *Any elementary interval of volume $q^{-(k-1)}$ has exactly q points of $P(\mathcal{C})$.*
2. *There exists an elementary interval $\prod_{i=1}^{n}\left[0,q^{-l_i}\right)$ of volume q^{-k} containing exactly q points and no smaller elementary intervals of this form containing exactly q points exist.*

Corollary 4.8 *An $[nr,k,d]$ NMDS code $\mathcal{C}$ in the ordered Hamming space forms a $(k-1,n,r,q)$ OOA of index q. The corresponding distribution $P(\mathcal{C})\subset U^n$ forms a $(k-r,k,n)$-net for $k-1\geq r$.*

Remark: Distributions of points in the unit cube obtained from NMDS codes have properties similar to those of distributions obtained from MDS codes. In particular, the points obtained from an $[nr,k,d]$ MDS code in $\overrightarrow{\mathcal{H}}(q,n,r)$ satisfy part 1 of Theorem 4.7 and give rise to a $(k-r,k,n)$-net for $k\geq r$ [78].

4.3 Weight distribution of NMDS codes

In this section we determine the weight distribution of NMDS codes. We will first determine the weight distribution for poset NMDS codes and then specialize to the ordered NMDS codes.

Let $\Omega(I)$ be the set of maximal elements of an ideal I and let $\tilde{I}\triangleq I\setminus\Omega(I)$. Let $\mathcal{C}$ be an NMDS $[n,k,d]$ linear poset code. Let $A_I\triangleq\{\boldsymbol{c}\in\mathcal{C}:\langle\operatorname{supp}\boldsymbol{c}\rangle=I\}$ be the number of codewords with l.a. support exactly I and let $A_s=\sum_{I:|I|=s}A_I$.

Theorem 4.9 *The weight distribution of* $\mathcal{C}$ *has the following form:*

$$A_s = \sum_{I\in\mathcal{I}_s}\sum_{l=0}^{s-d-1}(-1)^l\binom{|\Omega(I)|}{l}(q^{s-d-l}-1)+(-1)^{s-d}\sum_{I\in\mathcal{I}_s}\sum_{J\in\mathcal{I}_d(I),J\supseteq\tilde{I}}A_J,\quad n\geq s\geq d, \tag{4.2}$$

where $\mathcal{I}_s \triangleq \{I\subseteq \overrightarrow{\mathcal{P}} : |I|=s\}$ *and* $\mathcal{I}_s(I)\triangleq\{J : J\subseteq I, |J|=s\}$.

Proof: The computation below is driven by the fact that ideals are fixed by the sets of their maximal elements.

The number of codewords of weight s is given by $A_s = |\cup_{I\in\mathcal{I}_s}\mathcal{C}\cap S_I|$, where $S_I\triangleq\{\boldsymbol{x}\in\mathbb{F}_q^n : \langle\operatorname{supp}\boldsymbol{x}\rangle = I\}$ is the sphere with l.a. support exactly I. The above expression can be written as

$$\Big|\bigcup_{I\in\mathcal{I}_s}\mathcal{C}\cap S_I\Big| = \sum_{I\in\mathcal{I}_s}\Big(|\mathcal{C}\cap B_I^*| - |\bigcup_{J\in\mathcal{I}_{s-1}(I)}\mathcal{C}\cap B_J^*|\Big),$$

where $B_I \triangleq \{\boldsymbol{x}\in\mathbb{F}_q^n : \langle\operatorname{supp}\boldsymbol{x}\rangle_{\overrightarrow{\mathcal{P}}}\subseteq I\}$ and $B_I^*\triangleq B_I\setminus\mathbf{0}$. We determine the cardinality of the last term using the inclusion-exclusion principle.

$$\Big|\bigcup_{J\in\mathcal{I}_{s-1}(I)}\mathcal{C}\cap B_J^*\Big| = \sum_{J\in\mathcal{I}_{s-1}(I)}|\mathcal{C}\cap B_J^*| - \sum_{J_1\neq J_2\in\mathcal{I}_{s-1}(I)}|\mathcal{C}\cap B_{J_1}^*\cap B_{J_2}^*| + \cdots$$
$$+(-1)^{|\Omega(I)|-1}\sum_{J_1\neq\cdots\neq J_{|\Omega(I)|}\in\mathcal{I}_{s-1}(I)}\Big|\mathcal{C}\cap\Big(\bigcap_i B_{J_i}^*\Big)\Big|. \tag{4.3}$$

Since $\mathcal{C}^\perp$ has minimum distance k, $\mathcal{C}$ forms an orthogonal array of strength $k-1$ with respect to the dual poset $\overleftarrow{\mathcal{P}}$. This provides us with an estimate for each individual term in (4.3) as described below. For distinct $J_1,\dots,J_l\in\mathcal{I}_{s-1}(I)$, we let $J\triangleq\cap_{i=1}^l J_i$. Using the fact that J does not contain l maximal elements of I, we get

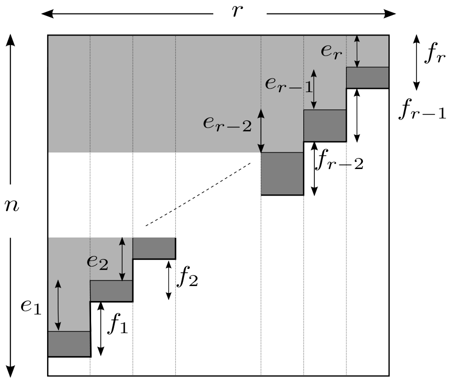
$$\Big|\Big\{\{J_1,\dots,J_l\} : J_i \text{ distinct}, J_i\in\mathcal{I}_{s-1}(I), i=1,\dots,l\Big\}\Big| = \binom{|\Omega(I)|}{l}.$$

For any $s\geq d+1$ consider the complement I^c of an ideal $I\in\mathcal{I}_s$. Since $|I^c|\leq n-d-1=k-1$, the code $\mathcal{C}$ supports an orthogonal array of strength $n-s$ and index q^{s-d} in the coordinates defined by I^c. Since $\cap_{i=1}^l B_{J_i}^* = B_J^*$ and since B_J^* does not contain the $\mathbf{0}$ vector, we obtain

$$\Big|\mathcal{C}\cap\big(\bigcap_{i=1}^l B_{J_i}^*\big)\Big| = q^{s-d-l}-1,\quad 1\leq l\leq s-d-1.$$

Finally, for $l=s-d$ we obtain $|\mathcal{C}\cap(\cap_{i=1}^l B_{J_i}^*)| = A_J$, and

$$\Big|\bigcup_{J\in\mathcal{I}_{s-1}(I)}\mathcal{C}\cap B_J^*\Big| = \sum_{l=1}^{s-d-1}(-1)^{l-1}\binom{|\Omega(I)|}{l}(q^{s-d-l}-1)+\sum_{J\in\mathcal{I}_d(I),J\supseteq\tilde{I}}(-1)^{s-d-1}A_J,$$



Light gray region: ideal J of shape e.
Light gray + Dark gray region: ideal I of shape f.

Figure 4.1: To the proof of Corollary 4.10

which implies

$$\sum_{I\in\mathcal{I}_s}|\mathcal{C}\cap S_I| = \sum_{I\in\mathcal{I}_s}\Bigg((q^{s-d}-1)-\Bigg(\sum_{l=1}^{s-d-1}(-1)^{l-1}\binom{|\Omega(I)|}{l}(q^{s-d-l}-1)$$
$$+\sum_{J\in\mathcal{I}_d(I),J\supseteq\tilde{I}}(-1)^{s-d-1}A_J\Bigg)\Bigg). \qquad \blacksquare$$

As a corollary of the above theorem, we obtain the weight distribution of NMDS codes in the ordered Hamming space $\overrightarrow{\mathcal{H}}(q,n,r)$. By definition, the number of vectors of ordered weight s in a code $\mathcal{C}\in\overrightarrow{\mathcal{H}}(q,n,r)$ equals $A_s=\sum_{e:|e|'=s}A_e$, where A_e is the number of codevectors of shape e.

Corollary 4.10 *The weight distribution of an ordered NMDS code $\mathcal{C}\subset\overrightarrow{\mathcal{H}}(q,n,r)$ is given by*

$$A_s=\sum_{l=0}^{s-d-1}(-1)^l\left(\sum_{e:|e|'=s}\binom{|e|}{l}\binom{n}{e_0,\dots,e_r}\right)(q^{s-d-l}-1)$$
$$+(-1)^{s-d}\sum_{e:|e|'=d}N_s(e)A_e, \quad s=d,d+1,\dots,n, \quad (4.4)$$

where

$$N_s(e)\triangleq\sum_{f:|f|'=s}\binom{e_{r-1}}{f_r-e_r}\binom{e_{r-2}}{(f_r+f_{r-1})-(e_r+e_{r-1})}\cdots\binom{e_0}{|f|-|e|}.$$

Proof: Recall that the shape of an ideal I is $\text{shape}(I)=e=(e_1,\dots,e_r)$, where $e_j, j=1,\dots,r$ is the number of chains of length j contained in I. We obtain $|\Omega(I)|=|e|$ and

$$\sum_{I\in\mathcal{I}_s}\binom{|\Omega(I)|}{l}=\sum_{e:|e|'=s}\binom{|e|}{l}\binom{n}{e_0,\dots,e_r}.$$

To determine the last term in (4.2), we rewrite it as

$$\sum_{I\in\mathfrak{I}_s}\sum_{J\in\mathfrak{I}_d(I),J\supseteq\tilde{I}} A_J = \sum_{J\in\mathfrak{I}_d} |\{I\in\mathfrak{I}_s : \tilde{I}\subseteq J\subseteq I\}|A_J$$
$$= \sum_{e:|e|'=d} N_s(e) \sum_{J:\text{shape}(J)=e} A_J,$$

where $N_s(e) = |\{I\in\mathfrak{I}_s : \tilde{I}\subseteq J\subseteq I, J \text{ fixed}, \text{shape}(J)=e\}|$.

Clearly, $\sum_{J:\text{shape}(J)=e} A_J = A_e$, and so we only need to determine the quantity $N_s(e)$ in the above summation. Let J be an ideal as shown in Fig. 4.1. The ideals I which satisfy the constraints in the set defined by $N_s(e)$ have the form as shown in Fig. 4.1. Letting $f = \text{shape}(I)$, we note that the components of the shape f must satisfy

$$\begin{aligned} f_r &\geq e_r, \\ f_r + f_{r-1} &\geq e_r + e_{r-1} \geq f_r, \\ &\vdots \\ f_1+\cdots+f_r = |f| &\geq |e| = e_1+\cdots+e_r \geq f_2+\cdots+f_s, \\ \text{and } |f|' &= s. \end{aligned}$$

It is now readily seen that the cardinality of the set

$$\{I\in\mathfrak{I}_s : \tilde{I}\subseteq J\subseteq I, J \text{ fixed}, \text{shape}(J)=e\}$$

is given by the formula for $N_s(e)$ as described in (4.4). ■

Remark: For $r=1$ we obtain $|e|=|e|'=e_1=d, |f|=f_1=s$ and $N_s(e)=\binom{n-d}{s-d}$. Thus we recover the expression for the weight distribution of an NMDS code in Hamming space [28]:

$$A_s = \sum_{l=0}^{s-d-1}(-1)^l\binom{s}{l}\binom{n}{s}(q^{s-d-l}-1)+(-1)^{s-d}\binom{n-d}{s-d}A_d. \tag{4.5}$$

Unlike the case of poset MDS codes [47], the weight distribution of NMDS codes is not completely known until we know the number of codewords with l.a. support J for every ideal of weight J of size d. In particular, for NMDS codes in the ordered Hamming space we need to know the number of codewords of every shape e with $|e|'=d$. This highlights the fact that the combinatorics of codes in the poset space (resp. ordered Hamming space) is driven by ideals (resp. shapes) and their support sizes, and that the weight distribution is a derivative invariant of those more fundamental quantities.

As a final remark we observe that, given that $d(\mathcal{C})=n-k$, the assumption $d(\mathcal{C}^\perp)=k$ (or the equivalent assumption $d_2(\mathcal{C})=n-k+2$) ensures that the only unknown components of the weight distribution of $\mathcal{C}$ correspond to ideals of size d. If instead we consider a code of defect s, i.e., a code with $d(\mathcal{C})=(n-k+1)-s$, $s\geq 2$, it will be possible to compute its weight distribution using the components $A_J, d\leq |J|\leq n-d(\mathcal{C}^\perp)$ (provided that we know $d(\mathcal{C}^\perp)$). In the case of the Hamming metric this was established in [35].

4.4 Constructions of NMDS codes

In this section we present some simple constructions of NMDS codes in the ordered Hamming space for the cases $n = 1, 2, 3$. We are not aware of any general code family of NMDS codes for larger n.

n=1: For $n = 1$ the construction is quite immediate once we recognize that an NMDS $[r, k, d]$ code is also an OOA of r.a. strength $k - 1$ and index q. Let I_l denote the identity matrix of size l. Let $\boldsymbol{x} = (x_1, \dots, x_r)$ be any vector of l.a. weight $d = r - k$, i.e. $x_d \neq 0$ and $x_l = 0$, $l = d+1, \dots, r$. Then the following matrix of size $k \times r$ generates an NMDS code with the above parameters

$$\begin{bmatrix} x_1 \dots x_d & 0 & \mathbf{0} \\ M & \mathbf{0} & I_{k-1} \end{bmatrix}, \tag{4.6}$$

where the $\mathbf{0}$s are zero vectors (matrices) of appropriate dimensions and $M \in \mathbb{F}_q^{(k-1)\times d}$ is any arbitrary matrix.

n=2: Let $D_l = \begin{bmatrix} 0 & \dots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \dots & 0 \end{bmatrix}$ be the $l \times l$ matrix with 1 along the inverse diagonal and 0 elsewhere. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be two vectors of length r in $\overrightarrow{\mathcal{H}}(q, 1, r)$ and l.a. weights $r - k_1$ and $r - k_2$ respectively and let $K = k_1 + k_2$. The following matrix generates a $[2r, K, 2r - K]$ linear NMDS code in $\overrightarrow{\mathcal{H}}(q, 2, r)$,

$$\left[\begin{array}{cccc|cccc} u_1 \dots u_{r-k_1-1} & u_{r-k_1} & 0 & \mathbf{0} & v_1 \dots v_{r-k_2-1} & v_{r-k_2} & 0 & \mathbf{0} \\ \mathbf{0} & 0 & 1 & 0 & \mathbf{0} & 0 & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & I_{k_1-1} & E_r(k_1, k_2) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ E_r(k_2, k_1) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & I_{k_2-1} \end{array}\right],$$

where $E_r(i, j)$ is an $(i - 1) \times (r - j - 1)$ matrix which has the following form:

$$E_r(i,j) = \begin{cases} \left[\begin{array}{c} D_{r-j-1} \\ \hline \mathbf{0}_{(i+j-r)\times(r-j-1)} \end{array}\right], & i + j > r, \\ \\ \left[\begin{array}{c|c} \mathbf{0}_{(i-1)\times(r-i-j)} & D_{i-1} \end{array}\right], & i + j \leq r. \end{cases}$$

From the form of the generator matrix it can be seen that any $K - 1$ r.a. columns of the above matrix are linearly independent. But the last k_1 and k_2 columns from the first and the second blocks respectively are linearly dependent. This implies that it forms an OOA of r.a. strength exactly $K - 1$. Hence the dual of the code has distance K. Finally, the minimum weight of any vector produced by this generator matrix is $2r - K$. Hence by Lemma 4.2, this matrix generates an NMDS code.

n=3: For $n = 3$, we have an NMDS code with very specific parameters. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \overrightarrow{\mathcal{H}}(q, 1, r)$ be three vectors of l.a. weight $r - 2$ each. Then the matrix shown below is the generator matrix of a $[3r, 6, d]$ code in base $q \geq 3$. It is formed of three blocks, corresponding to the three dimensions given by n. Here $\mathbf{0}$ is a $1 \times (r-6)$ zero vector.

$$\begin{bmatrix} u_1 \dots u_{r-6} & u_{r-5} & u_{r-4} & u_{r-3} & u_{r-2} & 0 & 0 \\ \mathbf{0} & 0 & 0 & 0 & 0 & 1 & 0 \\ \mathbf{0} & 0 & 1 & 0 & 0 & 1 & 0 \\ \mathbf{0} & 1 & 0 & 0 & 0 & 0 & 1 \\ \mathbf{0} & 0 & 1 & 0 & 0 & 0 & 1 \\ \mathbf{0} & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$
$$\begin{bmatrix} v_1 \dots v_{r-6} & v_{r-5} & v_{r-4} & v_{r-3} & v_{r-2} & 0 & 0 \\ \mathbf{0} & 0 & 0 & 0 & 0 & 1 & 0 \\ \mathbf{0} & 0 & 1 & 0 & 0 & 0 & 0 \\ \mathbf{0} & 0 & 0 & 1 & 0 & 0 & 0 \\ \mathbf{0} & 0 & 1 & 0 & 0 & 0 & 1 \\ \mathbf{0} & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} w_1 \dots w_{r-6} & w_{r-5} & w_{r-4} & w_{r-3} & w_{r-2} & 0 & 0 \\ \mathbf{0} & 0 & 1 & 0 & 0 & 0 & 0 \\ \mathbf{0} & 0 & 0 & 0 & 0 & 1 & 0 \\ \mathbf{0} & 0 & 1 & 0 & 0 & 0 & 1 \\ \mathbf{0} & 0 & 0 & 1 & 0 & 0 & 0 \\ \mathbf{0} & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

CHAPTER 5

Bounds on ordered codes and orthogonal arrays

This chapter is devoted to the derivation of new upper bounds on codes $\mathcal{C}$ in the ordered Hamming space $\overrightarrow{\mathcal{H}}(q, n, r)$. The duality between ordered codes and ordered orthogonal arrays (OOAs) (Section 2.2) and the relation between OOAs and (t, m, n)-nets (Theorem 4.5) imply that upper bounds on codes translate into lower bounds on OOAs and (t, m, n)-nets. Lower bounds on (t, m, n)-nets are of interest because of their use in determining the error of numerical integration of functions on the unit cube [67]. In particular, given the error of integration, we require the function to be sampled at as few points as possible. Lower bounds on (t, m, n)-nets provide limits on the minimum possible size of the set of sampling points.

Aside from the application to numerical integration, the ordered Hamming space has applications to wireless fading channels [40, 82], linear complexity of sequences [61] and to list decoding of Reed-Solomon codes [65]. Thus, bounds on the size of codes are of interest in determining the maximum size of the code that can be achieved given the minimum distance. If q grows with n then the exact asymptotic tradeoff between the size and the minimum distance of the code is given by the Singleton bound [72]. Several bounds are known for the (more interesting) case of fixed q, see [72, 19, 57]. Addressing this case, we prove a new upper estimate on codes for the ordered Hamming space. We also establish two asymptotic bounds obtained via the method of linear programming, in the context of association schemes (Sections 2.1–2.2). These new bounds provide the best known asymptotic estimates on the size of ordered codes and OOAs.

The known bounds on codes are presented in Section 5.2. After that we prove our first new result, a Bassalygo-Elias bound on codes. Section 5.3 is devoted to properties of multivariate Krawtchouk polynomials which correspond to the eigenvalues of the ordered Hamming scheme (see Section 2.2). A universal bound on codes is proved in Section 5.4 (here we employ the spectral method of Section 3.2). In Section 5.5 we establish another new bound by studying the location of extremal roots of certain bivariate Krawtchouk polynomials.

Publications: The results of this chapter are published in [11, 14].

5.1 Introduction

We recall some basic definitions from Section 2.2. Let $\mathcal{Q} = \{0, \dots, q-1\}$ be a finite alphabet. A vector $\boldsymbol{x}$ in the ordered Hamming space $\overrightarrow{\mathcal{H}} = \overrightarrow{\mathcal{H}}(q, n, r)$ is written as n blocks of r elements each: $\boldsymbol{x} = (x_{11}, \dots, x_{1r}; \dots; x_{n1}, \dots, x_{nr})$. The weight of $\boldsymbol{x}$ is given by $\mathrm{w}_r(\boldsymbol{x}) = |e|'$ where $e = \mathrm{shape}(\boldsymbol{x})$ is the shape of the vector.

Any subset $\mathcal{C} \subset \overrightarrow{\mathcal{H}}$ is called an ordered code. A code $\mathcal{C}$ of size M and minimum distance d is called an (nr, M, d) code. If $\mathcal{C}$ is a linear subspace of dimension k, then we write the parameters of the code as $[nr, k, d]$. If $\mathcal{C} \subset \overrightarrow{\mathcal{H}}$ is a linear code of dimension k, then the dual code $\mathcal{C}^\perp$ is a subspace of $\overleftarrow{\mathcal{H}} = \overleftarrow{\mathcal{H}}(q, n, r)$ of dimension $n-k$. An OOA of strength t in $\overrightarrow{\mathcal{H}}$ is denoted by (t, n, r, q). By Theorem 4.5, any $(m-t, n, m-t, q)$ OOA corresponds to a (t, m, n)-net. If $\mathcal{C} \subset \overrightarrow{\mathcal{H}}$ is a code of minimum distance d, then $\mathcal{C}^\perp$ is also an OOA of strength $d-1$ in $\overrightarrow{\mathcal{H}}$.

The valencies $v_e,\ e \in \Delta_{n,r}$, of the ordered Hamming scheme are given in (2.8). Valency v_e is also the cardinality of the sphere $\mathcal{S}_e$ that consists of all vectors of shape e. The cardinality of the sphere of weight d equals

$$S_d = \sum_{e:|e|'=d} v_e.$$

The asymptotic volume of the sphere of radius d was determined by Rosenbloom and Tsfasman [72] as described below. Let $A(z) = (q-1)z(z^r-1)/(q(z-1))$ and let $z_0 = z_0(x)$ satisfy the equation

$$xr(1 + A(z)) = \frac{q-1}{q} \sum_{i=1}^{r} iz^i. \tag{5.1}$$

Define the function

$$H_{q,r}(x) = x(1 - \log_q z_0) + \frac{1}{r} \log_q(1 + A(z_0)).$$

In the case $r = 1$ we write $h_q(x)$ instead of $H_{q,1}(x)$, where

$$h_q(x) = -x \log_q \frac{x}{q-1} - (1-x) \log_q(1-x)$$

is the usual q-ary entropy function. Let

$$\delta_{\text{crit}} = 1 - \frac{1}{r} \sum_{i=1}^{r} q^{-i} = 1 - \frac{1}{rq^r} \frac{q^r - 1}{q-1}. \tag{5.2}$$

The asymptotic volume of the sphere of weight d in $\overrightarrow{\mathcal{H}}$ is given in the next lemma.

Lemma 5.1 (Rosenbloom and Tsfasman [72])

1. *For $0 < x < 1$, equation (5.1) has a unique positive root $z_0(x) \in [0, r]$.*

2. Let $r \ge 1$ be fixed, $n \to \infty, d/n \to r\delta$, then

$$\lim_{n\to\infty} (nr)^{-1} \log_q \sum_{i=0}^{d} S_i = \begin{cases} H_{q,r}(\delta), & 0 \le \delta \le \delta_{crit}, \\ 1, & \delta_{crit} < \delta \le 1. \end{cases} \tag{5.3}$$

The asymptotic volume of the sphere $\mathcal{S}_e$ of shape e depends only on the asymptotic behavior of $|e|'/n$.

5.2 Bounds on ordered codes and OOAs

In this section we recall the known bounds on ordered codes and OOAs and derive a new bound on the size of codes.

5.2.1 Existence bounds

A bound analogous to the Gilbert-Varshamov bound of Theorem 3.1 is given in the next theorem.

Theorem 5.2 [72] *There exists an (nr, M, d) code in the space $\overrightarrow{\mathcal{H}}$ whose parameters satisfy*

$$M \sum_{i=0}^{d-1} S_i \ge q^{nr}.$$

If $\mathcal{Q}$ is a finite field, then there exists a linear code with the same parameters.

A slightly better result can be proved if we assume that the code is linear [17].

Theorem 5.3 [17] *Suppose that m and t satisfy the conditions*

$$\sum_{i=0}^{t-\tau} S_{i,n-1} < q^{m-(\tau-1)}, \quad \tau = 1, \dots, t-1.$$

Then there exists an $[nr, nr-m]$ linear code in $\overrightarrow{\mathcal{H}}$ of distance $\ge t+1$, and a (t, n, r, q) linear OOA in $\overrightarrow{\mathcal{H}}$ of dimension m.

5.2.2 Nonexistence bounds

While in general bounds on codes do not imply lower bounds on OOAs, there are two special cases when these two types of results are equivalent. First, if $\mathcal{C}$ is an $[nr, k, d]$ linear code in $\overrightarrow{\mathcal{H}}$ then the code $\mathcal{C}' := \{\boldsymbol{y} \in \overrightarrow{\mathcal{H}} : \forall\ \boldsymbol{c} \in \mathcal{C}\ \sum_{i=1}^{nr} c_i y_i = 0\}$ is a $(d-1, n, r, q)$ linear OOA in $\overrightarrow{\mathcal{H}}$. Next, if an upper (resp. lower) bound on codes (resp. OOAs) is obtained by linear programming as explained in Chapter 2 then the same solution of the LP problem gives a lower (resp. upper) bound on OOAs (resp. codes).

We next mention some upper bounds on codes and lower bounds on OOAs.

Singleton bound [72]: The parameters of any (nr, M, d) code satisfy

$$M \leq q^{nr-d+1}.$$

Plotkin bound: A Plotkin bound on codes was established in [72]. Namely, the following result holds true.

Theorem 5.4 [72] *Let $\mathcal{C} \subset \overrightarrow{\mathcal{H}}$ be a code of size M and distance $d > nr\delta_{crit}$. Then*

$$M \leq \frac{d}{d - nr\delta_{crit}}.$$

A dual Plotkin bound on OOAs was proved by Martin and Visentin [59].

Theorem 5.5 [59] *Let $\mathcal{C}$ be a (t, n, r, q) OOA. If $t > nr\delta_{crit} - 1$ then*

$$|\mathcal{C}| \geq q^{nr}\Big(1 - \frac{nr\delta_{crit}}{t+1}\Big).$$

Hamming-Rao bound: According to the Hamming bound, the parameters of any $(nr, M, d = 2\tau + 1)$ code satisfy

$$M \leq \frac{q^{rn}}{\sum_{i=0}^{\tau} S_i}.$$

A dual bound in this case is the Rao bound which for the ordered Hamming space was established by Martin and Stinson [56]: the size M of any $(t = 2\tau, n, r, q)$ OOA satisfies

$$M \geq \sum_{i=0}^{\tau} S_i.$$

A Bassalygo-Elias bound on codes: The next result is new.

Theorem 5.6 *Let $\mathcal{C}$ be an (nr, M, d) code. Then for any $w < nr\delta_{crit}\Big(1 - \sqrt{1 - \frac{d}{nr\delta_{crit}}}\Big)$,*

$$M \leq q^{rn} dn \frac{1}{S_w(dn - 2wn + \frac{w^2}{r\delta_{crit}})}.$$

We rely upon the next lemma, which is a generalization of the Johnson bound to the ordered Hamming space.

