On the optimum of Delsarte's linear program

Alex Samorodnitsky *

Abstract

We are interested in the maximal size $A(n,d)$ of a binary error-correcting code of length n and distance d, or, alternatively, in the best packing of balls of radius $(d-1)/2$ in the n-dimensional Hamming space.

The best known lower bound on $A(n,d)$ is due to Gilbert and Varshamov, and is obtained by a covering argument. The best know upper bound is due to McEliece, Rodemich, Rumsey and Welch, and is obtained using Delsarte's linear programming approach. It is not known, whether this is the best possible bound one can obtain from Delsarte's linear program.

We show that the optimal upper bound obtainable from Delsarte's linear program will strictly exceed the Gilbert-Varshamov lower bound. In fact, it will be at least as big as the average of the Gilbert-Varshamov bound and the McEliece, Rodemich, Rumsey and Welch upper bound.

Similar results hold for constant weight binary codes.

The average of the Gilbert-Varshamov bound and the McEliece, Rodemich, Rumsey and Welch upper bound might be the true value of Delsarte's bound. We provide some evidence for this conjecture.

*Institute for Advanced Study, Princeton, NJ 08540. E-mail: asamor@ias.edu. This work was done while the author was a student in the Hebrew University of Jerusalem, Israel.

1 Introduction

A *binary error-correcting code C of length n and distance d* is a subset of the Hamming cube $\{0,1\}^n$ such that $\|x-y\| \geq d$ for all $x \neq y \in C$. Here $\|x-y\| := \sum_{i=1}^{n} |x_i - y_i|$ is the Hamming distance between x and y.

Efficient and error-resistant communication channels require large codes with large minimal distance. This leads to one of the main problems in combinatorial coding theory, which is to determine the maximal size $A(n,d)$ of an error-correcting code C of length n and distance d. One is also interested in the asymptotic behaviour of $A(n,d)$, when $n \to \infty$ and d is proportional to n, namely $d = \delta n$, for $0 \leq \delta \leq 1$. The question then is to estimate the *asymptotic maximal rate* $R(\delta)$ of an error-correcting code with relative distance δ:

$$R(\delta) = \sup_{d_n} \quad \limsup_{n\to\infty} \frac{1}{n} \log_2 A(n, d_n),$$

where the supremum is over all the sequences d_n with $\frac{d_n}{n} \to \delta$.

An important subclass of binary codes are codes in which all the codewords have the same Hamming weight. *A constant weight* binary error-correcting code C of length n, weight w and distance d is a subset of the Hamming sphere $S(n,w) = \{x \in \{0,1\}^n : \|x\| = \sum_{i=1}^{n} x_i = w\}$, such that $\|x-y\| \geq d$ for all $x \neq y \in C$. Now the problem is to determine the maximal size $A(n,d,w)$ of an error-correcting code C of length n, weight w and distance d. The asymptotic maximal rate $R(\delta,\xi)$ of an error-correcting code with relative weight ξ and relative distance δ for $0 \leq \delta \leq 2\xi \leq 1$ is defined as:

$$R(\delta,\xi) = \sup_{d_n,w_n} \quad \limsup_{n\to\infty} \frac{1}{n} \log_2 A(n, d_n, w_n),$$

where the supremum is over all the sequences w_n, d_n with $\frac{d_n}{n} \to \delta$, $\frac{w_n}{n} \to \xi$.

This paper deals with bounds on $R(\delta)$ and $R(\delta,\xi)$. We start with a review of the known bounds.

The classical bounds on $R(\delta)$ and $R(\delta,\xi)$ are based on the observation that the Hamming cube and the Hamming sphere, equipped with the Hamming metric, are transitive metric spaces. Namely, for any two points x, y there is an isometry of the space taking x to y. In particular, the size of the ball of radius d centered at any point of the metric space $(\{0,1\}^n, \|\cdot\|)$ is that of the ball centered at zero, which is $\sum_{i=0}^{d} \binom{n}{i}$. The usual packing and covering arguments [10] now give the Gilbert-Varshamov lower bound [10]:

$$A(n,d) \geq \frac{2^n}{\sum_{i=0}^{d-1} \binom{n}{i}}.$$

and the Hamming upper bound [10]:

$$A(n,d) \leq \frac{2^n}{\sum_{i=0}^{\frac{d-1}{2}} \binom{n}{i}}.$$

The standard estimates for the binomial distribution give the corresponding bounds for $R(\delta)$ in the interval $0 \le \delta \le \frac{1}{2}$:

$$1 - H(\delta) \le R(\delta) \le 1 - H\left(\frac{\delta}{2}\right), \tag{1}$$

where $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function. It is also known [10] that $R(\delta) = 0$ for $\delta \ge \frac{1}{2}$.

A ball of radius d centered at any point of the metric space $(S(n,w), \|\cdot\|)$ is of size $\sum_{i=0}^{\frac{d}{2}} \binom{w}{i}\binom{n-w}{i}$. This leads to upper and lower bounds on $A(n,d,w)$. We will only need a following counterpart for constant weight codes of the Gilbert-Varshamov lower bound [6]:

$$R(\delta,\xi) \ge H(\xi) - \xi H\left(\frac{\delta}{2\xi}\right) - (1-\xi)H\left(\frac{\delta}{2(1-\xi)}\right), \tag{2}$$

where $\xi \le \frac{1}{2}$ and $0 \le \delta \le 2\xi(1-\xi)$. It is known [10] that $R(\delta,\xi) = 0$ for $\delta \ge 2\xi(1-\xi)$.

The Bassalygo-Elias inequality [3] establishes a connection between $A(n,d)$ and $A(n,d,w)$:

$$A(n,d) \le A(n,d,w) \cdot \frac{2^n}{\binom{n}{w}} \tag{3}$$

The best known upper bounds on $R(\delta), R(\delta,\xi)$ are due to McEliece, Rodemich, Rumsey and Welch [11] (see also [8]), and are obtained using Delsarte's linear programming approach. The starting point of this approach is that both the Hamming cube and the Hamming sphere, equipped with the Hamming metric, are in fact *doubly transitive* metric spaces. This means that for any two pairs of points x, y and x_1, y_1 with $\|x - y\| = \|x_1 - y_1\|$ there is an isometry of the space taking x to x_1 and y to y_1. Let (X, D) be a finite metric space with $n+1$ distinct distances $0 = d_0 < d_1 < d_2 < ... < d_n$. Consider a partition of $X \times X$ into $n+1$ relations $R_i = \{(x,y) \,:\, D(x,y) = d_i\}$, for $i = 0...n$. It turns out that if (X, D) is doubly transitive then $(X, R_0...R_n)$ is a symmetric *association scheme*. (See [2], [14] chapter 30 for an introduction to association schemes. Section 2 also provides some information.) In particular, starting from the Hamming cube, one obtains the Hamming association scheme $H(n,2)$, and the Hamming sphere of radius w leads to the Johnson association scheme $J(n,w)$.

Let $(X, \mathcal{R})$, where $\mathcal{R} = (R_0, .., R_n)$ is a partition of $X \times X$ and R_0 is the identity matrix, be a symmetric association scheme with $n+1$ classes. Let Q be the *second eigenmatrix* of $(X, \mathcal{R})$. The *inner distribution* $(a_0, ..., a_n)$ of $Y \subseteq X$ is given by

$$a_k = \frac{|R_k \cap (Y \times Y)|}{|Y|}$$

for $k = 0, ..., n$. Delsarte has shown [3] that the inner distribution of any subset Y of X satisfies

$$\sum_{k=0}^{n} a_k Q_{k,s} \ge 0 \quad \text{for } s = 0, ..., n.$$

Therefore, if $C \subseteq X$ is an error-correcting code of distance d, its inner distribution $a_0 ... a_n$ satisfies a system of linear constraints:

$$a_0 = 1, \quad a_1 = ... = a_{d-1} = 0,$$

$$\sum_{k=0}^{n} a_k = |C|$$

and

$$\sum_{k=0}^{n} a_k Q_{k,s} \geq 0 \ \text{ for } s = 0, ..., n.$$

It follows that if the second eigenmatrix Q is known, one can obtain upper bounds on $|C|$ by solving an explicitly given linear program.

The second eigenmatrices of the Hamming and the Johnson association schemes were determined by Delsarte [3]. Let Q be the second eigenmatrix of the Hamming scheme. Then for all $0 \leq k, s \leq n$ holds $Q_{s,k} = K_s(k)$. Here K_s is the *Krawtchouk polynomial* of degree s. Krawtchouk polynomials $K_0 ... K_n$ are a classical family of orthogonal polynomials of a discrete variable.
The second eigenmatrix $Q_{s,k}$ of the Johnson scheme is given by another family of orthogonal polynomials of a discrete variable. For the Johnson scheme, $Q_{s,k} = H_s(k)$ for $s, k = 0...w$. Here $H_0 ... H_w$ are a family of *Hahn polynomials*. (See section 2 for more on Krawtchouk and Hahn polynomials).

