On Delsarte's Linear Programming Bounds for Binary Codes

Michael Navon Alex Samorodnitsky
The Hebrew University of Jerusalem

Abstract

We prove two results about the value of Delsarte's linear program for binary codes.

Our main result is a new lower bound on the value of the program, which, in particular, is nearly tight for low rate codes.

We also give an easy proof of an upper bound, which coincides with the best known bound for a wide range of parameters.

1 Introduction

1.1 Background

Error-correcting codes

A *binary error-correcting code C of length n and distance d* is a set of binary sequences of length n such that any two distinct sequences in C differ in at least d places.

Large error-correcting codes with large minimal distance are interesting objects for a variety of reasons. We briefly mention two of them.

Information transmission over a noisy communication channel. Transmitting words from a code of minimal distance $d = 2e + 1$ allows the receiver to recover the transmitted word if the number of errors is at most e. The rate of a code C is $R(C) = \frac{1}{n} \log |C|$. [1] This is the average number of information bits transmitted per one bit of a codeword. Larger codes allow for more efficient information transmission.

A binary sequence of length n can be viewed as a vector in the n-dimensional vector space Z_2^n over the field with two elements. A linear code is a linear subspace of the vector space Z_2^n. Given a linear code of length n and distance d one can construct a sample space of size $2^n/|C|$ supporting n random variables which are $(d-1)$-wise independent. Small sample spaces with this property have important applications in theoretical computer science and combinatorics [1].

This leads to one of the main problems in combinatorial coding theory, which is to determine the maximal size $A(n, d)$ of a binary error-correcting code C of length n and distance d.

Bounds on error-correcting codes

It is convenient to view the set $\{0, 1\}^n$ of binary sequences of length n as a metric space. The Hamming distance $\|x - y\|$ between two sequences x and y is the number of places in which these two sequences differ. (The Hamming weight $\|x\|$ of x is the distance between x and 0.) A metric ball of radius r and center x is the set of all points of distance at most r from x.

Let C be a code of length n, distance d, and cardinality $A(n, d)$. Then balls of radius $\lfloor \frac{d-1}{2} \rfloor$ with centers in C are disjoint. On the other hand, balls of radius d with centers in C cover the whole space (if there is a point in $\{0, 1\}^n$ which is left uncovered we can add it to the code). This argument gives the two classical bounds on $A(n, d)$, the upper bound of Hamming and the lower bound of Gilbert-Varshamov [14].

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

Let the distance d be linear in the length n of the code, $d = \delta n$ for $0 < \delta < \frac{1}{2}$. Then the bounds above show $A(n, d)$ to be exponentially large in n, but exponentially smaller than the size of the entire space, which is 2^n. In other words, the rate of a code of cardinality $A(n, d)$ is bounded away both from 0 and from 1. An asymptotic family $\{C_n\}$ of codes with growing d and n but a fixed ratio $\delta = d/n$ is of interest because these codes can deal with the same type of noisy channels, in which at most a $\frac{\delta}{2}$-fraction of the bits is corrupted. The best possible information transmission rate supported by $\{C_n\}$ is given by

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

The bounds on $A(n, d)$ translate[2] to bounds on the asymptotic rate $R(\delta)$.

$$1 - H(\delta) \leq R(\delta) \leq 1 - H\left(\frac{\delta}{2}\right)$$

[1] All the logarithms in this paper are to base 2 unless stated otherwise.

[2] Using well-known asymptotics of binomial coefficients: $\frac{1}{n} \log \binom{n}{\delta n} = H(\delta) - o(1)$.

The Gilbert-Varshamov lower bound $R(\delta) \geq 1 - H(\delta)$ is the best lower bound known to date.

The best upper bounds on $R(\delta)$ were obtained using the linear programming approach of Delsarte [5]. This approach views the distance distribution of a code C of length n as an $(n+1)$-dimensional vector $\bar{a} = (a_0 ... a_n)$. The i-th coordinate of $\bar{a}$ counts the pairs of points in C at distance i.

The convex hull of all such vectors, as C ranges over all codes of length n and minimal distance d is a polytope $\mathcal{P}_{n,d}$. Finding $A(n,d)$ amounts to optimizing a linear function $\sum_{i=1}^{n} a_i$ over this polytope, i.e., to a solution of a linear program. However, a complete list of the facets (defining inequalities) of $\mathcal{P}_{n,d}$ is not known.

