Extremal properties of solutions for Delsarte's linear program
Preliminary version

Alex Samorodnitsky [*]

Abstract

The best known upper bounds on on the rate of binary codes are due to McEliece, Rodemich, Rumsey and Welch, and are obtained using Delsarte's linear programming approach.

We deal with optimality of known solutions for Delsarte's program, such as McEliece, Rodemich, Rumsey and Welch polynomials and the Lloyd polynomials.

We prove that the Lloyd polynomials are optimal in the class of polynomials vanishing on the same integer subset of $[0, n]$.
As a consequence we obtain tight bounds on the Lovasz's θ-function of certain subgraphs of the Hamming cube.

We also show that the polynomials of McEliece, Rodemich, Rumsey and Welch have essentially the smallest degree for which the constraints of Delsarte can be satisfied.

Our approach is based on the following idea:
To analyze the behaviour of a polynomial P on the interval $[0, n]$ endowed with the discrete binomial measure $\mu(i) = \frac{\binom{n}{i}}{2^n}$, it is convenient to look at the representation of P in the basis of polynomials orthogonal with respect to μ^2.

Similar ideas may be applied in a more general setting, that is concerned with extremal properties of polynomials. We consider, for instance, the following question: Let P a polynomial of degree $\leq n$. How far from zero can the global maximum of the function $|P(x)|e^{-\frac{x^2}{2}}$ be? We show, that the answer is $(1 + o(1))\sqrt{2n}$.

[*]Hebrew University

1 Introduction

This paper deals with extremal properties of known solutions for Delsarte's program, such as McEliece, Rodemich, Rumsey and Welch polynomials and the Lloyd polynomials.

First, we briefly describe the background.
Let Z_2^n be a set of 0,1 vectors of length n. Let $d_H(x,y)$ denote the Hamming distance between vectors x and y. The metric space (Z_2^n, d_H) is the n-dimensional cube.

Let $Q_{n,d}$ be a graph whose vertex set is Z_2^n, where vectors x and y are adjacent iff $d_H(x,y) < d$.

One of the main problems in combinatorial coding theory is to estimate the maximum size of any subset or "code" C of the cube or the Hamming sphere such that no two elements of C have distance less than d. In graph notation it means estimating $\alpha(Q_{n,d})$, where $\alpha(G)$ stands for the largest size of independent subsets in a graph G.

Let us observe, that $Q_{n,d}$ is a regular graph with degree $\sum_{i=1}^{d-1}\binom{n}{i}$, and therefore

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

This is the Gilbert-Varshamov lower bound on $\alpha(Q_{n,d})$.

Equally simple is the Hamming upper bound on $\alpha(Q_{n,d})$: Let C be an independent subset of $Q_{n,d}$. Then the Hamming balls of radius $\lfloor\frac{d-1}{2}\rfloor$ centered at the points of C are disjoint sets. The size of each ball is $\sum_{i=0}^{\lfloor\frac{d-1}{2}\rfloor}\binom{n}{i}$, implying:

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

For $0 \leq \delta \leq \frac{1}{2}$ denote

$$R(\delta) = \limsup_{n\to\infty} \frac{1}{n}\log_2 \alpha(Q_{n,d_n})$$

for all sequences d_n, $\frac{d_n}{n} \to \delta$.

Gilbert-Varshamov and Hamming bounds imply:

$$1 - H(\delta) \leq R(\delta) \leq 1 - H(\frac{\delta}{2}). \tag{1}$$

$H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function.

The best known upper bounds on $R(\delta)$ are due to McEliece, Rodemich, Rumsey and Welch [9] (see also [6]) and are obtained using Delsarte's linear programming approach which we will now briefly describe.
Let $C \subseteq Z_2^n$. Let us define the inner distribution $(a_0, ..., a_n)$ of C by

$$a_k = \frac{|\{(x,y) \mid x,y \in C, \|x-y\| = k\}|}{|C|},$$

for $k = 0, ..., n$; so $a_0 = 1$ and $\sum_{k=0}^{n} a_k = |C|$. Delsarte has shown [3] that the inner distribution of any subset C of Z_2^n satisfies:

$$\sum_{k=0}^{n} a_k K_k(s) \geq 0,$$

for $s = 0, ..., n$. K_k are *Krawchouk polynomials* - a family orthogonal with respect to a discrete measure μ supported on the integer points $0, ..., n$, and defined by: $\mu(i) = \frac{\binom{n}{i}}{2^n}$.

Therefore,

$$\alpha(Q_{n,d}) \leq max\{\sum_{k=0}^{n} a_k |\, a_k \geq 0;\ a_0 = 1; a_k = 0,\ 1 \leq k < d;\ \sum_{k=0}^{n} a_k K_s(k) \geq 0,\ 0 \leq s \leq n\} \quad (2)$$

$$= min\{\Lambda(0) |\ \Lambda = \sum_{s=0}^{n} b_s K_s, b_s \geq 0;\ b_0 = 1;\ \Lambda(i) \leq 0,\ d \leq i \leq n\}. \quad (3)$$

The last equality follows from the duality theorem of linear programming [13]. This bound on $\alpha(Q_{n,d})$ is called Delsarte's linear programming bound. Following [11] we denote this bound by $\theta'(Q_{n,d})$. We also define, in analogy with $R(\delta)$, $R(\delta, \xi)$:

$$\theta'(\delta) = \limsup_{n \to \infty} \frac{1}{n} \log_2 \theta'(Q_{n,d_n})$$

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

McEliece, Rodemich, Rumsey and Welch [9], used special properties of Krawchouk polynomials, and succeeded in constructing kernels Λ satisfying (3) with a small $\Lambda(0)$, thus proving:

$$\theta'(\delta) \leq H(\frac{1}{2} - \sqrt{\delta(1 - \delta)}). \quad (4)$$

Using the basic fact, $R(\delta) \leq \theta'(\delta)$ we obtain the *first linear programming bound* on $R(\delta)$:

$$R(\delta) \leq H(\frac{1}{2} - \sqrt{\delta(1 - \delta)}).$$

However, this bound is not always better than the Hamming bound. Indeed, for small values of δ, $1 - H(\frac{\delta}{2}) < H(\frac{1}{2} - \sqrt{\delta(1 - \delta)})$.
The best known upper bound on $R(\delta)$ was also obtained in [9], using Delsarte's linear programming approach for constant weight codes. This is the so called *second linear programming bound*.

In this paper we are interested in the first bound. (Let us only add, that the two bounds coincide for $0.273... \leq \delta \leq \frac{1}{2}$.) Interestingly enough there exists a family of kernels, called the Lloyd polynomials [8], satisfying (3) and yielding:

$$\theta'(\delta) \leq 1 - H(\frac{\delta}{2}).$$

To the best of our knowledge, the McEliece, Rodemich, Rumsey and Welch polynomials and the Lloyd polynomials are the optimal (in the corresponding ranges) known solutions of (3).

We turn to a short description of our results. The optimality of McEliece, Rodemich, Rumsey and Welch kernels has been studied by Levenshtein [6]. He proves that they are indeed optimal, if one augments (3) by two additional constraints: Let t_d denote the degree of the MRRW kernel satisfying (3), then Levenshtein's additional constraints are:

- $\Lambda(x) \leq 0;\ d \leq x \leq n,$
- $deg\ \Lambda \leq t_d.$

Belov has shown that even without the first constraint the conclusion remains valid. We prove a related (somewhat surprising) statement:

Proposition 1.1: *There exists a function $\epsilon(n) = o(n)$ such that for any $0 \leq d \leq \frac{n}{2}$ no kernel Λ with $deg\Lambda \leq t_d - \epsilon(n)$, Λ can satisfy the constraints (3).*

In other words, the kernels of McEliece, Rodemich, Rumsey and Welch have essentially the smallest possible degree for which the constraints (3) can be satisfied.