This leads [3] to linear programming upper bounds on $A(n,d)$ and $A(n,d,w)$:

$$A(n,d) \leq max \left\{ \sum_{k=0}^{n} a_k | \ a_k \geq 0; \ a_0 = 1; a_k = 0, \ 1 \leq k < d; \ \sum_{k=0}^{n} a_k K_s(k) \geq 0, \ 0 \leq s \leq n \right\} \quad (4)$$

$$= min \left\{ \Lambda(0) | \ \Lambda = \sum_{s=0}^{n} b_s K_s; \ b_s \geq 0; \ b_0 = 1; \ \Lambda(i) \leq 0, \ d \leq i \leq n \right\} \quad (5)$$

and, assuming w.l.o.g. that d is even [1]

$$A(n,d,w) \leq max \left\{ \sum_{k=0}^{w} a_k | \ a_k \geq 0; \ a_0 = 1; a_k = 0, \ 1 \leq k < d/2; \ \sum_{k=0}^{w} a_k H_s(k) \geq 0, \ 0 \leq s \leq w \right\} \quad (6)$$

$$= min \left\{ \Lambda(0) | \ \Lambda = \sum_{s=0}^{w} b_s H_s; \ b_s \geq 0; \ b_0 = 1; \ \Lambda(i) \leq 0 \ , d/2 \leq i \leq w \right\}. \quad (7)$$

The equalities (4) = (5) and (6) = (7) follow from the duality theorem of linear programming [12].

[1] Since all the distances in the Hamming sphere are even

These are the *Delsarte linear programming bounds.* We will denote these bounds by $A_{LP}(n,d)$ and $A_{LP}(n,d,w)$. We also define, in analogy with $R(\delta)$, $R(\delta,\xi)$:

$$R_{LP}(\delta) = \sup_{d_n} \quad \limsup_{n\to\infty} \frac{1}{n}\log_2 A_{LP}(n,d_n)$$

over all sequences d_n, $\frac{d_n}{n} \to \delta$, and

$$R_{LP}(\delta,\xi) = \sup_{d_n,w_n} \quad \limsup_{n\to\infty} \frac{1}{n}\log_2 A_{LP}(n,d_n,w_n)$$

over all sequences d_n, w_n, $\frac{w_n}{n} \to \xi$, $\frac{d_n}{n} \to \delta$.

McEliece, Rodemich, Rumsey and Welch, using special properties of Krawtchouk and Hanh polynomials, constructed solutions Λ satisfying (5) (correspondingly (7)) with a small $\Lambda(0)$. This gave upper bounds on $R_{LP}(\delta), R_{LP}(\delta,\xi)$ for $0 \le \delta,\xi \le \frac{1}{2}$: [2]

$$R_{LP}(\delta) \le H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right) \tag{8}$$

$$R_{LP}(\delta,\xi) \le \begin{cases} 0 & \delta \ge 2\xi(1-\xi) \\ H\left(\frac{1}{2} - \sqrt{\frac{1}{4} - (\sqrt{\xi(1-\xi) - \frac{\delta}{2}(1-\frac{\delta}{2})} - \frac{\delta}{2})^2}\right) & 0 \le \delta < 2\xi(1-\xi) \end{cases} \tag{9}$$

We denote the first bound by $m(\delta)$ and the second bound by $m(\delta,\xi)$. These two bounds, together with (3), imply the following two bounds on $R(\delta)$.

$$R(\delta) \le m(\delta)$$

$$R(\delta) \le min_{\delta^* \le \xi \le \frac{1}{2}} \left(1 + m(\delta,\xi) - H(\xi)\right)$$

Here $\delta^* = \frac{1-\sqrt{1-2\delta}}{2}$. We denote the second bound by $M(\delta)$. It is not hard to see that $m(\delta,\frac{1}{2}) = m(\delta)$ and therefore $M(\delta) \le m(\delta)$. Surprisingly, the two bounds coincide for $0.273... \le \delta \le \frac{1}{2}$.

In conclusion, the best known upper and lower bounds on $R(\delta)$ are:

$$1 - H(\delta) \le R(\delta) \le M(\delta). \tag{10}$$

Unfortunately, the two bounds never coincide: $1-H(\delta) < M(\delta)$ for all $0 < \delta < \frac{1}{2}$. Consequently, it is natural to wonder whether Delsarte's approach can provide a better upper bound on $R(\delta)$. Specifically, one could ask whether the inequalities (8), (9) are tight.

A partial answer was given by Rodemich ([4], page 27). It turns out that from any solution of (6) implying $R_{LP}(\delta,\xi) \le \nu(\delta,\xi)$, one can construct a solution of (4) that shows $R_{LP}(\delta) \le 1 + \nu(\delta,\xi) - H(\xi)$. Therefore, for any $0 < \delta < \frac{1}{2}$ holds

$$R_{LP}(\delta) \le min_{\delta^* \le \xi \le \frac{1}{2}} \left(1 + R_{LP}(\delta,\xi) - H(\xi)\right). \tag{11}$$

[2] This explicit form of the bound (9) is from [9].

In particular $R_{LP}(\delta) \le M(\delta)$. This implies that (8) is not everywhere tight, since $M(\delta) < m(\delta)$ for $\delta < 0.273....$

However, no improvement for inequality (9) and therefore for the upper bound in (10) is known.

In this paper we give lower bounds on $R_{LP}(\delta)$, $R_{LP}(\delta, \xi)$. Our main result is

Theorem 1.1:

1. *For any* $0 \le \delta \le \frac{1}{2}$*:*

$$R_{LP}(\delta) \ge r(\delta) = \frac{(1 - H(\delta)) + H(\frac{1}{2} - \sqrt{\delta(1-\delta)})}{2}.$$

2. *For any* $0 \le \xi \le 1/2$*,* $0 \le \delta \le 2\xi(1-\xi)$*:*

$$R_{LP}(\delta, \xi) \ge r(\delta, \xi) = \frac{1}{2} m(\delta, \xi) + \frac{1}{2}\left[H(\xi) - \xi H\left(\frac{\delta}{2\xi}\right) - (1-\xi) H\left(\frac{\delta}{2(1-\xi)}\right)\right].$$

Remark 1.2: Curiously enough, in both cases our lower estimates are the arithmetic mean of the McEliece, Rodemich, Rumsey and Welch upper bounds and the "Gilbert-Varshamov" lower bounds (1), (2). ∎

Theorem 1.1 is proved in section 3.

Since $r(\delta) > 1 - H(\delta)$ for all $0 < \delta < \frac{1}{2}$, we conclude that Delsarte's linear programming approach can not close the gap between the upper and the lower bounds in (10).

Section 4 contains a somewhat informal discussion on the putative properties of the Delsarte bounds. We focus on some of the properties of the functions $R(\delta)$, $R(\delta, \xi)$ and conjecture that these properties are shared by $R_{LP}(\delta)$, $R_{LP}(\delta, \xi)$. We also point out that not all of these properties hold for the McEliece, Rodemich, Rumsey and Welch bounds. On the other hand, it turns out that the bounds $r(\delta)$, $r(\delta, \xi)$ of theorem 1.1 do have these properties. This leads us to conjecture that the inequality signs in theorem 1.1 should be replaced by equalities.

Main Conjecture 1.3:

$$R_{LP}(\delta) = r(\delta)$$
$$R_{LP}(\delta, \xi) = r(\delta, \xi)$$

∎

In the course of the discussion we make the following observation (see remark 4.5).

Lemma 1.4: *Let* $0 \le \delta \le \frac{1}{2}$*, and let* $\xi_{min} \in [\delta^*, \frac{1}{2}]$ *be a point at which the function* $m(\delta, \xi) - H(\xi)$ *attains its minimum (as a function of* ξ*). Then for any* $\xi_{min} \le \tau \le \frac{1}{2}$ *holds*

$$R(\delta, \tau) \le \min_{\delta^* \le \xi \le \frac{1}{2}} (m(\delta, \xi) - H(\xi)) + H(\tau). \tag{12}$$

The RHS of (12) does not exceed the Delsarte bound $R_{LP}(\delta, \tau)$, and on a certain set of points $\{(\delta, \tau)\} \subseteq \mathbf{R}^2$ it is strictly smaller. This observation is an immediate corollary of two known results [11], [6]. It is surprising that it apparently had not been made before.

2 Preliminaries

This section contains some of the definitions, terminology and facts that are required later on.

