Delsarte's linear programming bound

15-859 Coding Theory, Fall '14

Jason Li

December 10, 2014

Contents

1 Introduction

A fundamental objective in coding theory is to find upper bounds on the size of a code in $\mathbb{F}_q^n$ with a certain distance d. Upper bounds make us aware of the best transmission rates possible for codes of a certain size and distance.

Delsarte's linear program is a method for upper bounding the sizes of codes. It establishes a series of linear constraints that every code in $\mathbb{F}_q^n$ with distance d must satisfy. Naturally, the objective function becomes to maximize the size of the code, subject to these linear constraints. Together, the constraints and the objective function form a linear program, which, when solved, gives an upper bound on the size of a code in $\mathbb{F}_q^n$ with distance d.

The linear program can be computed numerically for small n. For large n, the best we can hope for is to provide asymptotics on the value of the optimum.

We begin with some preliminaries required to understand the motivation behind Delsarte's linear program. We then discuss numerical results, followed by asymptotic bounds, and conclude with open problems to be worked on.

2 Preliminaries

2.1 Association schemes

Delsarte's linear program makes use of the notion of a Hamming scheme, which is a type of association scheme. The association scheme is defined as follows.

Definition 1. *A symmetric association scheme $A = \{X, \mathcal{R}\}$ is a finite set X and a set of relations $\mathcal{R} = \{R_0, R_1, \ldots, R_d\}$ on X such that the R_i satisfy the following properties:*

1. *$R_0 = \{(x, x) : x \in X\}$.*
2. *If $(x, y) \in R_i$, then $(y, x) \in R_i$. (Note: this condition is weaker in asymmetric association schemes.)*
3. *$\mathcal{R}$ partitions $X \times X$.*
4. *Fix values $h, i, j \in [0, d]$, and consider the relations R_h, R_i, and R_j. For each $(x, y) \in R_h$, the number of elements $z \in X$ such that $(x, z) \in R_i$ and $(z, y) \in R_j$ is always the same, regardless of (x, y). Let this number be $p_{i,j}^h$.*

Perhaps the most interesting and important property is the fourth condition. Unfortunately, the condition is not so motivating when stated in mathematical notation. For more intuition on this condition, we can think of the association scheme in terms of a graph.

Consider the graph with vertex set X, and undirected edges representing the values $(x, y) \in X \times X$. (Note that (x, x) is allowed, so this graph has self-loops.) We can think of the relations $R_0, \ldots, R_d$ as $d+1$ distinct colors, and for each pair (x, y) in R_i, we color the corresponding edge (x, y) the color of R_i. Since $\{R_0, \ldots, R_d\}$ partitions $X \times X$, we know that each edge is colored exactly one color. In other words, we get a coloring of the complete graph (with self-loops) into $d+1$ colors.

Under this representation of the association scheme, the fourth condition may seem more intuitive, especially when we view the triples $\{(x, y),\ (y, z),\ (z, x)\}$ as triangles in the graph. Then, the condition becomes the following:

4. *Fix values $h, i, j \in [0, d]$, and consider all triangles with the first edge colored R_h, the second edge colored R_i, and the third edge colored R_j. Then, every edge colored R_h takes part in the same number of such triangles.*

This property will become very important later on, so it is important to develop intuition for it.

2.2 Hamming scheme

We now move to the Hamming scheme, which is the association scheme relevant to Delsarte's linear program.

Definition 2. *For fixed n and q, consider the vector space $\mathbb{F}_q^n$. The Hamming scheme on $\mathbb{F}_q^n$ is defined as follows:*

1. *There are $n+1$ relations $R_0, \ldots, R_n$, which correspond to Hamming distances between pairs of points.*
2. *For two coordinates $x, y \in \mathbb{F}_q^n$, the pair (x, y) belongs to the relation indexed by the Hamming distance of x and y. That is, $(x, y) \in R_{\Delta(x,y)}$.*

Before we continue, it is helpful to verify that the Hamming scheme is indeed an association scheme.

Claim 1. *The Hamming scheme on $\mathbb{F}_q^n$ is an association scheme.*

Proof. We need to check all of the properties of a symmetric association scheme.