Lemma 5.7 (Johnson bound) *Let $\mathcal{C} \subset \overrightarrow{\mathcal{H}}, |\mathcal{C}| = M$ be a code all of whose vectors have weight w and are at least distance d apart. Then for $d \geq 2w - w^2/(nr\delta_{crit})$,*

$$M \leq \frac{dn}{dn - 2wn + \frac{w^2}{r\delta_{crit}}}.$$

PROOF: Let $\mathcal{C}^i$ be a projection of $\mathcal{C}$ on the ith block of coordinates. For a vector $\boldsymbol{z} \in \mathcal{Q}^r$ let $\boldsymbol{z}^h = (z_{r-h+1}, \dots, z_r)$ be its suffix of length h. Given $\boldsymbol{x} \in \mathcal{C}$, we denote by $\boldsymbol{x}_i \in \mathcal{C}^i$ its i-th block and write $\boldsymbol{x}_i^h$ to refer to the h-suffix of $\boldsymbol{x}_i$. For $i = 1, \dots, n;\ h = 1, \dots, r;\ \boldsymbol{c} \in \mathcal{Q}^h$ let

$$\lambda_{i,\boldsymbol{c}}^h = |\{\boldsymbol{x}_i \in \mathcal{C}^i :\ \boldsymbol{x}_i^h = \boldsymbol{c}\}|$$

be the number of vectors in the ith block whose h-suffix equals $\boldsymbol{c}$. Let $\delta(\boldsymbol{x}, \boldsymbol{y})$ denote the δ function which takes the value 1 if $\boldsymbol{x} = \boldsymbol{y}$ and 0 otherwise. We have

$$\begin{aligned} d_r(\boldsymbol{x}_i, \boldsymbol{y}_i) &= r - \sum_{h=1}^{r} \delta(\boldsymbol{x}_i^h, \boldsymbol{y}_i^h) \\ &= r - \sum_{h=1}^{r} \sum_{\boldsymbol{c} \in \mathcal{Q}^h} \delta(\boldsymbol{x}_i^h, \boldsymbol{c}) \delta(\boldsymbol{y}_i^h, \boldsymbol{c}). \end{aligned}$$

Compute the sum of all distances in the code as follows:

$$\begin{aligned} \sum_{\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}} d_r(\boldsymbol{x}, \boldsymbol{y}) &= nrM^2 - \sum_{i=1}^{n} \sum_{\boldsymbol{x}_i, \boldsymbol{y}_i \in \mathcal{C}^i} \sum_{h=1}^{r} \sum_{\boldsymbol{c} \in \mathcal{Q}^h} \delta(\boldsymbol{x}_i^h, \boldsymbol{c}) \delta(\boldsymbol{y}_i^h, \boldsymbol{c}) \\ &= nrM^2 - \sum_{i=1}^{n} \sum_{h=1}^{r} \sum_{\boldsymbol{c} \in \mathcal{Q}^h} (\lambda_{i,\boldsymbol{c}}^h)^2. \end{aligned} \tag{5.4}$$

To bound above the right-hand side, we need to find the minimum of the quadratic form

$$F = \sum_{i=1}^{n} \sum_{h=1}^{r} \sum_{\boldsymbol{c} \in \mathcal{Q}^h \backslash \{\boldsymbol{0}\}} (\lambda_{i,\boldsymbol{c}}^h)^2 + \sum_{i=1}^{n} \sum_{h=1}^{r} (\lambda_{i,\boldsymbol{0}}^h)^2$$

under the constraints

$$\sum_{i=1}^{n} \sum_{h=1}^{r} \lambda_{i,\boldsymbol{0}}^h = M(nr - w), \qquad \sum_{\boldsymbol{c} \in \mathcal{Q}^h} \lambda_{i,\boldsymbol{c}}^h = M \ (1 \le h \le r,\ 1 \le i \le n). \tag{5.5}$$

Critical points of F in the intersection of these hyperplanes, together with (5.5), satisfy the equations

$$\begin{array}{ll} 2\lambda_{i,\boldsymbol{c}}^h + \beta_{i,h} = 0, & 1 \le i \le n; 1 \le h \le r; \boldsymbol{c} \in \mathcal{Q}^h \backslash \{\boldsymbol{0}\} \\ 2\lambda_{i,\boldsymbol{0}}^h + \alpha + \beta_{i,h} = 0, & 1 \le i \le n; 1 \le h \le r \end{array} \qquad \alpha, \beta_{i,h} \in \mathbb{R}. \tag{5.6}$$

The system (5.5)-(5.6) has a unique solution for the variables $\lambda_{i,\boldsymbol{c}}^h, \beta_{i,h}, \alpha$; in particular,

$$\begin{aligned} \lambda_{i,\boldsymbol{0}}^h &= M\Big[\Big(\frac{1}{q^h} - 1\Big)\frac{w}{nr\delta_{\text{crit}}} + 1\Big], \quad h = 1, \dots, r, i = 1, \dots, n \\ \lambda_{i,\boldsymbol{c}}^h &= \frac{Mw}{q^h nr \delta_{\text{crit}}}, \qquad h = 1, \dots, r, i = 1, \dots, n, \boldsymbol{c} \in \mathcal{Q}^h \backslash \{\boldsymbol{0}\}. \end{aligned}$$

To verify that this critical point is in fact a minimum, observe that the form F is convex because its Hessian matrix is $2I$ and is positive definite (both globally and restricted to the intersection of the hyperplanes (5.5)). Substituting these values of the λs and taking account of the fact that $\sum_{h=1}^{r} q^{-h} = r(1-\delta_{\text{crit}})$, we get

$$F \geq \sum_{i=1}^{n}\sum_{h=1}^{r}\sum_{c\neq \mathbf{0}}\Big(\frac{Mw}{q^h nr\delta_{\text{crit}}}\Big)^2 + \sum_{i}\sum_{h} M^2\Big[\Big(\frac{1}{q^h}-1\Big)\frac{w}{nr\delta_{\text{crit}}}+1\Big]^2$$

$$= M^2 n\Big(\frac{w^2}{n^2 r\delta_{\text{crit}}} - \frac{2w}{n} + r\Big).$$

Then from (5.4) we obtain

$$dM(M-1) \leq \sum_{\boldsymbol{x},\boldsymbol{y}\in\mathcal{C}} d_r(\boldsymbol{x},\boldsymbol{y}) \leq \frac{M^2}{n}\Big(2wn - \frac{w^2}{r\delta_{\text{crit}}}\Big)$$

which gives the result. ∎

Proof: (of Theorem 5.6). Let $\mathcal{S}_w \subset \mathcal{Q}^{n,r}$ be a sphere of radius w around zero, and let $A_q(nr,d,w)$ be the maximum size of a distance-d code in $\mathcal{S}_w$. It is readily seen (cf. (3.4)) that,

$$|\mathcal{C}||\mathcal{S}_w| = \sum_{\boldsymbol{x}\in\overrightarrow{\mathcal{H}}} |(\mathcal{C}-\boldsymbol{x})\cap\mathcal{S}_w| \leq q^{nr}A_q(nr,d,w).$$

With the previous lemma, this finishes the proof. ∎

Remarks:

1. This theorem implies a lower bound on the size M of a linear $(d-1,n,r,q)$ OOA: for any $w \leq nr\delta_{\text{crit}}(1-\sqrt{1-d/(nr\delta_{\text{crit}})})$,
$$M \geq \frac{1}{dn} S_w\Big(dn - 2wn + \frac{w^2}{r\delta_{\text{crit}}}\Big) \qquad (5.7)$$
and in particular, a lower bound on linear $(m-r,m,n)$-nets, $m=\log_q M$.

2. Caution should be exercised in dealing with codes of a constant weight in the ordered Hamming space, i.e., codes on the sphere $\mathcal{S}_w$ in $\overrightarrow{\mathcal{H}}$. Indeed, the sphere $\mathcal{S}_w$ together with the metric d_r is not ball-homogeneous: in particular, the number of points in $\mathcal{S}_w$ located up to a given distance from a point $\boldsymbol{x}\in\mathcal{S}_w$ depends on $\boldsymbol{x}$. However, this does not cause any problem in the previous theorem.

3. The argument used in the proof of Lemma 5.7 can be also used to give a proof of the Plotkin bound, Theorem 5.4, that is simpler than the ones known in the literature. Indeed, let $\mathcal{C}\subset\overrightarrow{\mathcal{H}}$ be a distance-d code. Consider again expression (5.4) and note that this time there is no restriction on the weight of the codewords. Using the Cauchy-Schwarz inequality and the fact that $\sum_{\boldsymbol{c}\in\mathbb{F}_q^h}\lambda_{i,\boldsymbol{c}}^h = M$, we obtain
$$M(M-1)d \leq nrM^2 - \sum_{i=1}^{n}\sum_{h=1}^{r}\frac{M^2}{q^h} = nrM^2\delta_{\text{crit}}.$$
Solving for M concludes the proof.

5.2.3 Asymptotics

In this section we assume that $n \to \infty$ and r is a constant. For a code of size M let $R = \frac{1}{nr}\log_q M$ be the code *rate*. Given a sequence of (rn_i, M_i, d_i) codes we will say that its *asymptotic rate* is R and the *asymptotic relative distance* is δ if

$$\lim_{i\to\infty} \frac{1}{rn_i}\log_q M_i = R, \quad \lim_{i\to\infty}\frac{d_i}{rn_i} = \delta.$$

The Plotkin bound implies that the asymptotic rate and distance of any sequence of codes satisfy

$$\begin{aligned} &R \le 1 - \frac{\delta}{\delta_{\text{crit}}}, && 0 \le \delta \le \delta_{\text{crit}}, \\ &R = 0, && \delta \ge \delta_{\text{crit}}. \end{aligned}$$

To state the "sphere packing" or "volume" bounds on ordered codes we rely upon Lemma 5.1. Namely [72], there exists a sequence of $[rn_i, k_i, d_i]$ linear codes $\mathcal{C}_i, i = 1, 2, \dots$ such that $n_i \to \infty$, $k_i/(rn_i) \to R$, $d_i/(rn_i) \to \delta$ such that

$$R \ge 1 - H_{q,r}(\delta), \quad 0 \le \delta \le \delta_{\text{crit}} \qquad \text{(Gilbert-Varshamov bound).}$$

On the other hand, for any such sequence of codes,

$$R \le 1 - H_{q,r}(\delta/2), \quad 0 \le \delta \le 1 \qquad \text{(Hamming bound).}$$

The asymptotic version of Theorem 5.6 is as follows:

Theorem 5.8 (Asymptotic Bassalygo-Elias bound) *For $0 \le \delta \le \delta_{crit}$ the asymptotic rate and distance of any sequence of codes satisfy*

$$R \le 1 - H_{q,r}\Big(\delta_{crit}(1 - \sqrt{1 - \delta/\delta_{crit}})\Big). \tag{5.8}$$

This bound is better than the Hamming bound for all $\delta \in (0, \delta_{\text{crit}}]$. It is also often better than the Plotkin bound. For instance, for $q = 2, r = 2$ the bound (5.8) is better than the Plotkin bound for all $\delta \in (0, \delta_{\text{crit}})$. For larger q, r the improvement is attained only for low values of δ since the right-hand side of (5.8) becomes $\cap$-convex close to δ_{crit}. For instance, for $q = 3, r = 4$ this range is $(0, 0.54)$, etc.

Asymptotic bounds for digital (t, m, n)-nets

A (t, m, n)-net is called digital if the OOA that corresponds to it forms a linear subspace of $\mathbb{F}_q^{n,r}$. Therefore, bounds on linear OOAs apply to the special case of digital (t, m, n)-nets. However, studying asymptotics for this case requires a different normalization since the strength $m - t$ of the OOA that corresponds to the net equals r, and both approach infinity independently of n. Therefore, let $R = m/n$ denote the rate and $\delta = (m - t)/n$ denote the relative strength of the OOA that corresponds to the net. To state the bounds, we need to compute the asymptotic behavior of the volume of the sphere, which is different from (5.3). The next result is due to Bierbrauer and Schmid [18].

Theorem 5.9 [18] *There exist families of digital* (t, m, n)*-nets with* $n, (m-t) \to \infty$ *for which* (R, δ) *satisfy the bound* $R \le \Psi(\delta)$*, where*

$$\Psi(\delta) = \delta - 1 + \log_q \Big(\frac{q-1+\alpha}{\alpha}\Big) - \delta \log_q(1-\alpha),$$

and α *is defined by* $\delta\alpha(q-1+\alpha) = (q-1)(1-\alpha)$.

On the other hand, by the Rao bound, any family of (t, m, n)-nets satisfies $R \ge \Psi(\delta/2)$. Observe that Theorem 5.6 in this case gives the same result as the Rao bound because the increase of the packing radius in (5.7) over $\delta/2$ vanishes asymptotically. Indeed, taking $\omega = w/n$ and replacing d with $m-t$, we obtain from (5.7)

$$M \ge \frac{1}{\delta}\big(\delta - 2\omega + o(1)\big) S_{\omega n}.$$

The tightest bound is obtained if we take $\omega = \delta/2$ in this inequality.

Remark: We note that in the case that both $n \to \infty$ and $r \to \infty$ while $\delta = d/nr$ tends to a constant bounded away from 0 and 1, the lower and upper bounds on codes coincide [72] (the Gilbert-Varshamov bound converges to the Singleton bound).

5.3 Multivariate Krawtchouk polynomials in the ordered Hamming scheme

In this and the next sections we implement for the ordered Hamming space the program outlined in Section 3.2.

Recall from Chapter 2 that the association scheme of the ordered Hamming space is formally self dual. Its first and second eigenvalues are evaluations of certain multivariate orthogonal polynomials called the (multivariate) Krawtchouk polynomials. The aim of this section is to establish the properties of the multivariate Krawtchouk polynomials. These properties generalize the corresponding results for the usual Hamming space in two ways: first, in Section 3.2 we discussed only the binary case while here we consider an arbitrary alphabet (this requires only minor changes); second, we deal with several variables instead of one (this entails substantial complications). These properties will be used in the next section to satisfy the conditions in (2.10), thereby deriving new linear programming bounds on the size of ordered codes and OOAs.

Observe that the valencies $v_e = p^0_{e,e}$ of the scheme are given by (2.8). By self-duality and (2.9), the eigenvalues are orthogonal on the space of partitions $\Delta_{n,r}$ with weight v_e. Below it will be convenient to normalize the weight. Let $V_{n,r}$ be the space of real polynomials of r discrete variables $e = (e_1, e_2, \dots, e_r)$ defined on $\Delta_{n,r}$. Let us define a bilinear form acting on the space $V_{n,r}$ by

$$\langle u_1, u_2 \rangle = \sum_{e \in \Delta_{n,r}} u_1(e) u_2(e) w(e), \tag{5.9}$$

where $w(e) = q^{-nr}v_e$. Letting $p_i = q^{i-r-1}(q-1), i = 1, \dots, r$; $p_0 = q^{-r}$, we observe that

$$w(e) = n! \prod_{i=0}^{r} \frac{p_i^{e_i}}{e_i!}$$

forms a multinomial probability distribution on $\Delta_{n,r}$. Therefore, r-variate polynomials orthogonal with respect to this weight form a particular case of multivariate Krawtchouk polynomials.

For a partition $f \in \Delta_{n,r}$ denote by

$$K_f(e) = K_{f_1,\dots,f_r}(e_1, \dots, e_r)$$

the Krawtchouk polynomial that corresponds to it. Let $\kappa = |f|$ be the degree of K_f.

Our goal in this section is to derive properties of the polynomials K_f. In their large part, these properties are obtained by specializing to the current case general relations of the ordered Hamming scheme in Section 2.2. However, some work is needed to transform them to a concrete form that can be used in later calculations.

The following relations are useful below.

Lemma 5.10

$$\begin{aligned}
\langle f_i, 1\rangle &= n(q-1)q^{i-r-1}, && i = 1, \dots, r, && (5.10)\\
\langle f_i, f_j\rangle &= n(n-1)(q-1)^2 q^{i+j-2r-2}, && 1 \le i \ne j \le r, && (5.11)\\
\langle f_i, f_i\rangle &= n(q-1)q^{i-r-1}(1 + (n-1)(q-1)q^{i-r-1}), && i = 1, \dots, r. && (5.12)
\end{aligned}$$

Proof: To prove (5.10), compute

$$\begin{aligned}
\langle f_i, 1\rangle &= q^{-nr} \sum_e \left\{ e_i \binom{n}{e_0, e_1, \dots, e_r} \prod_{j=1}^{r} ((q-1)q^{j-1})^{e_j} \right\}\\
&= nq^{-nr} \sum_e \binom{n-1}{e_0, e_1, \dots, e_i - 1, \dots, e_r} \prod_{j=1}^{r} ((q-1)q^{j-1})^{e_j}.
\end{aligned}$$

The sum on e on the last line equals $(q-1)q^{i-1+(n-1)r}$ which finishes the proof. The remaining two identities are proved in a similar way. ■

5.3.1 Properties of the polynomials $K_f(e)$

(i) $K_f(e)$ is a polynomial in the variables $e_1, \dots, e_r$ of degree $\kappa = |f|$. There are $\binom{\kappa+r-1}{r-1}$ different polynomials of the same degree, each corresponding to a partition of κ.

(ii) (Orthogonality) Equation (2.9) is rewritten with normalized weights as

$$\langle K_f, K_g\rangle = v_f \delta_{f,g}, \quad \|K_f\| = \sqrt{v_f}. \tag{5.13}$$

In particular, let $F_i = (0^{i-1}10^{r-i-1}), i = 1, \ldots, r$ be a partition with one part. Using (2.8) we get

$$\|K_{F_i}\|^2 = \langle K_{F_i}, K_{F_i} \rangle = n(q-1)q^{i-1}, \quad i = 1, \ldots, r. \tag{5.14}$$

We take

$$K_{(0,\cdots,0)}(e) = 1, \quad e \in \Delta_{n,r}. \tag{5.15}$$

(iii) The next property is a special case of (2.2).

$$v_e K_f(e) = v_f K_e(f), \quad e, f \in \Delta_{n,r}.$$

In particular,

$$K_f(0) = v_f. \tag{5.16}$$

(iv) (Linear polynomials) For $i = 1, \ldots, r$,

$$K_{F_i}(e) = q^{i-1}(q-1)(n - e_r - \cdots - e_{r-i+2}) - q^i e_{r-i+1}. \tag{5.17}$$

Proof: This is shown by orthogonalizing the set of linear polynomials $\{1, e_1, e_2, \ldots, e_r\}$. Use Lemma 5.10 and (5.15) to compute

$$K_{F_1}(e) = c_1(e_r - \langle e_r, 1\rangle) = c_1(e_r - n(q-1)/q)$$

for some constant c_1. To find c_1, use (5.14):

$$n(q-1) = c_1^2 \left\| e_r - \frac{n(q-1)}{q} \right\|^2 = c_1^2 n(q-1)q^{-2}.$$

Hence $c_1 = \pm q$. We take $K_{F_1}(e) = n(q-1) - qe_r$ choosing $c_1 = -q$ so that $K_{F_1}(0) > 0$.

Next let us perform the induction step to compute $K_{F_{i+1}}(e)$:

$$K_{F_{i+1}}(e) = c_{i+1}\Big(e_{r-i} - \sum_{j=0}^{i} \|K_{F_j}\|^{-2} \langle e_{r-i}, K_{F_j}\rangle K_{F_j}(e)\Big), \tag{5.18}$$

where the polynomials $K_{F_j}(e), j = 0, \ldots, i$, have the form (5.17) by the induction hypothesis. Straightforward calculations using (5.10)-(5.12) show that

$$K_{F_{i+1}}(e) = c_{i+1}(e_{r-i} - ((q-1)/q)(n - e_r - \cdots - e_{r-i+1})).$$

Again using (5.14), we find that $c_{i+1} = \pm q^{i+1}$; as above, we choose the minus. ■

(v) For any $e, f, g \in \Delta_{n,r}$

$$K_f(e)K_g(e) = \sum_{h \in \Delta_{n,r}} p^h_{f,g} K_h(e), \tag{5.19}$$

where the linearization coefficients $p^h_{f,g}$ are the intersection numbers of the scheme, described in (2.7).

(vi) (Three-term relation) Let $\mathbb{K}_\kappa$ be a column vector of the polynomials $K_f(e)$ ordered lexicographically with respect to all f that satisfy $|f| = \kappa$. The three-term relation is obtained by expanding a product $P(e)\mathbb{K}_\kappa(e)$ in the basis $\{K_f\}$, where $P(e)$ is a first-degree polynomial. By orthogonality, the only nonzero terms in this expansion will be polynomials of degrees $\kappa+1, \kappa, \kappa-1$ [31, p.75].

We establish an explicit form of the three-term relation for $P(e) = \delta_{\text{crit}} rn - |e|'$. We have

$$P(e)\mathbb{K}_\kappa(e) = a_\kappa \mathbb{K}_{\kappa+1}(e) + b_\kappa \mathbb{K}_\kappa(e) + c_\kappa \mathbb{K}_{\kappa-1}(e), \tag{5.20}$$

where $a_\kappa, b_\kappa, c_\kappa$ are matrices of order $\binom{\kappa+r-1}{r-1} \times \binom{\kappa+s+r-1}{r-1}$ and $s = 1, 0, -1$, respectively. The nonzero elements of these matrices have the following form:

$$\begin{aligned}
(a_\kappa)_{f,h} &= L_i(f_i+1) && \text{if } h = (f_1, \dots, f_i+1, \dots, f_r), \\
(c_\kappa)_{f,h} &= L_i(n-\kappa+1)q^{i-1}(q-1) && \text{if } h = (f_1, \dots, f_i-1, \dots, f_r),
\end{aligned}$$

$$(b_\kappa)_{f,h} = \begin{cases} L_i f_i q^{i-1}(q-2) & \text{if } h = f, \\ L_i(f_k+1)q^{i-1}(q-1) & \text{if } h = (f_1, \dots, f_k+1, \dots, f_i-1, \dots, f_r), \\ & \qquad 1 \le k < i, \\ L_i(f_i+1)q^{k-1}(q-1) & \text{if } h = (f_1, \dots, f_k-1, \dots, f_i+1, \dots, f_r), \\ & \qquad 1 \le k < i, \end{cases} \tag{5.21}$$

where $L_i = \frac{q^{r-i+1}-1}{q^r(q-1)}$.

Proof: According to Property (v), the coefficients of the expansion of the product $K_{F_i}(e)K_f(e)$ into the basis $\{K_h(e)\}$ are given by the intersection numbers of the scheme:

$$K_{F_i}(e)K_f(e) = \sum_h p^h_{F_i,f} K_h(e). \tag{5.22}$$

The ordered Hamming scheme is *translation invariant*, i.e., $(\boldsymbol{x}, \boldsymbol{y}) \in R_e \Leftrightarrow (\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + \boldsymbol{z}) \in R_e$. Hence we can assume that $\boldsymbol{y} = \boldsymbol{0}$. So $p^h_{F_i,f}$ is the number of vectors $\boldsymbol{z}$ with $\text{shape}(\boldsymbol{z}) = f$ that satisfy $\text{shape}(\boldsymbol{z} - \boldsymbol{x}) = F_i$ for a fixed vector $\boldsymbol{x}$ with $\text{shape}(\boldsymbol{x}) = h$. In other words,

$$\boldsymbol{z} - \boldsymbol{x} = (0^r, \dots, 0^r, (u_1, \dots, u_{i-1}, u_i, 0, \dots, 0), 0^r, \dots, 0^r), \tag{5.23}$$

where the nonzero block is located in any of the n possible blocks, and $u_j \in \mathbb{F}_q, 1 \le j < i, u_i \ne 0$.

The numbers $p^h_{F_i,f}$ are nonzero only in the three following cases.

1. $|h| = |f|+1$. By the above we have that $h_j = f_j$ for $j \ne i$ and $h_i = f_i+1$. Hence $\boldsymbol{z}$ can be chosen so that its f_i blocks of weight i annihilate the corresponding blocks of $\boldsymbol{x}$, leaving one such block in any of the $h_i = f_i + 1$ locations. Thus,

$$p^h_{F_i,f} = \begin{cases} f_i + 1, & h = (f_1, \dots, f_i+1, \dots, f_r), \\ 0, & \text{otherwise.} \end{cases}$$

2. $|h| = |f|$. The following numbers are easily verified by (5.23):

$$p^h_{F_i,f} = \begin{cases} f_i(q-2)q^{i-1}, & h = f, \\ (f_k+1)(q-1)q^{i-1}, & h = (f_1, \dots, f_k+1, \dots, f_i-1, \dots, f_r), \\ & 1 \le k < i, \\ (f_i+1)(q-1)q^{k-1}, & h = (f_1, \dots, f_k-1, \dots, f_i+1, \dots, f_r), \\ & 1 \le k < i, \\ 0, & \text{otherwise.} \end{cases}$$

Other than these three cases, no other possibilities for h arise.

3. $|h| = |f| - 1$. Now we should add to $\boldsymbol{x}$ one block of weight i in any of the $n - |f| + 1$ all-zero blocks. Thus we obtain

$$p^h_{F_i,f} = (n - |f| + 1)q^{i-1}(q-1), \quad h = (f_1, \dots, f_i - 1, \dots, f_r)$$

and $p^h_{F_i,f} = 0$ for all other h.

To prove (5.20) we now need to represent $P(e)$ as a linear combination of the K_{F_i}s. Using (5.17) we find that

$$|e|' = \sum_{i=1}^r ie_i = \delta_{\text{crit}} rn - \sum_{i=1}^r L_i K_{F_i}(e),$$

hence

$$P(e) = \sum_{i=1}^r L_i K_{F_i}(e). \tag{5.24}$$

The proof is now concluded by using (5.22) together with the intersection numbers computed above. ■

Along with the polynomials $K_f(e)$ below we use their normalized version $\tilde{K}_f(e) = K_f(e)/\sqrt{v_f}$. The polynomials $\{\tilde{K}_f(e), f \in \Delta_{n,r}\}$ form an orthonormal basis of $V_{n,r}$.

Denote by $A_\kappa, B_\kappa, C_\kappa$ the coefficient matrices of the normalized form of relation (5.20). The new matrix elements are given by

$$(A_\kappa)_{f,h} = L_i\sqrt{(f_i+1)(n-\kappa)q^{i-1}(q-1)} \qquad \text{if } h = (f_1, \dots, f_i+1, \dots, f_r),$$

$$(C_\kappa)_{f,h} = L_i\sqrt{(n-\kappa+1)f_i q^{i-1}(q-1)} \qquad \text{if } h = (f_1, \dots, f_i-1, \dots, f_r),$$

$$(B_\kappa)_{f,h} = \begin{cases} L_i f_i q^{i-1}(q-2) & \text{if } h = f, \\ L_i \dfrac{q-1}{q}\sqrt{(f_k+1)f_i q^{k+i}} & \text{if } h = (f_1, \dots, f_k+1, \dots, f_i-1, \dots, f_r), \\ & \quad 1 \le k < i, \\ L_i \dfrac{q-1}{q}\sqrt{f_k(f_i+1)q^{k+i}} & \text{if } h = (f_1, \dots, f_k-1, \dots, f_i+1, \dots, f_r), \\ & \quad 1 \le k < i. \end{cases} \tag{5.25}$$

Let $V_\kappa \subset V_{n,r}$ be the set of polynomials of total degree $\le \kappa$. Let E_κ be the orthogonal projection of $V_{n,r}$ on V_κ. Define the operator

$$\begin{aligned} S_\kappa : & V_\kappa \to V_\kappa \\ & f \mapsto E_\kappa(Pf). \end{aligned}$$

Its matrix in the orthonormal basis has the form

$$\tilde{\mathbf{S}}_\kappa = \begin{bmatrix} B_0 & A_0 & \mathbf{0} & \dots & \mathbf{0} \\ C_1 & B_1 & A_1 & \dots & \mathbf{0} \\ \mathbf{0} & C_2 & B_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & C_\kappa & B_\kappa \end{bmatrix},$$

where the B_is are symmetric and $C_i = A_{i-1}^T,\ i = 1, \dots, \kappa$. On account of property **(v)** and (5.24), the matrix elements of $\tilde{\mathbf{S}}_\kappa$ are nonnegative.

The matrix of S_κ in the basis $\{K_f\}$ has the property

$$v_h(\mathbf{S}_\kappa)_{f,h} = v_f(\mathbf{S}_\kappa)_{h,f}, \quad f, h \in \Delta_{n,r}. \tag{5.26}$$

(vii) (Explicit expression)

$$K_f(e) = q^{|f|'-|f|} \prod_{i=1}^{r} K_{f_i}(e_{r-i+1}; n_i), \tag{5.27}$$

where K_{f_i} is a univariate Krawtchouk polynomial (2.21), $n_i = \sum_{j=0}^{r-i+1} e_j - \sum_{j=i+1}^{r} f_j$, and $f, e \in \Delta_{n,r}$. This form of the polynomial $K_f(e)$ was obtained in [19] (various other forms were found in [60, 29]). We remark that (5.27) can be proved by performing the Gram-Schmidt procedure (5.18) for monomials of higher degrees. It is known that the resulting system of polynomials is unique up to a constant factor once the polynomials of degrees 0 and 1 together with the three-term relation (5.20) have been fixed, see [31, Theorem 3.4.9].

(viii) The eigenvalues of the ordered Hamming scheme $\overrightarrow{\mathcal{H}}$ are given by $p_f(e) = q_f(e) = K_f(e)$. This follows from the previous property and (2.9) because the polynomials $\{K_f\}$ form a unique orthogonal family on $\Delta_{n,r}$ with respect to the weight $w(e)$.

(ix) (Fourier transform representation) Let ζ be a qth degree primitive root of unity, and $e, f \in \Delta_{n,r}$. Then

$$K_f(e) = \sum_{\boldsymbol{z}:\mathrm{shape}(\boldsymbol{z})=f} \zeta^{\boldsymbol{x}\cdot\boldsymbol{z}}, \tag{5.28}$$

where $\overline{\mathrm{shape}}(\boldsymbol{x}) = e$. In [19] this relation is taken as a definition of the polynomials $K_f(e)$. Under our approach, it follows from the well-known Fourier transform representation of the Krawtchouk polynomials $K_i(x; n)$ in the case $r = 1$ and Theorem 2.3.