2.1 Association schemes [2], [14]

Let X be a finite set. A partition $\mathcal{R} = R_0...R_k$ of $X \times X$ into $k+1$ nonempty symmetric binary relations, such that R_0 is the identity relation, is a *symmetric association scheme* on X if it has the following property: there exist nonnegative integers $p^l_{i,j}$, $0 \le l, i, j \le k$ such that given any $(x, y) \in R_l$, there are exactly $p^l_{i,j}$ elements $z \in X$ such that $(x, z) \in R_i$ and $(y, z) \in R_j$. This implies that for any $x \in X$ and $0 \le i \le k$, the number $v_i(x)$ of $y \in X$ such that $(x, y) \in R_i$ does not depend on x. The numbers $v_0...v_k$ are the *valences* of the scheme. The numbers v_i satisfy $v_0 = 1$, $v_0 + ... + v_k = K$, where $K := |X|$.

The *adjacency matrices* of an association scheme are $K \times K$ matrices $A_0...A_k$ with rows and columns indexed by the elements of X. The matrices are defined as follows: $A_i(x, y) = 1$ if $(x, y) \in R_i$, otherwise $A_i(x, y) = 0$. It is easy to check that the adjacency matrices commute. It can be shown that there exists an orthogonal decomposition of $\mathbf{R}^K$, the space of real vectors whose coordinates are indexed by the elements of X, into a direct sum of $k+1$ subspaces $V_0...V_k$, such that $V_0...V_k$ are the eigenspaces of $A_0...A_k$, and V_0 is a one-dimensional subspace spanned by the vector of all 1's. Let $m_j = \dim V_j$. The numbers $m_0...m_k$ are the *multiplicities* of the scheme. They satisfy $m_0 = 1$, $m_0 + ... + m_k = K$.

The *first eigenmatrix* P of the scheme is a $(k + 1) \times (k + 1)$ matrix defined by setting P_{il} to be the eigenvalue of A_l corresponding to the eigenspace V_i. It is easy to see that P is invertible. We will be more interested in the *second eigenmatrix* Q of the scheme, defined by $Q = KP^{-1}$. It has the following properties: $Q_{i,0} \equiv 1$, $Q_{0,l} = m_l$, the columns of Q are orthogonal with respect to the inner product defined by the weights $v_0...v_k$, namely

$$\sum_{i=0}^{k} v_i Q_{il} Q_{is} = \begin{cases} Km_l & \text{if } l = s \\ 0 & \text{otherwise} \end{cases}$$

If there are polynomials $Q_0...Q_k$ such that deg $Q_i = i$ and $Q_{il} = Q_l(i)$ for $0 \le i, l \le k$, such polynomials will be called the *Q-polynomials* of the scheme. In particular, Q-polynomials are orthogonal with respect to the discrete probability measure $\frac{v_0}{K} ... \frac{v_k}{K}$ on $0...k$.

We will encounter three association schemes:

The Hamming scheme $H(n, 2)$

The set X is $\{0, 1\}^n$, so $K = 2^n$. There are n classes, $(x, y) \in R_i$ iff the Hamming distance between x and y is i. The valences and the multiplicities are given by $v_i = m_i = \binom{n}{i}$. The Q-polynomials for the scheme are the Krawtchouk polynomials $K_s(x)$ described in the next subsection.

The Halved Hamming scheme $I(n, 2)$

The elements of X are $0, 1$ vectors of length n with an even number of 1's. Therefore $K = 2^{n-1}$. There are $\lfloor \frac{n}{2} \rfloor$ classes, $(x, y) \in R_i$ iff the Hamming distance between x and y is $2i$. The valences are $v_i = \binom{n}{2i}$. The multiplicities m_j are $\binom{n}{j}$ for $j = 0...\lfloor \frac{n}{2} \rfloor - 1$. If n is odd then $m_{\lfloor \frac{n}{2} \rfloor} =$

$\binom{n}{\lfloor \frac{n}{2} \rfloor}$, otherwise $m_{\lfloor \frac{n}{2} \rfloor} = \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}$. The Q-polynomials for the scheme are modified Krawtchouk polynomials: $Q_s(x) = K_s(2x)$ for $s = 0...\lfloor \frac{n}{2} \rfloor - 1$. If n is odd then $Q_{\lfloor \frac{n}{2} \rfloor}(x) = K_{\lfloor \frac{n}{2} \rfloor}(2x)$, otherwise $Q_{\lfloor \frac{n}{2} \rfloor}(x) = \frac{1}{2}K_{\lfloor \frac{n}{2} \rfloor}(2x)$.

The Johnson scheme $J(n, w)$
The elements of X are $0, 1$ vectors of length n with precisely w ones. So $K = \binom{n}{w}$. There are w classes, $(x, y) \in R_i$ iff the Hamming distance between x and y is $2i$. The valences of the scheme are $v_i = \binom{w}{i}\binom{n-w}{i}$. The multiplicities are $m_0 = 1$, $m_j = \binom{n}{j} - \binom{n}{j-1}$ for $1 \leq j \leq w$. The Q-polynomials for the scheme are the Hahn polynomials $H_s(x)$, described in subsection 2.3.

2.2 Properties of Krawtchouk polynomials [11], [13]

Let $\mu_K(i) = \frac{\binom{n}{i}}{2^n}$, $0 \leq i \leq n$, be a discrete probability measure on $0...n$. The orthogonal polynomials associated with μ_K are the Krawtchouk polynomials $K_s^n(x)$, $s = 0...n$. (For the remainder of this subsection, and whenever possible later on, we will omit the superscript n.) This subsection collects some of the properties of these polynomials. The second claim of lemma 2.3 seems to be new.

Definition:

$$K_s(x) = \sum_{k=0}^{s} (-1)^k \binom{x}{k}\binom{n-x}{s-k}. \tag{13}$$

Corollary 2.1:

$$K_0 \equiv 1,\ K_s(0) = \binom{n}{s},\quad K_{n-s}(i) = \begin{cases} K_s(i) & \textit{if } i \textit{ is even} \\ -K_s(i) & \textit{if } i \textit{ is odd} \end{cases}$$

Corollary 2.2: *$K_s(\frac{n}{2} - x)$ is an even (odd) function of x if s is even (odd).*

Reciprocity:

$$\binom{n}{s} K_i(s) = \binom{n}{i} K_s(i). \tag{14}$$

Orthogonality:

$$\sum_{i=0}^{n} \mu_K(i) K_s(i) K_t(i) = \begin{cases} \binom{n}{s} & \text{s = t} \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

Difference equation:

$$(n - i)K_s(i + 1) - (n - 2s)K_s(i) + iK_s(i - 1) = 0. \tag{16}$$

First root behaviour, as n grows to infinity:

$$x_s = n \cdot \left(\frac{1}{2} - \sqrt{\frac{s}{n}(1 - \frac{s}{n})} \right) + o(n). \tag{17}$$

Monotonicity:

Lemma 2.3: *Let $\Delta = \lceil n^{\frac{4}{5}} + n^{\frac{1}{2}} \rceil$. Let s be such that $x_s > \Delta$. Then, assuming n is large enough*
(1) $\mu_K(i)K_s(i)$ is an increasing function of i in the interval $[0, x_s - \Delta]$.
(2) A stronger statement is also true: There is a constant $c > 0$ such that for any $i \in [0, x_s - \Delta]$

$$\left(1 + cn^{-\frac{2}{5}}\right) \cdot \mu_K(i)K_s(i) \leq \mu_K(i+1)K_s(i+1).$$

Proof: The first claim is known [8]. (See the proof of (22) below.) We omit the proof of the second claim, since it is quite similar to the proof of lemma 2.5, which will be given below. We also refer to Section 5 of [5] in which the ratio $\frac{K_s(i)}{K_s(i-1)}$ is determined for $0 \leq i \leq x_s - o(n)$. ∎

2.3 Properties of Hahn polynomials [11], [13]

For a pair (n, w), $n \geq 2w$ define a discrete probability measure μ_H on $0...w$ in the following way: $\mu_H(i) = \frac{\binom{w}{i}\binom{n-w}{i}}{\binom{n}{w}}$, $0 \leq i \leq w$. The orthogonal polynomials associated with μ_H are the Hahn polynomials $H_s(x)$, $s = 0...w$. This section describes properties of these polynomials. Lemma 2.5 appears to be new.

Definition:

$$H_s = \frac{w_s}{2^s} \sum_{k=0}^{s} \frac{\binom{s}{k}}{\binom{w}{k}\binom{n-w}{s-k}} K_k^w K_{s-k}^{n-w}, \tag{18}$$

where $w_s = \binom{n}{s} - \binom{n}{s-1}$.

Corollary 2.4: $H_0 \equiv 1$, $H_s(0) = w_s$.