1. $R_0 = \{(x, x) : x \in X\}$.
   - Satisfied because $\Delta(x, y) = 0 \Leftrightarrow x = y$.
2. If $(x, y) \in R_i$, then $(y, x) \in R_i$.
   - Satisfied because the Hamming distance is symmetric.
3. $\mathcal{R}$ partitions $X \times X$.
   - Satisfied by definition.
4. Fix values $h, i, j \in [0, d]$, and consider the relations R_h, R_i, and R_j. For each $(x, y) \in R_h$, the number of elements $z \in X$ such that $(x, z) \in R_i$ and $(z, y) \in R_j$ is always the same, regardless of (x, y).
   - This is perhaps the most interesting to check. Intuitively, this is true because the Hamming distance has many symmetric properties. For example, it is invariant under coordinate shifts and coordinate permutations. It is also possible to compute $p_{i,j}^h$, the exact number of satisfying triples (x, y, z) for each $(x, y) \in R_h$. This number is

$$p_{i,j}^h = \sum_{\delta=0}^{\lfloor (i+j-h)/2 \rfloor} (q-2)^{i+j-h-2\delta} \binom{h}{j-\delta} \binom{j-\delta}{h-i+\delta} \binom{n-h}{(i+j-w)/2}.$$

This result is not important later on, so we omit the derivation.

□

2.3 Associate matrices

In an association scheme with set X and relations $R_0, \ldots, R_d$, we can define one associate matrix A_i for each R_i as follows:

Definition 3. *The associate matrices $A_0, \ldots, A_d$ have rows and columns indexed by elements in X. (So each A_i is an $|X|$-by-$|X|$ matrix.) The matrix A_i represents an "indicator" matrix of R_i, in that entry (x, y) of A_i is 1 if $(x, y) \in R_i$, and 0 otherwise.*

As an example, let us consider the Hamming scheme on $\mathbb{F}_2^3$, indexed by $\{000, 001, 010, 011, 100, 101, 110, 111\}$. Then, the associate matrices are as follows.

$$A_0 = \begin{bmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \end{bmatrix}, A_1 = \begin{bmatrix} 0&1&1&0&1&0&0&0 \\ 1&0&0&1&0&1&0&0 \\ 1&0&0&1&0&0&1&0 \\ 0&1&1&0&0&0&0&1 \\ 1&0&0&0&0&1&1&0 \\ 0&1&0&0&1&0&0&1 \\ 0&0&1&0&1&0&0&1 \\ 0&0&0&1&0&1&1&0 \end{bmatrix}, A_2 = \begin{bmatrix} 0&0&0&1&0&1&1&0 \\ 0&0&1&0&1&0&0&1 \\ 0&1&0&0&1&0&0&1 \\ 1&0&0&0&0&1&1&0 \\ 0&1&1&0&0&0&0&1 \\ 1&0&0&1&0&0&1&0 \\ 1&0&0&1&0&1&0&0 \\ 0&1&1&0&1&0&0&0 \end{bmatrix}, A_3 = \begin{bmatrix} 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \end{bmatrix}.$$

The associate matrices have several nice properties, listed below.

1. $A_0 = I$, since R_0 only has elements of the form (x, x).
2. $\sum_{i=0}^{d} A_i$ is the all-ones matrix, since $R_0, \ldots, R_d$ partition $X \times X$.
3. If we multiply two matrices A_i and A_j, then we get a linear combination of A_h for $h \in [0, d]$.
   - In particular, $A_j A_i = \sum_{h=0}^{d} p_{i,j}^h A_h$. This is where property 4 of association schemes comes into play. It is easy to verify through matrix multiplication that the value in entry (x, y) is exactly the number of triangles with first edge (x, y), second edge in R_i, and third edge in R_j. And since this number is the same for each element in R_h, we get the above equation.
4. Since $p_{i,j}^h = p_{j,i}^h$, we get $A_j A_i = A_i A_j$ from the above property, so the matrices are pairwise commutative.
5. If we think of the matrices as a vector space, then the A_i are linearly independent. This is because each of the $|X|^2$ matrix entries is 1 in exactly one A_i.

2.4 The Bose-Mesner algebra

Consider the vector space spanned by the associate matrices. Since multiplying two basis elements A_i and A_j gives us another element of the vector space, we know that multiplying any two linear combinations of the A_i gives another element in the vector space.

An algebra is a vector space equipped with a bilinear product. Therefore, the vector space spanned by the associate matrices forms an algebra, with the bilinear product being matrix multiplication. This algebra is called the Bose-Mesner algebra.

2.5 Orthogonal basis

It turns out that the Bose-Mesner algebra always has another basis of pairwise "orthogonal" matrices $E_0, \ldots, E_d$ such that

1. $E_i E_j$ is the zero matrix if $i \neq j$
2. $E_i^2 = E_i$ (such matrices are called idempotent.)