Lloyd polynomials have a full degree n, and therefore do not contradict Belov's theorem. However, they do have a special property of their own: $\Lambda(i) = 0; d \leq i \leq n.$ We prove the optimality of these polynomials within this class.

Let us define in analogy with (2), (3)

$$\theta(Q_{n,d}) = max\{\sum_{k=0}^{n} a_k \mid a_0 = 1; a_k = 0\ for\ 1 \leq k < n;\ \sum_{k=0}^{n} a_k K_k(s) \geq 0,\ 0 \leq s \leq n\} \quad (5)$$

$$= min\{\Lambda(0) |\ \Lambda(i) = \sum_{s=0}^{n} b_s K_s(i); b_s \geq 0;\ b_0 = 1;\ \Lambda(i) = 0,\ d \leq i \leq n\}. \quad (6)$$

We also define

$$\theta(\delta) = \limsup_{n\to\infty} \frac{1}{n} \log_2 \theta(Q_{n,d_n})$$

for all sequences d_n, $\frac{d_n}{n} \to \delta$.

Then the following proposition holds:

Proposition 1.2:

$$\theta(\delta) = 1 - H(\frac{\delta}{2}).$$

The notation $\theta(Q_{n,d})$ introduced in (5) is by no means arbitrary. Indeed, Schrijver had shown ([11]) that for a subgraph G of a symmetric association scheme the Lovasz's θ-function bound on G's Shannon capacity of is given by the value of a linear program analogous to (5), (6). (This also explains the somewhat cumbersome choice of θ' as our notation for Delsarte's bound).

Let $\Theta(G)$ denote the Shannon capacity of G, and define

$$\Theta(\delta) = \limsup_{n\to\infty} \frac{1}{n} \log_2 \Theta(Q_{n,d_n})$$

for all sequences d_n, $\frac{d_n}{n} \to \delta$. We have shown:

$$\Theta(\delta) \leq 1 - H(\frac{\delta}{2}).$$

For any graph G, $\Theta(G) \geq \alpha(G)$, implying $\Theta(\delta) \geq R(\delta)$.

Question 1.3: Does there exist $0 \leq \delta \leq \frac{1}{2}$ such that

$$R(\delta) < \Theta(\delta)?$$

∎

As a matter of fact, the proof of proposition 1.2 gives a rather accurate lower bound on $\theta(Q_{n,d})$, for natural numbers n, d, $d \leq \frac{n}{2}$. Together with a result of Lovasz ([7]) that for a vertex-transitive graph G, $\theta(G) \cdot \theta(\bar{G}) = |V(G)|$, this provides an upper bound on $\theta(\overline{Q_{n,d}}) \geq \Theta(\overline{Q_{n,d}}) \geq \alpha(\overline{Q_{n,d}})$. Kleitman ([5]) had determined $\alpha(\overline{Q_{n,d}})$ and our upper bound turns to be rather tight. In particular, $\Theta(\overline{Q_{n,d}}) \approx \alpha(\overline{Q_{n,d}})$.

Let us say a few words about our methods. Propositions 1.1 and 1.2 are concerned with extremal behaviour of polynomials Λsatisfying the constraints of (3) and (6). A major part of these constraints are given in the form of information on $\Lambda's$ expansion in the basis of the Krawchouk polynomials. Therefore, a natural and accepted approach to 1.1 and 1.2 lies in minute analysis of the properties of the Krawchouk polynomials, and of their linear combinations.

We present a new, and a convenient angle of view, in which a prominent role is played by another family of orthogonal polynomials - the Hahn polynomials, associated with the squared binomial measure on $0, 1, ..., n$. It turns out, that considering the expansion of Λ in this basis provides easy answers in 1.1 and 1.2. An explanation for this fenomenon is given in the next section.

In the proof of 1.1 we deal with a seemingly new extremal polynomial property: Assume we are given an unimodal even sharply decreasing weight function $w(x)$ We have to decide, how far from the origin the global maximum of the function $P(x)w(x)$ can be obtained, where P is a polynomial of a given degree d. In 1.1 w is a shifted binomial measure $\mu_n(i) = \frac{\binom{n}{i+\frac{n}{2}}}{2^n}$.

The connection between the asymptotic behaviour of the measures μ_n on $[-\frac{n}{2}, \frac{n}{2}]$, and the Gaussian measure on $\mathbf{R}$, given by the density $g(x) = \frac{1}{2\pi} e^{-\frac{x^2}{2}}$ makes it natural to ask, whether we can say something similar to 1.1, which has to do with the Gaussian measure. This is indeed the case.

Let π_n denote the set of all the polynomials of degree $\leq n$. For $P \in \pi_n$, consider the function $f_P(x) = P(x)e^{-\frac{x^2}{2}}$. Clearly, $f_P(x) \to 0$ at $\pm\infty$. Let $m(P)$ denote the farthest from the origin point at which a global maximum of $|f_P(x)|$ is obtained. (Note that $|f_P|$ has a finite number of local maxima, since they are necessarily roots of the polynomial $P'(x) - xP(x)$.) Let $M_n = sup\{|m(P)| \mid P \in \pi_n\}$. Then:

Proposition 1.4:

1. $$M_n \leq (1 + o(1))\sqrt{2n}.$$

2. *Let D_n be the n-th Hermite polynomial. Then*
$$m(D_n) = (1 + o(1))\sqrt{2n}.$$

Hermite polynomials [14] are a family of orthogonal polynomials associated with the "squared" Gaussian measure $\tilde{g}(x) = \frac{1}{\pi}e^{-x^2}$ on $\mathbf{R}$. (The relevant properties of $\{D_n\}$ are described in the Appendix.)

Remark 1.5: It turns out, that in the proofs of both 1.1 and 1.4, the crucial property of the orthogonal polynomials that we are dealing with, be it the Hahn or the Hermite polynomials, is, that these polynomials *obtain essentially all the l_2 norm in the interval between the first and the last roots*. It seems interesting to find out how general this phenomenon is. ∎

2 Proof of Proposition 1.1

We start with the description of a simple but rather useful idea, on which our approach is based. Analysis of the Delsarte bound (3) has to do with understanding the behaviour of a polynomial $\Lambda = \sum_{s=0}^{n} b_s P_s$, given as a linear combination of a family of orthogonal polynomials P_s, associated with a measure μ. As it happens, we may know a great deal about properties of the individual P_s, but it is very hard to see, given only the constraints of (3), what of these properties is inherited by Λ. In other words, it is hard to see what the function $\Lambda(i)$ may look like. However, thanks to the Parseval identity, one property of Λ that is easy to deal with, is the l_2 norm:
$$\|\Lambda\|_2 = \sqrt{\sum_{i=0}^{n} \mu(i)\Lambda^2(i)} = \sqrt{\sum_{s=0}^{n} b_s^2 \|P_s\|_2^2}.$$
Unfortunately, as we will see, the information we have or require about Λ, often has more to do with the l_1 norm $\|\Lambda\|_1 = \sum_{i=0}^{n} \mu(i)|\Lambda(i)|$.

Our main idea is, that, in some cases, one can overcome this difficulty by considering an expansion of Λ as a linear combination of a family of polynomials orthogonal with respect to a new measure $\mu' = C \cdot \mu^2$. (C is a normalization constant - all our measures are probability measures). Let $|||.|||_p$ denote the l_p norm in this new measure space. The point is:

$$|||\Lambda|||_2 = \sqrt{\sum_{i=0}^{n} \mu'(i)\Lambda^2(i)} = C^{\frac{1}{2}}\sqrt{\sum_{i=0}^{n} \mu^2(i)\Lambda^2(i)}$$

is equivalent, up to a factor of $O(n)$ (which, for our purposes, is negligible) to the original l_1 norm $\sum_{i=0}^{n} \mu(i)|\Lambda(i)|$.