Orthogonality:

$$\sum_{i=0}^{w} \mu_H(i) H_s(i) H_t(i) = \begin{cases} w_s & \text{s = t} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

Difference equation:

$$(w-i)(n-w-i)H_s(i+1) - (w(n-w) - i(n-2i) - s(n-s+1))\, H_s(i) + i^2 H_s(i-1) = 0 \tag{20}$$

First root behaviour as n, w grow to infinity:

$$x_s = n \cdot \frac{\frac{w}{n}(1 - \frac{w}{n}) - \frac{s}{n}(1 - \frac{s}{n})}{1 + 2\sqrt{\frac{s}{n}(1 - \frac{s}{n})}} + o(w). \tag{21}$$

Monotonicity:

Lemma 2.5: *Let $\Delta = \lceil w^{\frac{4}{5}} + w^{\frac{1}{2}} \rceil$. Let s be such that $x_s > \Delta$. Then, assuming w is large enough*
(1) $\mu_H(i)H_s(i)$ is an increasing function of i in the interval $[0, x_s - \Delta]$.
(2) A stronger statement is also true: There is a constant $c > 0$ such that for any $i \in [0, x_s - \Delta]$

$$\left(1 + cw^{-\frac{2}{5}}\right) \cdot \mu_H(i)H_s(i) \leq \mu_H(i+1)H_s(i+1).$$

The lemma is proved in the Appendix.

3 Proof of Theorem 1.1

We present two proofs for theorem 1.1. In the first subsection we construct a solution of the primal linear program (4) and prove the first part of the theorem. In the second subsection we work with Delsarte's dual linear program in a more general setting of association schemes. We analyze the solutions of the linear program for a certain class of association schemes. This includes the Hamming and the Johnson schemes. The results of the analysis for these two schemes provide the key to the proof of the theorem.

3.1 The primal approach.

Let n, d, $d \leq n/2$ be natural numbers. We assume w.l.o.g. that d is even. Set

$$\epsilon = \frac{1}{2n} \cdot \sqrt{\frac{\binom{n}{\lfloor x_d \rfloor}}{2^n \cdot \binom{n}{d}}},$$

where $x_s = x_s(n)$ is the first root of the Krawtchouk polynomial K_s.

Lemma 3.1: *Let numbers $a_0, ..., a_n$ be defined as follows:*

- $a_0 = 1;\ a_1 = ... = a_{d-1} = 0.$
- $a_d = (d+1) \cdot \binom{n}{d} \cdot \epsilon.$
- *For $d + 1 \leq k \leq n$, $a_k = \binom{n}{k} \cdot \epsilon$.*

Then:

1. $$\sum_{k=0}^{n} a_k \geq \frac{1}{4n} \cdot \sqrt{\frac{\binom{n}{\lfloor x_d \rfloor} 2^n}{\binom{n}{d}}}.$$

2. *For $0 \leq s \leq n$,*
$$\sum_{k=0}^{n} a_k K_s(k) \geq 0.$$

Recall (17), that $x_d = n \cdot \left(\frac{1}{2} - \sqrt{\frac{d}{n}(1-\frac{d}{n})} \right) + o(n)$. Therefore the lemma, together with the standard estimates on binomial coefficients, gives the first part of theorem 1.1.

Proof: 1 is immediate.

$$\sum_{k=0}^{n} a_k \geq \epsilon \sum_{k=d}^{n} \binom{n}{k} \geq \epsilon 2^{n-1} \geq \frac{1}{4n} \cdot \sqrt{\frac{\binom{n}{\lfloor x_d \rfloor} 2^n}{\binom{n}{d}}}.$$

It remains to prove 2. For $s = 0$ the claim is trivial. Assume $s \geq 1$. By the definition of a_k,

$$\sum_{k=0}^{n} a_k K_s(k) = K_s(0) + \epsilon \sum_{k=d}^{n} \binom{n}{k} K_s(k) + \epsilon d \cdot \binom{n}{d} K_s(d).$$

Krawtchouk polynomials K_s are orthogonal with respect to the measure $\mu_K(i) = \frac{\binom{n}{i}}{2^n}$, $0 \leq i \leq n$, implying

$$\sum_{k=d}^{n} \binom{n}{k} K_s(k) = \sum_{k=0}^{n} \binom{n}{k} K_s(k) - \sum_{k=0}^{d-1} \binom{n}{k} K_s(k) =$$

$$2^n \cdot \langle K_s, K_0 \rangle - \sum_{k=0}^{d-1} \binom{n}{k} K_s(k) = - \sum_{k=0}^{d-1} \binom{n}{k} K_s(k).$$

Here $<,>$ is the inner product with respect to μ_K. The last equality uses $s > 0$. Now, $K_s(0) = \binom{n}{s}$, and by the reciprocity property (14) of Krawtchouk polynomials: $\binom{n}{s} K_i(s) = \binom{n}{i} K_s(i)$. Consequently

$$\sum_{k=0}^{n} a_k K_s(k) = \binom{n}{s} - \epsilon \binom{n}{s} \sum_{k=0}^{d-1} K_k(s) + \epsilon \binom{n}{s} \cdot d K_d(s).$$

Therefore it suffices to show

$$1 - \epsilon \sum_{k=0}^{d-1} K_k(s) + \epsilon d K_d(s) \geq 0. \tag{22}$$

We prove (22) in two steps. First, we show that for s lying close enough to the endpoints of the interval $[0, n]$, the third summand in (22) is positive and is larger than the second summand. Then we will see that for the rest of the values of s, the last two summands in (22) are dominated by the first one.

We start with an identity ([8], (46)), which is valid for any integers $1 \leq s \leq n$ and for any real x:

$$K_s^n(x) - K_{s-1}^n(x) = K_s^{n+1}(x+1).$$

Recall that $\{K_s^m(x)\}_{s=0}^m$ with superscript m stands for the family of Krawtchouk polynomials orthogonal with respect to the measure $\mu_K^{(m)}(i) = \frac{\binom{m}{i}}{2^m}$ on $0...m$. Since $x_s(n+1) > x_s(n)$, and the sequence $x_s(n)$ decreases with s (see (67) in [8]), we have

$$K_d(s) > K_k(s) > 0 \quad \text{if } 0 \leq k \leq d-1 \text{ and } s \leq x_d - 1.$$

Therefore (22) holds for $s \le x_d - 1$. Now, by corollary 2.2, $K_k(\frac{n}{2} - x)$ is an even or an odd function of x, depending on the parity of k. We have taken d to be even, and therefore

$$K_d(s) > |K_k(s)| > 0 \ \text{ if } 0 \le k \le d-1 \text{ and } s \ge n - x_d + 1.$$

This implies (22) for $s \ge n - x_d + 1$. It remains to prove (22) for $s \in [x_d, n - x_d]$. By (15), $\frac{1}{2^n}\sum_{s=0}^{n}\binom{n}{s}K_k^2(s) = \binom{n}{k}$, implying

$$|K_k(s)| \le \sqrt{\frac{2^n\binom{n}{k}}{\binom{n}{s}}}.$$

Consequently, for $s \in [x_d, n - x_d]$ and $0 \le k \le d$,

$$|K_k(s)| \le \sqrt{\frac{2^n\binom{n}{d}}{\binom{n}{\lfloor x_d \rfloor}}}.$$

Therefore, for $s \in [x_d, n - x_d]$

$$1 - \epsilon\sum_{k=0}^{d-1}K_k(s) + \epsilon d K_d(s) \ge 1 - \epsilon\sum_{k=0}^{d-1}|K_k(s)| - \epsilon d|K_d|(s) > 1 - \epsilon \cdot 2n\sqrt{\frac{2^n\binom{n}{d}}{\binom{n}{\lfloor x_d \rfloor}}} \ge 0.$$

The last inequality follows from our choice of ϵ. ∎

3.2 The dual approach.

Let $(X, \mathcal{R})$ be a symmetric association scheme with k classes. Let $K := |X|$ be the size of the scheme, and let $v_0 ... v_k$ and $m_0 ... m_k$ be the valences and the multiplicities of the scheme.

We call $C \subseteq X$ *an error-correcting code of minimal distance* D, if $R_i \cap (C \times C) = \emptyset$ for $i = 1...D$. Delsarte [3] gave a linear programming upper bound on the maximal size of an error-correcting code in a scheme. This is the dual form of this bound:

$$|C| \le min\left\{\Lambda(0)|\Lambda(i) = \sum_{s=0}^{k} b_s Q_{i,s},\ i = 0, ..., k;\ b_s \ge 0;\ b_0 = 1;\ \Lambda(i) \le 0 \ for\ i = D...k\right\}. \tag{23}$$

Here $Q = (Q_{i,j})$ is the second eigenmatrix of the scheme. Note that bounds (5) and (7) are special cases of (23).

Assume that there exist *Q-polynomials* $Q_0 ... Q_k$, such that the degree of Q_s is s and $Q_s(i) = Q_{i,s}$ for $0 \le i, s \le k$. In particular, these polynomials are orthogonal with respect to the discrete probability measure $\frac{v_0}{K} ... \frac{v_k}{K}$ on the integer points $0...k$. Let x_s be the smallest root of Q_s. It is well known [13] that x_s is in $[0, k]$ and that $x_1 > x_2 > ... > x_k$. Since $Q_s(0) = m_s$ is positive, $Q_s(x) \ge 0$ for $0 \le x \le x_s$.