We will not prove this theorem, but think of it as analogous to the spectral theorem of linear algebra.

2.6 First and second eigenmatrices

The Bose-Mesner algebra has two special matrices P and Q, called the first eigenmatrix and the second eigenmatrix, respectively. These matrices are defined as follows:

1. The entries of P satisfy $A_i = \sum_{j=0}^{d} P_{ji} E_j$.
2. The entries of Q satisfy $E_i = \frac{1}{|X|} \sum_{j=0}^{d} Q_{ji} A_j$.

They are essentially change-of-basis matrices from the basis $A_0, \ldots, A_d$ to the basis $E_0, \ldots, E_d$ and back.

If we instead think of the A_i and E_i as individual entries of a matrix (i.e. pretend that they are numbers), then the definitions can be more concisely written as

1. $\begin{bmatrix} A_0 & A_1 & \cdots & A_d \end{bmatrix} = \begin{bmatrix} E_0 & E_1 & \cdots & E_d \end{bmatrix} \cdot \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ P_{*,0} & P_{*,1} & \cdots & P_{*,d} \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}$
2. $\begin{bmatrix} E_0 & E_1 & \cdots & E_d \end{bmatrix} = \frac{1}{|X|} \cdot \begin{bmatrix} A_0 & A_1 & \cdots & A_d \end{bmatrix} \cdot \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ Q_{*,0} & Q_{*,1} & \cdots & Q_{*,d} \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}$

If we combine the two equations, then we can see that

$$\begin{bmatrix} E_0 & E_1 & \cdots & E_d \end{bmatrix} = \frac{1}{|X|} \cdot \left(\begin{bmatrix} E_0 & E_1 & \cdots & E_d \end{bmatrix} \cdot P \right) \cdot Q.$$

Since the E_i are linearly independent, we must have

$$P \cdot Q = |X| \cdot I,$$

where I is the $(d+1)$-by-$(d+1)$ identity matrix. This is what we might expect, when viewing P and Q as change-of-basis matrices.

2.7 Eigenmatrices for the Hamming scheme

Even for the Hamming scheme, computing the eigenmatrices P and Q is highly non-trivial. Delsarte [2] showed that the eigenmatrix Q for the Hamming scheme on $\mathbb{F}_q^n$ can be represented in terms of the Krawtchouk polynomials, which are defined as

$$K_k(x) = \sum_{i=1}^{k} \binom{x}{i} \binom{n-x}{k-i} (-1)^i (q-1)^{k-i}.$$

In particular, $Q_{i,k} = K_k(i)$.

The particular values in Q are not particularly important, so we also skip this derivation.

3 Delsarte's linear programming bound

In this section, we present the main ideas behind Delsarte's linear program.

3.1 Distribution vectors

For a given association scheme $(X, \mathcal{R})$ and a subset Y of X, we can define the distribution vector of Y as follows:

Definition 4. *The distribution vector of Y is the vector $\mathbf{a}$ of length $d+1$ such that*

$$a_i = \frac{|(Y \times Y) \cap R_i|}{|Y|}.$$

In graph notation, we can think of the subgraph induced by the vertices in Y. Then, a_i is simply the average degree of a vertex in Y, where only edges in R_i are considered.

We can easily see that $a_i \geq 0$ for each i, and $\sum_{i=0}^{d} a_i = |Y|$.

3.2 Hamming scheme

Let us go back to the Hamming scheme, with $X = \mathbb{F}_q^n$. We can let $Y \subset X$ be a set of codewords. From before, we know that for the distribution vector $\mathbf{a}$ of Y, we have $a_i \geq 0$ and $\sum_{i=0}^{d} a_i = |Y|$.

In addition, suppose that the code has distance r. Then, we also know that $a_1 = a_2 = \cdots = a_{r-1} = 0$.

There is one more property of $\mathbf{a}$, which we prove next.

3.3 Main theorem

Here is the key theorem on distribution vectors that forms the basis for the linear programming bound.

Theorem 1. *If $\mathbf{a}$ is a distribution vector of a subset Y of an association scheme with second eigenmatrix Q, then $\mathbf{a}Q \geq \mathbf{0}$. (That is, the vector $\mathbf{a}Q$ has only non-negative entries.)*

Proof. Let $\mathbf{y}$ be the characteristic vector of Y. That is, $y_x = 1$ if $x \in Y$, and 0 otherwise. Then,

$$0 \leq \|\mathbf{y}E_i\|^2 = (\mathbf{y}E_i)(\mathbf{y}E_i)^T = \mathbf{y}E_iE_i^T\mathbf{y}^T = \mathbf{y}E_i\mathbf{y}^T,$$

where the last step is true because E_i is idempotent and symmetric.