Now, we need two things to proceed. First, to find a family of polynomials P'_s orthogonal with respect to μ'. Second, to be able to understand well enough the behavior of the individual P'_s.

In the case of the binomial measure $\mu(i) = \frac{\binom{n}{i}}{2^n}$, we are in luck.

Let n, N, be natural numbers, such that $2n \leq N$. Let $\mu_{N,n}$ be the measure supported on the integer points $0, ..., n$ and given by: $\mu_{N,n}(i) = \frac{\binom{n}{i}\binom{N-n}{i}}{\binom{N}{n}}$. It turns out that the family of orthogonal polynomials associated with this measure is a well known family of *Hahn* polynomials, [9].

If we take $N = 2n$, and define

$$\mu'(i) = \mu_{2n,n}(i) = \frac{\binom{n}{i}^2}{\binom{2n}{n}},$$

then the measure μ' is exactly the <u>squared binomial measure</u>.

Let H_s, $1 \leq s \leq n$ be the family of orthogonal polynomials associated with μ'.
The good news are, that we are rather well acquainted with the properties of the Hahn polynomials H_s. Several relevant properties of these polynomials, as well as of Krawchouk polynomials K_s are described in the Appendix.

To sum up, the plan is to use Hahn polynomials H_s orthogonal with respect to the measure $\mu'(i) = \frac{\binom{n}{i}^2}{\binom{2n}{n}}$, $0 \leq i \leq n$, in the study of the bound (3).

Now, we are ready to prove proposition 1.1.

But, first, since we are going to prove a statement that has to do with the *MRRW* solution to (3), it seems like a good place to give a brief description of this solution.

Let t, $0 \leq t \leq n$, be the minimal index, such that the first root $x_{t,1}$ of the Krawchouk polynomial K_t (see Appendix) is smaller than d. (A well-known property of a family P_s, $s = 0, ...$ of orthogonal polynomials is, that the roots of P_{s+1} separate these of P_s, and therefore first roots form a decreasing sequence.)
Let a be the point in the interval $[x_{t,1}, x_{t-1,1}]$, such that $K_{t-1}(a) = -K_t(a) = b$. Then the *MRRW* kernel Λ is given by:

$$\Lambda(x) = b^2 \cdot \frac{(K_{t-1}(x) + K_t(x))^2}{a - x}.$$

Note, that it is not immediately obvious, that Λ satisfies the constraints of (3). It is proven in [9], using the Christoffel-Darboux formula [14] for orthogonal polynomials.

Proof: Of Proposition 1.1.
Let $\Lambda = \sum_{s=0}^{n} b_s K_s$ be a kernel satisfying (3). The orthogonality of $\{K_s\}$ implies:

$$1 = b_0 =< \Lambda, K_0 >= \sum_{i=0}^{n} \mu(i)\Lambda(i).$$

Since $\Lambda(i) \leq 0$, $d \leq i \leq n$, this implies an existence of an integer $0 \leq i_0 \leq d$, such that

$$\mu(i_0)\Lambda(i_0) \geq \frac{1}{n+1}\sum_{i=0}^{n} \mu(i)|\Lambda(i)|.$$

We are going to prove the claim of the proposition by exhibiting a natural number $k = k(d)$ with the following properties:
(1) $k \geq t_d - \epsilon(n)$, where t_d is the degree of the MRRW kernel satisfying (3), $\epsilon(n) = o(n)$.
(2) For any polynomial Ψ of degree $\leq k$, and for all $0 \leq i \leq d$:

$$\mu(i)\Psi(i) < \frac{1}{n+1}\sum_{i=0}^{n} \mu(i)|\Psi(i)| = \frac{1}{n+1}\|\Psi\|_1. \tag{7}$$

We turn to the definition of k. Let $\{H_s\}$ be the family of Hahn polynomials H_s orthogonal with respect to the measure $\mu'(i) = \frac{\binom{n}{i}^2}{\binom{2n}{n}}$, $0 \leq i \leq n$. Let $x_{s,1}$ denote the first root of H_s. Let $\gamma(n)$ be the function defined in the corollary 5.3, i.e.: $\gamma(n) = o(n)$, and for any $0 \leq i \leq x_{s,1}$:

$$\mu(i)H_s^2(i) \leq \frac{1}{(n+1)^6}\sum_{i=0}^{n} \mu(i)H_s^2(i) = \|H_s\|_2^2. \tag{8}$$

We define $k = max\{s|x_{s,1} \geq d + \gamma(n)\}$. How big is k? According to (21), $x_{s,1} = n(\frac{1}{2} - \sqrt{\frac{s}{2n}(1 - \frac{s}{2n})}) + o(n)$. Using this and the fact that the function $\phi(y) = \frac{1}{2} - \sqrt{y(1-y)}$ is an involution on $[0, \frac{1}{2}]$, we have:

$$k \geq 2n\left(\frac{1}{2} - \sqrt{\frac{d}{n}(1 - \frac{d}{n})}\right) - o(n).$$

We would like to compare this with t_d. Let $y_{s,1}$ denote the first root of K_s. Set $r = max\{s|y_{s,1} \geq d\}$. Then, (see [9]), using (18):

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

Thus, the first property of k is proven.

Assume now, that Ψ is a polynomial of degree $\leq k$ and expand Ψ as a linear combination of $\{H_s\}$: $\Psi = \sum_{s=0}^k c_s H_s$. Let i be an integer, $0 \leq i \leq d$. Then:

$$\mu(i)|\Psi(i)| = \frac{\binom{n}{i}}{2^n}|\Psi(i)| = \frac{\sqrt{\binom{2n}{n}}}{2^n}\sqrt{\mu'(i)} \cdot |\Psi(i)| \leq \sqrt{\mu'(i)\Psi^2(i)} =$$

$$\sqrt{\mu'(i)(\sum_{s=0}^{k} c_s H_s(i))^2} \leq \sqrt{n+1}\sqrt{\sum_{s=0}^{k} c_s^2 \mu'(i) H_s^2(i)}.$$

By the choice of k, for any $s \leq k$ we have $d \leq x_{s,1} - \gamma(n)$. So, for any $s \leq k$, $0 \leq i \leq d$, (8) holds. Therefore, by the Parseval identity, we have:

$$\mu(i)|\Psi(i)| \leq \frac{1}{n^{\frac{5}{2}}}\sqrt{\sum_{s=0}^{k} c_s^2 |||H_s|||_2^2} = \frac{1}{n^{\frac{5}{2}}}|||\Psi|||_2,$$

where $|||.|||_2$ denotes the l_2 norm corresponding to μ'. Note, that for any function F:

$$|||F|||_2 = \sqrt{\sum_{i=0}^{n} \mu'(i)F^2(i)} = \frac{2^n}{\sqrt{\binom{2n}{n}}}\sqrt{\sum_{i=0}^{n} \mu^2(i)F^2(i)} \leq O(n^{\frac{1}{4}})\cdot\sum_{i=0}^{n} \mu(i)|F(i)| = O(n^{\frac{1}{4}})\cdot\|F\|_1. (9)$$

And so, at last, for any $0 \leq i \leq d$, (and n big enough):

$$\mu(i)|\Psi(i)| \leq O\left(\frac{1}{n^{\frac{9}{4}}}\right) \cdot \|F\|_1 < \frac{1}{n+1}\|\Psi\|_1,$$

proving the second property of k and the claim of the proposition. ∎

Note, that in no place have we relied on the non-negativity of b_s, $1 \leq s \leq n$ in (3). Therefore, a stronger statement is true:

Corollary 2.1: *There exists a function $\epsilon(n) = o(n)$ such that for any $0 \leq d \leq \frac{n}{2}$ and for any kernel Λ, such that $deg\Lambda \leq 2n\left(\frac{1}{2} - \sqrt{\frac{d}{n}(1-\frac{d}{n})}\right) - \epsilon(n)$, Λ can't satisfy the following constraints:*
(1) $\sum_{i=0}^n \mu(i)\Lambda(i) = \frac{1}{2^n}\sum_{i=0}^n \binom{n}{i}\Lambda(i) > 0$.
(2) $\Lambda(i) \leq 0$, $d \leq i \leq n$.