Assume also that the polynomials $\{Q_s\}_s$ satisfy the following monotonicity property: there exists an integer $\Delta > 0$ such that for any s with $x_s \ge \Delta$ and for any $i \in [0, x_s - \Delta]$ holds

$$v_i Q_s(i) \le v_{i+1} Q_s(i+1). \tag{24}$$

We define the *inverse root function* $r : [0, x_1 - \Delta] \to \{1...k\}$, by setting $r(x) = max\{1 \le s \le k \mid x_s \ge x + \Delta\}$. [3]

Now we are ready to formulate the main technical claim:

Lemma 3.2: *Let $1 \le D \le x_1 - \Delta$. In the above assumptions, the value of Delsarte's linear program (23) is at least*

$$\frac{1}{8k^2}\left(\frac{K \min_{i \ge r(D)} m_i}{\max_{j \le D} v_j}\right)^{\frac{1}{2}}. \tag{25}$$

Proof: Let $\Lambda = \sum_{s=0}^{k} b_s Q_s$ be a solution of (23). Let $\|\Lambda\|_1 = \frac{1}{K}\sum_{i=0}^{k} v_i |\Lambda(i)|$. We know two things about Λ. First

$$1 = \langle \Lambda, Q_0 \rangle = \frac{1}{K}\sum_{i=0}^{k} v_i \Lambda(i).$$

Second: $\Lambda(i) \le 0$ for $i = D...k$. Taken in conjunction, they imply that Λ is 'large' for some integer point j of the interval $[0, D-1]$. Specifically

$$\frac{v_j}{K}\Lambda(j) \ge \frac{\|\Lambda\|_1}{2k+2} \ge \frac{1}{2k+2}. \tag{26}$$

Write $\Lambda = \Lambda_1 + \Lambda_2$, where $\Lambda_1 = \sum_{s=0}^{r(D)} b_s Q_s$. We will show that either $|\Lambda_2(j)|$ or $|\Lambda_2(D)|$ is 'large'. Let $\ell \in \{j, D\}$ be such that $v_\ell |\Lambda_2(\ell)| = \max\{v_j|\Lambda_2(j)|, v_D|\Lambda_2(D)|\}$. Then we claim

$$\frac{v_\ell}{K}|\Lambda_2(\ell)| \ge \frac{1}{4k+4}. \tag{27}$$

If $\frac{v_j}{K}\Lambda_1(j) < \frac{1}{4k+4}$, then (27) follows from (26). Assume then that $\frac{v_j}{K}\Lambda_1(j) \ge \frac{1}{4k+4}$. By the definition of r, for all $1 \le s \le r(D)$ holds $x_s \ge D + \Delta$. Therefore by (24)

$$\frac{v_D}{K}\Lambda_1(D) = \frac{v_D}{K}\sum_{s=0}^{r(D)} b_s Q_s(D) \ge \frac{v_j}{K}\sum_{s=0}^{r(D)} b_s Q_s(j) = \frac{v_j}{K}\Lambda_1(j) \ge \frac{1}{4k+4}.$$

On the other hand, $\Lambda(D) \le 0$. Hence $\frac{v_D}{K}|\Lambda_2(D)| \ge \frac{1}{4k+4}$, proving (27).

Next, we show that if $\frac{v_\ell}{K}|\Lambda_2(\ell)|$ is large for $\ell \in [0, D]$ then $\Lambda(0)$ is large. Expand $\Lambda_2(\ell) = \sum_{s=r(D)+1}^{k} b_s Q_s(\ell)$.

Since $\Lambda(0) = \sum_{s=0}^{k} b_s Q_s(0) = \sum_{s=0}^{k} b_s m_s$, we get $b_s \le \frac{\Lambda(0)}{m_s}$. Since $\sum_{i=0}^{k} \frac{v_i}{K} Q_s^2(i) = m_s$, we have $|Q_s(\ell)| \le \sqrt{\frac{K m_s}{v_\ell}}$. Therefore

$$\frac{1}{4k+4} \le \frac{v_\ell}{K}|\Lambda_2(\ell)| \le \Lambda(0) \cdot \sqrt{\frac{v_\ell}{K}} \sum_{s=r(D)+1}^{k} m_s^{-\frac{1}{2}}, \tag{28}$$

[3] The function is well-defined for all $x \in [0, x_1 - \Delta]$, since $0 < x_k < 1$.

implying the claim of the lemma. ▮

We are ready to prove the second part of theorem 1.1. Fix $0 < \xi \leq \frac{1}{2}$ and $0 < \delta < 2\xi(1-\xi)$. Let $w = \xi n + o(n)$, $d = \delta n + o(n)$. Assume that d is even. We want to apply lemma 3.2 in order to obtain a lower bound on $A_{LP}(n, d, w)$. Indeed, (7) is a special case of (23) with $D = \frac{d}{2}$, and by lemma 2.5, the Hahn polynomials H_s satisfy (24) with $\Delta = \lceil w^{\frac{4}{5}} + w^{\frac{1}{2}} \rceil$. The parameters of the scheme are : $k = w$, $K = \binom{n}{w}$, $m_i = \binom{n}{i}$, $v_j = \binom{w}{j}\binom{n-w}{j}$. It remains to find the value of the inverse root function on $D = \frac{d}{2}$. The first root x_s of H_s is given by (21). The function $\phi(x) = \frac{\frac{w}{n}(1-\frac{w}{n}) - x(1-x)}{1+2\sqrt{x(1-x)}}$ is a decreasing mapping from $[0, \frac{w}{n}]$ to $[0, \frac{w}{n}(1-\frac{w}{n})]$. Its inverse is [9]:

$$\psi(x) = \frac{1}{2} - \sqrt{\frac{1}{4} - \left(\sqrt{\frac{w}{n}\left(1 - \frac{w}{n}\right) - x(1-x)} - x\right)^2}.$$

Therefore

$$r\left(\frac{d}{2}\right) = n \cdot \left[\frac{1}{2} - \sqrt{\frac{1}{4} - \left(\sqrt{\frac{w}{n}\left(1 - \frac{w}{n}\right) - \frac{\delta}{2}\left(1 - \frac{\delta}{2}\right)} - \frac{\delta}{2}\right)^2}\right] + o(n).$$

It is easy to check that the minimal value of the multiplicity m_i for $i \in [r\left(\frac{d}{2}\right), w]$ is attained at the left endpoint $r\left(\frac{d}{2}\right)$. The valences v_j increase for $j \in \left[0, \frac{d}{2}\right]$. Therefore

$$A(n, d, w) \geq \Omega\left(\frac{1}{n^2}\right) \cdot \sqrt{\frac{\binom{n}{w}\binom{n}{r\left(\frac{d}{2}\right)}}{\binom{w}{\frac{d}{2}}\binom{n-w}{\frac{d}{2}}}}$$

Substituting the value of $r\left(\frac{d}{2}\right)$, and using the standard estimates for the binomial distribution, we obtain the second claim of the theorem.

The first claim is slightly trickier. A straightforward application of lemma 3.2 for the Hamming scheme gives a trivial bound of $R_{LP}(\delta) \geq \frac{1-H(\delta)}{2}$. The problem lies in the fact that the multiplicities $m_i = \binom{n}{i}$ of the Hamming scheme $H(n, 2)$ decrease for $i \geq n/2$. In particular, $m_n = 1$.

To solve this problem, we bring in a different, but closely related association scheme - the scheme $I(n, 2)$ formed by vectors of even weight in $H(n, 2)$ (see section 2).

Recall that the scheme $I(n, 2)$ has $k = \lfloor \frac{n}{2} \rfloor$ classes.We will assume w.l.o.g. that n is odd. The parameters of the scheme are: $K = 2^{n-1}$, $v_i = \binom{n}{2i}$, $m_j = \binom{n}{j}$. The Q-polynomials for this scheme are $Q_s(i) = K_s(2i)$ for $0 \leq s, i \leq k$.

First, we link the Delsarte bounds for the two schemes.

Lemma 3.3: *Let $B_{LP}(n, D)$ be the solution of the linear program (23) for the scheme $I(n, 2)$ and distance D. Then*

$$B_{LP}(n, D) \leq A_{LP}(n, 2D).$$

Proof: Let $\Lambda_H = \sum_{s=0}^n b_s K_s$ be an optimal solution for (5) with $d = 2D$. We construct a solution Λ_I of (23) with $\Lambda_I(0) \leq \Lambda_H(0)$.