Recall that $E_i = \frac{1}{|X|}\sum_{j=0}^{d} Q_{ji}A_j$ and $a_i = \frac{\mathbf{y}A_i\mathbf{y}^T}{|Y|}$. Therefore,

$$0 \leq \mathbf{y}E_i\mathbf{y}^T = \frac{1}{|X|}\mathbf{y}\left(\sum_{j=0}^{d} Q_{ji}A_j\right)\mathbf{y}^T = \frac{1}{|X|}\left(\sum_{j=0}^{d} Q_{ji}\mathbf{y}A_j\mathbf{y}^T\right) = \frac{|Y|}{|X|}\sum_{j=0}^{d} a_jQ_{ji} = \frac{|Y|}{|X|}(\mathbf{a}Q)_i.$$

So for each i, $(\mathbf{a}Q)_i \geq 0$, as desired.

□

3.4 Formulation of Delsarte's linear program

Let us collect all of the conditions that $\mathbf{a}$ must satisfy:

1. $a_i = 1$.
2. $a_i = 0$ for $1 \leq i < r$.
3. $a_i \geq 0$ for $r \leq i \leq n$.
4. $\mathbf{a}Q \geq 0$. (This introduces $d+1$ linear inequalities.)

At the end, we know that $\sum_{i=0}^{d} a_i = |Y|$. Therefore, to upper bound the set Y of codewords, our objective of the linear program is to maximize $\sum_{i=0}^{d} a_i$.

That's it for Delsarte's linear program. Since any code in F_q^n of distance r must satisfy the above constraints, the optimum of the linear program is an **upper bound** on the maximum size of the code.

One thing of note is that Delsarte's linear program makes no assumptions on the structure of Y. Therefore, the upper bound holds for all codes, not just linear codes.

4 Numerical computations

For small n and q, we can numerically solve the linear program to find the upper bound for codes in F_q^n. Here is a table comparing the Hamming bound and the LP bound for codes in $\mathbb{F}_2^n$ with distance δ:

n	δ	Hamming Bound	Linear Programming Bound
11	3	170.7	170.7
11	5	30.6	24
11	7	8.8	4
12	3	315.1	292.6
12	5	51.9	40
12	7	13.7	5.3
13	3	585.1	512
13	5	89.0	64
13	7	21.7	8
14	3	1092.3	1024
14	5	154.6	128
14	7	34.9	16
15	3	2048	2048
15	5	270.8	256
15	7	56.9	32

Note that the tables suggest that the LP bound is always at most the Hamming bound. This is in fact true: Delsarte [2] showed how to establish the Hamming bound using the LP bound, so the LP bound is always at least as strong.

Also, note the perfect code with $n = 15 = 2^4 - 1$ and $\delta = 3$. As expected, both bounds achieve this perfect code.

5 Asymptotics

What about for higher n? We would like to find asymptotics for the linear programming bound. For the rest of this report, we will focus only on *binary* codes.

Let $A(n, \lfloor \delta n \rfloor)$ be the maximum size of a binary code with length n and distance δn. We can define the function $R(\delta) = \limsup_{n \to \infty} \frac{\log_2 A(n, \lfloor \delta n \rfloor)}{n}$. Intuitively, this is an asymptotic measure of the best rate possible for a binary code.

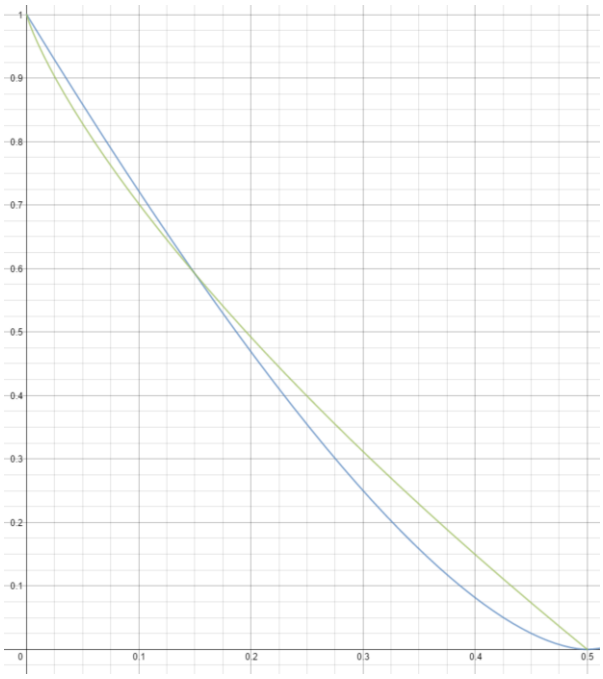
Similarly, let $A_{LP}(n, \lfloor \delta n \rfloor)$ to be the maximum value of $\sum_{i=0}^{d} a_i$ for some $\mathbf{a}$ that satisfies Delsarte's linear program. Since the LP bound is an upper bound, we have $A(n, \lfloor \delta n \rfloor) \leq A_{LP}(n, \lfloor \delta n \rfloor)$.