(x) (Christoffel-Darboux) Let $L \subset \Delta_{n,r}$ and for $a, e \in \Delta_{n,r}$ define

$$U_L(a,e) \triangleq \sum_{f\in L} v_f^{-1} K_f(a) K_f(e).$$

Let $\mathbf{S}$ be the matrix of the operator $S : V_{n,r} \to V_{(n+1),r}$ given by $f \mapsto Pf$, written in the basis $\{K_f\}$. The action of $P(e)$ on U_L is described as follows:

$$\begin{aligned}\big(P(e)-P(a)\big)U_L(a,e) &= \sum_{f\in L} v_f^{-1} \sum_{h\in\Delta_{n,r}} (\mathbf{S})_{f,h}\Big(K_h(e)K_f(a) - K_h(a)K_f(e)\Big)\\ &= \sum_{f\in L} v_f^{-1} \sum_{h\in\Delta_{n,r}\setminus L} (\mathbf{S})_{f,h}\Big(K_h(e)K_f(a) - K_h(a)K_f(e)\Big),\end{aligned}$$

the last equality justified by (5.26) as follows:

$$\begin{aligned}&\sum_{f,h\in L} v_f^{-1}(\mathbf{S})_{f,h}\Big(K_h(e)K_f(a) - K_h(a)K_f(e)\Big)\\ &= \sum_{f,h\in L} (\mathbf{S})_{f,h}\sqrt{\frac{v_h}{v_f}}\,\Big(\tilde{K}_h(e)\tilde{K}_f(a) - \tilde{K}_f(e)\tilde{K}_h(a)\Big)\\ &= 0.\end{aligned}$$

A particular case of the above is obtained when $L = \{f : |f| \le \kappa\}$. The kernel U_L, denoted in this case by U_κ, equals $U_\kappa = \sum\limits_{s=0}^{\kappa} \tilde{\mathbb{K}}_s(e)^T \tilde{\mathbb{K}}_s(a)$, and we obtain

$$\begin{aligned}\big(P(e)-P(a)\big)U_\kappa(a,e) &= \tilde{\mathbb{K}}_{\kappa+1}(e)^T A_\kappa^T \tilde{\mathbb{K}}_\kappa(a) - \tilde{\mathbb{K}}_\kappa(e)^T A_\kappa \tilde{\mathbb{K}}_{\kappa+1}(a) \qquad (5.29)\\ &= \sum_{f:|f|=\kappa} Q_f(e)\tilde{K}_f(a) - \tilde{K}_f(e)Q_f(a),\end{aligned}$$

where $Q_f(e) = \sum_{|h|=\kappa+1} \tilde{K}_h(e)(A_\kappa)_{f,h}$. This relation is called the Christoffel-Darboux formula.

(xi) The generating function of the polynomials K_f is given by

$$\sum_{f\in\Delta_{n,r}} K_f(e)z^f = \Big(1+(q-1)\sum_{i=1}^{r} q^{i-1}z_i\Big)^{n-|e|} \prod_{j=1}^{r}\Big(1+(q-1)\sum_{k=1}^{j-1} q^{k-1}z_k - q^{j-1}z_j\Big)^{e_{r-j+1}}.$$

In particular,

$$\sum_{f\in\Delta_{n,r}} K_f(e) = q^{nr}\delta_{e,0}.$$

Remarks:

1. The polynomials $K_f(e)$ were considered in [60, 29, 19]. However none of these papers constructed them from their definition as eigenvalues of the r-Hamming scheme (to be more precise, Martin and Stinson [60] mention this approach but pursue the path suggested in Theorem 2.3 which makes explicit calculations difficult). Under the approach taken above, many properties of the polynomials K_e follow as special cases of the general combinatorial results of Section 2.1.

2. Other generalizations of univariate Krawtchouk polynomials were considered earlier in [84, 71]. These papers study bi-orthogonal polynomials for the weight given by the multinomial probability distribution, resulting in polynomial families different from the one considered above.

3. Property **(xi)** implies a MacWilliams theorem for ordered codes. It was previously proved in [60, 29] using different means.

   Theorem 5.11 (**MacWilliams theorem**) *Let* $\mathcal{C} \subset \overrightarrow{\mathcal{H}}$ *and* $\mathcal{C}^\perp \subset \overleftarrow{\mathcal{H}}$ *be two linear codes that satisfy* $\sum_{i=1}^{nr} x_i y_i = 0$ *for every* $\boldsymbol{x} \in \mathcal{C}, \boldsymbol{y} \in \mathcal{C}^\perp$*. Let* $A(z_0, z) = \sum_e A_e \prod_{i=0}^{r} z_i^{e_i}$ *be the shape enumerator of* $\mathcal{C}$ *and let* $A^\perp(z_0, z)$ *be the same for* $\mathcal{C}^\perp$*. Then*

   $$A^\perp(z_0, z_1, \dots, z_r) = \frac{1}{|\mathcal{C}|} A(u_0, u_1, \dots, u_r)$$

   where

   $$u_0 = z_0 + (q-1)\sum_{i=1}^{r} q^{i-1} z_i,\ u_{r-j+1} = z_0 + (q-1)\sum_{k=1}^{j-1} q^{k-1} z_k - q^{j-1} z_j, 1 \le j \le r.$$

5.4 A linear programming bound on ordered codes and OOAs

In this section we prove one of our main results, an LP bound on the rate of codes. An LP bound for OOAs was first formulated in Martin and Stinson [60], and a numerical implementation of the LP bound for specific values of n was presented in Martin [57]. The LP bound on codes and OOAs was also re-formulated by Bierbrauer [19]. However asymptotic LP bounds derived in the context of Delsarte's theory from Theorem 2.4 were not known. In the rest of the chapter we provide new asymptotic LP bounds on codes and OOAs by determining the polynomials required to satisfy the conditions in Theorem 2.4.

5.4.1 The bound

We state the LP bound in the theorem below. Its proof uses a spectral method first employed in [7], and explained in Section 3.2 for the binary Hamming space.

Theorem 5.12 *Let κ be any degree such that $P(e) \le \lambda_{\kappa-1}$ for all shapes e with $|e|' \ge d$, where λ_i is the maximum eigenvalue of S_i and $d \ge 1$ is an integer.*

Let $\mathcal{C} \subset \overrightarrow{\mathcal{H}}$ be an (nr, M, d) code. Then

$$M \le \frac{4r\delta_{crit}(n-\kappa)(q^r-1)^\kappa}{\delta_{crit}rn - \lambda_\kappa}\binom{n}{\kappa}. \tag{5.30}$$

Let $\mathcal{C}$ be a $(t = d-1, n, r, q)$ OOA of size M. Then

$$M \ge \frac{q^{nr}}{\binom{n}{\kappa}} \frac{(\delta_{crit}rn - \lambda_\kappa)}{4r\delta_{crit}(n-\kappa)(q^r-1)^\kappa}. \tag{5.31}$$

PROOF: Consider the operator T_κ that equals S_κ on $V_{\kappa-1}$ and acts on a function $\varphi \in V_\kappa \backslash V_{\kappa-1}$ by

$$T_\kappa(\varphi) := S_\kappa\varphi - \sum_{f:|f|=\kappa} \epsilon_f \varphi_f \tilde{K}_f,$$

where $\epsilon_f > 0$ are some constants indexed by the partitions of weight κ (their values will be chosen later). The matrix of T_κ in the orthonormal basis equals

$$\tilde{\mathbf{T}}_\kappa = \tilde{\mathbf{S}}_\kappa - \begin{bmatrix} 0 & 0 \\ 0 & E \end{bmatrix},$$

where $E = \operatorname{diag}(\epsilon_f, |f| = \kappa)$ is a matrix of order $\binom{\kappa+r-1}{r-1}$. Let m be such that $\tilde{\mathbf{T}}_\kappa + mI > 0$. By Perron-Frobenius theorem (Theorem A.4), the spectral radius $\rho(T_\kappa + mI)$ is well defined and is an eigenvalue of (algebraic and geometric) multiplicity one of $T_\kappa + mI$. Moreover, using Perron-Frobenius and Lemma A.3,

$$\rho(S_{\kappa-1} + mI) < \rho(T_\kappa + mI) < \rho(S_\kappa + mI).$$

Then

$$\lambda_{\kappa-1} < \theta_\kappa < \lambda_\kappa, \tag{5.32}$$

where $\theta_\kappa = \rho(T_\kappa)$. Let $G > 0$ be the eigenfunction of T_κ with eigenvalue θ_κ. Let us write out the product $P(e)G$ in the orthonormal basis:

$$P(e)G = S_\kappa G + G_\kappa A_\kappa \tilde{\mathbb{K}}_{\kappa+1} = \theta_\kappa G + \sum_{f:|f|=\kappa} \epsilon_f G_f \tilde{K}_f + G_\kappa A_\kappa \tilde{\mathbb{K}}_{\kappa+1},$$

where G_κ is a projection of the vector G on the space $V_\kappa \backslash V_{\kappa-1}$. This implies the equality

$$G = \frac{\sum_{|f|=\kappa} G_f(\epsilon_f \tilde{K}_f + Q_f)}{P(e) - \theta_\kappa},$$

where $Q_f(e)$ is defined after (5.29). Now take $F(e) = (P(e) - \theta_\kappa)G^2(e)$. Let us verify (2.10). Since multiplication by a function is a self-adjoint operator, we obtain

$$F_0 = \langle F, 1 \rangle = \Big\langle \sum_{|f|=\kappa} G_f(\epsilon_f \tilde{K}_f + Q_f), G \Big\rangle = \sum_{|f|=\kappa} G_f^2 \epsilon_f > 0.$$

Using (5.19) one can easily check that $F_f \geq 0$ for all $f \neq 0$. The assumption of the theorem together with (5.32) implies that $F(e) \leq 0$ for $|e|' \geq d$. Hence

$$\begin{aligned} M &\leq \frac{F(0)}{F_0} \\ &= \frac{\Big(\sum_{|f|=\kappa} G_f(\epsilon_f \tilde{K}_f(0) + Q_f(0))\Big)^2}{(P(0)-\theta_\kappa)\sum_{|f|=\kappa} G_f^2 \epsilon_f} \\ &\leq \frac{1}{P(0)-\lambda_\kappa} \sum_{|f|=\kappa} \frac{(\epsilon_f \tilde{K}_f(0) + Q_f(0))^2}{\epsilon_f}, \end{aligned}$$

where in the last step we used the Cauchy-Schwarz inequality. Computing the minimum on ϵ_f, we obtain

$$M \leq \frac{4}{P(0)-\lambda_\kappa} \sum_{|f|=\kappa} Q_f(0)\sqrt{v_f}. \tag{5.33}$$

Next,

$$\sum_{|f|=\kappa} Q_f(0)\sqrt{v_f} = \sum_{f:|f|=\kappa} \sqrt{v_f} \sum_{h:|h|=\kappa+1} (A_\kappa)_{f,h}\sqrt{v_h}.$$

Let $h = (f_1, \ldots, f_i + 1, \ldots, f_r)$ for some $i, 1 \leq i \leq r$. Then using (2.8) we find

$$\begin{aligned} (A_\kappa)_{f,h}\sqrt{v_h} &= L_i\sqrt{(f_i+1)(n-\kappa)q^{i-1}(q-1)}\sqrt{v_h} \\ &= L_i\sqrt{(f_i+1)(n-\kappa)q^{i-1}(q-1)}\sqrt{v_f\frac{(n-\kappa)q^{i-1}(q-1)}{f_i+1}} \\ &= \Big(1-\frac{1}{q^{r-i+1}}\Big)(n-\kappa)\sqrt{v_f}. \end{aligned}$$

Thus we have

$$\sum_{|f|=\kappa} Q_f(0)\sqrt{v_f} = \sum_{|f|=\kappa}\sum_{i=1}^{r}(n-\kappa)\Big(1-\frac{1}{q^{r-i+1}}\Big)v_f$$

$$= (n-\kappa)r\delta_{\text{crit}}\sum_{|f|=\kappa} v_f = (n-\kappa)r\delta_{\text{crit}}\binom{n}{\kappa}(q^r-1)^\kappa.$$

Substitution of this expression into (5.33) concludes the proof of (5.30). The bound (5.31) follows by (2.12). ■

5.4.2 Spectral radius of $\mathbf{S}_\kappa$

In this section we derive an asymptotic lower bound on the spectral radius of $\mathbf{S}_\kappa$. This estimate will be later used to optimize the bound (5.30) on the choice of κ.

Theorem 5.13

$$\lim_{\substack{n\to\infty \\ \frac{\kappa}{n}\to\tau}} \frac{\lambda_\kappa}{n} \ge \max_{\substack{\tau_i\ge 0 \\ \sum_{i=1}^r \tau_i=\tau}} \Lambda(\tau_1,\dots,\tau_r),$$

where

$$\Lambda(\tau_1,\dots,\tau_r)=\sum_{i=1}^r L_i\Big(2\sqrt{(1-\tau)\tau_i(q-1)q^{i-1}} \\ +(q-2)\tau_i(q^r-q^{i-1})+2\frac{(q-1)}{q}\sum_{k=1}^{i-1}\sqrt{\tau_k\tau_i q^{i+k}}\Big). \qquad (5.34)$$

To prove this theorem, we will bound below the largest eigenvalue λ_κ of the matrix $\tilde{\mathbf{S}}_\kappa$. For any real vector $\boldsymbol{y}$ we have by Rayleigh-Ritz inequality,

$$\lambda_\kappa \ge \frac{\boldsymbol{y}^T\tilde{\mathbf{S}}_\kappa\boldsymbol{y}}{(\boldsymbol{y},\boldsymbol{y})}.$$

We will construct a suitable $(0,1)$-vector $\boldsymbol{y}$. Its coordinates are indexed by the partitions arranged in the increasing order of their length μ and lexicographically within a block of coordinates for each value of $\mu, 0\le\mu\le\kappa$. Let $\boldsymbol{y}=(\boldsymbol{y}_0,\boldsymbol{y}_1,\dots,\boldsymbol{y}_\kappa)$ where $\boldsymbol{y}_\mu=(\boldsymbol{y}_f,|f|=\mu)$.

Let $f, |f|=\mu$, be a shape vector. For an integer J consider the set

$$\mathcal{F}_\mu=\mathcal{F}_\mu(J,f)\triangleq\{(f_1+\mu-\kappa+j_1,\dots,f_r+j_r):\sum_{i=1}^r j_i=0;\ |j_i|\le J, i=1,\dots,r\}$$

and denote $m=|\mathcal{F}_\mu|$. Next, let

$$(\boldsymbol{y}_\mu)_h=1(h\in\mathcal{F}_\mu)$$

for $\mu=\kappa+1-s,\dots,\kappa$ where s will be chosen later, and $\boldsymbol{y}_\mu=\mathbf{0}$ otherwise.

In the next two lemmas we derive a lower bound on the part of the product $\boldsymbol{y}^T\tilde{\mathbf{S}}_\kappa\boldsymbol{y}$ that involves only the rows of $\tilde{\mathbf{S}}_\kappa$ that correspond to the shapes f of length μ. Let

$$E_h=\{(h_1,\dots,h_k\pm 1,\dots,h_l\mp 1,\dots,h_r), 1\le k<l\le r\}$$

be the index set of the nonzero off-diagonal elements in the row in B_μ which is indexed by $h=(h_1,\dots,h_r)$.

Lemma 5.14 *Let* $e=\operatorname{argmin}_{h\in\mathcal{F}_\mu}\big(\sum_{g\in E_h\cup\{h\}}(B_\mu)_{h,g}\big)$ *and let* $\psi_\mu=\sum_g(B_\mu)_{e,g}$. *Then*

$$\boldsymbol{y}_\mu^T B_\mu\boldsymbol{y}_\mu\ge\psi_\mu m(1-o_m(1)).$$

Proof: I. Since $|h|=\mu$ for every $h\in\mathcal{F}_\mu$, the quantity $|\mathcal{F}_\mu|$ equals the number of ordered partitions of 0 into at most r parts, each part bounded between $-J$ and J, or the number of ordered partitions

$$Jr=\sum_{i=1}^r j_i,\quad 0\le j_i\le 2J,\ i=1,\dots,r.$$

The number of such partitions is given by [42, p.1037]:

$$\pi(r, 2J, Jr) = \sum_{i=0}^{\lfloor \frac{r}{2} \rfloor} (-1)^i \binom{r}{i} \binom{r + Jr - (2J+1)i - 1}{r-1}.$$

Writing this expression as a polynomial in J, we find the coefficient of J^{r-1} to be

$$\frac{1}{(r-1)!} \sum_{i=0}^{\lfloor \frac{r}{2} \rfloor} (-1)^i \binom{r}{i} (r-2i)^{r-1}.$$

Since this is always positive[1], we conclude that m is a degree-$(r-1)$ polynomial in J; in particular, if $J \to \infty$, then also $m \to \infty$.

II. We have

$$\boldsymbol{y}_\mu^T B_\mu \boldsymbol{y}_\mu = \sum_{h,g \in \mathcal{F}_\mu} (B_\mu)_{h,g}.$$

To bound $\boldsymbol{y}_\mu^T B_\mu \boldsymbol{y}_\mu$ below we estimate the difference between the above sum and the sum of all the nonzero elements of B_μ in the rows $h \in \mathcal{F}_\mu$ which is obtained by replacing the range of column indices g above with $g \in E_h \cup \{h\}$. Therefore, for a given $h \in \mathcal{F}_\mu$ let us estimate the number $|E_h \backslash \mathcal{F}_\mu|$ of nonzero entries in $(B_\mu)_{h,\cdot}$ not included in the sum. Let $f = (f_1, \dots, f_r)$ and let h be of the form $h = (\dots, f_k + J, \dots) \in \mathcal{F}_\mu$ for some $1 \le k \le r$. Consider the column indices $g \in E_h$ given by

$$g = (f_1 + \mu - \kappa + j_1, \dots, f_k + J + 1, \dots, f_l + j_l - 1, \dots, f_r + j_r) \tag{5.35}$$

for any $k \neq l \in \{1, \dots, r\}$. For any pair h, g of this form, $(B_\mu)_{h,g} \neq 0$ but $g \notin \mathcal{F}_\mu$. The number of shapes h that result in shapes g of the form (5.35) equals the number of ordered partitions of $-J$ into at most $r-1$ parts of magnitude $\le J$; equivalently, this is the number of ordered partitions

$$J(r-2) = j_2 + \dots + j_r, \quad 0 \le j_i \le 2J, i = 2, \dots, r,$$

which equals $\Pi_+ \triangleq \pi(r-1, 2J, J(r-2))$.

Next consider the row indices $h = (\dots, f_k - J, \dots) \in \mathcal{F}_\mu$ and column indices $g \in E_h$ given by

$$g = (f_1 + \mu - \kappa + j_1, \dots, f_k - J - 1, \dots, f_l + j_l + 1, \dots, f_r + j_r)$$

which again account for $(B_\mu)_{h,g} \neq 0$ and $g \notin \mathcal{F}_\mu$. The number of such shapes h equals the number of ordered partitions of Jr into $r-1$ or fewer parts $0 \le j_i \le 2J$. Denote this number by $\Pi_- \triangleq \pi(r-1, 2J, Jr)$. Note that as $J \to \infty$, both Π_+ and Π_- grow proportionally to J^{r-2}.

[1] To prove positivity, observe that the numbers $S_{r,m} = \sum_{i=0}^{\lfloor \frac{r}{2} \rfloor} (-1)^i \binom{r}{i} (r-2i)^m$ satisfy the recurrence

$$S_{r,m} = r^2 S_{r,m-2} + 4r(r-1) S_{r-2,m-2}, \quad 3 \le m \le r-1$$

and then use induction to prove that $S_{r.m} > 0$ (< 0) according as $r - m \equiv 1$ or $3 \bmod 4$.

It is easy to verify that $E_h \backslash \mathfrak{F}_\mu \neq \emptyset$ if and only if h and g are of the described form. Observe that by (5.25), $|E_f| = r^2 - r$. We then obtain

$$\sum_{h,g \in \mathfrak{F}_\mu} (B_\mu)_{h,g} \geq \psi_\mu (m - r(r^2 - r)(\Pi_+ + \Pi_-)) = \psi_\mu m(1 - o_m(1)).$$

The lemma is proved. ∎

We now consider the part of the product $\boldsymbol{y}^T \tilde{\mathbf{S}}_\kappa \boldsymbol{y}$ that involves the matrix $C_\mu, \mu = \kappa - s + 2, \dots, \kappa$. For a shape h let

$$D_h = \{(h_1, \dots, h_k - 1, \dots, h_r), 1 \leq k \leq r\}.$$

The proof of the next lemma is very similar to the above proof and will therefore be omitted.

Lemma 5.15 *Let* $e = \operatorname{argmin}_{h \in \mathfrak{F}_\mu} \big(\sum_{g \in D_h} (C_\mu)_{h,g} \big)$ *and let* $\phi_\mu = \sum_g (C_\mu)_{e,g}$.

$$\boldsymbol{y}_\mu^T C_\mu \boldsymbol{y}_{\mu-1} \geq \phi_\mu m(1 - o_m(1)).$$

To complete the proof of Theorem 5.13, compute

$$\begin{aligned} \lambda_\kappa &\geq \frac{1}{ms} \boldsymbol{y}^T \tilde{\mathbf{S}}_\kappa \boldsymbol{y} \\ &= \frac{1}{ms} \Big(\sum_{\mu=\kappa+1-s}^{\kappa} \boldsymbol{y}_\mu^T B_\mu \boldsymbol{y}_\mu + 2 \sum_{\mu=\kappa+2-s}^{\kappa} \boldsymbol{y}_\mu^T C_\mu \boldsymbol{y}_{\mu-1} \Big) \\ &\geq \frac{1}{s} \Big(\sum_{\mu=\kappa+1-s}^{\kappa} \psi_\mu + 2 \sum_{\mu=\kappa+2-s}^{\kappa} \phi_\mu \Big) (1 - o_m(1)) \\ &\geq \psi^* + 2 \frac{s-1}{s} \phi^* (1 - o_m(1)), \end{aligned}$$

where ψ^* (ϕ^*) is the smallest of the numbers ψ_μ (ϕ_μ) above. Note that both ψ^* and ϕ^* are nonzero. Now let $n \to \infty, \kappa = \tau n$, and let us choose f in the definition of $\mathfrak{F}_\mu$ to be of the form $f = (f_1, \dots, f_r), f_i = n\tau_i, 1 \leq i \leq r$. We assume that none of the τ_i's approach 0 as n grows. Take $s = o(n), s \to \infty$. Using (5.25), and letting $J = o(n), J \to \infty$ we get

$$\lim_{n \to \infty} \frac{\psi^*}{n} = \sum_{i=1}^{r} L_i \Big((q-2)\tau_i (q^r - q^{i-1}) + 2 \frac{(q-1)}{q} \sum_{k=1}^{i-1} \sqrt{\tau_k \tau_i q^{i+k}} \Big)$$

$$\lim_{n \to \infty} \frac{\phi^*}{n} = \sum_{i=1}^{r} L_i \sqrt{(1-\tau)\tau_i (q-1) q^{i-1}}.$$

Then since $\kappa/n \to \tau$,

$$\lim_{n \to \infty} \frac{\lambda_\kappa}{n} \geq \lim_{n \to \infty} \frac{\boldsymbol{y}^T \tilde{\mathbf{S}}_\kappa \boldsymbol{y}}{msn} \geq \Lambda(\tau_1, \dots, \tau_r).$$

The theorem is proved.

5.4.3 Asymptotic estimate for codes and OOAs

Theorems 5.12 and 5.13 together enable us to prove one of the main results of the chapter.

Theorem 5.16 *Let $R_{\mathrm{LP}}(\delta)$ be the function defined parametrically by the relations*

$$R(\tau) = \frac{1}{r}\left(h_q(\tau) + \tau \log_q \frac{q^r-1}{q-1}\right), \tag{5.36}$$

$$\delta(\tau) = \delta_{crit} - \frac{1}{r} \max_{\substack{\tau_i \ge 0 \\ \sum_{i=1}^r \tau_i = \tau}} \Lambda(\tau_1, \dots, \tau_r), \quad 0 \le \tau \le 1. \tag{5.37}$$

Then the asymptotic rate of any code family of relative distance δ satisfies $R \le R_{\mathrm{LP}}(\delta)$ and the rate of any family of OOAs of relative strength δ satisfies $R \ge 1 - R_{\mathrm{LP}}(\delta)$.

To prove this theorem, take the logarithms in (5.30) and pass to the limit as $n \to \infty$. Using the standard asymptotics for the binomial coefficient, we find that the code rate is bounded above by the right-hand side of (5.36). The condition $P(e) \le \lambda_{\kappa-1}$ of the Theorem 5.12 will be satisfied for large n if

$$\delta_{\mathrm{crit}} - \delta \le \frac{\lambda_{\tau n}}{rn}.$$

This defines the function in (5.37). Thus, the proof is complete.

Remark: For $r = 1$ this bound reduces to the linear programming bound on the rate of codes in [2]. Just as that result, the bound of this theorem improves upon the asymptotic Plotkin bound for large values of the code distance.

5.5 The case $r = 2$

In this section we prove a bound for codes in $\mathcal{Q}^{n,2}$ which improves upon the general result of the previous section. The improvement is due to the fact that in the case $r = 2$ it is possible to work with the polynomials $K_f(e)$ in their explicit form, and base the bound on the behavior of their zeros instead of the spectral radius of the operator S_κ. Below we use the notation $K_k(x; n)$ to denote the univariate Krawtchouk polynomial of degree k in variable x. This should be differentiated from the bivariate Krawtchouk polynomial $K_f(e)$ which is indexed by a shape $f = (f_1, f_2)$ and has a shape $e = (e_1, e_2)$ as its argument.

Let $f = (f_1, f_2), e = (e_1, e_2)$. From (5.27) we have

$$K_f(e) = q^{f_2} K_{f_2}(e_1; n - e_2) K_{f_1}(e_2; n - f_2).$$

We also have

$$P(e) = n\Big(2 - \frac{q+1}{q^2}\Big) - e_1 - 2e_2.$$

In the following text, we use the properties of the univariate Krawtchouk polynomial and the behavior of the roots of the univariate Krawtchouk polynomial which are stated in Section 2.4.3.

Remark: Properties (2.22)-(2.25) in Section 2.4.3 are usually stated for integer n. This is related to the fact that the polynomials $K_k(x;n)$ represent the eigenvalues of the Hamming association scheme. As pointed out by M. Aaltonen [1], it is possible to prove these properties for any $n \in \mathbb{R}^+$ relying on the generating function of the Krawtchouk polynomials:

$$(1+(q-1)z)^{n-x}(1-z)^x = \sum_{k=0}^{\infty} K_k(x;n)z^k.$$

This remark is important because in this section we sometimes use a non-integer parameter for n in $K_k(x;n)$.

The main result of this section is given in the following theorem.

Theorem 5.17 *The asymptotic rate of any family of codes of relative distance δ satisfies $R \le \Phi(\delta)$, where*

$$\Phi(\delta) = \min_{\tau_1,\tau_2} {}^1\!/_2 \left\{ \tau_2 + h_q(\tau_1) + (1-\tau_1)h_q\left(\frac{\tau_2}{1-\tau_1}\right)\right\},$$

where the minimum is taken over all τ_1, τ_2 that satisfy

$$0 \le \tau_1 \le (q-1)/q^2, \quad 0 \le \tau_2 \le (q-1)/q,$$
$$\gamma(\tau_2) + (2-\gamma(\tau_2))(1-\tau_2)\gamma(\tau_1) \le 2\delta.$$

The asymptotic rate of any family of OOAs of relative strength δ satisfies $R \ge 1-\Phi(\delta)$.

The remainder of the section is devoted to the proof of this result. We note that the polynomials $K_f(e)$ are formed as products of two Krawtchouk polynomials. A similar situation arose in [3] which dealt with the non-binary Johnson association scheme whose second eigenvalues are equal to a product of a Krawtchouk and a Hahn polynomial. Therefore, we adopt some elements of the analysis in [3] in our proof below.