Set $\Lambda_I(i) = \frac{1}{b_0+b_n}\Lambda_H(2i)$ for $0 \leq i \leq k$. Then:

$$\langle \Lambda_I, Q_s \rangle = \sum_{i=0}^{k} \frac{v_i}{K} \Lambda_I(i) Q_s(i) = \frac{1}{b_0+b_n} \sum_{i=0}^{k} \frac{\binom{n}{2i}}{2^{n-1}} \Lambda_H(2i) K_s(2i) =$$

by corollary 2.1

$$\frac{1}{b_0+b_n} \sum_{i=0}^{n} \frac{\binom{n}{i}}{2^n} \Lambda_H(i)(K_s(i) + K_{n-s}(i)) = \frac{b_s + b_{n-s}}{b_0+b_n} \binom{n}{s}.$$

Therefore $\Lambda_I = \frac{1}{b_0+b_n} \sum_{s=0}^k (b_s + b_{n-s}) Q_s$. It satisfies the conditions of (23), and $\Lambda_I(0) \leq \Lambda_H(0)$, since $b_0 = 1$. ∎

Now we are ready to prove the first part of theorem 1.1. Fix $0 < \delta < \frac{1}{2}$. Let $d = \delta n + o(n)$. Assume that d is even. In order to obtain a lower bound on $A_{LP}(n,d)$, we apply lemma 3.2 to the scheme $I(n,2)$ with $D = \frac{d}{2}$. Note that by lemma 2.3, the Q-polynomials $Q_s(x) = K_s(2x)$ satisfy (24) with $\Delta = \lceil \frac{w^{\frac{4}{5}} + w^{\frac{1}{2}}}{2} \rceil$. The first root x_s of K_s is given by (17). The function $\phi(x) = \frac{1}{2} - \sqrt{x(1-x)}$ is a decreasing involution ($\phi \circ \phi(x) = x$) from $[0, \frac{1}{2}]$ to $[0, \frac{1}{2}]$. Therefore

$$r(D) = r\left(\frac{d}{2}\right) = n\left(\frac{1}{2} - \sqrt{\frac{d}{n}\left(1 - \frac{d}{n}\right)}\right) + o(n).$$

The multiplicities m_i increase for $i \in [0,k]$ and so do the valences v_j for $j \in [0, D]$. Therefore

$$A(n,d) \geq \Omega\left(\frac{1}{n^2}\right) \cdot \sqrt{\frac{2^n \binom{n}{r\left(\frac{d}{2}\right)}}{\binom{n}{d}}}$$

Substituting the value of $r\left(\frac{d}{2}\right)$, and applying the standard estimates for the binomial distribution, we conclude the proof of the first claim of the theorem. ∎

4 Discussion

The main goal of this section is to explain conjecture 1.3.

We view the rate functions $R(\delta)$ and $R(\delta, \xi)$ as an implicitly given pair of real functions, defined on an interval $[0, \frac{1}{2}]$ and on a square $[0, \frac{1}{2}] \times [0, \frac{1}{2}]$. These functions are known explicitly only in trivial 'boundary' cases. However, they are also known to satisfy several explicit properties, which reflect their 'geometric origins'.

We will compile a list of properties of the functions $R(\delta)$ and $R(\delta, \xi)$. This collection of properties will define a class $\mathcal{R}$ of functions $f(\delta)$, $f(\delta, \xi)$ with the same domain of definition, which share these properties. Next, we will check which of the bounds we have encountered so far, belong to this function class.

Recall that we have seen three bounds on $R(\delta)$ and $R(\delta,\xi)$. These are the "Gilbert-Varshamov" lower bound: $h(\delta)$ given by (1) and $h(\delta,\xi)$ given by (2); the Delsarte upper bound $R_{LP}(\delta)$ and $R_{LP}(\delta,\xi)$; the McEliece, Rodemich, Rumsey and Welch upper bound $m(\delta)$ and $m(\delta,\xi)$.

There is also the lower bound $r(\delta)$ and $r(\delta,\xi)$ on $R_{LP}(\delta)$ and $R_{LP}(\delta,\xi)$, given by theorem 1.1.

Among these, the Delsarte bound is an implicit one, defined as an optimum of a linear program. All the other bounds are explicit. The bounds are ordered: $h < r \leq R_{LP} \leq m$. The rate functions R lie between h and R_{LP}.

We will see that $r \in \mathcal{R}$, but $m \notin \mathcal{R}$. We won't be able to determine whether R_{LP} is in $\mathcal{R}$, but we will conjecture that it is. This, together with an observation that r seems to be a natural bound in the framework of association schemes, will lead us to conjecture 1.3, namely $R_{LP} = r$.

Before we define the class $\mathcal{R}$, let us focus on one of the properties of $R(\delta)$ and $R(\delta,\xi)$. This property will be crucial for the sake of this discussion, since it separates m from $\mathcal{R}$.

Consider the Bassalygo-Elias inequality (3), and let the radius w of the sphere decrease from $n/2$ to zero. Intuitively, the metric spaces $(S(n,w), \|\cdot\|)$ differ more and more from the whole Hamming space $(\{0,1\}^n, \|\cdot\|)$. Therefore, it seems reasonable to conclude that the information one can obtain on $A(n,d)$ through estimates on $A(n,d,w)$ and the Bassalygo-Elias inequality, should go down with w. This intuition is made precise by the following result [7], which we formulate in an asymptotic form:

Lemma 4.1: *$R(\delta,\xi) - H(\xi)$ is a non-increasing function of ξ in the interval $[0, \frac{1}{2}]$.*

Lemma 4.2: *The functions $R(\delta)$, $R(\delta,\xi)$ satisfy the following properties:*

- *Boundary values*
  (1) $R(0) = 1$, $R(\frac{1}{2}) = 0$.
  (2) $R(0,\xi) = H(\xi)$, $R(\delta,\frac{1}{2}) = R(\delta)$, $R(\delta,\xi) = 0$ for $0 \leq \xi \leq \frac{1-\sqrt{1-2\delta}}{2}$.
- *Monotonicity*
  (3) $R(\delta)$ is a non-increasing function of δ, $0 \leq \delta \leq \frac{1}{2}$.
  (4) For any $0 \leq \xi \leq \frac{1}{2}$:
  $R(\delta,\xi)$ is a non-increasing function of δ, $0 \leq \delta \leq \frac{1}{2}$.
  (5)For any $0 \leq \delta \leq \frac{1}{2}$:
  $R(\delta,\xi)$ is a non-decreasing function of ξ, $0 \leq \xi \leq \frac{1}{2}$.
- *(6) $R(\delta,\xi) - H(\xi)$ is a non-increasing function of ξ, $0 \leq \xi \leq \frac{1}{2}$.*

Proof: The only non-trivial property which remains to be verified is (5). It is proved in [6]. ∎

We define $\mathcal{R}$ to be the class of all real functions $f(\delta)$, $f(\delta,\xi)$, which share properties (1)-(6).

Now, consider the functions h, R_{LP}, m, r.

Proposition 4.3:
(a) The Gilbert-Varshamov bound h, given by

$$h(\delta) = 1 - H(\delta),$$

$$h(\delta,\xi) = \begin{cases} 0 & 0 \le \xi \le \frac{1-\sqrt{1-2\delta}}{2} \\ H(\xi) - \xi H(\frac{\delta}{2\xi}) - (1-\xi)H(\frac{\delta}{2(1-\xi)}) & \frac{1-\sqrt{1-2\delta}}{2} \le \xi \le \frac{1}{2} \end{cases}$$

is in $\mathcal{R}$.

(b) The bound m of McEliece, Rodemich, Rumsey and Welch is **not** *in $\mathcal{R}$.*

(c) The bound $r = \frac{h+m}{2}$ is in $\mathcal{R}$.

Proof:
(a) Properties (1)-(4) are immediate. Property (5) is proved in [6]. Property (6) follows from

$$\frac{\partial(h(\delta,\xi) - H(\xi))}{\partial\xi} = \log_2\left(1 - \frac{\delta(1-2\xi)}{\xi(2-2\xi-\delta)}\right)$$

for $\frac{1-\sqrt{1-2\delta}}{2} < \xi < \frac{1}{2}$.

(b) It is easy to see that properties (1)-(5) are satisfied. However, property (6) does not hold. Indeed, (2) and (6) would imply $min_{\delta^* \le \xi \le \frac{1}{2}}(1 + m(\delta,\xi) - H(\xi)) = m(\delta,\frac{1}{2}) = m(\delta)$, which is not true [11] for $\delta < 0.273....$

(c) Properties (1)-(5) are immediate, since they hold for h and m, and $r = \frac{h+m}{2}$. Property (6) is harder. It is given by the following lemma.

Lemma 4.4: *For any $0 \le \delta \le \frac{1}{2}$ the function*

$$r(\delta,\xi) - H(\xi) = \frac{1}{2}\left[m(\delta,\xi) - H(\xi) - \xi H\left(\frac{\delta}{2\xi}\right) - (1-\xi)H\left(\frac{\delta}{2(1-\xi)}\right)\right]$$

is a non-increasing function of ξ in the interval $[\frac{1-\sqrt{1-2\delta}}{2}, \frac{1}{2}]$.