3 Proof of Proposition 1.2

We start with an important and a well-known property of Krawchouk polynomials.

Lemma 3.1: *Define an $(n+1)\times(n+1)$ matrix $A = (a_{ij})$ by setting $a_{ij} = K_j(i)$, $0 \leq i, j \leq n$. Then:*

$$A^2 = 2^n \cdot I,$$

where I is an $(n+1)\times(n+1)$ identity matrix.

Proof: Let $B = A^2$, $B = (b_{ij})$, $0 \le i, j \le n$. Then:

$$b_{ij} = \sum_{k=0}^{n} a_{ik}a_{kj} = \sum_{k=0}^{n} K_k(i)K_j(k) = \frac{1}{\binom{n}{i}} \sum_{k=0}^{n} \binom{n}{k} K_i(k)K_j(k) = \begin{cases} 2^n & \text{i = j} \\ 0 & \text{otherwise} \end{cases}.$$

We have used the reciprocity (16) and the orthogonality (17) properties of the Krawchouk polynomials. ∎

Corollary 3.2:

$$\theta(Q_{n,d}) = min\{\Lambda(0) \mid \Lambda(i) = \sum_{s=0}^{n} b_s K_s(i);\ b_0, ..., b_n \ge 0;\ b_0 = 1;\ \Lambda(i) = 0,\ d \le i \le n\} =$$

$$min\{\Lambda(0) \mid b_s = \frac{1}{2^n} \sum_{i=0}^{n} \Lambda(i)K_i(s) \ge 0,\ s = 0, ..., n;\ b_0 = 1;\ \Lambda(i) = 0,\ d \le i \le n\} =$$

$$min\left\{\frac{\Lambda(0)}{b_0} \mid\ b_s = \frac{1}{2^n} \sum_{i=0}^{d-1} \Lambda(i)K_i(s) \ge 0,\ s = 0, ..., n\right\} =$$

$$min\{\frac{\sum_{i=0}^{n} \binom{n}{i} F(i)}{F(0)} | F(i) \ge 0,\ i = 0, ..., n; F\ polynomial\ of\ degree < d\}. \tag{10}$$

Proof: The first equality is a consequence of the preceding lemma. The second one is an obvious reformulation; the third one is obtained by taking $F(s) = \sum_{i=0}^{n} \Lambda(i)K_i(s)$, $0 \le s \le n$. ∎

Now, recall $\theta(\delta) \le 1 - H(\frac{\delta}{2})$. To see that, one uses the Lloyd polynomials: For $0 \le d \le \frac{n}{2}$, let $t = \lfloor \frac{d}{2} \rfloor$. Define polynomial

$$L_d(x) = (\sum_{j=0}^{t} K_j(x))^2.$$

L_d clearly satisfies the conditions of (10). Now,

$$L_d(0) = (\sum_{j=0}^{t} K_j(0))^2 = (\sum_{j=0}^{t} \binom{n}{j})^2 \ge \binom{n}{t}^2.$$

On the other hand,

$$\sum_{i=0}^{n} \binom{n}{i} L_d(i) = 2^n < L_d, K_0 >= 2^n < (\sum_{j=0}^{t} K_j(x))^2, K_0 >= 2^n < \sum_{j=0}^{t} K_j(x), \sum_{j=0}^{t} K_j(x) >=$$

$$2^n \sum_{j=0}^{t} < K_j, K_j >= 2^n \sum_{j=0}^{t} \binom{n}{j} \le n2^n \binom{n}{t}.$$

We will require a well known fact on binomial coefficients. Let c be a constant, $0 \leq c \leq 1$, then

$$\binom{n}{cn} = 2^{H(c)\cdot n + o(n)}. \tag{11}$$

Therefore: $\frac{\sum_{i=0}^{n} \binom{n}{i} L_d(i)}{L_d(0)} \leq 2^{(1-H(\frac{\delta}{2}))\cdot n + o(n)}$, implying:

$$\theta(\delta) \leq 1 - H(\frac{\delta}{2}).$$

We turn to the proof of the other direction of the inequality, i.e.,

$$\theta(\delta) \geq 1 - H(\frac{\delta}{2}).$$

This claim will follow from

Lemma 3.3: *Let F be a polynomial of degree $< d$, with $F(i) \geq 0$, $0 \leq i \leq n$, $F(0) = 1$. Then:*

$$\sum_{i=0}^{n} \binom{n}{i} F(i) = 2^n \|F\|_1 \geq 2^{(1-H(\frac{\delta}{2}))\cdot n - o(n)}.$$

Proof: We use the notation and the results of proposition 1.1. Expand F as a linear combination of H_s, $F = \sum_{s=0}^{d-1} c_s H_s$. Note, by (5.2) and by (20), that $|||H_s|||_2^2 = H_s(0) = w_s = \begin{cases} \binom{2n}{s} - \binom{2n}{s-1} & s > 0 \\ 1 & s = 0 \end{cases}$. With the help of the Parseval identity and of (9), we obtain:

$$\|F\|_1 \geq \Omega(n^{-\frac{1}{4}})|||F|||_2 = \Omega(n^{-\frac{1}{4}})\sqrt{\sum_{s=0}^{d-1} c_s^2 |||H_s|||_2^2} = \Omega(n^{-\frac{1}{4}})\sqrt{\sum_{s=0}^{d-1} w_s c_s^2}.$$

On the other hand,

$$1 = F(0) = \sum_{s=0}^{d-1} c_s H_s(0) = \sum_{s=0}^{d-1} w_s c_s.$$

We have here a simple maximization problem: Let $\{w_s\}$ $0 \leq s \leq d-1$, be a given positive sequence. Let $g(c_0, ..., c_{d-1})$ be a real function of d variables, given by: $g(c_0, ..., c_{d-1}) = \sum_{s=0}^{d-1} w_s c_s^2$. Find $max\{g(c_0, ..., c_{d-1}) \mid \sum_{s=0}^{d-1} w_s c_s = 1\}$. Lagrange multipliers provide the answer: The maximum is obtained for

$$c_s^{max} = \frac{1}{\sum_{s=0}^{d-1} w_s} = \frac{1}{\binom{2n}{d-1}},$$

$0 \leq s \leq d-1$;
the value of the maximum is also given by: $\frac{1}{\binom{2n}{d-1}}$. Therefore, using (11):

$$\sum_{i=0}^{n} \binom{n}{i} F(i) = 2^n \|F\|_1 \geq \Omega(n^{-\frac{1}{4}}) \cdot \sqrt{\frac{1}{\binom{2n}{d-1}}} \geq 2^{(1-H(\frac{\delta}{2}))\cdot n - o(n)},$$

and we are done.