In quest of an LP bound, we require a polynomial $F(e) = F(e_1, e_2)$ that satisfies conditions (2.10). Consider the polynomial of the form

$$F(e) = \big(P(e) - P(a)\big)\big(U_L(a,e)\big)^2 \tag{5.38}$$

for some $a = (\alpha, \beta)$ and a subset L. For brevity below we write S_{fh} instead of $(\mathbf{S}_\kappa)_{f,h}$ and denote $\bar{L} = \Delta_{n,2}\backslash L$. We find

$$\begin{aligned} F_0 = \langle F, 1\rangle &= \big\langle \big(P(e)-P(a)\big)U_L(a,e), U_L(a,e)\big\rangle \\ &= \Big\langle \sum_{f\in L}\frac{1}{v_f}\sum_{h\in\bar{L}} S_{fh}\big(-K_h(a)\big)K_f(e), U_L\Big\rangle \\ &= -\sum_{f\in L}\sum_{h\in\bar{L}} S_{fh}K_h(a)\frac{K_f(a)}{v_f}. \end{aligned}$$

In order to ensure that $F_0 > 0$ we will choose L and a so that

$$K_h(a) \le 0 \quad \text{if } h \in \bar{L}; \qquad K_f(a) > 0 \quad \text{if } f \in L. \tag{5.39}$$

Let $s = (s_1 - 1, s_2) \in \Delta_{n,2}$ be a shape that satisfies $\{(s_1 - 1, s_2 + 1), (s_1, s_2)\} \subset \Delta_{n,2}$. Let $a = (\alpha, \beta)$ satisfy

$$\beta = x_1(n - s_2, s_1), \quad x_1(n - \beta, s_2 + 1) < \alpha < x_1(n - \beta, s_2), \quad \alpha + 2\beta \le d. \tag{5.40}$$

For any $f_2, 0 \le f_2 \le s_2 + 1$ denote by $\phi(f_2)$ the degree such that

$$x_1(n - f_2, \phi(f_2) + 1) \le \beta < x_1(n - f_2, \phi(f_2)). \tag{5.41}$$

By (2.22), $\phi(\cdot)$ is well defined and implies the following:

$$[(x_1(n - u, w) > \beta) \;\Rightarrow\; (w \le \phi(u))], \quad [(x_1(n - u, w) \le \beta) \;\Rightarrow\; (w \ge \phi(u) + 1)].$$

We choose the region L to be given by

$$L = \{(f_1, f_2) : f_2 = 0, \dots, s_2; f_1 = 0, \dots, \phi(f_2)\}.$$

For the moment this choice is not unique because there are many possibilities for s. This ambiguity will be later removed by optimizing the bound on the choice of s.

To argue about the sign of F_0 we need to establish some properties of the region L. First, we claim that for a fixed f_2,

$$\phi(f_2) - 1 \le \phi(f_2 + 1) \le \phi(f_2). \tag{5.42}$$

Indeed, by (2.22),

$$\beta < x_1(n - f_2, \phi(f_2)) < x_1(n - f_2 - 1, \phi(f_2) - 1)$$

which implies the left-hand side of (5.42). On the other hand,

$$\beta \ge x_1(n - f_2, \phi(f_2) + 1) > x_1(n - f_2 - 1, \phi(f_2) + 1)$$

which implies the right-hand side.

The values of f, h for which $S_{fh} \ne 0$ are given in (5.25). In particular, if $f \in L$, then the set H of the shape vectors h that index the nonzero matrix elements of $\mathbf{S}$ and that lie outside the region L is as follows:

$$H = \{(\phi(f_2) + 1, f_2), f_2 = 0, 1, \dots, s_2\} \cup \{(f_1, s_2 + 1), f_1 = 0, 1, \dots, s_1 - 1\}.$$

The region L and the corresponding set H are shown in Fig. 5.1. By our choice of the parameters,

$$\begin{aligned}
K_{f_2}(\alpha; n - \beta) &> 0, && 0 \le f_2 \le s_2,\\
K_{s_2+1}(\alpha; n - \beta) &< 0, &&\\
K_{f_1}(\beta; n - f_2) &> 0, && 0 \le f_1 \le \phi(f_2),\ 0 \le f_2 \le s_2 + 1,\\
K_{\phi(f_2)+1}(\beta; n - f_2) &\le 0, && 0 \le f_2 \le s_2 + 1.
\end{aligned}$$

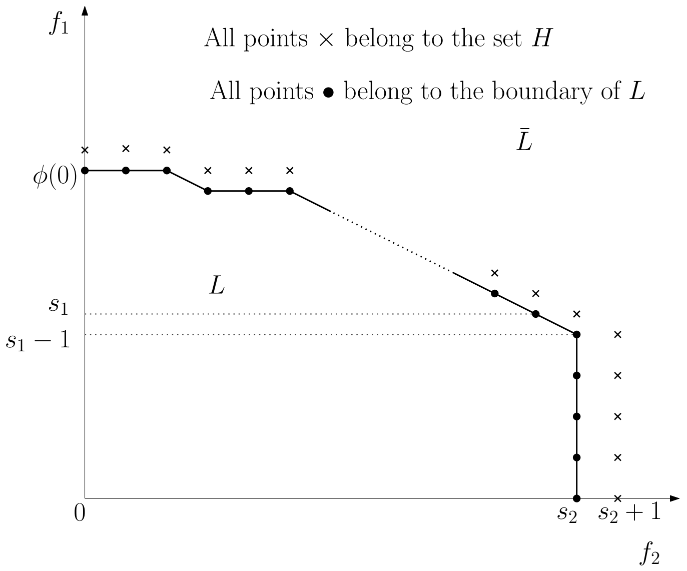

f_1
All points × belong to the set H
All points ● belong to the boundary of L
$\bar{L}$
$\phi(0)$
L
s_1
$s_1 - 1$
0
s_2
s_2+1
f_2

Figure 5.1: The region L

Then

$$K_{(f_1,f_2)}(a) = q^{f_2} K_{f_2}(\alpha; n-\beta) K_{f_1}(\beta; n-f_2) > 0, \quad f \in L, \tag{5.43}$$

$$K_{(\phi(f_2)+1,f_2)}(a) = q^{f_2} K_{f_2}(\alpha; n-\beta) K_{\phi(f_2)+1}(\beta; n-f_2) \le 0, \quad 0 \le f_2 \le s_2, \tag{5.44}$$

$$K_{(f_1,s_2+1)}(a) = q^{s_2+1} K_{s_2+1}(\alpha; n-\beta) K_{f_1}(\beta; n-s_2-1) < 0, \quad 0 \le f_1 \le s_1 - 1. \tag{5.45}$$

Thus,

$$\begin{aligned} K_f(a) &\le 0, \quad f \in \bar{L}, \\ K_f(a) &> 0, \quad f \in L. \end{aligned} \tag{5.46}$$

This proves that $F_0 > 0$.

Let us show that $F_e \ge 0$ for all e. For this rewrite F as follows:

$$\begin{aligned} F(e) &= \big(P(e) - P(a)\big) U_L(a,e)^2 \\ &= \sum_{f\in L} \frac{1}{v_f} \sum_{h \in \bar{L}} S_{fh} \big(K_h(e)K_f(a) - K_h(a)K_f(e)\big) \sum_{g\in L} \frac{K_g(a)K_g(e)}{v_g} \\ &= \Big(\sum_{h\in\bar{L}} K_h(e) \sum_{f\in L} \frac{S_{fh}K_f(a)}{v_f} - \sum_{f\in L} K_f(e) \sum_{h\in\bar{L}} \frac{K_h(a)S_{fh}}{v_f} \Big) \sum_{g\in L} \frac{K_g(a)K_g(e)}{v_g} \\ &= \sum_{h\in\bar{L},\, g\in L} \frac{K_g(a)}{v_g} K_h(e) K_g(e) \sum_{f\in L} \frac{S_{fh}K_f(a)}{v_f} \\ &\qquad - \sum_{f,g\in L} \frac{K_g(a)}{v_g} K_f(e) K_g(e) \sum_{h\in\bar{L}} \frac{K_h(a)S_{fh}}{v_f}. \end{aligned}$$

By (5.19), the products $K_h K_g$ and $K_f K_g$ are expanded in the basis $\{K_f\}$ with nonnegative coefficients. Moreover, the other terms in the above formula also have the needed signs on account of (5.43)-(5.45). This establishes our claim.

Finally, because of the third condition in (5.40), $F(e) \le 0$ for all e with $|e|' \ge d$.

We are now able to formulate the bound on codes and OOAs.

Theorem 5.18 *Let $\mathcal{C}$ be an $(2n, M, d)$ code $\mathcal{C} \subset \overrightarrow{\mathcal{H}}(q, n, 2)$. Then*

$$M \le \frac{4(n-\beta-s_2)(n-s_2-s_1+1)^2(q-1)^3(\alpha+2\beta)}{q^3\alpha^2\beta^2} v_s, \qquad (5.47)$$

where $s = (s_1 - 1, s_2)$ satisfies $\{(s_1-1, s_2+1), (s_1, s_2)\} \subset \Delta_{n,2}$ and $a = (\alpha, \beta)$ is chosen to fulfill conditions (5.40).

Let $\mathcal{C}$ be a $(t = d-1, n, 2, q)$ OOA of size M. Then

$$M \ge \frac{q^{nr}}{v_s} \frac{q^3\alpha^2\beta^2}{4(n-\beta-s_2)(n-s_2-s_1+1)^2(q-1)^3(\alpha+2\beta)}. \qquad (5.48)$$

Proof: Let us compute $F_0 = \langle F, 1\rangle$. Denote $\sigma_1 = (s_1-1, s_2+1), \sigma_2 = (s_1-2, s_2+1)$. By (5.46) and (5.21) we have

$$\begin{aligned} F_0 &= -\sum_{f\in L} \frac{K_f(a)}{v_f} \sum_{h\in \bar{L}} S_{fh} K_h(a) \\ &\ge -\frac{K_s(a)}{v_s}(S_{s,\sigma_1}K_{\sigma_1}(a) + S_{s,\sigma_2}K_{\sigma_2}(a) + S_{s,(s_1,s_2)}K_{(s_1,s_2)}(a)) \\ &= -\frac{K_s(a)}{q^2 v_s}\Big((s_2+1)K_{\sigma_1}(a) + (s_2+1)(q-1)K_{\sigma_2}(a) + s_1(q+1)K_{(s_1,s_2)}(a)\Big) \\ &= -\frac{K_s(a)(s_2+1)q^{s_2+1}}{q^2 v_s} K_{s_2+1}(\alpha; n-\beta) K_{s_1-1}(\beta; n-s_2). \qquad (5.49) \end{aligned}$$

Let us now evaluate $U_L(a, 0)$.

$$U_L(a,0) = \sum_{f\in L} K_f(a) = \sum_{f_2=0}^{s_2} q^{f_2} K_{f_2}(\alpha; n-\beta) \sum_{f_1=0}^{\phi(f_2)} K_{f_1}(\beta; n-f_2).$$

Let us bound above the last sum. We shall prove that

$$\sum_{f_1=0}^{\phi(f_2)} K_{f_1}(\beta; n-f_2) \le q^{s_2-f_2} \sum_{f_1=0}^{\phi(s_2)} K_{f_1}(\beta; n-s_2). \qquad (5.50)$$

Indeed, using (2.24), we obtain

$$\begin{aligned} \sum_{f_1=0}^{\phi(f_2)} K_{f_1}(\beta; n-f_2) &= \sum_{f_1=0}^{\phi(f_2)} \big(K_{f_1}(\beta; n-f_2-1) + (q-1)K_{f_1-1}(\beta; n-f_2-1)\big) \\ &= K_{\phi(f_2)}(\beta; n-f_2-1) + q \sum_{f_1=0}^{\phi(f_2)-1} K_{f_1}(\beta; n-f_2-1). \end{aligned}$$

Recall that $\phi(f_2+1)=\phi(f_2)$ or $\phi(f_2+1)=\phi(f_2)-1$. In the former case,

$$K_{\phi(f_2)}(\beta;n-f_2-1)+q\sum_{f_1=0}^{\phi(f_2)-1}K_{f_1}(\beta;n-f_2-1)\le q\sum_{f_1=0}^{\phi(f_2+1)}K_{f_1}(\beta;n-f_2-1);$$

in the latter, $K_{\phi(f_2)}(\beta;n-f_2-1)=K_{\phi(f_2+1)+1}(\beta;n-f_2-1)\le 0$ on account of (5.41) and (2.22). Repeating this procedure s_2-f_2 times, we arrive at (5.50).

Note that $\phi(s_2)=s_1-1$. Therefore

$$U_L(a,0)\le q^{s_2}\sum_{f_2=0}^{s_2}K_{f_2}(\alpha;n-\beta)\sum_{f_1=0}^{s_1-1}K_{f_1}(\beta;n-s_2).$$

By (2.25), (5.40), and (5.16) we have

$$\sum_{f_1=0}^{s_1-1}K_{f_1}(\beta;n-s_2)=\frac{s_1\binom{n-s_2}{s_1}(q-1)}{q\beta\binom{n-s_2}{s_1-1}}K_{s_1-1}(\beta;n-s_2),$$

$$\begin{aligned}\sum_{f_2=0}^{s_2}K_{f_2}(\alpha;n')&=\frac{(s_2+1)\big(K_{s_2+1}(0;n')K_{s_2}(\alpha;n')-K_{s_2+1}(\alpha;n')K_{s_2}(0;n')\big)}{q\alpha K_{s_2}(0;n')}\\&=\frac{s_2+1}{q\alpha}K_{s_2}(\alpha;n')\big(W(0)-W(\alpha)\big),\end{aligned}$$

where $n'=n-\beta$ and $W(x)=K_{s_2+1}(x;n')/K_{s_2}(x;n')$. Using these expressions, we can bound $U_L(a,0)$ as

$$U_L(a,0)\le\frac{(s_2+1)(n-s_2-s_1+1)(q-1)}{q^2\alpha\beta}K_s(a)\big(W(0)-W(\alpha)\big).$$

Hence using (2.11), (5.38), and (5.49) we can write

$$M\le v_s\frac{(s_2+1)(n-s_2-s_1+1)^2(q-1)^2(\alpha+2\beta)}{q^3\alpha^2\beta^2}\frac{\big(W(0)-W(\alpha)\big)^2}{-W(\alpha)}.$$

Since $W(0)=(q-1)(n'-s_2)/(s_2+1)>0>W(\alpha)>-\infty$ as α ranges in between the bounds in (5.40), it is possible to find α such that $W(\alpha)=-W(0)$. With this choice and (2.8) the last expression turns into (5.47). The estimate (5.48) follows from (2.12). ■

The proof of Theorem 5.17 is obtained by passing to asymptotics in (5.47). Namely, let $n\to\infty, d/nr\to\delta, s_1/n\to\tau_1, s_2/n\to\tau_2$. By (2.23),

$$\limsup_{n\to\infty}\frac{\beta}{n}=\gamma(\tau_1)(1-\tau_2),\quad \limsup_{n\to\infty}\frac{\alpha}{n}=\gamma(\tau_2)\big(1-\gamma(\tau_1)(1-\tau_2)\big).$$

Computing the logarithm on the right-hand side of (5.47), we observe that the only term of exponential growth arises from v_s. Using standard estimates we obtain

$$\begin{aligned}\log_q v_s &= \log_q \binom{n}{s_1-1}\binom{n-s_1+1}{s_2} q^{s_2}(q-1)^{s_1+s_2-1} \\ &\le n\Big\{\tau_2 + h_q(\tau_1) + (1-\tau_1)h_q\Big(\frac{\tau_2}{1-\tau_1}\Big)\Big\}.\end{aligned}$$

The tightest bound is obtained by computing the minimum of this expression on τ_1, τ_2. The range of the variables τ_1, τ_2 is obtained on observing that $n(q-1)/q^2$ and $n(q-1)/q$ are the maximizing values of s_1, s_2 for large n (by a direct calculation from the above expression; or, specializing from a general result in [72]). The third restriction in the statement of the theorem is implied by $\alpha + 2\beta \le d$. This completes the proof.

Asymptotic bounds for ordered codes are shown in several plots in Figures 5.2 and 5.3.

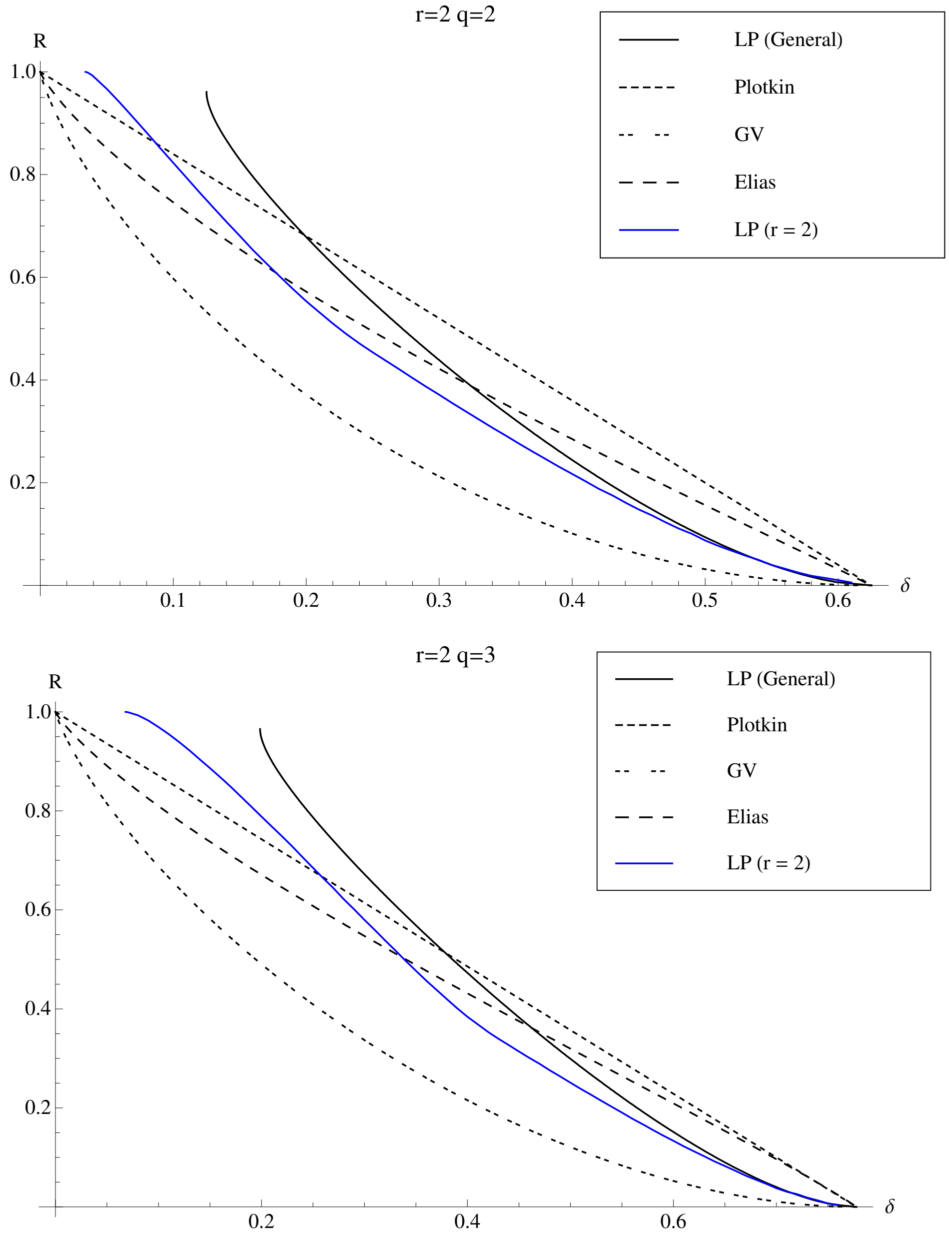


Figure 5.2: Bounds on codes for $r = 2$.

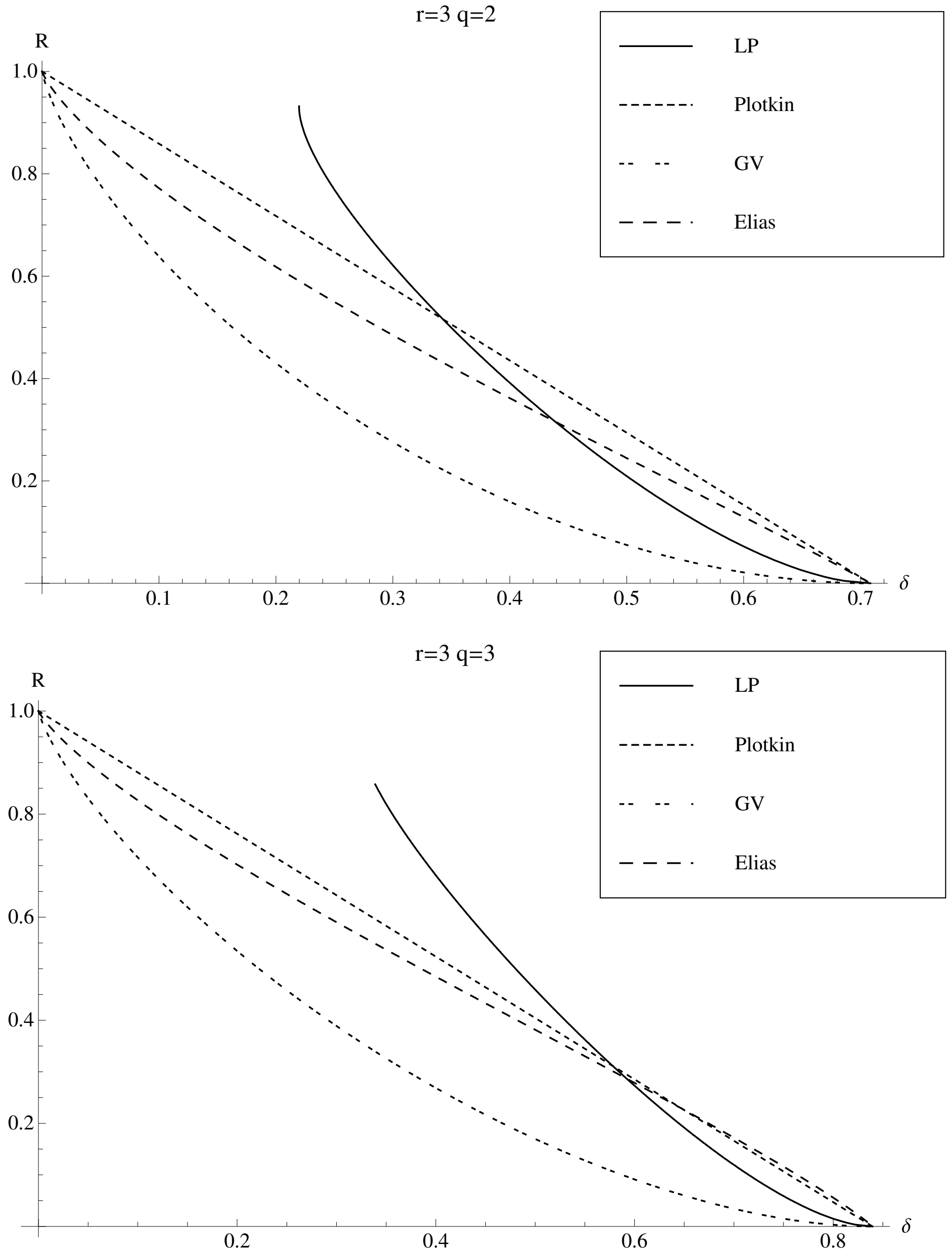


Figure 5.3: Bounds on codes for $r = 3$.

CHAPTER 6

Bounds on sets with few intersections

Let $\mathcal{F} = \{U_1, U_2, \dots\} \subset 2^{[n]}$ be a family of subsets of $[n]$. Suppose that $|U_i \cap U_j| \in \{r_1, \dots, r_l\}$ for all $i \neq j$, The problem that we are considering in this chapter is to find the maximum size of the family $\mathcal{F}$ under this condition. The study of problems of this kind dates back to the work of Fisher [36] on experimental designs, and to the celebrated Erdös-Ko-Rado theorem [34]. Several versions of this basic problem have been considered over the years and in some cases the extremal families which achieve the bounds have been determined [70, 27, 37, 38, 5].

An upper bound on $|\mathcal{F}|$ was provided by Frankl and Wilson [37].

Theorem 6.1 [37] *Let $\mathcal{F} \subset 2^{[n]}$ be a family of subsets of $[n]$ with l intersections. Then*

$$|\mathcal{F}| \leq \sum_{i=0}^{l} \binom{n}{i}.$$

As shown by Qian and Ray-Chaudhuri [69], the Frankl-Wilson bound is achieved if and only if $\mathcal{F}$ consists of all the subsets of $[n]$ of size $\leq l$. An improvement to the bound was proved by Füredi and Katona [39].

Theorem 6.2 [39] *Let $\mathcal{F} \subset 2^{[n]}$ be a family of subsets of $[n]$ such that $|U_i \cap U_j| \in \{a, a+1, \dots, b\}$ $(i \neq j)$, with $b \geq 2a$. Then for $n > n_0(a,b)$, where $n_0(a,b)$ is a constant depending only on a and b,*

$$|\mathcal{F}| \leq \sum_{i=0}^{b-a} \binom{n-a}{i} + \binom{n-a}{b+1-a}.$$

In this chapter we prove another improvement of the Frankl-Wilson bound for the case that $\{r_1, \dots, r_l\} = \{n-1, \dots, n-l\}$, $l \leq \left\lfloor \frac{n}{2} \right\rfloor$, where the Füredi-Katona theorem is not applicable. The bound is obtained by studying certain polynomials $E_{k,s,t}$ (intertwining functions of the symmetric group) associated with a refinement of the Bose-Mesner algebra of the Hamming scheme, called the Terwilliger algebra [83, 74] (see Section 2.3). The extension considered above uses the fact that sets with few intersections can be described in terms of triples rather than pairs of points. Such triples form precisely the type of objects described by the Terwilliger algebra and the functions associated with it.

This chapter is organized as follows. In the next section we study the decomposition of functions on the binary Johnson space, refining the decomposition in Section 2.4. The objective is to show how the polynomials $E_{k,s,t}$ arise from this decomposition. In Section 6.1.3 we recall the expansion of the Krawtchouk polynomials in terms of $E_{k,s,t}$. Finally, in Section 6.2 we will use the properties of these polynomials to derive the new bound.

Publications: The results in this chapter will be submitted for publication.

6.1 Intertwining functions of the symmetric group

The material in this section is a summary of the detailed results in [32, 85], presented in a way convenient for our goals.

6.1.1 Decomposition of the Johnson space

Recall from Section 2.4 that the isometry group of the binary Hamming space $\mathcal{H} = \mathcal{H}(2, n)$ is given by the group $\mathcal{H} \rtimes S_n$, where $\mathcal{H}$ is considered as an additive group and S_n is the symmetric group on n letters. The action of this group on $\mathcal{H}$ results in the decomposition of the $\mathcal{H}$ into disjoint spheres $\mathcal{S}_s,\ s = 0, \dots, n$. Next, the group action on the space of square-integrable functions $L^2(\mathcal{H})$ results in the decomposition of the space into an orthogonal direct sum of irreducible subspaces $H_s,\ s = 0, \dots, n$, where each H_s is a space of S_n-invariant functions. In this section we are concerned with further decomposition of the subspaces $\mathcal{S}_s, H_s$ under the action of its isometry group $G = S_n$ (the stationary subgroup of the origin in $\mathcal{H}$).

The group G acts transitively on $\mathcal{S}_s$, since any vector of weight s can be mapped to any other vector of weight s by a permutation of its coordinates. We can identify $\mathcal{S}_s$ as a homogeneous space for G as follows. Fix the point $\boldsymbol{b} = (1, \dots, 1, 0, \dots, 0) \in \mathcal{S}_s$ that has 1 in the first s coordinates and 0 elsewhere. The subgroup $H = S_s \times S_{n-s}$ is the stationary subgroup of $\boldsymbol{b}$. Hence we can realise $\mathcal{S}_s$ as the (right) cosets $H\backslash G$ of G by identifying $H\sigma$ with the vector $\boldsymbol{u} \in \mathcal{S}_s$ such that $\sigma \cdot \boldsymbol{u} = \boldsymbol{b}$. The orbit of this action is formed by the vectors that correspond to a double coset of H in G, i.e., to an element of $H\backslash G/H$. This enables us to define the zonal spherical functions for the Johnson space.

6.1.2 Decomposition of functions on the Johnson space

In analogy with the Hamming space, one can study the action of $G = S_n$ on the vector space of functions $L^2(\mathcal{S}_s) = L^2(H\backslash G)$. Under the action of G this space decomposes into pairwise orthogonal irreducible subspaces [32, 85],

$$L(\mathcal{S}_s) = H_s = \begin{cases} H_{0,s} \perp \cdots \perp H_{s,s}, & 0 \le s \le \left\lfloor \frac{n}{2} \right\rfloor, \\ H_{0,s} \perp \cdots \perp H_{n-s,s}, & \text{otherwise,} \end{cases}$$

where the subspaces $H_{k,s}$ have the following properties:

1. $\dim H_{k,s} = h_k \triangleq \binom{n}{k} - \binom{n}{k-1}$;

2. Any two distinct subspaces $H_{k,s}, H_{j,t}$ are pairwise orthogonal;

3. For any k, j, s, t such that $k \le s, t \le n-k$ the spaces $H_{k,s}$ and $H_{k,t}$ are isometric.

The subspace $H_{k,s}$ has an orthonormal basis of spherical harmonics $\{e_{k,s,l}(\boldsymbol{u}) : l = 0, \dots, h_k\}$. The explicit expression for the basis functions is given, for instance, in [25, Theorem 4].