Delsarte discovered a system of linear inequalities satisfied by every point in $\mathcal{P}_{n,d}$. This system of inequalities defines a polytope $\mathcal{D}_{n,d}$ that properly contains $\mathcal{P}_{n,d}$. The problem of optimizing $\sum_{i=1}^{n} a_i$ over this polytope leads to Delsarte's linear program, which we present in its dual form:

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

The functions $\{K_s\}_{s=0}^{n}$ are the Krawchouk polynomials. This is a family of orthogonal polynomials (with respect to the binomial measure $\mu(i) = \frac{\binom{n}{i}}{2^n}$ on $\{0...n\}$).

It is convenient to denote the value of this program by $A_{LP}(n,d)$. Since we will be mostly interested in asymptotic bounds let us define $R_{LP}(\delta)$ similarly to $R(\delta)$ with $A_{LP}(n,d)$ replacing $A(n,d)$. Clearly $R(\delta) \leq R_{LP}(\delta)$.

In [15] McEliece, Rodemich, Rumsey and Welch construct feasible solutions to (1) yielding

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

This is the best bound for $\delta \geq 0.273...$ The same paper gives the best bounds for smaller values of δ as well. This is more complicated and we do not go into details. [3]

Related work

We briefly describe several results especially relevant to our work.

The first result is an alternative proof of Delsarte's linear bound (1) due to Kalai and Linial [11]. The proof uses the language of Fourier analysis on the group Z_2^n. This approach turns out to be very useful.

[16] proves a lower bound

$$R_{LP}(\delta) \geq \frac{(1 - H(\delta)) + H(\frac{1}{2} - \sqrt{\delta(1-\delta)})}{2}. \tag{3}$$

This means, in particular, that the linear programming approach is limited, in the sense that the best upper bounds it can provide are far from the best known lower bounds. There are two proofs of this inequality. One is by construction. A point $\bar{a} \in \mathcal{D}_{n,d}$ with a large sum of the coefficients is given explicitly. (It is easy to see that $\bar{a} \notin \mathcal{P}_{n,d}$). The second proof shows that a feasible solution of the dual linear program has to have a large value in 0. The proof exploits various properties of Krawchouk polynomials.

In [2] Barg and Jaffe compute the values of $A(n,d)$ for $n = 1000$ and $1 \leq d \leq 500$. The obtained results argue convincingly that the bounds in [15] give the correct value of $R_{LP}(\delta)$.

[3] This involved dealing with a different but closely related linear program for *constant weight codes*.

Friedman and Tillich [8] suggest a different approach. Given a linear code of length n, minimal distance d, and cardinality k they construct a Cayley graph of degree n on Z_2^k whose second eigenvalue λ_2 is $n-2d$. Random walk techniques are then used to obtain lower bounds on λ_2. This approach leads to a new proof of (2) for linear codes. Two questions are left open: is it possible to extend the bound to general codes, and is this approach stronger than the linear programming approach, at least for linear codes.

1.2 Our results

Our main result is a strengthening of a lower bound of [16] on $R_{LP}(\delta)$. We view this result as a step in the following program:

1. Give a sequence of increasingly stronger lower bounds on $R_{LP}(\delta)$ culminating in proving the bound of [15] on $R_{LP}(\delta)$ to be tight. (Recall that this is supported by the numerical evidence from [2].) In particular show $R_{LP}(\delta) = H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right)$ for $\delta \geq 0.273...$

2. Construct an optimal solution of the primal linear program of Delsarte. Namely, a point $\bar{a} \in \mathcal{D}_{n,d}$ with $\frac{1}{n}\log\sum a_i = R_{LP}(\delta) - o(1)$. Note that this is not necessarily provided at the first step, since the optimality of bounds in [15] can be shown via the dual linear program. This is in fact the way we obtain our new bound. Compare to [16] where the lower bound on $R_{LP}(\delta)$ was first proved using the dual linear program, and then a matching primal solution was constructed.

3. Find a hyperplane (a linear inequality) separating $\bar{a}$ and $\mathcal{P}_{n,d}$. Add it to Delsarte's linear inequalities.

If this program works through, the augmented linear program has a good chance to be stronger than that of Delsarte. We note that it is not hard to come up with inequalities separating $\mathcal{P}_{n,d}$ and *some* subset of $\mathcal{D}_{n,d}$ (see [12], section 3.1). The real challenge is to separate $\mathcal{P}_{n,d}$ from the *bad* subset of $\mathcal{D}_{n,d}$.

This is our new lower bound.