We can also define $R_{LP}(\delta) = \limsup_{n \to \infty} \frac{\log_2 A_{LP}(n, \lfloor \delta n \rfloor)}{n}$. We want bounds on $R_{LP}(\delta)$, which is an upper bound for $R(\delta)$.

5.1 Upper bound

We are most interested in an upper bound for $R_{LP}(\delta)$, since this will also be an upper bound for $R(\delta)$. McEliece, Rodemich, Rumsey, and Welch [MRRW '77] showed that $R_{LP}(\delta) \leq H(\frac{1}{2} - \sqrt{\delta(1-\delta)})$. **This is the best bound known for** $\delta \geq 0.273$.

Here is a plot of this upper bound with the Elias-Bassalygo bound $R(\delta) \leq 1 - H\left(\frac{1 - \sqrt{1-2\delta}}{2}\right)$, which we saw in class:



5.2 Proof of $R_{LP}(\delta) \leq H(\frac{1}{2} - \sqrt{\delta(1-\delta)})$

Navon and Samorodnitsky [5] established a simpler proof of the same bound, $R_{LP}(\delta) \leq H(\frac{1}{2} - \sqrt{\delta(1-\delta)})$. We will present this proof below, which uses basic properties of Fourier analysis on $\mathbb{F}_2^n$.

5.2.1 Fourier analysis on $\mathbb{Z}_2^n$

In this section, we present some preliminaries on the Fourier analysis on $\mathbb{Z}_2^n$. To begin, we can think of $\mathbb{Z}_2^n$ in multiple ways. Instead of thinking of $\mathbb{Z}_2$ as the group $\{0,1\}$ under addition, we can think of $\mathbb{Z}_2$ as the group $\{-1,1\}$ under multiplication. In addition, since $\mathbb{Z}_2$ is boolean, we can think of $\mathbb{Z}_2$ as the set of subsets of $[n]$. For each $T \in \{-1,1\}$, we will also attribute to T the subset of $[n]$ that contains all elements i for which $T_i = -1$. We will be using the $\{-1,1\}^n$ and the subset definitions simultaneously for elements of $\mathbb{Z}_2^n$.

For each $T \in \mathbb{Z}_2^n$, define the function W_T from $\mathbb{Z}_2$ to the set of functions on n variables as $W_T(x_1, \ldots, x_n) = \prod_{i \in T} x_i$. In other words, W_T is the monomial composed of the x_i's for which $T_i = -1$. From here, it is easy to see that if $S \in \mathbb{Z}_2^n$, then $W_T(S) = (-1)^{|T \cap S|}$.

We would like to introduce a measure on each of the functions W_T. In particular, we consider the measure $\mu(W_T) = \mathbb{E}_{S \sim \mathbb{Z}_2^n}[W_T(S)]$, which is the expected value of $W_T(S)$ for a uniformly random $S \in \mathbb{Z}_2^n$. It is easy to see that $\mu(W_T) = 1$ for $T = \emptyset$, and 0 otherwise. In addition, observe that $\mu(W_S W_T) = \mu(W_{S \oplus T})$, because any squared term x_i^2 in the product becomes 1 when x_i is restricted to $\{-1,1\}$. This means that $\mu(W_S W_T) = 1$ if $S = T$, and 0 otherwise. Therefore, with respect to the measure μ, the functions W_T are orthonormal. In particular, they form an orthonormal basis for all functions from $\mathbb{Z}_2^n$ to $\mathbb{R}$.