Remark 3.4: Let A be the adjacency matrix of $Q_{n,d}$. $Q_{n,d}$ is a vertex-transitive graph, and therefore, [7], we can estimate $\theta(Q_{n,d})$ using the eigenvalues of A. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{2^n}$ be A's eigenvalues. Then the following inequality holds:

$$\theta(Q_{n,d}) \leq \frac{2^n(-\lambda_{2^n})}{\lambda_1 - \lambda_{2^n}}. \tag{12}$$

It is natural to ask, whether this upper bound on $\theta(Q_{n,d})$ does not already imply proposition 1.2. It turns out, that the answer is "No". Indeed, it is easily verified, that (12) gives only: $\theta(\delta) \geq 1 - H(\delta)$. ∎

Let $\overline{G}$ denote the graph complementary to G: $V(\overline{G}) = V(G)$, $E(\overline{G}) = \overline{E(G)}$. Lovasz ([7]) had shown, that for a vertex-transitive graph G, $\theta G \cdot \theta(\overline{G}) = |V(G)|$. Therefore, $\theta(Q_{n,d}) \cdot \theta(\overline{Q_{n,d}}) = |Z_2^n| = 2^n$. In this way, proposition 1.2 provides an upper bound on $\theta(\overline{Q_{n,d}})$.
While the original purpose of Lovasz in introducing the θ-function ([7]) was to provide an upper bound on the Shannon capacity $\Theta(G)$ of a graph G, he had also used $\theta(G)$ directly as an upper bound on $\alpha(G)$, the size of the maximal independent subset of G.
Let N, n, t be natural numbers ($N \geq n \geq t \geq 1$) and let $K_{N,n,t}$ be the graph whose vertices are all the n-subsets of a fixed N-set, two of them being adjacent iff their intersection contains less that t elements. Erdös, Ko and Rado ([4]) had proven, that $\alpha(K_{N,n,t}) = \binom{N-t}{n-t}$, if N is large enough compared to n. In [7], Lovasz had given a new proof of the Erdös-Ko-Rado theorem for $t = 1$, by proving a stronger statement: $\Theta(K_{N,n,1}) = \binom{N-1}{n-1}$, if $N \geq 2n$.

Schrijver in [12] had used his characterization of $\theta(G)$ as the value of a linear program [11], to extend the result of Lovasz: $\Theta(K_{N,n,t}) = \binom{N-t}{n-t}$, for N *very* large compared to n.
It turns out, that for graphs $K_{N,n,t}$ the Shannon capacity equals to the maximal size of an independent subset. Not surprisingly, a similar situation occurs in dealing with graphs $\overline{Q_{n,d}}$, the Hamming scheme counterparts of $K_{N,n,t}$. An exact value of $\alpha(\overline{Q_{n,d}})$ is due to Kleitman ([5]). He had shown, for an even d: $\alpha(\overline{Q_{n,d}} = \sum_{i=0}^{\frac{d}{2}} \binom{n}{i}$. Our approach has little chance of obtaining an exact value of $\theta(\overline{Q_{n,d}})$, due, for example, to a possible loss of a factor of $O(n^{\frac{1}{4}})$ in (9). Nevertheless, we show:

Corollary 3.5:

$$\theta(\overline{Q_{n,d}}) \leq O(\sqrt{n}) \cdot \binom{n}{\lfloor \frac{d}{2} \rfloor}.$$

Proof: From the proof of the preceding lemma, we have:

$$\theta(Q_{n,d}) \geq 2^n \cdot \frac{\Omega(n^{-\frac{1}{4}})}{\binom{2n}{d-1}}.$$

The following fact is an easy consequence of the Stirling formula. For all natural numbers $0 \leq s \leq n$:

$$\binom{2n}{2s} \leq O\left(\sqrt{\frac{s(n-s)}{n}}\right) \cdot \binom{n}{s}^2.$$

Therefore,

$$\theta(Q_{n,d}) \geq 2^n \cdot \frac{\Omega(n^{-\frac{1}{4}})}{\sqrt{O\left(\sqrt{\frac{s(n-s)}{n}}\right) \cdot \binom{n}{\lfloor \frac{d}{2} \rfloor}}} = \frac{1}{O((s(n-s))^{\frac{1}{4}})} \cdot \frac{2^n}{\binom{n}{\lfloor \frac{d}{2} \rfloor}},$$

implying

$$\theta(\overline{Q_{n,d}}) \leq O((s(n-s))^{\frac{1}{4}}) \cdot \binom{n}{\lfloor \frac{d}{2} \rfloor} \leq O(\sqrt{n}) \cdot \binom{n}{\lfloor \frac{d}{2} \rfloor}.$$

∎

Remark 3.6:
(1) Note, that our bound for $\theta(\overline{Q_{n,d}})$ is necessarily tight upto a factor of $O(\sqrt{n})$, and implies

$$\alpha(\overline{Q_{n,d}}) \leq \Theta(\overline{Q_{n,d}}) \leq O(\sqrt{n}) \cdot \alpha(\overline{Q_{n,d}}).$$

(2) We conjecture $\Theta(\overline{Q_{n,d}}) = \alpha(\overline{Q_{n,d}})$, for n big enough, compared to d.
(3) Note, that for graphs $Q_{n,d}$, $\theta(Q_{n,d})$ and $\alpha(Q_{n,d})$ are very far apart. One one hand $\alpha(Q_{n,d}) \leq 2^{m_2(\delta)\cdot n+o(n)}$, where $m_2(\delta)$ is the unspecified *second linear programming bound*, mentioned in the Introduction. On the other hand, by proposition 1.2, $\theta(Q_{n,d}) \geq 2^{(1-H(\frac{\delta}{2})\cdot n-o(n)}$. Recall $m_2(\delta) < 1 - H(\frac{\delta}{2})$, $0 \leq \delta \leq \frac{1}{2}$. We see, that $\theta(Q_{n,d})$ behaves as $\alpha(Q_{n,d})$ to some power greater than 1. ∎

Remark 3.7: The approach of proposition 1.2 can be applied in a more general situation. Let d, n be natural numbers. Let $(X, \mathcal{R})$, $\mathcal{R} = (R_0, .., R_n)$, be a symmetric association scheme with $n+1$ classes. (See [1] for an introduction to association schemes). Let G_d be a graph with $V(G_d) = X$ and $E(G) = \cup_{0 \leq i < d} R_i$. Note, that $Q_{n,d}$ is a special cases of graphs G_d, where $(X, \mathcal{R})$ is the the Hamming association scheme. Assume $(X, \mathcal{R})$ to be a P-polynomial scheme, where the family $\{P_s\}$ of P-polynomials for $(X, \mathcal{R})$, is orthogonal with respect to a discrete measure μ supported on a n-subset $S \subseteq \mathbf{R}$. Let $\{T_s\}$ be a family of polynomials orthogonal with respect to μ^2. We use Schrijver's characterization of $\theta(G)$, to obtain:

$$\Omega(n^{-1}) \cdot \frac{|X|}{\sqrt{\sum_{s=0}^{d-1} \frac{T_s^2(0)}{\|T_s\|_2^2}}} \leq \theta(G_d) \leq \frac{|X|}{\sqrt{\sum_{s=0}^{\lfloor \frac{d-1}{2} \rfloor} P_s^2(0)}}. \tag{13}$$

A counterpart of Lloyd polynomials gives the lower bound, while lemma 3.3 is easily generalized to give the upper bound.

To apply the lower bound in (13) we have to study the behaviour of $\frac{T_s^2(0)}{\|T_s\|_2^2}$, $0 \leq s \leq n$. For a general ν the only approach we are aware of, is by using the following explicit representation (see [14], section 2.2). Let $c_k = \int x^k d\nu(x)$ denote the k-th moment of ν. Then:

$$\frac{T_s^2(0)}{\|T_s\|_2^2} = \frac{A_s}{\sqrt{D_{s-1} D_s}},$$

where $A_s = \det([c_{i+j+1}])$, $0 \leq i, j \leq s-1$; and $D_t = \det([c_{i+j}])$, $0 \leq i, j \leq t$.