The zonal spherical kernel of degree k is given by

$$E_{k,s,s}(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{|\mathcal{H}|} \sum_{i=1}^{h_k} e_{k,s,i}(\boldsymbol{u}) e_{k,s,i}(\boldsymbol{v}).$$

The corresponding zonal spherical function ϕ_k is constant on the double coset $H \backslash G / H$. The zonal spherical kernel is invariant under the action of H:

$$E_{k,s,s}(\sigma \cdot \boldsymbol{u}, \sigma \cdot \boldsymbol{v}) = E_{k,s,s}(\boldsymbol{u}, \boldsymbol{v}), \;\; \sigma \in H.$$

It is in fact a function of only one variable as described below. For any vector $\boldsymbol{u} \in \mathcal{S}_s$ let

$$r = |\{\operatorname{supp}(u_{s+1}, \dots, u_n)\}|.$$

In other words, if $\sigma \in S_n$ is applied to the fixed vector $\boldsymbol{b}$, then r is the number of nonzero entries of $\sigma \cdot \boldsymbol{b}$ outside the first s coordinates:

$$r = s - |\sigma(\{1, \dots, s\}) \cap \{1, \dots, s\}|.$$

The function $E_{k,s,s}(\boldsymbol{u}, \boldsymbol{v})$ depends only on the parameter r, and is hence a univariate polynomial (this happens because the action of S_n on $\mathcal{S}_s$ is distance-transitive). An explicit expression for $E_{k,s,s}$ will be provided in the next section.

6.1.3 Intertwining functions on the Hamming space

Let K be a subgroup of G different from H. A function on G that is constant on the double cosets $H \backslash G / K$ is called an *intertwining function* of G. In this section we describe such functions. The main aim here is to state the properties of the intertwining functions and define the expansion of the Krawtchouk polynomial into the basis of intertwining functions.

We identify the Johnson space $\mathcal{S}_s$ with the set of right cosets $H \backslash G$, where $H = S_s \times S_{n-s}$ and $G = S_n$. The action of the subgroup $K = S_t \times S_{n-t}$ on $\mathcal{S}_s$ partitions the space $\mathcal{S}_s$ into orbits which are in one-to-one correspondence with the double cosets $H \backslash G / K$. The orbits are parameterized by the quantity

$$r = |\{\operatorname{supp}(u_{t+1}, \dots, u_n)\}|, \text{ where } \boldsymbol{u} \in \mathcal{S}_s.$$

In other words, r is the number of nonzero entries of the vector $\sigma \cdot \boldsymbol{b}, \sigma \in S_n$, outside the first t coordinates:

$$r = s - |\sigma(\{1, \dots, s\}) \cap \{1, \dots, t\}|.$$

A function ϕ on G is called intertwining if it is H-K invariant, i.e., constant on the double cosets $H \backslash G / K$:

$$\phi(hgk) = \phi(g), \quad h \in H,\ g \in G,\ k \in K.$$

These functions were found by Dunkl in [32]. The kernels (cf. Section 2.4) that correspond to them are denoted by $E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v})$. They are univariate polynomials in r, where $r = |\operatorname{supp} \boldsymbol{u} \cap \operatorname{supp} \boldsymbol{v}|$ and up to a scaling coincide with the Hahn polynomials. Below we state some properties of the functions $E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v})$ [32, 33, 85].

1. For $k = 0, \dots, \left\lfloor \frac{n}{2} \right\rfloor, s = k, \dots, n-k$, let $\{e_{k,s,i}(\boldsymbol{u}) : i = 1, \dots, h_k\}$ be the orthonormal basis of spherical harmonics in $H_{k,s}$. Then for $\boldsymbol{u} \in \mathcal{S}_s$ and $\boldsymbol{v} \in \mathcal{S}_t$, the H-K invariant (intertwining) function $E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v})$ is given by

$$E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{|\mathcal{H}|} \sum_{i=1}^{h_k} e_{k,s,i}(\boldsymbol{u}) e_{k,t,i}(\boldsymbol{v}). \tag{6.1}$$

   The function $E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v}) = 0$ if either $\boldsymbol{u} \notin \mathcal{S}_s$ or $\boldsymbol{v} \notin \mathcal{S}_t$.

2. Let $U = \operatorname{supp} \boldsymbol{u}$ and $V = \operatorname{supp} \boldsymbol{v}$. $E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v})$ is a polynomial of degree k in $r = |U \cap V|$, and it is given by the expression

$$E_{k,s,t}(r) = \frac{1}{b_k(s,t)} Q_k(s-r; -(n-t)-1, -t-1, s), \tag{6.2}$$

   where $Q_k(s-r; -(n-t)-1, -t-1, s)$ is the Hahn polynomial of degree k in the variable $(s-r)$ normalized by $Q_k(0; -(n-t)-1, -t-1, s) = 1$ and

$$b_k(s,t) = \frac{\sqrt{\binom{n}{s}\binom{n}{t}}}{h_k} \sqrt{\prod_{i=0}^{k-1} \frac{(-t+i)(s-n+i)}{(-s+i)(t-n+i)}}.$$

   The explicit form of the Hahn polynomial is as follows:

$$Q_k(s-r; -(n-t)-1, -t-1, s) = \frac{1}{\binom{s}{k}} \sum_{j=0}^{k} (-1)^j \frac{\binom{t-k+j}{j}}{\binom{n-t}{j}} \binom{r}{k-j} \binom{s-r}{j}.$$

3. By (6.1), $E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v}) = E_{k,t,s}(\boldsymbol{v}, \boldsymbol{u})$. Hence the matrix $E_{k,s,t} = (E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v}))_{\boldsymbol{u},\boldsymbol{v} \in \mathcal{H}}$ is symmetric.

4. The matrices $E_{k,s,t}$ satisfy the following relation

$$E_{k,s,t} E_{k',s',t'} = E_{k,s,t'} \delta_{k,k'} \delta_{t,s'}. \tag{6.3}$$

5. $\mathrm{Tr}(E_{k,s,s}) = h_k = \binom{n}{k} - \binom{n}{k-1}$.

We now state the expansion of the Krawtchouk polynomial into the basis of the intertwining functions $E_{k,s,t}(r)$ (the "addition formula" for the Krawtchouk polynomials) [32]. Let $K_k(x;n)$ be the Krawtchouk polynomial of degree k (2.15) *normalized by* $K_k(0;n) = 1$,

$$K_k(x;n) = \frac{1}{\binom{n}{k}} \sum_{i=0}^{k} (-1)^i \binom{x}{i} \binom{n-x}{k-i}.$$

Let $\boldsymbol{u} \in \mathcal{S}_s$ and $\boldsymbol{v} \in \mathcal{S}_t$. Recall from Section 2.4 that $K_k(\boldsymbol{u}, \boldsymbol{v})$ depends only the distance $d_H(\boldsymbol{u}, \boldsymbol{v})$. Let $0 \le s \le t \le n$ and let $s - \min\{s, n-t\} \le r \le s$. Then

$$\begin{aligned} K_k(s+t-2r;n) = \sum_{j=0}^{\min\{s,n-k,k,n-t\}} & 2^{2j} \frac{\binom{n-2j}{k-j}\binom{s}{j}\binom{n-t}{j}}{\binom{n}{k}\binom{n-j+1}{j}} b_j(s,t) \\ & \times K_{k-j}(s-j;n-2j) K_{k-j}(t-j;n-2j) E_{j,s,t}(r). \quad (6.4) \end{aligned}$$

Relation of $E_{k,s,t}(r)$ to the Terwilliger algebra: The polynomials $E_{k,s,t}(r)$ are related to the Terwilliger algebra $\mathbb{T}$ of the Hamming space. Recall from Section 2.3 that the matrices in the Terwilliger algebra can be simultaneously block-diagonalized by a unitary matrix. This transformation establishes the following isomorphism of algebras:

$$\mathbb{T} \cong \bigoplus_{k=0}^{\lfloor n/2 \rfloor} I_{h_k} \otimes \mathbb{C}^{(n-2k+1)\times(n-2k+1)}.$$

It is possible to choose a basis on the right-hand side so that the entries of the blocks are (up to scale factors) given by the evaluations of the polynomials $E_{k,s,t}(r)$ described above (see [74], Equations (7),(8)).

6.2 New bound on sets with few intersections

In this section we return to the problem stated in the introduction to this chapter. An improvement of the Frankl-Wilson bound (Theorem 6.1) is given in the following theorem.

Theorem 6.3 *Let* $2 \le l \le \left\lfloor \frac{n}{2} \right\rfloor$*, and let* $\{r_1, \ldots, r_l\} = \{n-1, \ldots, n-l\}$*, with* $r_i = n - i$*. If* $\mathcal{F}$ *is a collection of* M *subsets of* $[n]$ *such that* $|U \cap V| \in \{n-1, \ldots, n-l\}$ *for any distinct* $U, V \in \mathcal{F}$*, then*

$$M \le 2\binom{n}{l-1} + \binom{n}{l-2} + \cdots + \binom{n}{0}.$$

Note that the Füredi-Katona theorem is not applicable in this range of values of $\{r_1, \ldots, r_l\}$.

We proceed with the proof of this theorem. Below we work with characteristic vectors of the subsets in $\mathcal{F}$. Let $U, |U| = s$ and $V, |V| = t$ be two such subsets, and let $\boldsymbol{u}, \boldsymbol{v}$ be their characteristic vectors in $[n]$. We have

$$d_H(\boldsymbol{u}, \boldsymbol{v}) = s + t - 2|U \cap V|. \tag{6.5}$$

The driving idea of the proof is that given s, t and $r = |U \cap V|$ we can compute $d_H(\boldsymbol{u}, \boldsymbol{v})$. These three parameters exactly match the parameters of the functions $E_{k,s,t}(r)$.

The annihilator polynomial of the family $\mathcal{F}$ is defined by

$$f(r) = (r - r_1)(r - r_2) \dots (r - r_l).$$

Remark: Two sets U, V can have intersection $|U \cap V| \geq n - l$ only for $|U|, |V| \geq n - l$. Moreover, for sets U, V with $|U|, |V| \geq n - l$ we have that $E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v}) = 0$ for all $k \leq s, t < n - l$, where $\boldsymbol{u}, \boldsymbol{v}$ are the characteristic vectors of the sets in $[n]$.

Lemma 6.4 *For all sets U, V such that $|U|, |V| \geq n - l$, $|U \cap V| = r$, and $n - l \leq r \leq n - 1$, we have*

$$\begin{aligned} f(|U \cap V|) &= \sum_{k=0}^{l} \sum_{s,t=n-l}^{n-k} A_{k,s,t} E_{k,s,t}(\boldsymbol{u}, \boldsymbol{v}), \quad \textit{or} \\ f(r) = 0 &= \sum_{k=0}^{l} \sum_{s,t=n-l}^{n-k} A_{k,s,t} E_{k,s,t}(r), \quad A_{k,s,t} \in \mathbb{R}, \end{aligned} \tag{6.6}$$

and for any $k = 0, \dots, l,\ s, t = n - l, \dots, n - k$, we have $A_{k,s,t} = A_{k,t,s}$.

Proof: We expand the polynomial $f(r)$ into the basis of Krawtchouk polynomials of degree at most l. For fixed s, t we have

$$f(r) = \sum_{k=0}^{l} f_{k,s,t} K_k(s + t - 2r), \quad f_{k,s,t} \in \mathbb{R}. \tag{6.7}$$

First we consider the range $n \geq t \geq s \geq n - l$. We rewrite the addition formula (6.4) as

$$K_k(s + t - 2r; n) = \sum_{j=0}^{\min\{k, n-t\}} c_{j,k,s,t,n} E_{j,s,t}(r),$$

where $s + t - n \leq r \leq s$ and

$$\begin{aligned} c_{j,k,s,t,n} &= 2^{2j} \frac{\binom{n-2j}{k-j}\binom{s}{j}\binom{n-t}{j}}{\binom{n}{k}\binom{n-j+1}{j}} b_j(s,t) K_{k-j}(s - j; n - 2j) K_{k-j}(t - j; n - 2j) \\ &= 2^{2j} \frac{\binom{n-2j}{k-j}}{h_k \binom{n}{k}\binom{n-j+1}{j}} \sqrt{\binom{s}{j}\binom{t}{j}\binom{n-s}{j}\binom{n-t}{j}\binom{n}{s}\binom{n}{t}} \times \\ &\quad K_{k-j}(s - j; n - 2j) K_{k-j}(t - j; n - 2j) \end{aligned} \tag{6.8}$$

Using the fact that $l \leq \lfloor n/2 \rfloor$, and using the above addition formula, we get for $s+t-n \leq r \leq s,\ n \geq t \geq s$,

$$\begin{aligned} f(r) &= \sum_{k=0}^{l} f_{k,s,t} \sum_{j=0}^{\min\{k,n-t\}} c_{j,k,s,t,n} E_{j,s,t}(r) \\ &= \sum_{j=0}^{\min\{l,n-t\}} \left(\sum_{k=j}^{l} f_{k,s,t} c_{j,k,s,t,n} \right) E_{j,s,t}(r) \\ &= \sum_{k=0}^{\min\{l,n-t\}} g_{k,s,t} E_{k,s,t}(r), \quad g_{k,s,t} \in \mathbb{R}, \end{aligned} \tag{6.9}$$

where we collect all the terms independent of r in the last expression as $g_{k,s,t}$.

By definition $E_{k,s,t}(r)$ is identically zero for all values of s,t in the case that no two sets U,V of sizes $|U|=s,\ |V|=t$ give rise to an intersection $|U \cap V| = r$. Thus, given the value of r, we require the valid ranges of s and $t \geq s$ for which the term $E_{k,s,t}(r)$ is not identically zero. This can be determined as follows. For a fixed r, and given $s \geq n-l$, we obviously have $s \geq \max\{n-l, r\}$ since no set of size less than r can give rise to an intersection of size r. The size of the intersection between sets $U, |U| = s$, and $V, |V| = t$ is exactly r only if

$$\begin{aligned} t-r &\leq n-s, \quad \text{or} \\ s-r \leq t-r &\leq n-s, \quad \text{or} \\ s &\leq \frac{n+r}{2}. \end{aligned}$$

Therefore, for $E_{k,s,t}(r)$ not identically zero and for vectors of weight at least $n-l$, the valid ranges for s,t are

$$\begin{aligned} \max\{n-l, r\} \leq s &\leq \frac{n+r}{2}, \\ s \leq t &\leq n+r-s. \end{aligned} \tag{6.10}$$

For values of s,t not in the range above and for $n-l \leq r \leq n-1$, equation (6.9) still holds since $f(r) = 0$ and $E_{k,s,t}(r) = 0$. Hence, for a fixed r, taking the sum of (6.9) over all $n-l \leq s \leq n,\ s \leq t \leq n$, we get

$$\begin{aligned} 0 &= \sum_{s=n-l}^{n} \sum_{t=s}^{n} \sum_{k=0}^{\min\{n-t,l\}} g_{k,s,t} E_{k,s,t}(r) \\ &= \sum_{k=0}^{l} \sum_{s=n-l}^{n-k} \sum_{t=s}^{n-k} g_{k,s,t} E_{k,s,t}(r). \end{aligned} \tag{6.11}$$

Interchanging the summations between s and t in the above equation, and using $E_{k,s,t}(r) = E_{k,t,s}(r)$, we get

$$\begin{aligned} 0 &= \sum_{k=0}^{l} \sum_{t=n-l}^{n-k} \sum_{s=n-l}^{t} g_{k,s,t} E_{k,s,t}(r) \\ &= \sum_{k=0}^{l} \sum_{t=n-l}^{n-k} \sum_{s=n-l}^{t} g_{k,s,t} E_{k,t,s}(r). \end{aligned}$$

Now, from (6.7) we note that for every k, the coefficient $f_{k,s,t}$ is symmetric in s, t and from (6.8) we can deduce that $c_{j,k,s,t,n}$ is also symmetric in s, t for every j, k. Hence $g_{k,s,t} = g_{k,t,s}$. Thus we can rewrite the above equation as

$$0 = \sum_{k=0}^{l} \sum_{t=n-l}^{n-k} \sum_{s=n-l}^{t} g_{k,t,s} E_{k,t,s}(r) \tag{6.12}$$

Combining equations (6.11) and (6.12) (after relabelling s and t) we get (6.6), with

$$A_{k,s,t} = \begin{cases} g_{k,s,t}, & n-l \le s \ne t \le n-k, \\ 2g_{k,s,s}, & n-l \le s \le n-k. \quad \blacksquare \end{cases} \tag{6.13}$$

Let $\mathcal{F}' = \mathcal{F} \setminus \{[n]\}$ and let $M' = |\mathcal{F}'|$. For $k = 0, \ldots, l$, and $i = 1, \ldots, h_k$ define

$$\begin{aligned} &\mathcal{E}_{k,i}(\boldsymbol{u}) = [e_{k,n-l,i}(\boldsymbol{u}) \ e_{k,n-l+1,i}(\boldsymbol{u}) \cdots e_{k,n-k,i}(\boldsymbol{u})], && (1 \times (l-k+1) \text{ row vector}) \\ &h(\boldsymbol{u}) = [\mathcal{E}_{0,1}(\boldsymbol{u}) \cdots \mathcal{E}_{0,h_0}(\boldsymbol{u}) \cdots \mathcal{E}_{l,1} \cdots \mathcal{E}_{l,h_l}(\boldsymbol{u})], && \Big(1 \times \sum_{i=0}^{l} \binom{n}{i} \text{ row vector}\Big) \\ &H(\mathcal{F}) = \big(h(\boldsymbol{u})\big)_{\boldsymbol{u} \in \mathcal{F}'}, && \Big(M' \times \sum_{i=0}^{l} \binom{n}{i} \text{ matrix}\Big) \\ &F = \big(f(|U \cap V|)\big)_{U,V \in \mathcal{F}'}, && (M' \times M' \text{ matrix}). \end{aligned} \tag{6.14}$$

For $k = 0, \ldots, l$ let $A_k = (A_{k,s,t})_{s,t \in \{n-l,\ldots,n-k\}}$ denote the matrix of the coefficients in (6.6). Then we have the following

Lemma 6.5 *Let F be an $M' \times M'$ all-zero matrix. Then*

$$F = \frac{1}{|\mathcal{H}|} H(\mathcal{F}') \mathbf{A} H(\mathcal{F}')^T, \tag{6.15}$$

where $\mathbf{A} = \big(\bigoplus_{k=0}^{l} (I_{h_k} \otimes A_k)\big)$.

Proof: Using the expansion (6.1) we obtain the following sequence of equalities for the function $f(|U \cap V|)$:

$$\begin{aligned} f(|U\cap V|) &= \sum_{k=0}^{l}\sum_{s,t=n-l}^{n-k} A_{k,s,t}\frac{1}{|\mathcal{H}|}\sum_{i=1}^{h_k} e_{k,s,i}(\boldsymbol{u})e_{k,t,i}(\boldsymbol{v}) \\ &= \frac{1}{|\mathcal{H}|}\sum_{k=0}^{l}\sum_{i=1}^{h_k} \operatorname{Tr}\left(A_k\mathcal{E}_{k,i}^T(\boldsymbol{v})\mathcal{E}_{k,i}(\boldsymbol{u})\right) \\ &= \frac{1}{|\mathcal{H}|}\sum_{k=0}^{l}\sum_{i=1}^{h_k} \mathcal{E}_{k,i}(\boldsymbol{u})A_k\mathcal{E}_{k,i}^T(\boldsymbol{v}) \\ &= \frac{1}{|\mathcal{H}|}\sum_{k=0}^{l}[\mathcal{E}_{k,1}(\boldsymbol{u})\cdots\mathcal{E}_{k,h_k}(\boldsymbol{u})]\begin{bmatrix} A_k & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & A_k \end{bmatrix}\begin{bmatrix} \mathcal{E}_{k,1}^T(\boldsymbol{v}) \\ \vdots \\ \mathcal{E}_{k,h_k}^T(\boldsymbol{v}) \end{bmatrix} \\ &= \frac{1}{|\mathcal{H}|}h(\boldsymbol{u})\Big(\bigoplus_{k=0}^{l}(I_{h_k}\otimes A_k)\Big)h(\boldsymbol{v})^T. \end{aligned}$$

This relation describes the matrix on the right-hand side of (6.15). Finally note that $f(|U\cap V|)=0$ for all $U,V\in\mathcal{F}'$. This completes the proof. ■

The next step is to prove that the rows of $H(\mathcal{F}')$ are linearly independent. We use the notation $\binom{[n]}{k}$ to denote the family of all subsets of $[n]$ of size k.

Lemma 6.6 *Let*

$$\mathcal{N}=\binom{[n]}{n-l}\cup\cdots\cup\binom{[n]}{n}$$

and let $H(\mathcal{N})=(h(\boldsymbol{u}))_{\boldsymbol{u}\in\mathcal{N}}$ be an $|\mathcal{N}|\times|\mathcal{N}|$ matrix constructed as in (6.14). Then the matrix $\widehat{H}=\frac{1}{\sqrt{|\mathcal{H}|}}H(\mathcal{N})$ is orthogonal.

Proof: We will show that $\widehat{H}\widehat{H}^T=I$. We have

$$(\widehat{H}\widehat{H}^T)_{\boldsymbol{u},\boldsymbol{v}}=\sum_{k=0}^{l}\sum_{s=n-l}^{n-k}E_{k,s,s}(\boldsymbol{u},\boldsymbol{v}),$$

or in the matrix form,

$$\widehat{H}\widehat{H}^T=\sum_{k=0}^{l}\sum_{s=n-l}^{n-k}\widehat{E}_{k,s,s},$$

where $\widehat{E}_{k,s,s}=\big(E_{k,s,s}(\boldsymbol{u},\boldsymbol{v})\big)_{\boldsymbol{u},\boldsymbol{v}\in\mathcal{N}}$. Form a $2^n\times 2^n$ matrix $E_{k,s,s}=(E_{k,s,s}(\boldsymbol{u},\boldsymbol{v}))_{\boldsymbol{u},\boldsymbol{v}\in\mathcal{H}}$. Since $s\ge n-l$, we have

$$E_{k,s,s}=\begin{bmatrix}\widehat{E}_{k,s,s} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}, \tag{6.16}$$

and hence $\mathrm{Tr}(\widehat{E}_{k,s,s}) = \mathrm{Tr}(E_{k,s,s}) = h_k$. Further,

$$\mathrm{Tr}(\widehat{H}\widehat{H}^T) = \sum_{k=0}^{l}\sum_{s=n-l}^{n-k} h_k = \sum_{k=0}^{l}\binom{n}{k} = |\mathcal{N}|.$$

By (6.3), $E_{k,s,s}E_{j,t,t} = E_{k,s,s}\delta_{k,j}\delta_{s,t}$ and hence

$$\widehat{E}_{k,s,s}\widehat{E}_{j,t,t} = \widehat{E}_{k,s,s}\delta_{k,j}\delta_{s,t}.$$

Therefore,

$$\begin{aligned}(\widehat{H}\widehat{H}^T)^2 &= \Big(\sum_{k=0}^{l}\sum_{s=n-l}^{n-k}\widehat{E}_{k,s,s}\Big)\Big(\sum_{j=0}^{l}\sum_{t=n-l}^{n-j}\widehat{E}_{j,t,t}\Big)\\ &= \sum_{k=0}^{l}\sum_{s=n-l}^{n-k}\widehat{E}_{k,s,s} = \widehat{H}\widehat{H}^T.\end{aligned}$$

Thus the matrix $\widehat{H}\widehat{H}^T$ is idempotent and therefore, its eigenvalues are 0 or 1. Since its trace equals $|\mathcal{N}|$, all its eigenvalues are equal to 1, and thus $\mathrm{rank}\,(\widehat{H}\widehat{H}^T) = |\mathcal{N}|$. We conclude that $\widehat{H}\widehat{H}^T = I$, so $\widehat{H}$ is orthogonal. ■

Note that the A_k's are of sizes $(l-k+1)\times(l-k+1)$, and hence A_l is actually a scalar quantity. Next, we prove that $A_l \neq 0$.

Lemma 6.7 $A_l \neq 0$.

Proof: Note that $A_l = A_{l,n-l,n-l}$. Using (6.9) and (6.13), we can deduce that A_l is proportional to $f_{l,n-l,n-l}\ c_{l,l,n-l,n-l,n}$.

Now, for any s, t, $f_{l,s,t}$ is the coefficient of the Krawtchouk polynomial of degree l in (6.7) and is hence non-zero. For $s = t = n-l,\ j = k = l$, we have

$$c_{l,l,n-l,n-l,n} = 2^{2l}\frac{\binom{n-l}{l}}{\binom{n}{l}\binom{n-l+1}{l}}b_l(n-l,n-l)\big(K_0(n-2l;n-2l)\big)^2,$$

which is strictly positive. Thus, $A_l \neq 0$. ■

Denote by $i(D) = (i_+(D), i_-(D), i_0(D))$ the inertia (the number of positive, negative, and zero eigenvalues, respectively) of a symmetric matrix D. The matrix F has inertia $(0,0,M')$.

By Ostrowski's theorem (Theorem A.5), we have

$$\begin{aligned} i_+(F) &\le i_+(\mathbf{A}),\\ i_-(F) &\le i_-(\mathbf{A}).\end{aligned} \tag{6.17}$$

To account for the zero eigenvalues we use the following perturbation argument. If $A_l > 0$, consider the matrix $F_\epsilon = F - \epsilon I_{M'},\ \epsilon > 0$ and let $\widehat{H}(\mathcal{F}') = |\mathcal{H}|^{-1/2}H(\mathcal{F}')$.

Lemma 6.6 implies that the rows of the matrix $\widehat{H}(\mathcal{F}')$ are pairwise orthogonal. We have

$$\begin{aligned}F_\epsilon &= \widehat{H}(\mathcal{F}')\mathbf{A}\widehat{H}(\mathcal{F}')^T - \epsilon I_{M'}\\ &= \widehat{H}(\mathcal{F}')\big(\mathbf{A} - \epsilon I_{|\mathcal{N}|}\big)\widehat{H}(\mathcal{F}')^T.\end{aligned}$$

For small ϵ we have $i_-(F_\epsilon) = M'$. By Lemma 6.7, A_l is an eigenvalue of $\mathbf{A}$ of multiplicity at least h_l, so for $A_l > 0$ and $0 < \epsilon < A_l$,

$$i_-(\mathbf{A} - \epsilon I_{|\mathcal{N}|}) \le \sum_{k=0}^{l-1}(l-k+1)h_k.$$

By (6.17)

$$M' \le \sum_{k=0}^{l-1}(l-k+1)h_i,$$

or

$$M \le M' + 1 \le 2\binom{n}{l-1} + \binom{n}{l-2} + \cdots + \binom{n}{0}.$$

For $A_l < 0$, we consider $F_\epsilon = F + \epsilon I_{M'}$ for some $0 < \epsilon < |A_l|$, and use the upper bound $i_+(F_\epsilon) \le i_+(\mathbf{A} + \epsilon I_{|\mathcal{N}|})$. This concludes the proof of Theorem 6.3.

6.3 Concluding remarks

The method employed in the proof extends the use of the annihilator polynomial initiated by Delsarte in [26]. Earlier results that relied on this method are related to estimates of the size of codes with a small number of distances $(d_1, \dots, d_l)$, between the codewords in the Hamming space. Delsarte's approach and its extensions [13] use the expansion of $f(x) = \prod_{i=1}^{l}(d_i - x)$ into the basis of Krawtchouk polynomials. This ties the distances between pairs of vectors and matrices formed of evaluations of spherical harmonics (2.14). Linear-algebraic considerations are then used to establish bounds on the size of sets with few distances.

CHAPTER 7

Constant weight codes

Upper bounds on constant weight codes (in the Johnson space) is a very well studied area. One of the oldest upper bounds present is the Johnson bound (see Theorem 3.3). Numerous other upper bounds have been provided over the course of the last several decades, see Agrell et. al. [4] for an extensive survey and new results. In this chapter we derive some new bounds on constant weight codes by extending the averaging argument of the Johnson bound (see Section 3.1). In Sections 7.2 and 7.3 we consider constant weight codes in the binary Hamming space $\mathcal{H}(2,n)$. The new bounds derived here are interesting in light of the fact that they hold in some of the cases where the Johnson bound is vacuous. One of the ideas in our calculation is performing a "second-order" averaging, i.e., studying the average frequency of letters over pairs of columns of the codematrix. Earlier results are based on averaging over entries in individual columns of the matrix. In Section 7.5 we generalize the above bounds to the nonbinary Hamming space $\mathcal{H}(q,n)$.