For each $f : \mathbb{Z}_2^n \to \mathbb{R}$, define $\widehat{f} : \mathbb{Z}_2^n \to \mathbb{R}$ so that $\widehat{f}(T)$ equals the coefficient of the monomial with elements in T in the expansion of $f(x)$. From the orthogonal properties of W_T, we can observe that $\widehat{f}(T) = \frac{1}{2^n} \sum_{x \in \mathbb{Z}_2^n} f(x) W_T(x)$. The function $\widehat{f}$ is the Fourier transform of f.

Finally, we will use two well-known properties of the Fourier transform. First is Parseval's identity, $\frac{1}{2^n} \sum_{x \in \mathbb{Z}_2^n} f(x)g(x) = \sum_{S \in \mathbb{Z}_2^n} \widehat{f}(S)\widehat{g}(S)$. Next, for two functions $f, g : \mathbb{Z}_2^n \to \mathbb{R}$, let $h = f * g$ be their convolution, defined as $h(x) = \frac{1}{2^n} \sum_{y \in \mathbb{Z}_2^n} f(y)g(x \oplus y)$. Then, $\widehat{h} = \widehat{f} \cdot \widehat{g}$.

5.2.2 Proof of $R_{LP}(\delta) \leq H(\frac{1}{2} - \sqrt{\delta(1-\delta)})$

Define $L_1 : \mathbb{Z}_2^n \to \mathbb{R}$ so that $L_1(S) = 1$ if $|S| = 1$, and 0 otherwise. Then, define the function $\Delta : \mathbb{Z}_2^n \to \mathbb{R}$ as $\Delta = 2^n L_1$. Observe that $\widehat{\Delta}(S) = n - 2|S|$. Note that for any $f : \mathbb{Z}_2^n \to \mathbb{R}$, the function $f * \Delta$ evaluated at x returns the sum of the values of f at all points at distance 1 from x.

First, we show that any nonnegative function Γ with a nonnegative Fourier transform provides an upper bound on the cardinalities of codes. Note that here, we will consider the $\{0,1\}$ additive representation of $\mathbb{Z}_2$.

Lemma 1. *Let Γ be a function on $\{0,1\}^n$ with $\Gamma \geq 0$, $\widehat{\Gamma} \geq 0$, and assume that for some $0 < \lambda < 1$, we have $\Gamma * \Delta \geq \lambda n \cdot \Gamma$. Then, if C is a code with distance d such that*

$$d > \frac{1-\lambda}{2} \cdot n,$$

then we have

$$|C| \leq \frac{2^n}{2d/n - (1-\lambda)} \cdot \frac{\widehat{\Gamma}(0)}{\Gamma(0)}.$$

Proof. Let $\delta = \frac{d}{n} = \frac{1-\lambda}{2} + \epsilon$.

Let $f = \mathbb{1}_C$, the indicator function for C. Let $F = \frac{2^n}{|C|} f * f$, and note that $F(x) = 0$ for $0 < |x| < d$. Let

g be the function such that $\widehat{g} = F$. It can be verified that $g = \frac{2^{2n}}{|C|}\widehat{f}^2$, so that $g \geq 0$. We compute the inner product $\langle \Gamma * \Delta, g \rangle = \sum_{S \in \mathbb{Z}_2^n} (\Gamma * \Delta)(S) g(S)$ in two ways.

On one hand,

$$\langle \Gamma * \Delta, g \rangle \geq \lambda n \cdot \langle \Gamma, g \rangle = \lambda n \cdot \sum_S \widehat{\Gamma}(S) F(S).$$

On the other hand,

$$\langle \Gamma * \Delta, g \rangle = \langle \widehat{\Gamma} \cdot (n - 2|S|), F \rangle = \sum_S \widehat{\Gamma}(S) F(S)(n - 2|S|)$$

$$\leq n\widehat{\Gamma}(0)F(0) + \sum_{|S| \geq d} \widehat{\Gamma}(S)F(S)(n - 2|S|) \leq n\widehat{\Gamma}(0)F(0) + (n - 2d) \cdot \sum_S \widehat{\Gamma}(S)F(S)$$

$$= n\widehat{\Gamma}(0) + (\lambda - 2\epsilon)n \cdot \sum_S \widehat{\Gamma}(S)F(S).$$

Combining the two computations, we get

$$2\epsilon \langle \Gamma, g \rangle \leq \widehat{\Gamma}(0).$$

Since

$$\langle \Gamma, g \rangle \geq \frac{1}{2^n}\Gamma(0)g(0) = \frac{1}{2^n}|C|\Gamma(0),$$

we have

$$|C| \leq \frac{2^n}{2\epsilon} \cdot \frac{\widehat{\Gamma}(0)}{\Gamma(0)}.$$

□

Next, we show how to construct a function Γ that satisfies the conditions of the lemma.