Let us also mention, without going into details, that the graphs G_d for the *Johnson association scheme* are complementary to the Kneser graphs $K_{N,n,t}$. More precisely $G_d = \overline{K_{N,n,n-d}}$ and since these graphs are vertex-transitive, lower bounds on $\theta(G_d)$ give upper bounds on $\theta(K_{N,n,n-d})$. It seems interesting to understand, what can be learned about $\theta(K_{N,n,t})$ with the help of (13). The P-polynomials for the Johnson scheme are the Eberlein polynomials (see [1]). These polynomials are orthogonal with respect to a measure μ, supported on the points $x_i = i^2 - (N+1)i + n(N-n)$, $0 \leq i \leq n$, and given by

$$\mu(x_i) = \begin{cases} \frac{\binom{N}{i} - \binom{N}{i-1}}{\binom{N}{n}} & i > 0 \\ 1 & i = 0 \end{cases}$$

To apply (13), one will have to study the behaviour of $\frac{T_s^2(0)}{\|T_s\|_2^2}$ for polynomials $\{T_s\}$ orthogonal with respect to μ^2 on $x_0, ..., x_n$. ∎

4 Proof of Proposition 1.4

The key step in the proof of the proposition 1.1 was to show, that for a polynomial P of a bounded degree, and for a binomial measure $\mu(i) = \frac{\binom{n}{i}}{2^n}$ on $[0, n]$, the maximum of $\mu(i)|P(i)|$ can not be obtained far from the middle of the interval (see (7). Of course, the binomial measure is strongly concentrated around $\frac{n}{2}$.

Here we are in a very similar situation, which is, as we are about to find out, a little more complicated since we are dealing with a continuous measure. However, a fairly straightforward application of the ideas of the second Section suffices for the proof of 1.4.

Proof: Of Proposition 1.4 We start with the first part.

Let us assume, for the contrary, an existence of an $\epsilon > 0$, and a sequence $\{P_{n_k}\}_k$ of polynomials, such that $P_{n_k} \in \pi_{n_k}$ and $|m(P_{n_k})| > (1+\epsilon)\sqrt{2n_k}$. To simplify matters, we also assume $\{n_k\} = \{n\}$.

Our plan to reach a contradiction is quite similar to one in the proof of proposition 1.1. We will prove two following statements, which are clearly contradicting. (In the following discussion we will always assume, without loss of generality, that $m = m(P)$ and $P(m)$ are positive, and $m > 1$.)

Lemma 4.1: *For $P \in \pi_n$,*

$$\int_{-\infty}^{-m+1} P^2(x)e^{-x^2}dx + \int_{m-1}^{\infty} P^2(x)e^{-x^2}dx \geq \Omega(\frac{1}{n^2}) \int_{-\infty}^{\infty} P^2(x)e^{-x^2}dx = \|P\|_2^2.$$

Lemma 4.2: *For $P \in \pi_n$,*

$$\int_{-\infty}^{-(1+\epsilon)\sqrt{2n}} P^2(x)e^{-x^2}dx + \int_{(1+\epsilon)\sqrt{2n}}^{\infty} P^2(x)e^{-x^2}dx \leq O(e^{-C_\epsilon \cdot n})\|P\|_2^2,$$

where the constant C_ϵ is defined in corollary 5.9

But first we need a following auxiliary result.

Lemma 4.3: *For all integer n, $M_n < 6\sqrt{2n}$.*

Proof: Assume, that $P \in \pi_n$ is such, that $m(P) \geq 6\sqrt{2n}$. Let $m = m(P)$. Then, for all $x \in [-\frac{m}{2}, \frac{m}{2}]$,

$$|P(m)| \geq |P(x)|e^{\frac{m^2-x^2}{2}} \geq |P(x)|e^{\frac{m^2}{4}} \geq |P(x)|e^{18n}.$$

Set $Q(x) = P(\frac{m}{2}x)$. Then $Q \in \pi_n$ and $|Q(2)| \geq |Q(x)|e^{18n}$, for all $x \in [-1, 1]$.
Now, by (25), for any $Q \in \pi_n$,

$$|Q(2)| \leq |T_n(2)| \max_{x\in[-1,1]} |Q(x)| < 4^n \max_{x\in[-1,1]} |Q(x)|,$$

where T_n is the n-th Chebyshev polynomial. So, a contradiction is reached, completing the proof. ∎

Proof: of Lemma 4.1 Let $f(x) = P(x)e^{-\frac{x^2}{2}}$. Since the inequality of the Lemma clearly does not change if we multiply f by a scalar, we may and will assume that $\|f\|_\infty = 1$.

We start with an estimate on $f'(x)$ near m.

$$|f'(x)| = |(P' - xP)e^{-\frac{x^2}{2}}| \leq (|P'| + |x||P|)e^{-\frac{x^2}{2}}. \tag{14}$$

The following inequality is due to Markov ([15], page 218): For any $Q \in \pi_n$, and for all $x \in [-1, 1]$, $|Q'(x)| \leq n^2 \max_{x\in[-1,1]} |Q(x)|$. Taking $Q(x) = P(mx)$, we obtain, that for all $x \in [-m, m]$, $|P'(x)| \leq \frac{n^2}{m} \max_{x\in[-m,m]} |P(x)|$.
We know that for all $x \in [-m, m]$, $|P(x)|e^{-\frac{x^2}{2}} \leq |P(m)|e^{-\frac{m^2}{2}}$. On the other hand $e^{-\frac{x^2}{2}}$ decreases as $|x| \to \infty$, and therefore, for all $x \in [-m, m]$, $|P(x)| \leq |P(m)|$. We conclude therefore, that for all $x \in [-m, m]$, $|P'(x)| \leq \frac{n^2}{m}|P(m)|$.

Returning to (14) we have, that for all $x \in [m - \frac{1}{m}, m]$,

$$|f'(x)| \leq O(\frac{n^2 + m^2}{m})|P(m)|e^{-\frac{m^2}{2}} \leq O(\frac{n^2 + m^2}{m})|f(m)|.$$

Lemma 4.3 tells us that $m = O(\sqrt{n})$, and therefore for $x \in [m - \frac{1}{m}, m]$, $|f'(x)| \leq O(n^2)|f(m)|$.

We have obtained the desired bound on f', which implies, that in a small neighbourhood of m, $f(x) \approx f(m)$. More precisely, for $x \in [m - \Omega\left(\frac{1}{n^2}\right), m]$, $|f(x)| > \frac{|f(m)|}{2} = \frac{1}{2}$. We are almost done.