Theorem 1.1:

$$R_{LP}(\delta) \geq \frac{1}{2}H(1-2\sqrt{\delta(1-\delta)}). \tag{4}$$

This bound is better than (3) for $\delta \geq 0.22...$

The new bound is especially nice for $\delta \approx \frac{1}{2}$. Let $\delta = \frac{1}{2} - \epsilon$, for a small ϵ. Then the lower bound of Gilbert-Varshamov gives $R(\delta) \geq 1 - H(\delta) \approx \frac{2}{\ln 2}\epsilon^2$. The bound (2) is optimal in this range and returns $H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right) \approx H\left(\epsilon^2\right) \approx 2\epsilon^2 \log\frac{1}{\epsilon} + \frac{1}{\ln 2}\epsilon^2$, which is clearly very far from the lower bound. The new bound is $\frac{1}{2}H(1-2\sqrt{\delta(1-\delta)}) \approx \frac{1}{2}H\left(2\epsilon^2\right) \approx 2\epsilon^2\log\frac{1}{\epsilon} + \left(\frac{1}{\ln 2} - 1\right)\epsilon^2$, which is very close to (2), and quite far from the Gilbert-Varshamov bound. It follows that the linear programming bound cannot give essentially new results for δ close to $\frac{1}{2}$.

The proof of the theorem makes extensive use of a family of orthogonal polynomials, the Hahn polynomials, which are related to the Krawchouk polynomials. Technically speaking, working with these polynomials allowed us to translate questions about l_1 norm of functions with respect to one measure to questions about l_2 norm with respect to a different measure. Euclidean norms are easier to handle and this advantage eventually leads to the improved bound.

Our second result is a new proof of an *upper* bound on $R_{LP}(\delta)$. We combine the ideas of Friedman and Tillich with the approach of Kalai and Linial and give a Fourier-analytic proof of (2) for general

codes. (In effect, this gives an affirmative answer to the first question of [8].) The original proof of this inequality in [15] is rather complicated, and makes an extensive use of the theory of orthogonal polynomials. Our proof requires only a basic familiarity with the properties of Fourier transform on Z_2^n.

The following proposition is the main step in the proof. (It is a direct corollary of lemmas 3.1 and 3.2.)

Proposition 1.2: *Let $0 < \lambda < 1$ be a constant. Let Λ be a nonnegative function on $\{0,1\}^n$ such that for any $x \in \{0,1\}^n$ holds*

$$\sum_{y:\ \|y-x\|=1} \Lambda(y) \geq \lambda n \Lambda(x). \tag{5}$$

Then if C is a code with minimal distance $d > \frac{1-\lambda}{2} \cdot n$

$$|C| \leq \frac{1}{2d/n - (1-\lambda)} \cdot \Big|supp(\Lambda)\Big|.$$

$\Big|supp(\Lambda)\Big|$ *is the cardinality of the support of Λ.*

This proposition draws a connection between two problems. One is an 'isoperimetric' problem of constructing nonnegative functions with a small support satisfying (5). The other - obtaining upper bounds for error-correcting codes. It would be interesting to see whether this connection can lead to better bounds on $A(n,d)$ and $R(\delta)$. In particular, whether one can take this route to prove the bounds of [15] for small values of δ.

Remark 1.3: It is not hard to see that the best bound on $R(\delta)$ obtainable via proposition 1.2 can not be better than $R_{LP}(\delta)$. (This gives a negative answer to the second question of [8].) We omit the details. ∎

The paper is organized as follows: the bound (2) is proved in Section 3. Theorem 1.1 is proved in Section 4. Section 2 is a reference section, containing useful definitions and facts. Several technical lemmas are proved in the Appendix, which also contains additional reference material, especially on orthogonal polynomials.

2 Preliminaries

2.1 Norms

Several norms of functions on $\{0,1\}^n$ will be used throughout the paper.

- $\|f\|_p^p = \left(\frac{1}{2^n}\sum_{x\in\{0,1\}^n}|f(x)|^p\right)^{1/p}$.
  A function f is *symmetric* if $f(y)$ depends only on $\|y\|$. For a symmetric function f:
- $\|f\|_p^p = \left(\frac{1}{2^n}\sum_{i=0}^n \binom{n}{i}|f(i)|^p\right)^{1/p}$.
  This may be viewed as a norm on a space of functions on $\{0...n\}$. We will also encounter two other (semi)-norms on this space of functions.

- The *even* l_1 norm: $\|f\|_{1,even} = \sum_{i\ even} \binom{n}{i} |f(i)|$.

- Norms with respect to the measure $\eta(i) = \frac{\binom{n}{i}^2}{\binom{2n}{n}}$ on $\{0...n\}$
  $|||f|||_p^p = \left(\frac{1}{\binom{2n}{n}} \sum_{i=0}^n \binom{n}{i}^2 |f(i)|^p \right)^{1/p}$.
  $|||f|||_{p,even}^p = \left(\frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 |f(i)|^p \right)^{1/p}$.