The above bounds were the result of a study on bounds under list decoding. Bounds on the size of codes under list decoding were derived by Blinovskii [20]. The bound in [20] is also based on an averaging argument on the frequency of letters over individual columns of the codematrix. Our objective was to provide improvements to the asymptotic bounds on the size of codes under list decoding. Although improvements were not obtained, the new ideas that arose in this study, led to the bounds in Sections 7.2–7.5. To illustrate this link, we also present an application of our methods to list decoding bounds, arriving at the same estimate as [20].

The results in this chapter are an outcome of a study which provides new methods and bounds on constant weight codes. They are not intended for further publication.

7.1 Introduction

Let $\mathcal{C} \subset \mathcal{H}(2,n)$ be an (n,M,d) code in which each codeword has weight w. Let $A(n,d,w) = M$ denote the maximum possible size of $\mathcal{C}$. The Johnson bound (Theorem 3.3) states that

$$A(n,d,w) \le \left\lfloor \frac{dn}{dn - 2wn + 2w^2} \right\rfloor$$

as long as w does not exceed the *Johnson radius*

$$J(n,d) = \Big\lfloor \frac{n}{2}\Big(1 - \sqrt{1 - \frac{2d}{n}}\Big)\Big\rfloor. \tag{7.1}$$

An improvement of the Johnson bound was presented in [4, Corollary 5].

Theorem 7.1 (Agrell et. al. [4])

$$A(n,d,w) \le \begin{cases} \big\lfloor \frac{dn}{dn-2wn+2w^2} \big\rfloor, & dn - 2wn + 2w^2 \ge \frac{d}{2}, \\ n, & 0 < dn - 2wn + 2w^2 \le \frac{d}{2}, \\ 2n-2, & dn - 2wn + 2w^2 = 0. \end{cases} \tag{7.2}$$

The following theorem gives a new upper bound on $A(n,d,w)$.

Theorem 7.2 *Let* $D = (d-2w)(1-2\frac{w}{n}) + \big(\frac{w}{n}\big)^2\big(4\frac{(n-w)^2}{n-1} - 2(n-d)\big)$. *Then for* $D > 0$*,*

$$A(n,d,w) \le \left\lfloor \frac{d}{D} \right\rfloor. \tag{7.3}$$

The bound in (7.3) is better than the Johnson bound (3.1), (7.2) for values of w close to the Johnson radius and sometimes holds for values of w beyond the Johnson radius. For $n = 2w = 2d$, the bound in (7.3) reduces to $2n-2$, and so it is equal to the bound in (7.2).

Notation: We use the following abbreviations (and their combinations) for readability: $\sum_l$ for $\sum_{l=1}^n$, $\sum_{l,k>l}$ for $\sum_{l=1}^n \sum_{k=l+1}^n$, $\sum_i$ for $\sum_{i=0}^{q-1}$, $\sum_{j<j'}$ for $\sum_{0\le j<j'\le q-1}$, and $\sum_{i\ne i'}$ for $\sum_{0\le i\ne i'\le q-1}$.

7.2 Bound from weighted averages

We now proceed to prove Theorem 7.2. The main idea in the following proof is to use weighted average instead of an average using uniform weights in deriving the bound on constant weight codes (also see Section 3.1).

Let $\boldsymbol{u}, \boldsymbol{v}$ be binary vectors where we write $a_0 = 1$ for 0 and $a_1 = -1$ for 1. We have $d_H(\boldsymbol{u}, \boldsymbol{v}) = n - {}^1\!/_2 \sum_{l=1}^n |u_l + v_l|$, so the minimum distance of the code equals

$$d = n - {}^1\!/_2 \max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C} \\ \boldsymbol{u}\ne\boldsymbol{v}}} \sum_{l=1}^n |u_l + v_l|. \tag{7.4}$$

For any vector $\boldsymbol{x} \in \mathbb{R}^N$ and any probability distribution $g(i)$ on N points we can write $\max_i x_i \ge \sum_i g(i) x_i$, so

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C} \\ \boldsymbol{u}\ne\boldsymbol{v}}} \sum_{l=1}^n |u_l + v_l| \ge \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C} \\ \boldsymbol{u}\ne\boldsymbol{v}}} g(\boldsymbol{u},\boldsymbol{v}) \sum_{l=1}^n |u_l + v_l|. \tag{7.5}$$

Taking a uniform distribution $g(\boldsymbol{u},\boldsymbol{v}) = \binom{M}{2}^{-1}$ we get back the Johnson bound. To improve upon it, we consider the distribution

$$g(\boldsymbol{u},\boldsymbol{v}) = \frac{\sum_{l=1}^{n} 2^{s|u_l+v_l|}}{\sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} 2^{s|u_l+v_l|}}, \quad \text{for } s \geq 1.$$

The weights are chosen such that the largest weights are assigned to the pair of codewords with the largest value of $\sum_l |u_l + v_l|$. In the rest of this section we will focus on computing a lower bound on this maximum, which will lead to the upper bound of Theorem 7.2.

Exchanging the order of summation on the right-hand side (7.5) we can write

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \frac{\sum_{l=1}^{n} f_l(s) + \sum_{\substack{l,k=1\\ l\neq k}}^{n} f_{l,k}(s)}{\sum_{l=1}^{n} g_l(s)}, \tag{7.6}$$

where

$$f_l(s) = \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} 2^{s|u_l+v_l|}|u_l+v_l|, \quad f_{l,k}(s) = \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} 2^{s|u_l+v_l|}|u_k+v_k|,$$

$$g_l(s) = \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} 2^{s|u_l+v_l|}.$$

Denote by ν_1^l the number of a_1's in the l-th column of the codematrix of $\mathcal{C}$, and denote by $\lambda_{1,0}^{l,k}, \lambda_{0,1}^{l,k}$, and $\lambda_{1,1}^{l,k}$ the number of pairs $(a_1, a_0), (a_0, a_1)$, and (a_1, a_1), respectively, in the (l, k) column pair of the codematrix. Then

$$f_l(s) = \left[\binom{\nu_1^l}{2} + \binom{M-\nu_1^l}{2}\right] 2^{2s+1},$$

$$g_l(s) = \sum_{i=0}^{2} \binom{\nu_1^l}{i}\binom{M-\nu_1^l}{2-i} 2^{s|2-2i|},$$

$$f_{l,k}(s) = \sum_{\substack{i_1+i_2+i_3\leq 2\\ i_1,i_2,i_3\geq 0}} \binom{M-\lambda_{1,0}^{l,k}-\lambda_{0,1}^{l,k}-\lambda_{1,1}^{l,k}}{2-i_1-i_2-i_3}\binom{\lambda_{1,0}^{l,k}}{i_1}\binom{\lambda_{0,1}^{l,k}}{i_2}$$
$$\times \binom{\lambda_{1,1}^{l,k}}{i_3} |2-2(i_1+i_3)| 2^{s|2-2(i_2+i_3)|}.$$

For the next step, note that $\lambda_{1,0}^{l,k} = \lambda_{0,1}^{k,l}$ and $\lambda_{1,1}^{l,k} = \lambda_{1,1}^{k,l}$. Letting $s \to \infty$ (so that only the terms involving 2^{2s} in (7.6) stay) we obtain the following inequality:

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l+v_l| \geq \frac{1}{{}^1\!/_2\left(\sum_l (\nu_1^l)^2 + (M-\nu_1^l)^2 - M\right)}\Big[\Big(\sum_l (\nu_1^l)^2 + (M-\nu_1^l)^2 - M\Big)$$
$$+ 2\Big(\sum_{l,k>l} (M-\lambda_{1,0}^{l,k}-\lambda_{0,1}^{l,k}-\lambda_{1,1}^{l,k})^2 + (\lambda_{1,0}^{l,k})^2 + (\lambda_{0,1}^{l,k})^2 + (\lambda_{1,1}^{l,k})^2 - M\Big)\Big]. \tag{7.7}$$

We now minimize the above expression over all possible values of $\nu_1^l, \lambda_{1,0}^{l,k}, \lambda_{0,1}^{l,k}, \lambda_{1,1}^{l,k}$ satisfying the following set of constraints, which hold for all $l = 1, \ldots, n$ and $k = 1, \ldots, n$, $k \neq l$.

$$0 \le \nu_1^l, \lambda_{1,0}^{l,k}, \lambda_{0,1}^{l,k}, \lambda_{1,1}^{l,k} \le M, \qquad \lambda_{1,0}^{l,k} + \lambda_{0,1}^{l,k} + \lambda_{1,1}^{l,k} \le M, \tag{7.8}$$

$$\lambda_{1,0}^{l,k} + \lambda_{1,1}^{l,k} = \nu_1^l, \qquad \lambda_{0,1}^{l,k} + \lambda_{1,1}^{l,k} = \nu_1^k, \tag{7.9}$$

$$\sum_{l=1}^{n} \nu_1^l = Mw, \tag{7.10}$$

$$\sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \lambda_{1,1}^{l,k} = M\binom{w}{2}. \tag{7.11}$$

The constraints given by (7.8)–(7.9) are clear from the definition of the variables $\nu_1^l, \lambda_{1,0}^{l,k}, \lambda_{0,1}^{l,k}$ and $\lambda_{1,1}^{l,k}$. (7.10) counts the total number of 1's in the codematrix of $\mathcal{C}$. Equation (7.11) counts the total number of $(1,1)$ pairs in the codematrix of $\mathcal{C}$.

As a first step of the minimization we relax the variables $\nu_1^l, \lambda_{1,0}^{l,k}, \lambda_{0,1}^{l,k}, \lambda_{1,1}^{l,k}$ from integers to reals while still satisfying the constraints defined in (7.8)–(7.11). Under this relaxation, the minimization of (7.7) is a case of *fractional programming.* In order to proceed with the minimization we first note some characteristics of the functions in the numerator and in the denominator. The numerator in (7.7) can be expanded to

$$n(M^2 - M) - 2M\Big(\sum_{l=1}^{n} \nu_1^l\Big) + 2\Big[\sum_{l=1}^{n} (\nu_1^l)^2\Big] + n(n-1)(M^2 - M) +$$
$$\sum_{l=1}^{n} \sum_{k=l+1}^{n} -2M(\lambda_{1,0}^{l,k} + \lambda_{0,1}^{l,k} + \lambda_{1,1}^{l,k}) + (\lambda_{1,0}^{l,k} + \lambda_{1,1}^{l,k})^2 + (\lambda_{0,1}^{l,k} + \lambda_{1,1}^{l,k})^2 + (\lambda_{1,0}^{l,k} + \lambda_{0,1}^{l,k})^2.$$

Let $\zeta_{l,k} = \lambda_{1,0}^{l,k} + \lambda_{0,1}^{l,k}$, $l = 1, \ldots, n-1,\ k = l+1, \ldots, n$, and let

$$\mathbf{X} \triangleq \big((\nu_1^l)_l, (\zeta_{l,k})_{l,k}\big) \tag{7.12}$$

denote the vector consisting of all the variables ν_1^l, $l = 1, \ldots, n$ and $\zeta_{l,k}$, $l = 1, \ldots, n-1,\ k = l+1, \ldots, n$. We use the constraints in (7.9)–(7.11) to simplify the linear terms in the above expanded version of the numerator. Denote the resulting expressions for the numerator and denominator of (7.7) by $f(\mathbf{X})$ and $g(\mathbf{X})$, respectively. Then

$$f(\mathbf{X}) = M^2((n-2w)^2 - 2w^2) - n^2M + 2n\sum_{l=1}^{n} (\nu_1^l)^2 + 2\sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2, \tag{7.13}$$

$$g(\mathbf{X}) = {}^1\!/\!{}_2\Big(M^2(n-2w) - nM + 2\sum_{l=1}^{n} (\nu_1^l)^2\Big). \tag{7.14}$$

Using the variables $\zeta_{l,k}$ the constraints in (7.8)–(7.11) take the form

$$\begin{aligned} 0 \le \nu_1^l, \zeta_{l,k} \le M, &\qquad \zeta_{l,k} \le \nu_1^l + \nu_1^k, \quad \forall\, l, k \\ \sum_{l=1}^{n} \nu_1^l = Mw, &\qquad \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k} = Mw(n-w). \end{aligned} \tag{7.15}$$

Proposition 7.3 *The expression in (7.7) attains its minimum at the point* $\mathbf{X}_0$ *given by* $\nu_1^l = \nu,\ l = 1, \ldots, n$ *and* $\zeta_{l,k} = \zeta,\ l = 1, \ldots, n-1,\ k = l+1, \ldots, n,$ *where*

$$\nu = M\frac{w}{n}, \quad \zeta = 2M\frac{w(n-w)}{n(n-1)}.$$

A proof of this proposition is given in Section 7.8.1 below.

Substituting the value of the minimum for the RHS in (7.4) we obtain the following lower bound on $n - d$,

$$n - d \geq \frac{Mn\big((1-2\frac{w}{n})^2 + 4\big(\frac{w}{n}\big)^2\frac{(n-w)^2}{n(n-1)}\big) - n}{M\big(1 - 2\frac{w}{n} + 2\left(\frac{w}{n}\right)^2\big) - 1}.$$

Solving this inequality for M we obtain the bound of Theorem 7.2.

7.3 Bound obtained by using L^2 norm

In this section we derive another bound on $A(n, d, w)$ which in certain cases provides an improvement of Theorem 7.2. To do so, we replace inequality (7.5) with another inequality and constrain the variables $\nu_1^l, \zeta_{l,k}$ of the previous section to integers.

Theorem 7.4 *Let* $\mathcal{C}(n, M, d)$ *be a constant weight code in* $\mathcal{S}_w \subset \mathcal{H}(2, n)$. *Let*

$$E = (n-2w)^2 + 4\frac{w^2(n-w)^2}{n(n-1)} - (n-d)^2 + \frac{1}{M^2}\big(2n^2\{\nu\}(1-\{\nu\}) + \\ n(n-1)\{\zeta\}(1-\{\zeta\})\big),$$

where ν *and* ζ *are as above, and* $\{a\} = a - \lfloor a \rfloor$ *is the fractional part of* a. *Then for* $E > 0$,

$$M \leq \left\lfloor \frac{d(2n-d)}{E} \right\rfloor. \tag{7.16}$$

In particular if a value of M violates the relation in (7.16) then we can decrease the value of M by 1.

Proof: We use the following inequality in lieu of (7.5):

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^n |u_l + v_l| \geq \Big(\frac{1}{\binom{M}{2}} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \Big(\sum_{l=1}^n |u_l + v_l|\Big)^2\Big)^{1/2} \tag{7.17}$$

$$= \Big(\frac{2}{\binom{M}{2}} f(\mathbf{X})\Big)^{1/2},$$

where $f(\mathbf{X})$ is defined in (7.13). The proof is a modification of the proof of Proposition 7.3. We wish to find the minimizing point on the right-hand side of (7.13) in the region given by (7.15) under the additional constraint that ν_1^l and $\zeta_{l,k} = \lambda_{1,0}^{l,k} + \lambda_{0,1}^{l,k}$ are integers. From (7.13) we obtain

$$\min_{\substack{\nu_1^l, \zeta_{l,k} \in \mathbb{N} \\ \text{subject to (7.15)}}} f(\mathbf{X}) \geq \text{constant} + \min_{\substack{0 \leq \nu_1^l \leq M,\ \nu_1^l \in \mathbb{N} \\ \sum_l \nu_1^l = Mw}} 2n \sum_l (\nu_1^l - \nu)^2 + \\ \min_{\substack{0 \leq \zeta_{l,k} \leq M,\ \zeta_{l,k} \in \mathbb{N} \\ \sum_{l,k>l} \zeta_{l,k} = Mw(n-w)}} 2 \sum_{l,k>l} (\zeta_{l,k} - \zeta)^2.$$

Note that we get a lower bound since we have dropped the constraints $\zeta_{l,k} \leq \nu_1^l + \nu_1^k$. This form helps us in determining the solution to the minimization problem on the RHS. Let $\epsilon_\nu = n\{\nu\}$ and $\epsilon_\zeta = \binom{n}{2}\{\zeta\}$. The minimizing assignment of the variables $\nu_1^l,\ 1 \leq l \leq n$ is given by the integer vector $(\nu_1^1, \ldots, \nu_1^n)$ closest in Euclidean distance to the n-length vector $(\nu, \ldots, \nu)$. Up to a permutation of the indices we obtain

$$\nu_1^l = \begin{cases} \lfloor \nu \rfloor + 1, & 1 \leq l \leq \epsilon_\nu, \\ \lfloor \nu \rfloor, & \epsilon_\nu < l \leq n. \end{cases}$$

Similarly, the minimizing vector with coordinates $\zeta_{l,k}$ will contain ϵ_ζ coordinates equal to $\lfloor \zeta \rfloor + 1$ and the remaining coordinates equal to $\lfloor \zeta \rfloor$. Thus we obtain

$$\begin{aligned} \sum_l (\nu_1^l)^2 &= (\lfloor \nu \rfloor + 1)^2 \epsilon_\nu + \lfloor \nu \rfloor^2 (n - \epsilon_\nu) \\ &= n\nu^2 + n\{\nu\}(1 - \{\nu\}), \\ \sum_{l,k>l} \zeta^2 &= \binom{n}{2}\zeta^2 + \binom{n}{2}\{\zeta\}(1 - \{\zeta\}), \end{aligned}$$

and

$$f(\mathbf{X}) \geq M^2 \Big[(n - 2w)^2 + 4\frac{w^2(n-w)^2}{n(n-1)} + \frac{1}{M^2}\Big(2n^2\{\nu\}(1 - \{\nu\}) \\ + n(n-1)\{\zeta\}\big(1 - \{\zeta\}\big)\Big)\Big] - n^2 M.$$

We now use $4(n-d)^2 = \left(\max_{\boldsymbol{u} \neq \boldsymbol{v} \in \mathcal{C}} \sum_l |u_l + v_l|\right)^2$ and (7.17), along with the above minimum, to arrive at the upper bound of Theorem 7.4. ∎

7.4 Numerical results

The two new bounds given by Theorems 7.2 and 7.4 meet the table of bounds for constant weight codes [4] at several values of $A(n, d, w)$. These points are given in Fig. 7.1. In each table, the first row contains the value of w and the first column contains the value of n.

$d=4$	3	4
6	4	
7	7	
8		14

$d=6$	4	5	6
10	5	6	
11	6	11	
12	9		22
13	13		

$d=8$	5	6	7	8
12		4		
14		7	8	
15	6	10	15	
16		16		30
19	12			
20	16			
21	21			

$d=10$	6	9	10
19		19	
20			38
21	7		
26	13		

$d=12$	7	9	11	12
14	2			
18		4		
23			23	
24				46
25		25		
28	8			

$d=14$	8	13	14
16	2		
27		27	
28			54

Figure 7.1: Values of (n, w, d) for which Theorems 7.2 and 7.4 give exact values. The first column is the value of n and the first row is the value of w.

7.5 Bounds on constant weight codes in $\mathcal{H}(q, n)$

In this section, we generalize the bound in Theorem 7.2, and a bound similar to Theorem 7.4 to the nonbinary Johnson space. The emphasis of this section is to show that the method of bounding the size of constant weight codes proposed in the previous sections extends to the case of an arbitrary q-ary alphabet. For this reason some of the computations are omitted from the presentation.

The main technical part of the calculation is to perform explicit optimization of the quadratic functions that arise in the proof. As before in (7.7), our bounds are based on performing a "second-order" averaging, i.e., looking at frequencies of pairs of letters in two columns of the codematrix.

To derive the bound in the q-ary case, let us first map the q-ary points to the real space. A mapping which is convenient for our purpose is described in Dunkl [32]. Under this mapping, the letters $0, \dots, q-1$ are represented as vertices $\boldsymbol{a}_0, \dots, \boldsymbol{a}_{q-1}$ of a simplex in $\mathbb{R}^{q-1}$ such that $\|\boldsymbol{a}_i\| = 1$ for all i. Suppose that $i \mapsto \boldsymbol{a}_i,\ i = 0, \dots, q-1$. The points $\boldsymbol{a}_i$ satisfy the properties

$$\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = \begin{cases} 1, & i = j, \\ -\frac{1}{q-1}, & i \neq j, \end{cases}$$

$$\|\boldsymbol{a}_i + \boldsymbol{a}_j\|^2 = \begin{cases} 4, & i = j, \\ 2\frac{q-2}{q-1}, & i \neq j. \end{cases} \tag{7.18}$$

The Hamming distance between two $q-$ary symbols can be written as

$$\begin{aligned} d_H(\boldsymbol{a}_i, \boldsymbol{a}_j) &= \frac{1 - \langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle}{1 + \frac{1}{q-1}} \\ &= \frac{q-1}{q}(2 - {}^1\!/_2 \|\boldsymbol{a}_i + \boldsymbol{a}_j\|^2), \end{aligned}$$

where we have used the relation $\|\boldsymbol{a}_i + \boldsymbol{a}_j\|^2 = 2 + 2\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle$. Extending this relation to the Hamming distance between vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{H}(q, n)$, we obtain

$$d_H(\boldsymbol{u}, \boldsymbol{v}) = \sum_{l=1}^{n} d_H(u_l, v_l) = \frac{q-1}{2q}\Big(4n - \sum_{l=1}^{n} \|\boldsymbol{u}_l + \boldsymbol{v}_l\|^2\Big),$$

where $\boldsymbol{u}_l, \boldsymbol{v}_l \in \{\boldsymbol{a}_0, \ldots, \boldsymbol{a}_{q-1}\}$. Let $\mathcal{C} \subset \mathcal{H}(q, n)$ be a code with minimum distance d. We have

$$d = \min_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} d_H(\boldsymbol{u}, \boldsymbol{v}) = \frac{q-1}{2q}\Big(4n - \max_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} \sum_{l=1}^{n} \|\boldsymbol{u}_l + \boldsymbol{v}_l\|^2\Big). \tag{7.19}$$

A bound similar to the one established in Theorem 7.4 is given in the following theorem:

Theorem 7.5 *Let $\mathcal{C} \subset \mathcal{H}(q, n)$ be an (n, M, d) code in which all its vectors have weight w. Let*

$$\begin{aligned} E = \Big[n^2\Big(\frac{q-2}{q-1}\Big)^2 + \frac{1}{n}\Big(4 - \Big(\frac{q-2}{q-1}\Big)^2 + \frac{2(n-1)q(q-2)}{(q-1)^2}\Big)\Big((n-w)^2 + \frac{w^2}{q-1}\Big) \\ + \frac{q^2}{(q-1)^2 n(n-1)}\Big(\big((n-w)(n-w-1)\big)^2 + \frac{2}{q-1}\big(w(n-w)\big)^2 \\ + \frac{1}{(q-1)^2}\big(w(w-1)\big)^2\Big)\Big] - \Big(2n - \frac{dq}{q-1}\Big)^2. \end{aligned}$$

For $E > 0$, the size M of the code is upper bounded by

$$M \leq \left\lfloor \frac{\frac{dq}{q-1}\big(4n - \frac{dq}{q-1}\big)}{E} \right\rfloor. \tag{7.20}$$

Proof: We first use the following lower bound on the maximization term in (7.19).

$$\begin{aligned} \Big(\max_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} \sum_{l=1}^{n} \|\boldsymbol{u}_l + \boldsymbol{v}_l\|^2\Big)^2 \geq \frac{1}{\binom{M}{2}}\Big(\sum_{l} \sum_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} \|\boldsymbol{u}_l + \boldsymbol{v}_l\|^4 \\ + 2\sum_{l,k>l} \sum_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} \|\boldsymbol{u}_l + \boldsymbol{v}_l\|^2 \|\boldsymbol{u}_k + \boldsymbol{v}_k\|^2\Big). \end{aligned} \tag{7.21}$$

The first term can be estimated as follows. Consider the $M \times n$ codematrix of the code $\mathcal{C}$. For each $l = 1, \ldots, n$ and for each $i = 0, \ldots, q-1$, let ν_i^l denote the number of terms in the l-th column which take the value $\boldsymbol{a}_i$. The parameters of the code imply the following constraints on the values of $\nu_i^l,\ i=0,\dots,q-1,\ l=1,\dots,n$.

$$0 \le \nu_i^l \le M, \quad i=0,\dots,q-1,\ l=1,\dots,n$$

$$\sum_i \nu_i^l = M, \quad l=1,\dots,n \tag{7.22}$$

$$\sum_l \sum_{i>0} \nu_i^l = Mw. \tag{7.23}$$

Eq. (7.22) follows from the fact that the total count of all possible alphabet symbols in the l-th column equals M, and (7.23) follows because the sum of all nonzero alphabets in the codematrix is Mw. We obtain

$$\sum_l \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \|\boldsymbol{u}_l+\boldsymbol{v}_l\|^4 = -8nM + 2n\Big(\frac{q-2}{q-1}\Big)^2 M^2 + \Big(8-2\Big(\frac{q-2}{q-1}\Big)^2\Big)\sum_l\sum_i(\nu_i^l)^2. \tag{7.24}$$

The second term in (7.21) can be estimated as follows. For each $l=1,\dots,n$, $k=l+1,\dots,n$ and for each $i,j\in\{0,\dots,q-1\}$, let $\lambda_{i,j}^{l,k}$ denote the number of $(\boldsymbol{a}_i,\boldsymbol{a}_j)$ pairs in the (l,k)-th column pair of the codematrix of $\mathcal{C}$. Then we have the following constraints on the variables $\lambda_{i,j}^{l,k}$:

$$0 \le \lambda_{i,j}^{l,k} \le M, \quad i,j\in\{0,\dots,q-1\},\ l,k>l,$$

$$\sum_{l,k>l}\ \sum_{i>0,j>0} \lambda_{i,j}^{l,k} = M\binom{w}{2}, \tag{7.25}$$

$$\sum_{l,k>l}\sum_{i>0} \lambda_{0,i}^{l,k} + \lambda_{i,0}^{l,k} = Mw(n-w), \tag{7.26}$$

$$\sum_{i,j} \lambda_{i,j}^{l,k} = M, \quad l,k>l, \tag{7.27}$$

$$\sum_j \lambda_{i,j}^{l,k} = \nu_i^l, \quad i=0,\dots,q-1,\ l,k>l, \tag{7.28}$$

$$\sum_i \lambda_{i,j}^{l,k} = \nu_j^k, \quad j=0,\dots,q-1,\ l,k>l. \tag{7.29}$$

Equation (7.25) follows from the fact that the total number of pairs of nonzero elements in each row of the codematrix is $\binom{w}{2}$, and hence the overall total of pairs is $M\binom{w}{2}$. Equation (7.26) is obtained similarly. Equation (7.27) counts the total number of pairs of letters in any (l,k) column pair.

Using (7.18), and (7.25)–(7.29) we obtain

$$\sum_{l,k>l}\sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \|\boldsymbol{u}_l+\boldsymbol{v}_l\|^2\|\boldsymbol{u}_k+\boldsymbol{v}_k\|^2 = -8\binom{n}{2}M + 2\binom{n}{2}\left(\frac{q-2}{q-1}\right)^2 M^2 +$$

$$2(n-1)\frac{q(q-2)}{(q-1)^2}\sum_{l,i}(\nu_i^l)^2 + \left(8-\frac{2(q-2)(3q-2)}{(q-1)^2}\right)\sum_{l,k>l}\sum_{i,j}(\lambda_{i,j}^{l,k})^2. \tag{7.30}$$

To get a universal lower bound, we determine the minimum in (7.24) and (7.30) under the constraints (7.22)–(7.23) and (7.25)–(7.29). This minimization is performed in Section 7.8.2. The optimum assignments of variables are given in (7.48) and (7.49). Upon substituting these values, using the relation

$$4\Big(2n - \frac{dq}{q-1}\Big)^2 = \Big(\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} \|\boldsymbol{u}_l + \boldsymbol{v}_l\|^2\Big)^2$$

derived from (7.19), and rearranging for M, we obtain the required bound of Theorem 7.5. ■

Next we consider a generalization of the bound of Theorem 7.2 to the non-binary Johnson space.