Let I denote the interval $[m - \Omega\left(\frac{1}{n^2}\right), m]$, in which $|f(x)| > \frac{1}{2}$. Then,

$$\int_I f^2(x)dx \geq \frac{1}{4}|I| \geq \frac{|I|}{8m} \int_{-m}^{m} f^2(x)dx \geq \Omega\left(\frac{1}{n^3}\right) \int_{-m}^{m} f^2(x)dx.$$

Adding $\int_{-\infty}^{-m} f^2(x)dx + \int_m^\infty f^2(x)dx$ to both sides of this inequality, we obtain the statement of the Lemma. ∎

Proof: of Lemma 4.2 We expand P in the basis of the Hermite polynomials D_k, (see Appendix for definition and properties of the Hermite polynomials).
Let $P = \sum_{k=0}^{n} \lambda_k D_k$. Then,

$$\int_{-\infty}^{(1+\epsilon)\sqrt{2n+1}} P^2(x) e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{(1+\epsilon)\sqrt{2n+1}} (\sum_{k=0}^{n} \lambda_k D_k)^2 e^{-\frac{x^2}{2}} dx \le n \cdot \int_{-\infty}^{(1+\epsilon)\sqrt{2n+1}} \sum_{k=0}^{n} \lambda_k^2 D_k^2 e^{-\frac{x^2}{2}} dx = n \cdot \sum_{k=0}^{n} \lambda_k^2 \int_{-\infty}^{(1+\epsilon)\sqrt{}}$$

Now, by Corollary 5.9, for all $0 \le k \le n$, $\int_{-\infty}^{(1+\epsilon)\sqrt{2n+1}} D_k^2(x) e^{-\frac{x^2}{2}} dx \le e^{-C_\epsilon \cdot n} \|D_k\|_2^2$, and we obtain:

$$\int_{-\infty}^{(1+\epsilon)\sqrt{2n+1}} P^2(x) e^{-\frac{x^2}{2}} dx \le n e^{-C_\epsilon \cdot n} \sum_{k=0}^{n} \lambda_k^2 \|D_k\|_2^2 =$$

by the orthogonality of D_k,

$$n e^{-C_\epsilon \cdot n} \|P\|_2^2.$$

The inequality

$$\int_{-(1+\epsilon)\sqrt{2n+1}}^{\infty} P^2(x) e^{-\frac{x^2}{2}} dx \le n e^{-C_\epsilon \cdot n} \|P\|_2^2,$$

is proved similarly, and we are done. ∎Lemmas 4.1 and 4.2 proven, the first part of the Proposition follows.

The second part of the Proposition follows from Lemma 5.4 in the Appendix. ∎

5 Appendix

5.1 Properties of the Krawchouk polynomials

Let $\mu(i) = \frac{\binom{n}{i}}{2^n}$ be a discrete probability measure on $[0, n]$. The orthogonal polynomials for μ are the Krawchouk polynomials $\{K_s\}$, and in this subsection we will describe the properties of the measure and the polynomials that have been used in this paper. See [9] for details.

Definition:

$$K_s(x) = \sum_{k=0}^{s} (-1)^k \binom{x}{k} \binom{n-x}{j-k}. \tag{15}$$

Corollary 5.1:

$$K_0 \equiv 1, \ K_s(0) = \binom{n}{s}.$$

Reciprocity:

$$\binom{n}{s} K_i(s) = \binom{n}{i} K_s(i). \tag{16}$$

Orthogonality:

$$\sum_{i=0}^{n} \mu(i) K_s(i) K_t(i) = \begin{cases} \binom{n}{s} & \text{s = t} \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

The first root behaviour:

$$x_{s,1} = n \cdot \left(\frac{1}{2} - \sqrt{\frac{s}{n}(1 - \frac{s}{n})} \right) + o(n). \tag{18}$$

5.2 Properties of Hahn polynomials

Let us define a discrete probability measure μ on $[0, n]$ in the following way: $\mu(i) = \frac{\binom{n}{i}^2}{\binom{2n}{n}}$, and let $\{H_s\}$ be the associated Hahn polynomials. We once more refer to [9] for details.

Definition:

$$H_s = \frac{w_s}{2^s} \sum_{k=0}^{s} \frac{\binom{s}{k}}{\binom{n}{k}\binom{n}{s-k}} K_k K_{s-k}, \tag{19}$$

K_s are the Krawchouk polynomials, and set $w_s = \binom{2n}{s} - \binom{2n}{s-1}$.

Corollary 5.2: $H_0 \equiv 1,\ H_s(0) = w_s$

Orthogonality:

$$\sum_{i=0}^{n} \mu(i) H_s(i) H_t(i) = \begin{cases} w_s & \text{s = t} \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

The first root behaviour:

$$x_{s,1} = 2n \cdot \frac{\frac{1}{4} - \frac{s}{2n}(1 - \frac{s}{2n})}{1 + 2\sqrt{\frac{s}{2n}(1 - \frac{s}{2n})}} + o(n). \tag{21}$$

We will also require a following property of H_s, which states, that the contribution of $i < x_{s,1} - o(n)$ to the l_2 norm $\|H_s\|_2 = \sqrt{\sum_{i=0}^{n} \mu(i) H_s^2(i)}$ is negligible. (See [10]).

Lemma 5.3: *Let $s = s_n$ be a sequence of natural numbers between zero and $\frac{n}{2}$, such that $\frac{s}{n}$ is bounded away from $0, \frac{1}{2}$. Then, there exists a function $\gamma(n) = o(n)$, such that in the interval $0 \le i \le x_{s,1} - \gamma(n)$:*

$$\mu(i) H_s^2(i) \le \frac{1}{(n+1)^6} \mu(i + \gamma(n)) H_s^2(i + \alpha(n)) \le \sum_{i=0}^{n} \mu(i) H_s^2(i) = \|H_s\|_2^2.$$

5.3 Properties of Hermite polynomials

Let μ be a probability measure on $\mathbf{R}$, given by the density $w(x) = \frac{1}{pi}e^{-x^2}$. The orthogonal polynomials $\{D_n\}$ associated with μ are the classical Hermite polynomials. See [14] for a very extensive treatment.

Definition:

$$D_n = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \frac{2^n n!}{k!(n-2k)!} x^{n-2k}.$$

Orthogonality:

$$\int_{-\infty}^{+\infty} w(x) D_n(x) D_m(x) dx = \begin{cases} \pi^2 2^n n! & \text{m = n} \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

The first root behaviour:

$$x_{n,1} = -(1+o(1))\sqrt{2n+1}. \tag{23}$$

The behaviour of relative maxima of $|D_n(x)|w^{\frac{1}{2}}(x)$ (see [14], Theorem 7.6.3):

Lemma 5.4: *The successive relative maxima of* $|D_n(x)|e^{-\frac{x^2}{2}}$ *form an increasing sequence for* $x \geq 0$*.*

We will also need a property of D_n, which states that for any $\epsilon > 0$ the contribution of the interval $(-\infty, (1+\epsilon)x_{n,1}]$ to the l_2 norm $\|D_n\|_2 = \sqrt{\int_{-\infty}^{+\infty} w(x) D_n^2(x) d(x)}$ in negligible:

Proposition 5.5: *For any* $\epsilon > 0$*, there is a constant* $C_\epsilon > 0$*, such that*

$$\int_{-\infty}^{-(1+\epsilon)\sqrt{2n+1}} w(x) D_n^2(x) dx + \int_{(1+\epsilon)\sqrt{2n+1}} w(x) D_n^2(x) dx \leq e^{-C_\epsilon \cdot n} \ \|D_n\|_2^2.$$

Proof: Observe (see the Definition), that the polynomials D_n are even functions for n even, and odd functions for n odd. Since the weight function $w(x)$ is even, in the following discussion we will be interested in the absolute value of the argument, and not in it's sign. In addition, note that we only have to prove the one-sided inequality, say

$$\int_{-\infty}^{-(1+\epsilon)\sqrt{2n+1}} w(x) D_n^2(x) dx \leq e^{-C_\epsilon \cdot n} \ \|D_n\|_2^2,$$

and the Proposition will follow by symmetry.

Let us start with the fact, that the asymptotic behaviour of $D_n(x)$, for $\sqrt{2n+1} < |x| = O(\sqrt{n})$ is known.

Lemma 5.6: *([14], Theorem 8.22.9) Let $\alpha < \omega$ be fixed positive numbers. We have, for $x = (2n+1)^{\frac{1}{2}} cosh(\phi)$, $\alpha \leq \phi \leq \omega$:*

$$D_n(x)e^{-\frac{x^2}{2}} = O(n^{-\frac{1}{4}} 2^{\frac{n}{2}} (n!)^{\frac{1}{2}} (sinh(\phi))^{-\frac{1}{2}} \cdot \exp[(\frac{n}{2} + \frac{1}{4})(2\phi - sinh(2\phi))]), \tag{24}$$

where the constant term implicit in the $O(.)$ notation depends only on α and ω.