2.2 Fourier analysis on Z_2^n

Z_2^n is a finite Abelian group, therefore its characters $\{W_T\}_{T \in Z_2^n}$ constitute a group (the *dual group* which is isomorphic to Z_2^n.) The character W_T is a function from Z_2^n to $\{-1, 1\}$, defined as: $W_T(S) = (-1)^{|T \cap S|}$. The characters $\{W_T\}_{T \in Z_2^n}$ form an orthonormal basis in the space of real-valued functions on Z_2^n, equipped with uniform probability distribution.

For $f : Z_2^n \to \mathbf{R}$, define $\hat{f} : Z_2^n \to \mathbf{R}$, as $\hat{f}(T) = \frac{1}{2^n} \sum_{x \in Z_2^n} f(x) W_T(x)$. The function $\hat{f}$ is the *Fourier Transform* of f.

For $0 \leq s \leq n$ define $K_s : Z_2^n \to \mathbf{R}$ as $K_s(y) = \sum_{|T|=s} W_T(y)$. Clearly, $K_s(y)$ depends only on $\|y\|$, and can therefore be viewed as a function on integers $0 \leq i \leq n$. With some abuse of notation we may, therefore, view K_s as $K_s(x) = \sum_{j=0}^k (-1)^j \binom{x}{j} \binom{n-x}{k-j}$, the s-th *Krawchouk polynomial*. The orthogonality of the characters W_T naturally leads to orthogonality of the Krawchouk polynomials, viewed here as functions of real variable, with respect to the binomial measure $\mu(i) = \frac{\binom{n}{i}}{2^n}$ on $\{0...n\}$.

For $0 \leq s \leq n$ define $L_s : Z_2^n \to \mathbf{R}$ to be the characteristic function of the *s-th level* of $\{0,1\}^n$, namely $L_s(x) = 1$ iff the Hamming weight of x is s, and $L_s(x) = 0$ otherwise. Observe that $\widehat{2^n L_s} = K_s$.

We will need two additional generic properties of the transform. First, the Parseval identity: $\frac{1}{2^n} \sum_x f(x)g(x) = \sum_S \hat{f}(S)\hat{g}(S)$. Next, for two functions f, g on $\{0,1\}^n$ let $h = f * g$ be their convolution, given by $h(x) = \frac{1}{2^n} \sum_y f(y) g(x \oplus y)$. Then $\hat{h} = \hat{f} \cdot \hat{g}$.

3 The upper bound

In this section we prove the bound (2).

We define a function Δ on $\{0,1\}^n$ by $\Delta = 2^n L_1$ (i.e. $\hat{\Delta}(S) = K_1(S) = n - 2|S|$). Note that for any f on $\{0,1\}^n$, the function $f * \Delta$ sums at each point of $\{0,1\}^n$ the values of its neighbours.

We first show that a nonnegative function on $\{0,1\}^n$ whose transform is also nonnegative provides upper bounds on cardinalities of codes, if it is 'superharmonic', in the appropriate sense.

Lemma 3.1: *Let Γ be a function on $\{0,1\}^n$ with $\Gamma \geq 0$, $\widehat{\Gamma} \geq 0$, and assume that for some $0 < \lambda < 1$ holds $\Gamma * \Delta \geq \lambda n \cdot \Gamma$. Then if C is a code with minimal distance d, such that*

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

we have

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

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

Let $f = 1_C$, and let $F = \frac{2^n}{|C|} f * f$. Note $F(x) = 0$ for $0 < |x| \leq d$. Let g be a function such that $\hat{g} = F$. Then $g = \frac{2^{2n}}{|C|}\hat{f}^2 \geq 0$. We compute the inner product $\langle \Gamma * \Delta, g \rangle$ in two ways.

On one hand

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

On the other hand

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

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

Therefore

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

Since

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

We get

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

∎

Next, we show how to construct a function Γ required by the lemma.

Lemma 3.2: *Let Λ be a function on $\{0,1\}^n$ with $\Lambda \geq 0$, and assume that for some $0 < \lambda < 1$ holds $\Lambda * \Delta \geq \lambda n \cdot \Lambda$. Let $\Gamma = \Lambda * \Lambda$. Then Γ satisfies the conditions of lemma 3.1 with the same λ. Moreover*

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

Proof: The first part of the lemma is immediate. $\widehat{\Gamma} = \widehat{\Lambda}^2 \geq 0$, and $\Gamma * \Delta = \Lambda * (\Lambda * \Delta) \geq \lambda n \Lambda * \Lambda = \lambda n \Gamma$.