The lemma is proved in the Appendix. ∎

Remark 4.5: Lemma 1.4 is an immediate consequence of (9) and lemma 4.1. The second part of proposition 4.3 implies that on a certain set of points $\{(\delta,\tau)\} \subseteq \mathbf{R}^2$, the RHS of (12) is strictly smaller than $R_{LP}(\delta,\tau)$. ∎

We are not able to determine whether $R_{LP} \in \mathcal{R}$. Properties (1)-(4) are easily verified. However, it is not clear whether properties (5) and (6) hold. We do conjecture that the Delsarte bounds are a sufficiently good approximation of $R(\delta)$, $R(\delta,\xi)$ in order to "inherit" these properties.

Conjecture 4.6: $R_{LP} \in \mathcal{R}$. ∎

This conjecture is partially validated by (11).

Let us consider once again the lower bounds $r(\delta)$, $r(\delta, \xi)$ on $R_{LP}(\delta)$, $R_{LP}(\delta, \xi)$. These bounds are special cases of (25) and therefore are naturally expressed by the parameters of the Hamming and Johnson association schemes. On the other hand, by proposition 4.3, $r \in \mathcal{R}$. This, taken together with conjecture 4.6, leads us to conjecture 1.3, namely that these bounds are in fact *the Delsarte bounds.*

Added in revision:
Recent numerical estimates [1] on $R_{LP}(\delta)$ strongly indicate that the McEliece, Rodemich, Rumsey and Welch bounds give (asymptotically) the right answer to the Delsarte problem for the Hamming cube, namely $R_{LP}(\delta) = M(\delta)$. In particular, conjecture 1.3 is probably false.

5 Appendix

5.1 Proof of lemma 2.5.

We start with an estimate on x_s which quantifies the $o(w)$ expression on the RHS of (21). The following claim is implicitly contained in [9].

Lemma 5.1: *For all $1 \leq s \leq w \leq \frac{n}{2}$ holds:*

$$|x_s - F| \leq \sqrt{w}, \tag{29}$$

where $F = n \cdot \frac{\frac{w}{n}\left(1-\frac{w}{n}\right) - \frac{s}{n}\left(1-\frac{s}{n}\right)}{1+2\sqrt{\frac{s}{n}\left(1-\frac{s}{n}\right)}}$.

Proof: (Sketch) Use lemma 5.15 of [9] with m_i and C defined by (5.51), taking $l = \sqrt{w}$ in (5.38). It is easy to check that this gives $|x_s - F| \leq \sqrt{w}$, unless $s \geq \frac{n}{2} - \frac{\sqrt{n}}{2}$. But in that case $F \leq 1$ and $s > \frac{1}{2}\left(n + 1 - \sqrt{(n+1)^2 - 4w(n-w)}\right)$. Therefore, by lemma 5.18, $x_s \leq 1$, completing the proof. ∎
In particular, $x_s \geq \Delta$ implies $F \geq w^{\frac{4}{5}}$, which implies $(w - s)(n - w - s) \geq nw^{\frac{4}{5}}$.

From now on we assume that $x_s \geq \Delta$, that $i \in [0, x_s - \Delta]$ and that w is large.

Next, similarly to [5, 11], we obtain a quadratic equation for the ratio $\rho = \rho(i) = \frac{H_s(i)}{H_s(i-1)}$. It is easy to see that $\rho(i+1) = \rho(i) \cdot (1 + \epsilon(i))$, where $\epsilon(i)$ is negative and small, $|\epsilon(i)| = O\left(\frac{w}{\Delta^2}\right)$. (We use the asymptotic O, Ω, o notation under the assumption $w \to \infty$). Let us write the difference equation (20) as

$$(w - i)(n - w - i) \cdot \frac{H_s(i+1)}{H_s(i)} \cdot \frac{H_s(i)}{H_s(i-1)} -$$

$$- ((n - w) - i(n - 2i) - s(n - s + 1)) \cdot \frac{H_s(i)}{H_s(i-1)} + i^2 = 0.$$

This gives

$$(w-i)(n-w-i)\cdot\rho^2\cdot(1+\epsilon)-(w(n-w)-i(n-2i)-s(n-s+1))\cdot\rho+i^2=0. \quad (30)$$

Here $\epsilon=\epsilon(i)$ is the "error term". Set $A=(w-i)(n-w-i)$, $B=w(n-w)-i(n-2i)-s(n-s+1)$, $C=i^2$. Then:

$$\rho=\frac{B\pm\sqrt{B^2-4AC(1+\epsilon)}}{2A(1+\epsilon)}. \quad (31)$$

In particular, $B>0$, and the discriminant $D^2:=B^2-4AC(1+\epsilon)$ is nonnegative. We will show that $\frac{\mu(i)H_s(i)}{\mu(i-1)H_s(i-1)}=\frac{\mu(i)}{\mu(i-1)}\cdot\rho(i)=1+\delta(i)$, with $\delta(i)\geq 1+\Omega\left(w^{-\frac{2}{5}}\right)$.

We start with estimating the discriminant. We have $D^2>B^2-4AC$ (recall $\epsilon<0$). The expression B^2-4AC is, as observed in [11], quadratic in i and can be written in the following form:

$$B^2-4AC=(n-2s)^2\cdot i^2-2n(w-s)(n-w-s)\cdot i+\ (w-s)^2(n-w-s)^2.$$

The roots of this quadratic are:

$$i_{1,2}=\frac{(w-s)(n-w-s)}{(n-2s)^2}\cdot\left(n\pm 2\sqrt{s(n-s)}\right)=n\cdot\frac{\frac{w}{n}(1-\frac{w}{n})-\frac{s}{n}(1-\frac{s}{n})}{1\pm 2\sqrt{\frac{s}{n}(1-\frac{s}{n})}}.$$

Note that $i_1=F$. Consequently, for $i\leq x_s-\Delta$,

$$D^2\geq(n-2s)^2(i_1-i)(i_2-i)\geq w^{\frac{4}{5}}\left((n-2s)^2w^{\frac{4}{5}}+4\sqrt{s(n-s)}(w-s)(n-w-s)\right)\geq\Omega(n^2w^{\frac{8}{5}}).$$

Now we can resolve the problem of the choice of sign in (31). Observe that the values $\rho(i)$ and $\rho(i-1)$ are very close, indeed $\frac{\rho(i-1)}{\rho(i)}=1+O\left(\frac{w}{\Delta^2}\right)=1+O\left(w^{-\frac{3}{5}}\right)$. On the other hand, $\frac{D}{B}=\Omega\left(\frac{nw^{\frac{4}{5}}}{nw}\right)\geq\Omega\left(w^{-\frac{1}{5}}\right)$. It follows that the choice of the sign must be uniform throughout the interval $i\in[1,x_s-\Delta]$. It is not hard to check that for $i=1$ the choice is "+" and therefore:

$$\rho=\rho(i)=\frac{B+\sqrt{B^2-4AC(1+\epsilon)}}{2A(1+\epsilon)}\geq\frac{B}{2A}.$$

Let us return to the ratio $\frac{\mu(i)}{\mu(i-1)}\cdot\rho(i)$. Note that $\frac{\mu(i)}{\mu(i-1)}=\frac{(w-i)(n-w-i)}{i^2}=\frac{A}{C}$. This implies that $\frac{\mu(i)}{\mu(i-1)}\cdot\rho(i)\geq\frac{B}{2C}$. We will conclude by showing that $\frac{B}{2C}\geq 1+\Omega\left(w^{-\frac{2}{5}}\right)$.

One checks easily that $i<x_s\leq x_1=\frac{w(n-w)}{n}$ implies $A\geq C$. Therefore $B^2-4C^2\geq B^2-4AC=\Omega(n^2w^{\frac{8}{5}})$. It follows $\frac{B}{2C}\geq 1+\Omega\left(\frac{n^2w^{\frac{8}{5}}}{2C(B+2C)}\right)=1+\Omega\left(w^{-\frac{2}{5}}\right)$. ∎

5.2 Proof of lemma 4.4

Consider a domain $D \subseteq \mathbf{R}^2$, $D = \left\{(\delta, \xi) \mid 0 \le \delta \le \frac{1}{2}; \frac{1-\sqrt{1-2\delta}}{2} \le \xi \le \frac{1}{2}\right\}$.
Let $f : D \to \mathbf{R}$ be given by:

$$f(\delta, \xi) = H(t(\delta, \xi)) - H(\xi) - \xi H\left(\frac{\delta}{2\xi}\right) - (1-\xi)H\left(\frac{\delta}{2(1-\xi)}\right),$$

where $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function, and $t : D \to \mathbf{R}$ is given by

$$t(\delta, \xi) = \frac{1}{2} - \sqrt{\frac{1}{4} - \left(\sqrt{\xi(1-\xi) - \frac{\delta}{2}\left(1 - \frac{\delta}{2}\right)} - \frac{\delta}{2}\right)^2}.$$

We have to prove that for any fixed $0 \le \delta \le \frac{1}{2}$, $f(\delta, \xi)$ is a non-increasing function of ξ.
For $\delta = 0$, $f(0, \xi) \equiv 0$, and the claim is valid.
Now let $\delta_0 > 0$. We have to show that for all $\frac{1-\sqrt{1-2\delta_0}}{2} \le \xi_1 \le \xi_2 \le \frac{1}{2}$ holds $f(\delta_0, \xi_1) \ge f(\delta_0, \xi_2)$.
Clearly it suffices to show

$$\frac{\partial}{\partial\delta} f(\delta, \xi_1) \ge \frac{\partial}{\partial\delta} f(\delta, \xi_2) \quad \text{for all } 0 \le \delta \le \delta_0. \tag{32}$$

From now on we verify (32). All the logarithms are to base 2.