Theorem 7.6 *Let $\mathcal{C} \subset \mathcal{H}(q,n)$ be an (n, M, d) code with each codeword having weight w. Let*

$$\begin{aligned} D =& \frac{1}{n}\Big(2 - \frac{1}{n} + (n-1)\frac{q-2}{q-1}\Big)\Big((n-w)^2 + \frac{w^2}{q-1}\Big) \\ &+ \frac{q}{n(n-1)(q-1)}\Big(\big((n-w)(n-w-1)\big)^2 + \frac{2(w(n-w))^2}{q-1} + \frac{(w(w-1))^2}{(q-1)^2}\Big). \end{aligned}$$

Then for $D > 0$, the size M of $\mathcal{C}$ can be upper bounded by

$$M \leq \left\lfloor \frac{ndq/(q-1)}{D} \right\rfloor.$$

Proof: We consider an inequality similar to (7.5):

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{i=1}^{n} \|\boldsymbol{u}_i + \boldsymbol{v}_i\|^2 \geq \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} g(\boldsymbol{u},\boldsymbol{v}) \sum_{i=1}^{n} \|\boldsymbol{u}_i + \boldsymbol{v}_i\|^2. \tag{7.31}$$

The weights are chosen as

$$g(\boldsymbol{u},\boldsymbol{v}) = \frac{\sum_{l=1}^{n} 2^{s\|\boldsymbol{u}_l+\boldsymbol{v}_l\|^2}}{\sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} 2^{s\|\boldsymbol{u}_l+\boldsymbol{v}_l\|^2}}, \qquad \text{where } s \geq 1.$$

Following the reasoning in Section 7.2, we can write the numerator in (7.31) as

$$\sum_{l=1}^{n} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} 2^{s\|\boldsymbol{u}_l+\boldsymbol{v}_l\|^2} \|\boldsymbol{u}_l + \boldsymbol{v}_l\|^2 + \sum_{\substack{l,k=1\\ l\neq k}}^{n} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} 2^{s\|\boldsymbol{u}_l+\boldsymbol{v}_l\|^2} \|\boldsymbol{u}_k + \boldsymbol{v}_k\|^2. \tag{7.32}$$

The denominator can be similarly determined. We first note that under the condition $s \to \infty$, and given (7.18) only certain terms containing 2^{4s} in the numerator and denominator will survive. In particular, we obtain the following expression for the numerator under $s \to \infty$.

$$f(\mathbf{X}) \triangleq -2n^2M + \left(2 + (n-1)\frac{q-2}{q-1}\right)\sum_l \sum_i (\nu_i^l)^2 + \frac{2q}{q-1}\sum_{l,k>l}\sum_{i,j}(\lambda_{i,j}^{l,k})^2, \quad (7.33)$$

where

$$\mathbf{X} = \Big((\nu_i^l)_{i,l}, (\lambda_{i,j}^{l,k})_{i,j,\ l,k>l}\Big)$$

is the vector of all the variables $\nu_i^l,\ i = 0,\dots,q-1,\ l = 1,\dots,n$ and $\lambda_{i,j}^{l,k},\ i,j \in \{0,\dots,q-1\},\ l = 1,\dots,n,\ k = l+1,\dots,n$.

The denominator in (7.31) is given by

$$g(\mathbf{X}) \triangleq \frac{1}{2}\Big(-nM + \sum_l \sum_i (\nu_i^l)^2\Big). \quad (7.34)$$

The optimum point of $f(\mathbf{X})$ and $g(\mathbf{X})$ under the constraints in (7.22)–(7.23) and (7.25)–(7.29) is given by (7.48) and (7.49).

From Lemma 7.9 we know that the optimum of $f(\mathbf{X})/g(\mathbf{X})$ also occurs at the same points. Using (7.19) and rearranging for M gives us the required bound of Theorem 7.6. ■

7.6 Bounds for codes under list decoding

In this section we document an attempt to improve the list decoding bound in the Hamming space. The ideas in the previous sections of this chapter resulted from this attempt. The computation presented below reproduces the result of [20] in a different way.

The list decoding problem can be described as follows. The condition that a code $\mathcal{C} \subset \mathcal{H}(2,n)$ corrects t errors under unique decoding corresponds to the fact that the metric ball of radius $\le 2t$ drawn about any point $\boldsymbol{x} \in \mathcal{H}(2,n)$ contains at most one vector from $\mathcal{C}$. A code is said to correct t errors under *decoding into a list of size* $L \ge 2$ if any such ball contains at most L vectors from $\mathcal{C}$. The maximum value of the list-of-L decoding radius for a given code is denoted by $r(L)$. Clearly, $r(1) \le r(2) \le \dots$.

In [20], Blinovskii determined the maximum size of a code in the binary Hamming space $\mathcal{H}(2,n)$ given the value of the decoding radius. His results are asymptotic in nature and are stated in the following theorem.

Theorem 7.7 [20] *Let $\mathcal{C}_i \subset \mathcal{H}(2,n_i), i = 1,2,\dots$ be a sequence of codes of size M_i and with list decoding radius $r_i(L) = n_i\rho_i(L)$. Suppose that $n_i \to \infty$, $\rho(L) = \lim_{i\to\infty}\rho_i(L)$ and let*

$$R = \limsup_{i\to\infty}\frac{1}{n_i}\log_2 M_i.$$

Then the values $\rho(L)$ and R are related by the following parametric equations:

$$R \le 1 - h_2(\omega), \tag{7.35}$$

$$\rho(L) = \sum_{i=1}^{\lceil L/2 \rceil} \frac{1}{i}\binom{2i-2}{i-1}\big(\omega(1-\omega)\big)^i, \quad 0 \le \omega \le 1/2, \tag{7.36}$$

where $h_2(\cdot)$ is the binary entropy function.

The proof proceeds by deriving a bound which is akin to the Johnson bound, but under list decoding. This provides (7.36). Next, the Bassalygo-Elias inequality of Lemma 3.4 is used to obtain (7.35). Although the original proof in the above theorem proceeds in a different manner, we can prove it by mapping the points in the Hamming space to the reals. For the simplest list decoding case of $L = 2$, we show that with an inequality similar to (7.17) we can recover equation (7.36).

In order to proceed with the derivation we first define a quantity related to the list decoding radius $r(L) = n\rho(L)$, but easier to work with. For a code $\mathcal{C}$ let

$$r_{\mathcal{C}}(L) \triangleq \min_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} \min_{\boldsymbol{y}\in\mathcal{H}(2,n)} \frac{1}{3}\big(d_H(\boldsymbol{u},\boldsymbol{y}) + d_H(\boldsymbol{v},\boldsymbol{y}) + d_H(\boldsymbol{z},\boldsymbol{y})\big) \tag{7.37}$$

be the minimum "average radius"[1] of the code $\mathcal{C}$. According to [20], this quantity is related to $\rho(L)$ in the asymptotics :

$$\lim_{i\to\infty} \frac{r_{\mathcal{C}_i}(L)}{n_i} = \rho(L).$$

Below we assume that the vectors are written over the $(1, -1)$ alphabet. Using $d_H(\boldsymbol{u},\boldsymbol{y}) = \frac{1}{2}(n - \langle \boldsymbol{u},\boldsymbol{y}\rangle)$, we have the following set of equalities:

$$\begin{aligned} r_{\mathcal{C}}(L) &= \frac{1}{2} \min_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} \min_{\boldsymbol{y}\in\{\pm1\}^n} \Big(n - \frac{1}{3}\langle \boldsymbol{u}+\boldsymbol{v}+\boldsymbol{z},\boldsymbol{y}\rangle\Big) \\ &= \frac{1}{2}\Big(n - \frac{1}{3} \max_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} \sum_{l=1}^{n} |u_l + v_l + z_l|\Big). \end{aligned}$$

Our objective will be to bound below the value of the maximum in the previous line. For this we use an inequality similar to (7.17):

$$\begin{aligned} \Big(\max_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} \sum_{l=1}^{n} |u_l + v_l + z_l|\Big)^2 &\ge \frac{1}{\binom{M}{3}} \sum_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} \Big(\sum_{l=1}^{n} |u_l + v_l + z_l|\Big)^2 \\ &= \frac{1}{\binom{M}{3}} \sum_{l} \sum_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} |u_l + v_l + z_l|^2 + \\ &\qquad 2\sum_{l,k>l} \sum_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} |u_l + v_l + z_l||u_k + v_k + z_k|. \end{aligned} \tag{7.38}$$

[1] For fixed $\boldsymbol{u},\boldsymbol{v},\boldsymbol{z} \in \mathcal{C}$ the minimizing $\boldsymbol{y} \in \mathcal{H}$ can be considered as the centroid of these three vectors.

Below we neglect the terms which are irrelevant in the asymptotics of $n_i \to \infty$. The first of the two sums in (7.38) (i.e., the sum $\sum_l \sum_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C}}$) contains only $\Theta(n)$ terms whereas the second sum contains $\Theta(n^2)$ terms. Hence we concentrate on estimating only the second sum.

Let $\lambda_{i,j}^{l,k}$ be as defined in Section 7.2. We obtain for each $l, k > l$

$$\sum_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} |u_l + v_l + z_l||u_k + v_k + z_k| = \sum_{\substack{i_1+i_2+i_3\leq 3\\ i_1,i_2,i_3\geq 0}} \binom{M - \lambda_{1,0}^{l,k} - \lambda_{0,1}^{l,k} - \lambda_{1,1}^{l,k}}{3 - i_1 - i_2 - i_3} \binom{\lambda_{1,0}^{l,k}}{i_1} \times$$
$$\binom{\lambda_{0,1}^{l,k}}{i_2} \binom{\lambda_{1,1}^{l,k}}{i_3} |3 - 2(i_1 + i_3)||3 - 2(i_2 + i_3)|.$$

Let

$$\frac{\lambda_{1,0}^{l,k}}{M} = \alpha^{l,k}, \quad \frac{\lambda_{0,1}^{l,k}}{M} = \beta^{l,k}, \quad \frac{\lambda_{1,1}^{l,k}}{M} = \gamma^{l,k}.$$

Using the inequality $\binom{m}{j} \geq \frac{(m-j+1)^j}{j!}$, we obtain

$$\sum_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} |u_l + v_l + z_l||u_k + v_k + z_k| \geq M^3 \sum_{\substack{i_1+i_2+i_3\leq 3\\ i_1,i_2,i_3\geq 0}} \frac{(1 - \alpha^{l,k} - \beta^{l,k} - \gamma^{l,k})^{3-i_1-i_2-i_3}}{(3 - i_1 - i_2 - i_3)!} \times$$
$$\frac{\alpha^{l,k}}{i_1!} \frac{\beta^{l,k}}{i_2!} \frac{\gamma^{l,k}}{i_3!} |3 - 2(i_1 + i_3)||3 - 2(i_2 + i_3)| + o_M(1).$$

Denote the summation on the RHS of the above inequality by $g(\alpha^{l,k}, \beta^{l,k}, \gamma^{l,k})$. To get a universal lower bound on the RHS of (7.38) consider the vector of all the variables,

$$\mathbf{X} \triangleq \left((\alpha^{l,k}, \beta^{l,k}, \gamma^{l,k})_{l,k>l}\right),$$

and define the function

$$f(\mathbf{X}) = \sum_{l,k>l} M^3 g(\alpha^{l,k}, \beta^{l,k}, \gamma^{l,k}). \tag{7.39}$$

Section 7.8.3 is devoted to the minimization of $f(\mathbf{X})$. The minimum is obtained at

$$\gamma = \frac{w(w-1)}{n(n-1)},$$
$$\alpha = \beta = \frac{w(n-w)}{n(n-1)}.$$

For $\lim_{n\to\infty} \frac{w}{n} = \omega$ we obtain,

$$\begin{aligned}\gamma &= \omega^2,\\ \alpha = \beta &= \omega(1-\omega).\end{aligned} \tag{7.40}$$

Substituting the values of α, γ in $f(\mathbf{X})$ we obtain,

$$\sum_{l,k>l} \sum_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} |u_l + v_l + z_l||u_k + v_k + z_k| \geq \frac{3}{2}M^3 n^2 (1 - 2\omega - 2\omega^2)^2 + o_{Mn^2}(1).$$

Substituting this value in (7.38), and then in the expression for $r_{\mathcal{C}}(L)$ we obtain the following inequalities:

$$\begin{aligned} \max_{\substack{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}\in\mathcal{C} \\ \boldsymbol{u}\neq\boldsymbol{v}\neq\boldsymbol{z}}} \sum_{l=1}^{n} |u_l + v_l + z_l| &\geq 3n(1 - 2\omega + 2\omega^2) + o_{Mn^2}(1) \\ \frac{r_{\mathcal{C}}(L)}{n} &\leq \omega(1-\omega) + o_{Mn^2}(1) \\ \Rightarrow \lim_{i\to\infty} \frac{r_{\mathcal{C}_i}(L)}{n_i} &= \rho(L) \leq \omega(1-\omega). \end{aligned}$$

This is the same expression for $\rho(L)$ obtained from (7.36) with list size $L = 2$.

7.7 Discussion

The following set of remarks contains some intuitive comments on the results of this chapter and outlines obstacles in the way of improving the list bounds.

1. The optimum in Prop. 7.3 occurs at a point which is what we would expect if we were to start with an $M \times n$ codematrix and write 0 or 1 for each entry independently with probabilities $1 - w/n$ and w/n, respectively. Thus, a random constant weight code satisfies the optimality condition.

   In [76], Sidelnikov proved an inequality called the *inequality in the mean* for subsets of $\mathbb{R}^n$. It was generalized by Kabatyanskii and Levenshtein to any metric space with a measure [52]. A simplified version of Sidelnikov's inequality is as follows. Consider the sphere $\mathcal{S}_w \subset \mathcal{H}(2,n)$ mapped to S^{n-1} by $U : \{0,1\} \to \{\frac{1}{\sqrt{n}}, -\frac{1}{\sqrt{n}}\}$ and let $U_w = U(\mathcal{S}_w)$ be the image of this mapping.

   Lemma 7.8 [76] *Let $C \subset U_w$ be a code of size M. Then*

   $$\frac{1}{M^2} \sum_{\boldsymbol{x},\boldsymbol{y}\in C} \langle \boldsymbol{x}, \boldsymbol{y} \rangle^t \geq \frac{1}{|U_w|^2} \sum_{\boldsymbol{x},\boldsymbol{y}\in U_w} \langle \boldsymbol{x}, \boldsymbol{y} \rangle^t \qquad (t \in \mathbb{N}).$$

   For a random code $\mathcal{C} \subset \mathcal{S}_w$ discussed above the inequality in the above lemma is satisfied with equality. This suggests that we can replace the minimizations in Sections 7.2 and 7.5 with the average over pairs of points on $\mathcal{S}_w$ (mapped to S^{n-1}). This indeed turns out to be true and we get the same value of the optimum.

Using the inequality in the mean and the following inequality, we can obtain a lower bound on the LHS of (7.5). For any set of points $c_1, \dots, c_N$, and for any $s, t \in \mathbb{N}$ with $t > s$,

$$\max_i \ c_i \geq \Big(\sum_{i=1}^{N} \frac{1}{N} c_i^t\Big)^{\frac{1}{t}} \geq \Big(\sum_{i=1}^{N} \frac{1}{N} c_i^s\Big)^{\frac{1}{s}}.$$

This essentially reduces to Sidelnikov's result in [76]. For increasing t the bound remains valid beyond the Johnson radius but the size of the code increases exponentially. In [76] this exponential growth is carefully controlled by the parameter t to provide asymptotic improvements on the Bassalygo-Elias bound.

The same comment holds for the nonbinary Johnson space. In this case, the exponential growth is used by Sidorenko to provide an improvement on the Bassalygo-Elias bound for the non-binary Hamming space [77].

2. Similar techniques could be also applied to the derivation on list decoding bounds. However, there are several problems with the generalization to t-norms. First, one needs to determine a suitable multilinear form. Secondly, the inequality in the mean is not known for anything but pairs (i.e., triples, etc.) of points in the space. Should it be possible to overcome the above two hurdles, one could potentially use this technique to derive asymptotic improvements on the list decoding bound, as was done in [76, 77] in the case of unique decoding.

7.8 Appendix

7.8.1 Minimization on the RHS of (7.7)

In this section we show that the functions $f(\mathbf{X})$ and $g(\mathbf{X})$ defined in (7.13) and (7.14) have a common minimum point under the constraints (7.15), and and that their ratio is also minimized at the same point.

The part of the expression for $f(\mathbf{X})$ that involves the variables $\nu_1^l, \zeta_{l,k}$ has the form

$$2n \sum_{l} (\nu_1^l)^2 + 2 \sum_{l,k>l} \zeta_{l,k}^2.$$

Both sums in this expression are strongly convex and symmetric functions of their variables. Therefore, under the conditions $0 \leq \nu_1^l \leq M, \sum_l \nu_1^l = Mw$, and $0 \leq \zeta_{l,k} \leq M, \sum_{l,k>l} \zeta_{l,k} = Mw(n-w)$, the minimum of the sum of squares is attained at the point where all the ν_1^l are equal to $\nu = M\frac{w}{n}$ and all the $\zeta_{l,k}$ are equal to $\zeta = 2M\frac{w(n-w)}{n(n-1)}$. It is readily checked that this assignment also satisfies the inequalities $\zeta_{l,k} \leq \nu_1^l + \nu_1^k$. We conclude that the minimum of $f(\mathbf{X})$ under the constraints (7.15) is attained at the point $\mathbf{X}_0$.

The same considerations apply to $g(\mathbf{X})$ and show that its minimum under (7.15) is also attained at $\mathbf{X}_0$.

The difficult part is to show that the ratio of the the functions $f(\mathbf{X})$ and $g(\mathbf{X})$ is minimized at the point $\mathbf{X}_0$. A proof of this relies on the next lemma which gives a general condition for the location of the minimum of the ratio of two quadratic functions when they are strongly convex and have a common minimum.

Lemma 7.9 *Let $\phi : \mathbb{R}^n \to \mathbb{R}$ and $\psi : \mathbb{R}^n \to \mathbb{R}$ be two quadratic functions defined by $\phi(\boldsymbol{x}) = (\boldsymbol{x} - \widetilde{\mathbf{x}})^T F(\boldsymbol{x} - \widetilde{\mathbf{x}}) + a$ and $\psi(\boldsymbol{x}) = (\boldsymbol{x} - \widetilde{\mathbf{x}})^T G(\boldsymbol{x} - \widetilde{\mathbf{x}}) + b$, such that $F \succ G \succ 0$ and $a > b > 0$, where $\succ$ indicates positive definiteness.*

If $S \triangleq \{\boldsymbol{x} : \|\boldsymbol{x}\| \leq B\}$ and $\widetilde{\mathbf{x}} \in S$, then the minimum of $\phi(\boldsymbol{x})/\psi(\boldsymbol{x})$ is attained either at $\boldsymbol{x} = \widetilde{\mathbf{x}}$, or at a boundary point $\boldsymbol{x} \in \partial S$, i.e. $\|\boldsymbol{x}\| = B$.

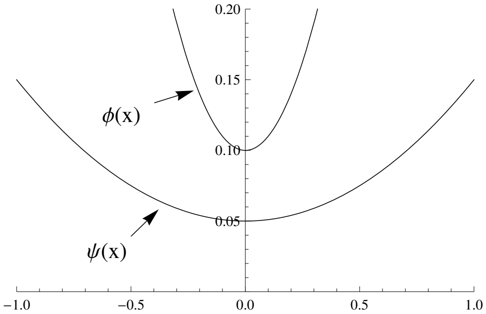


Figure 7.2: An example illustrating Lemma 7.9. Here $\phi(x) = x^2 + 0.1$, $\psi(x) = 0.1x^2 + 0.05$, $S = [-1, 1]$. All the functions $\phi(x), \psi(x), \phi(x) - \psi(x)$, and $\phi(x)/\psi(x)$ attain their minimum at $x = 0$.

Proof: The functions $\phi(\boldsymbol{x})$ and $\psi(\boldsymbol{x})$ are strongly convex, positive in S, have a common minimum at $\widetilde{\mathbf{x}}$. Their difference $\phi(\boldsymbol{x}) - \psi(\boldsymbol{x})$ is also strongly convex, positive in S and attains its minimum at $\widetilde{\mathbf{x}}$. Without loss of generality, we may also assume that F and G are symmetric. Let $\hat{\mathbf{x}} = \arg\min_{\boldsymbol{x} \in S} \phi(\boldsymbol{x})/\psi(\boldsymbol{x})$ and let $\hat{\alpha} = \phi(\hat{\mathbf{x}})/\psi(\hat{\mathbf{x}})$ be the value of the minimum. Then

$$\phi(\boldsymbol{x}) - \hat{\alpha}\psi(\boldsymbol{x}) \geq 0, \ \forall \boldsymbol{x} \in S.$$

Clearly $\hat{\alpha} > 1$ and for any $1 < \alpha < \hat{\alpha}$, $\phi(\boldsymbol{x}) - \alpha\psi(\boldsymbol{x}) > 0$ for any $\boldsymbol{x} \in S$. We consider the expression $\phi(\boldsymbol{x}) - \alpha\psi(\boldsymbol{x})$ and let α grow from 1. The following two cases arise:

Case 1: $F - \hat{\alpha}G \succ 0$ and $a - \hat{\alpha}b = 0$. It is readily seen that $\phi(\boldsymbol{x}) - \hat{\alpha}\psi(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in S$ with equality at $\boldsymbol{x} = \widetilde{\mathbf{x}}$. Hence $\hat{\mathbf{x}} = \widetilde{\mathbf{x}}$.

Case 2: Let α' be such that $F - \alpha' G \succeq 0$ and $a - \alpha' b > 0$. Suppose that $F - \alpha G \succ 0$ for all $1 < \alpha < \alpha'$. Then $\phi(\boldsymbol{x}) - \alpha'\psi(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in S$. Consider $\hat{\alpha} > \alpha'$ such that $\phi(\boldsymbol{x}) - \hat{\alpha}\psi(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in S$. Let $\hat{\mathbf{x}} \in S$ be such that

$$\phi(\hat{\mathbf{x}}) - \hat{\alpha}\psi(\hat{\mathbf{x}}) = (\hat{\mathbf{x}} - \widetilde{\mathbf{x}})^T (F - \hat{\alpha}G)(\hat{\mathbf{x}} - \widetilde{\mathbf{x}}) + a - \hat{\alpha}b = 0. \tag{7.41}$$

We will show that $\hat{\mathbf{x}} \in \partial S$, i.e. $\|\hat{\mathbf{x}}\| = B$. Assume the contrary. Then there exists a vector $\bar{\varepsilon} = \epsilon(\hat{\mathbf{x}} - \widetilde{\mathbf{x}})$ for some $\epsilon > 0$ such that

$$\phi(\hat{\mathbf{x}} + \bar{\varepsilon}) - \hat{\alpha}\psi(\hat{\mathbf{x}} + \bar{\varepsilon}) \geq 0.$$

Then

$$(\hat{\mathbf{x}}+\bar{\varepsilon}-\widetilde{\mathbf{x}})^T(F-\hat{\alpha}G)(\hat{\mathbf{x}}+\bar{\varepsilon}-\widetilde{\mathbf{x}})+a-\hat{\alpha}b \geq 0,$$

$$(\hat{\mathbf{x}}+\bar{\varepsilon}-\widetilde{\mathbf{x}})^T(F-\hat{\alpha}G)\bar{\varepsilon}+\bar{\varepsilon}^T(F-\hat{\alpha}G)(\hat{\mathbf{x}}+\bar{\varepsilon}-\widetilde{\mathbf{x}}) \geq 0,$$

and finally

$$(1+\epsilon)2\epsilon(\hat{\mathbf{x}}-\widetilde{\mathbf{x}})^T(F-\hat{\alpha}G)(\hat{\mathbf{x}}-\widetilde{\mathbf{x}}) \geq 0. \tag{7.42}$$

To show a contradiction we first claim that $a-\hat{\alpha}b > 0$. Assume the contrary, i.e., that $a-\hat{\alpha}b \leq 0$. If $a-\hat{\alpha}b = 0$, let $\boldsymbol{x}' \in S$ be a point at which $(\boldsymbol{x}'-\widetilde{\mathbf{x}})^T(F-\alpha' G)(\boldsymbol{x}'-\widetilde{\mathbf{x}}) = 0$. Then for $\hat{\alpha} > \alpha'$, we would have $(\boldsymbol{x}'-\widetilde{\mathbf{x}})^T(F-\hat{\alpha}G)(\boldsymbol{x}'-\widetilde{\mathbf{x}}) < 0$, i.e. $\phi(\boldsymbol{x}')-\hat{\alpha}\psi(\boldsymbol{x}') < 0$, which is a contradiction. If $a-\hat{\alpha}b < 0$, then there exists $\bar{\alpha}$ such that $\alpha' < \bar{\alpha} < \hat{\alpha}$ and $a-\bar{\alpha}b = 0$, so we can repeat the above argument with α' replaced by $\bar{\alpha}$ and again arrive at a contradiction. Thus, $a-\hat{\alpha}b > 0$. From (7.41) we obtain

$$(\hat{\mathbf{x}}-\widetilde{\mathbf{x}})^T(F-\hat{\alpha}G)(\hat{\mathbf{x}}-\widetilde{\mathbf{x}}) < 0,$$

which is a contradiction to (7.42). Thus, $\hat{\mathbf{x}} \in \partial S$. ■

Corollary 7.10 *Let ϕ and ψ be as in Lemma 7.9 with the modification that $G \succeq 0$. Assume in addition that $\phi(\boldsymbol{x})-\psi(\boldsymbol{x})$ is strongly convex and positive with a minimum at $\widetilde{\mathbf{x}}$. Let S be a convex set in $\mathbb{R}^n$, with $\widetilde{\mathbf{x}} \in S$. Let $\hat{\alpha}$ be a positive real number such that $\phi(\boldsymbol{x})-\hat{\alpha}\psi(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in S$ and $\phi(\widetilde{\mathbf{x}})-\hat{\alpha}\psi(\widetilde{\mathbf{x}}) = 0$. Then*

$$\arg\min_{\boldsymbol{x}\in S} \frac{\phi(\boldsymbol{x})}{\psi(\boldsymbol{x})} = \widetilde{\mathbf{x}}$$

and $\hat{\alpha} = \phi(\widetilde{\mathbf{x}})/\psi(\widetilde{\mathbf{x}}) = a/b$.

Put $\hat{\alpha} = f(\mathbf{X}_0)/g(\mathbf{X}_0)$ so that $f(\mathbf{X}_0)-\hat{\alpha}g(\mathbf{X}_0) = 0$. By the last corollary, it suffices to show that for $\mathbf{X}$ satisfying (7.15),

$$f(\mathbf{X})-\hat{\alpha}g(\mathbf{X}) \geq 0. \tag{7.43}$$

Define

$$a = M^2-M-2M(\zeta+z)+2\nu^2+\zeta^2 \text{ and } b = (M-\nu)^2+\nu^2-M,$$

where $z = M\frac{w(r-1)}{n(n-1)}$. Then $\hat{\alpha} = 2(1+\frac{(n-1)a}{b})$ and

$$f(\mathbf{X})-\hat{\alpha}g(\mathbf{X}) = \frac{1}{b}\Big(2b\Big(\sum_{l,k>l} M^2-M-2M(\zeta+z)+(\nu_1^l)^2+(\nu_1^k)^2+\zeta^2\Big)- \\ (n-1)a\Big(\sum_l (M-\nu_1^l)^2+(\nu_1^l)^2-M\Big)\Big).$$

To show (7.43) consider a deviation of the vector $\mathbf{X}$ from $\mathbf{X}_0$. Let $\nu_1^l = \nu+\epsilon_{\nu_1^l}$, $\zeta_{l,k} = \zeta+\epsilon_{\zeta_{l,k}}$. From (7.15) we obtain the following constraints on the deviations:

$$\sum_l \epsilon_{\nu_1^l} = 0, \qquad \sum_{l,k>l} \epsilon_{\zeta_{l,k}} = 0. \tag{7.44}$$

Using (7.11), (7.15), (7.44), and the identity

$$(n-1)\sum_{l=1}^{n}\epsilon^2_{\nu_1^l}=\sum_{l=1}^{n-1}\sum_{k=l+1}^{n}\epsilon^2_{(\nu_1^l)}+\epsilon^2_{(\nu_1^k)},$$

we obtain the following sequence of equalities:

$$\begin{aligned}
&2b\Big(\sum_{l,k>l}M^2-M-2M(\zeta+z)+(\nu_1^l)^2+(\nu_1^k)^2+\zeta^2\Big)\\
&\quad-(n-1)a\Big(\sum_{l}(M-\nu_1^l)^2+(\nu_1^l)^2-M\Big)\\
&=2b\Big(\binom{n}{2}a+\sum_{l,k>l}2\nu(\epsilon_{\nu_1^l}+\epsilon_{\nu_1^k})+2\zeta\epsilon_\zeta+\epsilon^2_{(\nu_1^l)}+\epsilon^2_{(\nu_1^k)}+\epsilon^2_\zeta\Big)\\
&\quad-(n-1)a\Big(nb+\sum_{l}-2(M-\nu)\epsilon_{(\nu_1^l)}+2\nu\epsilon_{(\nu_1^l)}+2\epsilon^2_{(\nu_1^l)}\Big)\\
&=2b\Big(\sum_{l,k>l}\epsilon^2_{(\nu_1^l)}+\epsilon^2_{(\nu_1^k)}+\epsilon^2_\zeta\Big)-2a\Big(\sum_{l,k>l}\epsilon^2_{(\nu_1^l)}+\epsilon^2_{(\nu_1^k)}\Big),
\end{aligned}$$

We note that $b\geq a$ because we can rewrite a as $a=(M-\zeta-z)^2+2(\nu-z)^2+z^2-M$. Thus, we obtain

$$\begin{aligned}
(M-\zeta-z)^2+(\nu-z)^2&\leq(M-\zeta-z+\nu-z)^2=(M-\nu)^2,\\
(\nu-z)^2+z^2&\leq\nu^2.
\end{aligned}$$

This establishes (7.43) and proves Proposition 7.3.