Corollary 5.7: *For any fixed $\epsilon > 0$, and for any fixed positive ω there is a constant $C = C(\epsilon, \omega) > 0$, such that*

$$\int_{-\omega\sqrt{2n+1}}^{(1+\epsilon)\sqrt{2n+1}} w(x) D_n^2(x) d(x) \leq e^{-C \cdot n} \, \|D_n\|_2^2.$$

Proof: Note, that the multiple of $(2^n n!)^{\frac{1}{2}} = \|D_n\|_2$ in the RHS of (24) is very small, since the exponent $2\phi - sinh(2\phi)$ is negative for $\phi > 0$. (and it's absolute value is exponential in ϕ !) This, together with some straightforward manipulation of the integrals, gives the statement of the Corollary. We do not go into details. ∎

Our plan now is to find $\omega_0 = \omega_0(\epsilon)$, with value to be specified later, and to show that

$$\int_{-\infty}^{-\omega_0\sqrt{2n+1}} w(x) D_n^2(x) d(x) \leq e^{-K \cdot n} \, \|D_n\|_2^2,$$

where K depends only on ω_0. This, together with the Corollary will immediately imply the Proposition.

For that purpose, we introduce the Chebyshev polynomials $T_n(x) = \frac{1}{2}[(x + \sqrt{x^2-1})^n + (x - \sqrt{x^2-1})^n$. A well known property of these polynomials [15] is, that $\|T_n\|_\infty$ on the interval $[-1, 1]$ is minimal, among all polynomials of degree $\leq n$ with leading coefficient 1. This property implies, among other things, the following estimate ([15], p. 88, Problem 14).

Lemma 5.8: *If a is real and $|a| > 1$, than any polynomial $P_n(x)$ of degree $\leq n$, with the property $\max_{-1 \leq x \leq 1} |P_n(x)| \leq 1$, satisfies the inequality*

$$|P_n(a)| \leq |T_n(a)|. \tag{25}$$

Now, fix $\omega = \omega(\epsilon)$, and let I denote the interval $[-\omega\sqrt{2n+1}, -(1+\epsilon)\sqrt{2n+1}]$. Let $x_0 < -\omega\sqrt{2n+1}$. We would like to estimate $|f(x_0)| = |D_n(x_0)e^{-\frac{x_0^2}{2}}|$ from above, using (25). Let i_0 denote the middle point of I, and let $Q(x) = D_n(\frac{|I|}{2} \cdot x + i_0)$. Then, by (25),

$$|D_n(x_0)| = |Q(\frac{2(x_0 - i_0)}{|I|})| \leq |T_n(\frac{2(x_0 - i_0)}{|I|})| \cdot max_{-1 \leq x \leq 1} |Q(x)| =$$

$$|T_n(\frac{2(x_0 - i_0)}{|I|})| \cdot max_{x \in I} |D_n(x)|.$$

Now, for $|x_0| > 2\omega\sqrt{2n+1}$, using the fact that for $|x| > 1$, $|T_n(x)| < (2|x|)^n$ and that, by (24) $max_{x\in I}|D_n(x)e^{-\frac{x^2}{2}}| < O(\|D_n\|_2)$, (where the constant hidden in the $O(.)$ notation depends only on ϵ), we get

$$|D_n(x_0)e^{-\frac{x_0^2}{2}}| \le \|D_n\|_2 \cdot O\left(\left(\frac{2|x_0|}{|I|}\right)^n e^{-\frac{x_0^2}{4}}\right).$$

Now, using the fact, that $|I| = \Omega(\sqrt{n})$, we see, that for there exists a constant B, such that for any x, $|x| > 2\omega$,

$$|D_n(x)e^{-\frac{x^2}{2}}| \le \|D_n\|_2 \cdot O\left(\left(\frac{B|x|}{\sqrt{n}}\right)^n e^{-\frac{x^2}{4}}\right). \tag{26}$$

Now, the function $g(x) = (\frac{B|x|}{\sqrt{n}})^n e^{-\frac{x^2}{4}}$ is rapidly descending to zero, and it is easy to find constants ω_0 and K, such that

$$\int_{-\infty}^{-\omega_0\sqrt{2n+1}} g^2(x)d(x) \le e^{-K\cdot n},$$

implying the Proposition. We skip the details. ▮

Corollary 5.9: *For any $\epsilon > 0$, there is a constant $C_\epsilon > 0$, such that for all $0 \le k \le n$,*

$$\int_{-\infty}^{-(1+\epsilon)\sqrt{2n+1}} w(x)D_k^2(x)dx + \int_{(1+\epsilon)\sqrt{2n+1}} w(x)D_k^2(x)dx \le e^{-C_\epsilon\cdot n}\ \|D_k\|_2^2.$$

Proof: (Sketch) Pick $A > 1$, with precise value to be specified later. For $k \ge \frac{n}{A}$, the Corollary follows directly from the Proposition, with different constant.

Now, assume $k < \frac{n}{A}$. Consider the proof of the proposition. In the proof, we have divided the interval $[-\infty, (1+\epsilon)\sqrt{2n+1}]$ into two parts. The first one was $I = [-\omega_0(\epsilon)\sqrt{2n+1}, (1+\epsilon)\sqrt{2n+1}]$, and for it precise estimates were given by Lemma 24.

The second one was I^c, and to treat it, we have used the brute force estimates of (26). Now, for A big enough, we can deal with the second case only, and we can use the estimates of (26), obtaining for $|x| > \sqrt{2n+1}$,

$$|D_k(x)e^{-\frac{x^2}{2}}| \le \|D_k\|_2 \cdot O\left(\left(\frac{B|x|}{\sqrt{k}}\right)^k e^{-\frac{x^2}{4}}\right). \tag{27}$$

Consider the function $g(x) = (\frac{B|x|}{\sqrt{k}})^k e^{-\frac{x^2}{4}}$. It's not difficult to check, that for $A = \frac{n}{k}$ big enough,

$$\int_{-\infty}^{-\sqrt{2n+1}} g^2(x)dx \le e^{-\Omega(n)},$$

completing the proof. ▮

6 Acknowledgement

I am grateful to my advisor Nathan Linial for generous help and advice. I would also like to thank Simon Litsyn and Yakov Itin for valuable conversations.

References

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

[2] Belov, preprint.

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

[4] P. Erdös, C. Ko and R. Rado, *Intersection theorem for systems of finite sets*, Quart. J. Math., Oxford (2) 12, 1961, 313-320.

[5] D. J. Kleitman, *On a combinatorial conjecture of Erdoš*, Journal of Comb. Theory 1, 1966, 209-214.

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

[7] L. Lovasz, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory, vol. IT-25, 1979, 1-7.

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

[9] 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.

[10] A. Samorodnitsky, *On the optimum of Delsarte's linear program*, preprint.

[11] A. Schrijver, *A comparison of the Delsarte and Lovasz bounds*, IEEE Trans. Inform. Theory, vol. IT-25, 1979, 425-429.

[12] A. Schrijver, *Association schemes and the Shannon capacity: Eberlein polynomials and the Erdös-Ko-Rado theorem*, in Colloquia Mathematica Societatis Janos Bolyai 25, Algebraic methods in graph theory, Szeged (Hungary) 1978, 671-688.

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

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

[15] A. F. Timan, **Theory of Approximation of Functions of a Real Variable**, Pergamon Press, 1963.