Lemma 2. *Let Λ be a function on $\{0,1\}^n$ with $\Lambda \geq 0$, and assume that for some $0 < \lambda < 1$ we have $\Lambda * \Delta \geq \lambda n \cdot \Lambda$. Let $\Gamma = \Lambda * \Lambda$. Then, Γ satisfies the conditions of lemma 1 with the same λ. Moreover,*

$$\frac{\widehat{\Gamma}(0)}{\Gamma(0)} \leq \frac{|\mathrm{supp}(\Lambda)|}{2^n}.$$

Proof. We have $\widehat{\Gamma} = \widehat{\Lambda}^2 \geq 0$, which means $\Gamma \geq 0$ as well. Also, $\Gamma * \Delta = \Lambda * (\Lambda * \Delta) \geq \lambda n \Lambda * \Lambda = \lambda n \Gamma$, so the conditions of lemma 1 are satisfied.

For the second condition of lemma 2, we have, by the Cauchy-Schwarz inequality,

$$\widehat{\Gamma}(0) = \widehat{\Lambda}^2(0) = \frac{1}{2^{2n}} \left(\sum_x \Lambda(x) \right)^2 \leq \frac{1}{2^{2n}} \cdot |\mathrm{supp}(\Lambda)| \cdot \sum_x \Lambda^2(x)$$

$$= \frac{1}{2^n} \cdot |\mathrm{supp}(\Lambda)| \cdot \frac{1}{2^n} \sum_x \Lambda(x)\Lambda(0 \oplus x) = \frac{1}{2^n} \cdot |\mathrm{supp}(\Lambda)| \cdot \Gamma(0).$$

□

It remains to construct a function Λ. This function will be symmetric, in that its value at a point only depends on the Hamming weight at the point. Therefore, we can think of Λ as taking in the integers from 0 to n. Note that to determine a function Λ, it suffices to determine $\Lambda(0), \ldots, \Lambda(n)$, which is the approach we will take.

Note that a symmetric function f taking on the integers from 0 to n satisfies $(f * \Delta)(r) = rf(r-1) + (n-r)f(r+1)$, where we take $f(-1) = f(n+1) = 0$. Motivated by this recurrence, we start by choosing $f(-1) = 0$ and $f(0) = 1$, and defining $f(r)$ for $1 \leq r \leq n$ so that the relation $\lambda n f(r) = rf(r-1) + (n-r)f(r+1)$ holds for all r. It turns out that this function becomes negative at some point in the interval. To more precisely determine when the function becomes negative, we use the following lemma, whose proof is omitted because it is rather technical. (The proof of the lemma can be found in the appendix of the corresponding paper.)

Lemma 3. *Let $0 < \lambda < 1$ and let $\epsilon > 0$ be arbitrarily small. There exists a sufficiently large $n_0 = n_0(\lambda, \epsilon)$ such that for any $n > n_0$, the function f defined above becomes negative in the interval $\left[0, \dfrac{1-\sqrt{1-\lambda^2}}{2}(1+\epsilon) \cdot n\right]$.*

Let r_0 be the such that f is positive on $[0, r_0]$ and non-positive at $r_0 + 1$. Define $\Lambda = f$ for $r \in [0, r_0]$ and $\Lambda = 0$ for $r > r_0$. We claim that this function works.

Lemma 4. *The function Λ satisfies the conditions for lemma 2.*

Proof. We need to check that $(\Lambda * \Delta)(r) \geq \lambda n \cdot \Lambda(r)$ for all $0 \leq r \leq n$. We know that $(\Lambda * \Delta)(r) = \lambda n \Lambda(r)$ for $r \leq r_0 - 1$ and for $r > r_0 + 1$ by definition of Λ. It remains to verify the inequality for $r = r_0$ and $r = r_0 + 1$. For $r = r_0 + 1$, we have $(\Lambda * \Delta)(r) \geq 0 = \lambda n \Lambda(r)$. For $r = r_0$, we have

$$(\Lambda * \Delta)(r) = r\Lambda(r-1) + (n-r)\Lambda(r+1) = r\Lambda(r-1) = rf(r-1)$$

$$\geq rf(r-1) + (n-r)f(r+1) = \lambda n f(r) = \lambda n \Lambda(r),$$

where the inequality above holds because $f(r+1) \leq 0$.