Hence the minimum of (7.7) under the constraints in (7.15) is attained at $\mathbf{X}=\mathbf{X}_0$.

Remark: The fact that $f(\mathbf{X})-\hat{\alpha}g(\mathbf{X})\geq 0$ can be intuitively proved by directly applying Lemma 7.9 as explained below. Since $g(\mathbf{X})$ is strongly convex in $(\nu_1^1,\ldots,\nu_1^n)$, we fix the vector $(\zeta_{l,k})_{l,k>l}$. Along the directions defined by the constraints (7.15) of ν_1^l's we immediately get from Lemma 7.9 that

$$\frac{f\big((\nu_1^1,\ldots,\nu_1^n),(\zeta_{l,k})_{l,k>l}\big)}{g\big((\nu_1^1,\ldots,\nu_1^n),(\zeta_{l,k})_{l,k>l}\big)}\geq\frac{f\big((\nu,\ldots,\nu),(\zeta_{l,k})_{l,k>l}\big)}{g\big(\nu,\ldots,\nu\big)}.\tag{7.45}$$

With $(\nu_1^1,\ldots,\nu_1^n)$ fixed and for changes in the $\zeta_{l,k}$, the value of the function $g(\mathbf{X})$ does not change while $f(\mathbf{X})$ remains strongly convex and so we get by using Lemma 7.9 and (7.45) in successive steps,

$$\frac{f\big((\nu_1^1,\ldots,\nu_1^n),(\zeta_{l,k})_{l,k>l}\big)}{g\big(\nu_1^1,\ldots,\nu_1^n\big)}\geq\frac{f\big((\nu_1^1,\ldots,\nu_1^n),(\zeta,\ldots,\zeta)\big)}{g\big(\nu_1^1,\ldots,\nu_1^n\big)}\geq\frac{f\big((\nu,\ldots,\nu),(\zeta,\ldots,\zeta)\big)}{g\big(\nu,\ldots,\nu\big)}.$$

7.8.2 Minimization of (7.24) and (7.30)

To get a universal lower bound for (7.24) and (7.30) under the constraints (7.22)–(7.23) and (7.25)–(7.29), we perform the following two minimizations:

$$\min_{\text{s.t. (7.22)–(7.23) holds}} \sum_{l}\sum_{i}(\nu_i^l)^2, \tag{7.46}$$

$$\min_{\text{s.t. (7.25)–(7.27) holds}} \sum_{l,k>l}\sum_{i,j}(\lambda_{i,j}^{l,k})^2. \tag{7.47}$$

Note that the equations (7.28)–(7.29) are not used in the minimizations above, but the optimum points satisfy those constraints as will be evident later in this section.

The optimum for (7.46) can be obtained directly from the proof of Lemma 5.7, by setting $r = 1$. Thus we get

$$\begin{aligned} \nu_0^l &= \Omega_0 = \frac{M(n-w)}{n}, \quad l = 1, \ldots, n, \\ \nu_i^l &= \Omega_1 = \frac{Mw}{(q-1)n}, \quad l = 1, \ldots, n,\ i > 0. \end{aligned} \tag{7.48}$$

Expression (7.47) is optimized as follows. We consider the Lagrangian and differentiate it with respect to $\lambda_{i,j}^{l,k}$ for all $i, j \in \{0, \ldots, q-1\}$ and for all $l = 1, \ldots, n$, $k = l+1, \ldots, n$. This leads to the following set of equations that the optimum points must satisfy. For each $l, k > l$,

$$\begin{aligned} 2\lambda_{0,0}^{l,k} + \gamma^{l,k} &= 0, \\ 2\lambda_{0,j}^{l,k} + \gamma^{l,k} + \tau &= 0, \quad j > 0, \\ 2\lambda_{i,0}^{l,k} + \gamma^{l,k} + \tau &= 0, \quad i > 0, \\ 2\lambda_{i,j}^{l,k} + \gamma^{l,k} + \mu &= 0, \quad i > 0, j > 0. \end{aligned}$$

This set of equations implies the conditions

$$\begin{aligned} \lambda_{0,j}^{l,k} &= \lambda_{i,0}^{l,k}, \quad i, j > 0, \\ \lambda_{i,j}^{l,k} &= \lambda_{i',j'}^{l,k}, \quad i, j, i', j' > 0,\ (i,j) \neq (i',j'). \end{aligned}$$

Finally, the objective function in (7.47) is convex, and the objective function and the constraints are symmetric vector-wise in $\big((\lambda_{i,j}^{l,k})_{i,j}\big)$ for each $l, k > l$. Hence at the point of optimum the following set of equalities must hold true:

$$\big((\lambda_{i,j}^{l,k})_{i,j}\big) = \big((\lambda_{i,j}^{l',k'})_{i,j}\big), \quad (l,k) \neq (l',k'),$$

that is, $\lambda_{i,j}^{l,k} = \lambda_{i,j}^{l',k'}$ for all $i, j \in \{0, \ldots, q-1\}$ and for all $(l,k) \neq (l',k')$. Combined with the constraints (7.25)–(7.27) we obtain that at the optimum

$$\begin{aligned} \lambda_{0,j}^{l,k} &= \lambda_{i,0}^{l,k} = \Lambda_1, \quad l, k > l,\ i, j > 0, \\ \lambda_{i,j}^{l,k} &= \lambda_{i',j'}^{l',k'} = \Lambda_2, \quad l, k > l,\ l', k' > l',\ i, j, i', j' > 0, \\ \lambda_{0,0}^{l,k} &= \lambda_{0,0}^{l',k'} = \Lambda_0, \quad l, k > l,\ l', k' > l'. \end{aligned}$$

Solving for $\Lambda_0, \Lambda_1, \Lambda_2$ using (7.25)–(7.27) gives

$$\begin{aligned}\Lambda_0 &= \frac{M(n-w)(n-w-1)}{n(n-1)},\\ \Lambda_1 &= \frac{1}{q-1}\frac{Mw(n-w)}{n(n-1)},\\ \Lambda_2 &= \frac{1}{(q-1)^2}\frac{Mw(w-1)}{n(n-1)}.\end{aligned} \tag{7.49}$$

7.8.3 Minimization of (7.39)

We minimize $f(\mathbf{X})$ over the following constraints, which are obtained from (7.8) and (7.11):

$$0 \le \alpha^{l,k}, \beta^{l,k}, \gamma^{l,k} \le 1, \qquad \alpha^{l,k} + \beta^{l,k} + \gamma^{l,k} \le 1, \tag{7.50}$$

$$\sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \gamma^{l,k} = \binom{w}{2}, \tag{7.51}$$

$$\sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \alpha^{l,k} + \beta^{l,k} = w(n-w). \tag{7.52}$$

First we show the following property.

Lemma 7.11 *The function $g(\alpha^{l,k}, \beta^{l,k}, \gamma^{l,k})$ is convex in the region defined by the set of constraints (7.50)–(7.52).*

Proof: Let $a = 4 - 4\alpha^{l,k} - 4\beta^{l,k} - 4\gamma^{l,k}$. To show convexity, differentiate the function twice to obtain

$$D^2 g = A = \begin{bmatrix} 2 + 4\alpha^{l,k} + a & a & 2 + a \\ a & 2 + 4\beta^{l,k} + a & 2 + a \\ 2 + a & 2 + a & 4 + 4\gamma^{l,k} + a \end{bmatrix}.$$

From (7.50) we see that $a \ge 0$. We obtain a non-negativity condition

$$A \ge B = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

and the condition

$$A - B = a \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + 4 \begin{bmatrix} \alpha^{l,k} & 0 & 0 \\ 0 & \beta^{l,k} & 0 \\ 0 & 0 & \gamma^{l,k} \end{bmatrix} \succeq 0.$$

This implies that $A - B$ is positive semidefinite (under the constraints specified by (7.50)–(7.52)). It can be easily verified that B itself is positive semidefinite. Hence, A is positive semidefinite. This proves the convexity of the function g. ■

Since the function $f(\mathbf{X})$ and the constraints are symmetric vector-wise in the variables $(\alpha^{l,k}, \beta^{l,k}, \gamma^{l,k})$, the optimum is obtained at the point where all the variables are equal vector-wise:

$$(\alpha^{l,k}, \beta^{l,k}, \gamma^{l,k}) = (\alpha^{l',k'}, \beta^{l',k'}, \gamma^{l',k'}), \quad (l,k) \neq (l',k').$$

Moreover, $g(\alpha^{l,k}, \beta^{l,k}, \gamma^{l,k})$ and the constraints are symmetric in $\alpha^{l,k}$ and $\beta^{l,k}$, which further implies that $\alpha^{l,k} = \beta^{l,k}$ for each $l, k > l$. The optimum values can be now determined to be

$$\begin{aligned} \gamma &= \frac{w(w-1)}{n(n-1)}, \\ \alpha = \beta &= \frac{w(n-w)}{n(n-1)}. \end{aligned} \tag{7.53}$$

APPENDIX A

Here we collect some results and concepts from linear algebra that are used in the thesis. The proofs of these results appear, for instance, in Horn and Johnson [46].

For an $n \times n$ real symmetric matrix A denote its largest eigenvalue by $\lambda_{\max}(A)$. We write $A \geq 0$ if all the entries of A are nonnegative.

Theorem A.1 (Rayleigh-Ritz inequality) [46, Theorem 4.2.2] *Let A be a symmetric matrix. For any $\boldsymbol{y} \neq 0$,*

$$\lambda_{\max}(A) \geq \frac{(A\boldsymbol{y}, \boldsymbol{y})}{(\boldsymbol{y}, \boldsymbol{y})}. \tag{A.1}$$

Lemma A.2 [46, Thm. 8.1.22] *Let $A \geq 0$ be symmetric. Then*

$$\lambda_{\max}(A) \leq \max_{1 \leq i \leq n} \sum_j A_{ij}. \tag{A.2}$$

Lemma A.3 [46, p. 491] *If $0 \leq B \leq A$ for some matrix B, or if B is a principal minor of A, then $|\lambda_{\max}(B)| \leq \lambda_{\max}(A)$.*

Definition A.1 *An $n \times n$ matrix $A \geq 0$ is called* irreducible *if for any partition of the set of indices $\{1, 2, \ldots, n\} = \{i_1, i_2, \ldots i_s\} \cup \{j_1, j_2, \ldots, j_t\}$ into two disjoint subsets with $s + t = n$, the matrix $(A_{i_\alpha, j_\beta})_{1 \leq \alpha \leq s, 1 \leq \beta \leq t}$ is nonzero.*

Theorem A.4 (Perron-Forbenius) [46, Theorem 8.4.4] *Let $A \geq 0$ be a $n \times n$ irreducible symmetric matrix. Its largest eigenvalue $\lambda_{\max}(A)$ is positive and has multiplicity one. There exists a vector $\boldsymbol{y} > 0$ such that $A\boldsymbol{y} = \lambda_{\max}(A)\boldsymbol{y}$.*

Theorem A.5 (Ostrowski) [46, Theorem 4.5.9] *Let A, S be $n \times n$ matrices with A symmetric. Let the eigenvalues of A and SS^T be arranged in increasing order, i.e., $\lambda_i(A) \leq \lambda_j(A)$, $\lambda_i(SS^T) \leq \lambda_j(SS^T)$, $i < j$. For each $k = 1, \ldots, n$ there exists a real number θ_k such that*

1. *if S is nonsingular then $\lambda_1(SS^T) \leq \theta_k \leq \lambda_n(SS^T)$, with $\theta_k > 0$,*

2. *if S is singular then $0 \leq \theta_k \leq \lambda_n(SS^T)$,*

and

$$\lambda_k(SAS^T) = \theta_k \lambda_k(A).$$

In particular, the number of positive (resp. negative) eigenvalues of SAS^T is at most the number of positive (resp. negative) eigenvalues of A.

Index

(t, m, n)-net, 33, 46

annihilator polynomial, 74
association scheme, 7
  duality, 9
  eigenvalue of, 8
  kernel scheme, 13
  polynomial scheme, 9

Bassalygo-Elias inequality, 23
Bose-Mesner algebra, 7, 16
bound
  Bassalygo-Elias, 23, 43, 46
  Gilbert-Varshamov, 21, 42, 46
  Hamming, 43, 46
  Johnson, 22, 43, 80
  linear programming, 11, 15, 24, 54
  list decoding, 90
  Plotkin, 22, 43, 46
  Singleton, 27, 29, 43

Christoffel-Darboux
  formula, 53
  kernel, 24, 53
code, 9
  almost-MDS, 27, 31
  constant weight, 80
  dual, 10, 14, 28
  linear, 10
  MDS, 27
  minimum distance, 9
  NMDS, 27, 31
  ordered, 14, 29
codematrix, 22

design, 10
dual distribution, 10, 11

Frankl-Wilson bound, 69

generalized poset weight, 29

Hahn polynomial, 12, 16, 72
Hamming
  distance, 11
  scheme, 11
  space, 11, 21
  weight, 11

ideal, 28
inequality in the mean, 93
intersection number, 7, 14
intertwining function, 71
irreducible matrix, 101
isometry group, 16, 70

Johnson
  radius, 81
  scheme, 12
  space, 12, 80
Johnson space, 70

Krawtchouk polynomial, 73
  multivariate, 14, 47
  univariate, 12, 19, 24

left-adjusted, 14, 28

MacWilliams theorem, 11, 54

optimal distribution, 33
ordered Hamming
  distance, 12
  scheme, 13
  space, 12, 29, 41
  weight, 12
orthogonal array, 10, 28
  ordered, 14, 29
Ostrowski theorem, 101

Perron-Forbenius Theorem, 101
poset, 28
    dual, 28

rate, 23, 46
Rayleigh-Ritz inequality, 101
relative distance, 23, 46
right-adjusted, 28

shape, 13
spectral method, 24, 54
support, 16

Terwilliger algebra, 15, 73

weight distribution, 34

zonal spherical
    function, 18, 71
    kernel, 18, 71

Bibliography

[1] M. Aaltonen, personal communication (2007).

[2] M. Aaltonen, *Linear programming bounds for tree codes*, IEEE Trans. Inform. Theory **25** (1977), 85–90.

[3] M. Aaltonen, *A new upper bound on nonbinary block codes*, Discrete Mathematics **83** (1990), no. 2-3, 139–160.

[4] E. Agrell, A. Vardy, and K. Zeger, *Upper bounds for constant-weight codes*, IEEE Trans. Inform. Theory **46** (2000), 2373–2395.

[5] R. Ahlswede and L. H. Khachatrian, *The complete intersection theorem for systems of finite sets*, European J. Combin. **18** (1997), no. 2, 125–136.

[6] N. Alon, O. Goldreich, J. Håstad and R. Peralta *Simple constructions of almost k-wise independent random variables*, **3** (1992), 289–304.

[7] C. Bachoc, *Linear programming bounds for codes in Grassmannian spaces*, IEEE Trans. Inform. Theory **52** (2006), 2111–2126.

[8] C. Bachoc, *Semidefinite programming, harmonic analysis and coding theory*, http://arxiv.org/abs/0909.4767 (2009).

[9] E. Bannai and T. Ito, *Algebraic combinatorics I. Association schemes*, Benjamen/Cummings, London e. a. (1984).

[10] A. Barg and D. Nogin, *Spectral approach to linear programming bounds on codes*, Problems of Information Transmission **42** (2006), 12–25.

[11] A. Barg and P. Purkayastha, *Bounds on ordered codes and orthogonal arrays*, IEEE International Symposium on Inform. Th. (2007), 331–335, 24–29 June 2007.

[12] A. Barg and D. Nogin, A functional view of upper bounds on codes, in Y. Li et al., Eds., *Coding and Cryptology*, Singapore: World Scientific (2008), pp. 15-24.

[13] A. Barg and O. R. Musin, *Bounds on sets with few distances*, preprint, http://arxiv.org/abs/0905.2423 (2009).

[14] A. Barg and P. Purkayastha, *Bounds on ordered codes and orthogonal arrays*, Moscow Math. Journal **9** (2009), no. 2, 211–243.

[15] A. Barg and P. Purkayastha, *Near MDS poset codes and distributions*, preprint, accepted for publication in *Error-Correcting Codes, Cryptography and Finite Geometries*, Editors: A. Bruen and D. Wehlau, AMS series in Contemporary Mathematics.

[16] A. Barg and P. Purkayastha, *Near MDS poset codes and distributions*, preprint, accepted for publication in IEEE International Symposium on Inform. Th. (2010)

[17] J. Bierbrauer, Y. Edel, and W. Ch. Schmid, *Coding-theoretic constructions of (t, m, s)-nets and ordered orthogonal arrays*, J. Combin. Des. **10** (2002), no. 6, 403–418.

[18] J. Bierbrauer and W. Ch. Schmid, *An asymptotic Gilbert-Varshamov bound for (T, M, S)-nets*, Integers **5** (2005), no. 3, A4, 11 pp. (electronic).

[19] J. Bierbrauer, *A direct approach to linear programming bounds*, Des. Codes Cryptogr. **42** (2007), 127–143.

[20] V. M. Blinovskii, *Bounds on codes decodable into a list of a finite size*, Probl. Inform. Trans. (1986), no. 1.

[21] S. Bochner, *Hilbert distances and positive definite functions*, Ann. of Math., **42** (1941), 647–656.

[22] M. de Boer, *Almost MDS codes*, Des. Codes Cryptogr. **9** (1996), 143–155.

[23] A. E. Brouwer, A. M. Cohen, and A. Neumaier, *Distance-regular graphs*, Springer-Verlag, Berlin e. a. (1989).

[24] R. A. Brualdi, J.S. Graves, and K. M. Lawrence, *Codes with a poset metric*, Discrete Math. **147** (1995), no. 1-3, 57–72.

[25] A. R. Calderbank, P. Delsarte and N. J. A. Sloane, *A strengthening of the Assmus-Mattson theorem.* IEEE Trans. Inform. Theory **37** (1991), no. 5, 1261–1268.

[26] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Research Repts. Suppl. **10** (1973), 1–97.

[27] M. Deza, P. Erdös, and P. Frankl, *Intersection properties of systems of finite sets*, Proc. London Math. Soc., **36** (1978), no. 3, 369–384.

[28] S. Dodunekov and I. Landgev, *Near-MDS codes*, J. of Geometry **54** (1995), no. 1, 30–43.

[29] S. T. Dougherty and M. M. Skriganov, *MacWilliams duality and the Rosenbloom-Tsfasman metric*, Mosc. Math. J. **2** (2002), no. 1, 81–97.

[30] S. T. Dougherty and M. M. Skriganov, *Maximum distance separable codes in the ρ metric over arbitrary alphabets*, Journal of Algebraic Combinatorics **16** (2002), 71–81.

[31] C. F. Dunkl and Y. Xu, *Orthogonal polynomials of several variables*, Cambridge University Press (2001).

[32] C. F. Dunkl, *A Krawtchouk polynomial addition theorem and wreath products of symmetric groups*, Indiana University Math. J., **25** (1976), no. 4, 335–358

[33] C. F. Dunkl, *Spherical functions on compact groups and applications to special functions*, Symposia Mathematica, Vol. XXII, Academic Press, London (1977), 145–161.

[34] P. Erdös, C. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quart J. Math., Oxford **12** (1961), no. 2, 313–320.

[35] A. Faldum and W. Willems, *A characterization of MMD codes*, IEEE Trans. Inform. Theory **44** (1998), no. 4, 1555–1558.

[36] R. A. Fisher, *An examination of the different possible solutions of a problem in incomplete blocks*, Annals of Eugenics, **10** (1940), 52–75.

[37] P. Frankl and R. M. Wilson, *Intersection theorems with geometric consequences*, Combinatorica, **1** (1981), no. 4, 357–368.

[38] P. Frankl and Z. Füredi, *Families of finite sets with missing intersections*, Colloquia Mathematica Societatis Janos Bolyai **37** (1981), 305–320.

[39] Z. Füredi and Z. Katona, *Multiply intersecting families of sets*, J. of Combinatorial Th. Ser. A, **106** (2004), no. 2, 315–326.

[40] A. Ganesan and P. O. Vontobel, *On the existence of universally decodable matrices*, IEEE Trans. Inform. Theory **53** (2007), no. 7, 2572–2575.

[41] I. M. Gelfand, *Spherical functions in symmetric Riemann spaces*, Doklady Akad. Nauk SSSR (N.S.) **70** (1950). 5–8.

[42] I. M. Gessel and R. P. Stanley, *Algebraic enumeration*, Handbook of combinatorics, Vol. 2, Elsevier, Amsterdam (1995), 1021–1061.

[43] D. Gijswijt, A. Schrijver and H. Tanaka, *New upper bounds for nonbinary codes based on the Terwilliger algebra and semidefinite programming*, J. Combin. Theory Ser. A **113** (2006), no. 8, 1719–1731.

[44] C. D. Godsil, *Algebraic combinatorics*, Chapman & Hall, New York (1993).

[45] E. Hlawka, *Funktionen von beschränkter Variation in der Theorie der Gleichverteilung*. (German) Ann. Mat. Pura Appl. **54** (1961), no. 4, 325–333.

[46] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge (1990).

[47] J. Y. Hyun and H. K. Kim, *Maximum distance separable poset codes*, Des. Codes Cryptogr. **28** (2008), no. 3, 247–261.

[48] S. M. Johnson, *A new upper bound for error-correcting codes*, IRE Trans. Inform. Th. **8** (1962), 203–207.

[49] H. K. Kim and D. Y. Oh, *A classification of posets admitting MacWilliams identity*, IEEE Trans. Inform. Theory **51** (2005), 1424–1431.

[50] T. Y. Lam, *A first course in noncommutative rings*, Second edition. Graduate Texts in Mathematics, 131. Springer-Verlag, New York (2001).

[51] K. M. Lawrence, *A combinatorial characterization of* (t, m, s)*-nets in base* b, J. Combin. Designs **4** (1996), 275–293.

[52] G. A. Kabatyanskii and V. I. Levenshtein, *Bounds for packings on a sphere and in space*, (Russian) Problemy Peredachi Informacii, **14** (1978), no. 1, 3–25.

[53] M. G. Kreĭn, *Hermitian positive kernels on homogeneous spaces* I, II, Ukrainian Math. Journal, **1** (1949), no. 4, 64–98 and **2** (1950), no. 2, 10–59.

[54] V. I. Levenshtein, *Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces*, IEEE Trans. Inform. Theory **41** (1995), no. 5, 1303–1321.

[55] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*, North-Holland, Amsterdam (1991).

[56] W. J. Martin and D. R. Stinson, *A generalized Rao bound for ordered orthogonal arrays and* (t, m, s)*-nets*, Canad. Math. Bull. **42** (1999), no. 3, 359–370.

[57] W. J. Martin, *Linear programming bounds for ordered orthogonal arrays and* (T, M, S)*-nets*, Monte Carlo and quasi-Monte Carlo methods 1998 (Claremont, CA), Springer, Berlin (2000), 368–376.

[58] W. J. Martin, (t, m, s)*-nets*, Handbook of Combinatorial Designs (C. J. Colbourn and J. H. Dinitz, eds.), CRC Press, Boca Raton, FL (2006), 639–643.

[59] W. J. Martin and T. I. Visentin, *A dual Plotkin bound for* (T, M, S)*-nets*, IEEE Trans. Inform. Theory **53** (2007), no. 1, 411–415.

[60] W. J. Martin and D. R. Stinson, *Association schemes for ordered orthogonal arrays and* (T, M, S)*-nets*, Canad. J. Math. **51** (1999), no. 2, 326–346.

[61] J. L. Massey and S. Serconek, *Linear complexity of periodic sequences: a general theory*, Advances in cryptology—CRYPTO '96 (Santa Barbara, CA), Lecture Notes in Comput. Sci., vol. 1109, Springer, Berlin (1996), 358–371.

[62] R. J. McEliece, E. R. Rodemich, H. Jr. Rumsey, and L. R. Welch, *New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities*, IEEE Trans. Information Theory **IT-23** (1977), no. 2, 157–166.

[63] G. L. Mullen and W. Ch. Schmid, *An equivalence between* (t, m, s)*-nets and strongly orthogonal hypercubes*, Journal of Combin. Theory, Ser. A **76** (1996), 164–174.

[64] H. Niederreiter and C. Xing, *Rational points on curves over finite fields*, Cambridge University Press (2001).

[65] R. R. Nielsen, *A class of Sudan-decodable codes*, IEEE Trans. Inform. Theory **46** (2000), no. 4, 1564–1572.

[66] H. Niederreiter, *Digital nets and coding theory*, Coding Theory, Cryptography, and Combinatorics (K. Feng, H. Niederreiter, and C. Xing, eds.), Birkhäuser, Basel e.a., (2004), 247–257.

[67] H. Niederreiter, *Low-discrepancy point sets*, Monatsh. Math. **102** (1986), no. 2, 155–167.

[68] H. Nozaki, *New upper bound for the cardinalities of s-distance sets on the unit sphere*, preprint, http://arxiv.org/abs/0906.0195 (2009).

[69] J. Qian and D. K. Ray-Chaudhuri, *Extremal case of Frankl-Ray-Chaudhuri-Wilson Inequality*, J. of Statistical Planning and Inference, **95** (2001), 293–306.

[70] D. K. Ray-Chaudhuri and R. M. Wilson, *On t-designs*, Osaka J. Math., **12** (1975), no. 3, 737–744.

[71] B. Roos, *Multinomial and Krawtchouk approximations to the generalized multinomial distribution*, Theory Probab. Appl. **46** (2002), no. 1, 103–117

[72] M. Yu. Rosenbloom and M. A. Tsfasman, *Codes for the m-metric*, Problems of Information Transmission **33** (1997), no. 1, 45–52.

[73] R. Roth, *Introduction to coding theory*, Cambridge University Press, Cambridge (2006).

[74] A. Schrijver, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Trans. Inform. Theory **51** (2005), no. 8, 2859–2866.

[75] I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–107.

[76] V. M. Sidelnikov, *Upper bounds on the cardinality of a binary code with a given minimum distance*, Translated from the Russian by A. M. Odlyzko (Problemy Peredaci Informacii **10** (1974), no. 2, 43–51). Information and Control **28** (1975), no. 4, 292–303.

[77] V. R. Sidorenko, *An upper bound on the length of q-ary codes*, (Russian) Problemy Peredaci Informacii 11 (1975), no. 3, 14–20.

[78] M. M. Skriganov, *Coding theory and uniform distributions*, Algebra i Analiz **13** (2001), no. 2, 191–239, English translation in *St. Petersburg Math. J.* vol. 13 (2002), no. 2, 301–337.

[79] M. M. Skriganov, *On linear codes with large weights simultaneously for the Rosenbloom-Tsfasman and Hamming metrics*, J. of Complexity **23** (2007), 926–936

[80] I. M. Sobol, *Distribution of points in a cube and approximate evaluation of integrals*, Ž. Vyčisl. Mat. i Mat. Fiz. **7** (1967), 784–802.

[81] G. Szegö, *Orthogonal polynomials*, Fourth edition. American Mathematical Society, Colloquium Publications, Vol. XXIII. Providence, R. I. (1975).

[82] S. Tavildar and P. Viswanath, *Approximately universal codes over slow-fading channels*, IEEE Trans. Inform. Theory **52** (2006), no. 7, 3233–3258.

[83] P. Terwilliger, *The subconstituent algebra of an association scheme. I,* J. Algebraic Combin. **1** (1992), no. 4, 363–388.

[84] M. V. Tratnik, *Multivariable Meixner, Krawtchouk, and Meixner-Pollaczek polynomials*, J. Math. Phys. **30** (1989), no. 12, 2740–2749.

[85] F. Vallentin, *Symmetry in semidefinite programs*, Linear Algebra and Applications, **430** (2009), 360–369.

[86] N. Ja. Vilenkin, *Special functions and the theory of group representations*, Translated from the Russian by V. N. Singh. Translations of Mathematical Monographs, Vol. 22 American Mathematical Society, Providence, R. I. (1968).

[87] V. Wei, *Generalized Hamming weights for linear codes*, IEEE Trans. Inform. Theory **37** (1991), no. 5, 1412–1418.