By the definition of f:

$$\frac{\partial}{\partial\delta} f(\delta, \xi) = \frac{\partial}{\partial\delta} H(t(\delta, \xi)) - \log\left(\sqrt{\xi(1-\xi) - \frac{\delta}{2}\left(1 - \frac{\delta}{2}\right)}\right) + \log\left(\frac{\delta}{2}\right).$$

Let $k = k(\delta, \xi) = \sqrt{\xi(1-\xi) - \frac{\delta}{2}\left(1 - \frac{\delta}{2}\right)}$. Then $t = \frac{1}{2} - \sqrt{\frac{1}{4} - (k - \frac{\delta}{2})^2}$ and therefore,

$$\frac{\partial}{\partial\delta} t(\delta, \xi) = \frac{\left(k - \frac{\delta}{2}\right)\left(\frac{\partial k}{\partial\delta} - \frac{1}{2}\right)}{\sqrt{\frac{1}{4} - (k - \frac{\delta}{2})^2}}.$$

Recall that $\frac{dH(x)}{dx} = \log\left(\frac{1-x}{x}\right)$, implying $\frac{\partial H(t)}{\partial\delta} = \log\left(\frac{1-t}{t}\right)\frac{\partial t}{\partial\delta}$.
Observe also that $\frac{\partial k}{\partial\delta} = -\frac{\frac{1}{2}-\delta}{2k}$. Consequently

$$\frac{\partial}{\partial\delta} f(\delta, \xi) = -\frac{(2k-\delta)(2k-\delta+1)}{4k\sqrt{1-(2k-\delta)^2}} \cdot \log\left(\frac{1+\sqrt{1-(2k-\delta)^2}}{1-\sqrt{1-(2k-\delta)^2}}\right) - \log(k) + \log\left(\frac{\delta}{2}\right).$$

For a fixed δ, $k(\delta, \xi)$ is an increasing one to one function of ξ from $[\frac{1-\sqrt{1-2\delta}}{2}, \frac{1}{2}]$ to $[\frac{\delta}{2}, \frac{1-\delta}{2}]$.
Therefore, we may change variables, and consider RHS as a function of δ and k.
We have to show that RHS is non-increasing in ξ, $\xi \in [\frac{1-\sqrt{1-2\delta}}{2}, \frac{1}{2}]$, which is the same as to show RHS is non-increasing in k, $k \in [\frac{\delta}{2}, \frac{1-\delta}{2}]$.

For this purpose, let us introduce one additional change of variable: $m = 2k - \delta$. Then we are left with a following claim: For every $0 \le \delta \le \frac{1}{2}$ the function

$$v(m) = \frac{m(m+1)}{2(m+\delta)\sqrt{1-m^2}} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right) + \log(m+\delta) - \log\left(\frac{\delta}{2}\right)$$

is non-decreasing in m, for $m \in [0, 1-2\delta]$.

Considering the first summand as a product of two factors: $\frac{m(m+1)}{2(m+\delta)}$ and $\frac{1}{\sqrt{1-m^2}} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right)$, we obtain:

$$\frac{dv}{dm} = \left(\frac{1}{2} + \frac{\delta(1-\delta)}{2(m+\delta)^2}\right) \cdot \left(\frac{1}{\sqrt{1-m^2}} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right)\right) +$$

$$\frac{m(m+1)}{2(m+\delta)} \cdot \left(\frac{m}{(\sqrt{1-m^2})^3} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right) - \frac{2\log(e)}{m(1-m^2)}\right) + \frac{\log(e)}{m+\delta}.$$

We want to show $\frac{dv}{dm} \ge 0$. Multiplying by $2(m+\delta)$ and moving the last summand to the other side, we see that this is equivalent to:

$$\left((m+\delta) + \frac{\delta(1-\delta)}{(m+\delta)}\right) \cdot \left(\frac{1}{\sqrt{1-m^2}} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right)\right) +$$

$$m(m+1)\left(\frac{m}{(\sqrt{1-m^2})^3} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right) - \frac{2\log(e)}{m(1-m^2)}\right) \ge -2\log(e).$$

The function $(m+\delta) + \frac{\delta(1-\delta)}{(m+\delta)}$ increases in δ, so it is enough to take $\delta = 0$, arriving at

$$\frac{m}{\sqrt{1-m^2}} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right) + m(m+1)\left(\frac{m}{(\sqrt{1-m^2})^3} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right) - \frac{2\log(e)}{m(1-m^2)}\right) \ge -2\log(e).$$

Opening brackets and rearranging, we get the following equivalent inequalities:

$$\frac{m}{\sqrt{1-m^2}} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right) + \frac{m^2(m+1)}{(\sqrt{1-m^2})^3} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right) \ge -2\log(e) + \frac{2\log(e)}{1-m} = \frac{2\log(e)m}{1-m},$$

or

$$\frac{1}{\sqrt{1-m^2}} \cdot \log\left(\frac{1+\sqrt{1-m^2}}{1-\sqrt{1-m^2}}\right) \ge 2\log(e).$$

Setting $s = \sqrt{1-m^2}$, it remains to verify that for all $0 < s < 1$,

$$l(s) = \frac{1}{s}\log\left(\frac{1+s}{1-s}\right) \ge 2\log(e).$$

This is quite simple. At the endpoints of the interval $(0,1)$, $l(s)$ is going to infinity. The minimum of $l(s)$ is attained at the only zero s_0 of the derivative, which, easily, satisfies:

$$l(s_0) = \frac{1}{s_0}\log\left(\frac{1+s_0}{1-s_0}\right) = \frac{2\log(e)}{1-s_0^2} \ge 2\log(e),$$

and we are done. ∎

6 Acknowledgements

I am very grateful to my advisor Nathan Linial for his generous help and advice and for much needed encouragement and discouragement. Numerous comments and corrections by Alexander Barg and Vladimir Levenshtein significantly improved and simplified the presentation, especially in section 3.1. I would also like to thank Simon Litsyn and Yoav Kirsch for sharing their knowledge of coding theory and of graph theory.

References

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

[2] E. Bannai and T. Ito, **Algebraic Combinatorics I. Association Schemes**, London, UK: Benjamin/Cummings, 1984.

[3] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep., Suppl., vol. 10, 1973.

[4] P. Delsarte, *Application and generalization of the MacWilliams transform in coding theory*, Proc. 15th Sympos. Inform. Theory in the Benelux (1994), 9–44.

[5] G. Kalai and N. Linial, *On the distance distribution of codes*, IEEE Trans. Inform. Theory, vol. IT-41, 1995, 1467-1472.

[6] V. I. Levenshtein, *Upper-bound estimates for fixed-weight codes*, Problems of Info. Trans, 7, No. 4, 1971, 281-287.

[7] V. I. Levenshtein, *On the minimum redundancy of binary error-correcting codes*, Problems of Info. Trans, 10, No.2, 1974, 110–123; and Inform. and Control, 28, 1975, 268-291.

[8] V. I. Levenshtein, *Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces*, IEEE Trans. Inform. Theory, vol. IT-41, 1995, 1303-1321.

[9] V. I. Levenshtein, *Universal bounds for codes and designs*, in Handbook of Coding Theory, V.S. Pless and W.C. Huffman Eds., Amsterdam, Elsevier, 1998.

[10] J. MacWilliams and N. J. A. Sloane, **The Theory of Error Correcting Codes**, Amsterdam, North-Holland, 1977.

[11] R. J. McEliece, E. R. Rodemich, H. Rumsey, Jr., and L. R. Welch, *New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities*, IEEE Trans. Inform. Theory, vol. IT-23, 1977, 157-166.

[12] A. Schrijver, **Theory of Linear and Integer Programming**, John Wiley and sons, New York, 1986.

[13] G. Szegö, **Orthogonal Polynomials**, American Mathematical Society, 1939.

[14] J. H. van Lint and R. M. Wilson, **A Course in Combinatorics**, Cambridge University Press, 1992.