□

Armed with this function Λ, we are ready to prove the bound.

Fix a distance parameter $\delta < \dfrac{1}{2}$, so that $d = \delta n$. Choose $\lambda = 1 - \dfrac{2d-2}{n}$, and let Λ be defined according to this λ. Let $\Gamma = \Lambda * \Lambda$. Then, by the preceding lemmas, we have, for sufficiently large n, every code C in $\mathbb{F}_2^n$ with distance δn satisfies

$$|C| \leq \frac{2^n}{2d/n - (1-\lambda)} \cdot \frac{\widehat{\Gamma}(0)}{\Gamma(0)} = \frac{2^n}{2/n} \cdot \frac{\widehat{\Gamma}(0)}{\Gamma(0)} \leq \frac{n}{2} \cdot |\mathrm{supp}(\Lambda)| \leq \frac{n}{2} \cdot \sum_{i=0}^{r_0} \binom{n}{i} \leq \frac{n}{2} \cdot 2^{n \cdot H(r_0/n)}.$$

Therefore, asymptotically, $R_{LP}(\delta) \leq H\left(\dfrac{r_0}{n}\right) \leq H\left(\dfrac{1-\sqrt{1-\lambda^2}}{2}\right)(1+\epsilon)$. Since $\lambda = 1 - \dfrac{2\delta n - 2}{n} \to 1 - 2\delta$, we have $R_{LP}(\delta) \leq H\left(\dfrac{1-\sqrt{1-(1-2\delta)^2}}{2}\right)(1+\epsilon) = H\left(\dfrac{1}{2} - \sqrt{\delta(1-\delta)}\right)(1+\epsilon)$. This holds for any $\epsilon > 0$, so we have the desired $R_{LP}(\delta) \leq H\left(\dfrac{1}{2} - \sqrt{\delta(1-\delta)}\right)$. □

5.3 Lower bound

We might also be interested in a lower bound for $R_{LP}(\delta)$, which gives an accurate measure of how powerful the LP bound actually is. It is essentially a cap on the strength of the bound.

Navon and Samorodnitsky [5] showed the lower bound $R_{LP}(\delta) \geq \frac{1}{2}H(1 - 2\sqrt{\delta(1-\delta)})$, which improved upon the previous best known bound. Their proof for the lower bound also uses Fourier analysis on $\mathbb{Z}_2^n$, but is a lot more complicated, so we omit it in this report.

Here is a plot of the lower bound with the upper bound.

6 Open problems

6.1 Improved asymptotic bounds

The major success of Delsarte's linear programming bound is the asymptotic upper bounds for the rate of binary codes, which are the best known for $\delta \geq 0.273$. Therefore, any improvement to the upper bound with δ in this range improves upon the best known upper bound for all binary codes.

However, Navon and Samorodnitsky [5] cite some numerical results from [1] that suggest that the bound $H\left(\frac{1}{2} - \sqrt{\delta(1-\delta}\right)$ might actually be tight. That is, they believe that $R_{LP}(\delta) = H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right)$. Therefore, there is work to be done on increasing the lower bound, to determine if the upper bound is indeed tight.

6.2 Generalizing asymptotics to $q > 2$

Note that for the asymptotic upper and lower bounds for $R_{LP}(\delta)$, we assumed that the codes are binary. Unfortunately, many nice properties of Fourier analysis on $\mathbb{Z}_2^n$ do not generalize in larger alphabets, so the techniques used in proving these bounds do not generalize, either.

However, Delsarte's linear program holds for all q. Therefore, an open question remains to show good asymptotic bounds for $R_{LP}(\delta)$ for $q > 2$.

References

[1] Barg, A., Jaffe D.B.: Numerical results on the asymptotic rate of binary codes, Codes and Association Schemes, AMS, 2001

[2] Delsarte, P.: An algebraic approach to the association schemes of coding theory. Philips Res. Rep., Suppl. 10 (1973)

[3] McEliece, R.J., Rodemich, E.R., Rumsey, H. Jr., Welch, L.R.: New upper bounds on the rate of a code via the DelsarteMacWilliams inequalities. IEEE Trans. Inf. Theory IT-23, 157166 (1977) of FOCS 46

[4] McKinley, S.: The Hamming Codes and Delsarte's Linear Programming Bound, available at http://www.mth.pdx.edu/ caughman/thesis.pdf

[5] Navon, M., Samorodnitsky, A.: On Delsarte's linear programming bounds for binary codes. In: Proceedings of FOCS 46