As for the second part, by the Cauchy-Schwarz inequality:

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

$$\frac{1}{2^n} \cdot |\mathrm{supp}(\Lambda)| \cdot \|\Lambda\|_2^2 = \frac{1}{2^n} \cdot |\mathrm{supp}(\Lambda)| \cdot \Gamma(0).$$

To see the last equality recall that $(f * f)(0) = \|f\|_2^2$ for any function f on $\{0,1\}^n$. ∎

Now we construct a function Λ. This function will be *symmetric*, namely its value at a point will depend only on the Hamming weight of the point. Such a function, of course, is fully defined by its values $\Lambda(0), ..., \Lambda(n)$ at Hamming weights $0...n$. For a symmetric function f on $\{0,1\}^n$ holds $(f * \Delta)(r) = rf(r-1) + (n-r)f(r+1)$, where we take $f(-1) = f(n+1) = 0$.

We start with a preliminary construction of a symmetric function f, setting $f(-1) = 0$, $f(0) = 1$, and defining $f(r)$ for $1 \leq r \leq n$ so that the relation $\lambda n f(r) = rf(r-1) + (n-r)f(r+1)$ is satisfied throughout[4]. It turns out that this function becomes negative at some point of the interval. A more precise information is given in the next lemma, proved in the Appendix.

[4]The function we have defined is actually a normalized Krawchouk polynomial.

Lemma 3.3: *Let $0 < \lambda < 1$ and let ϵ be an arbitrarily small positive number. There exists a sufficiently large $n_0 = n_0(\lambda, \epsilon)$ such that for any $n > n_0$ the function f defined above becomes negative in the interval $\left[0, \frac{1-\sqrt{1-\lambda^2}}{2}(1+\epsilon)\cdot n\right]$.*

Let r_0 be such that f is positive on $[0, r_0]$ and non-positive at r_0+1. Now define $\Lambda = f$ for $r \in [0, r_0]$ and $\Lambda = 0$ otherwise.

Lemma 3.4: *The function Λ satisfies the conditions of lemma 3.2.*

Proof: By the definition $(\Lambda * \Delta)(r) = \lambda n \Lambda(r)$, for for $r \leq r_0 - 1$ and for $r > r_0 + 1$. It remains to check these two boundary values. For $r = r_0 + 1$, $(\Lambda * \Delta)(r) \geq 0 = \lambda n \Lambda(r)$. For $r = r_0$,

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

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

Here the inequality holds since $f(r+1) \leq 0$. ∎

Now we are in position to prove the bound (2). Let $1 < d < \frac{n}{2}$ be given. Assume that d grows linearly with n, and is bounded away from 0 and $n/2$.

Choose $\lambda = 1 - \frac{2d-2}{n}$. Let Λ be the function defined above. Let $\Gamma = \Lambda * \Lambda$. Then by the preceding lemmas, for a code C with minimal distance d (and for a sufficiently large n)

$$|C| \leq n2^{n-1} \cdot \frac{\widehat{\Gamma}(0)}{\Gamma(0)} \leq \frac{n}{2} \cdot |supp\Lambda| \leq \frac{n}{2} \cdot \sum_{i=0}^{r_0} \binom{n}{i} \leq \frac{n}{2} \cdot 2^{n \cdot H\left(\frac{r_0}{n}\right)} \leq \frac{n}{2} \cdot 2^{n \cdot \frac{1-\sqrt{1-\lambda^2}}{2}(1+\epsilon)}.$$

Substituting the value of λ, we obtain, asymptotically, $R_{LP}(\delta) \leq \left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right) \cdot (1+\epsilon)$. This holds for any constant $\epsilon > 0$, and therefore

$$R_{LP}(\delta) \leq H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right),$$

completing the proof.

4 The lower bound

Consider the dual linear program of Delsarte (1). Let $A_{LP}(n,d)$ be the solution of the program. We want to show $A_{LP}(n,d)$ is large. [5]

It will be convenient to modify the program somewhat, introducing symmetry constraints on the coefficients b_s. Consider

$$min\left\{\Lambda(0) |\ \Lambda = \sum_{s=0}^{n} b_s K_s;\ b_s \geq 0;\ b_0 = 1;\ b_s = b_{n-s};\ \Lambda(i) \leq 0,\ d \leq i \leq n\right\} \tag{6}$$

We claim that we do not lose much by this modification.

Lemma 4.1: *([16]) Let $A_{SLP}(n,d)$ be the solution of the program. Then*

$$A_{LP}(n,d) \leq A_{SLP}(n,d) \leq 2 \cdot A_{LP}(n,d).$$

[5] Several of the steps below were already made in [16]. When the proofs are short we show them, for completeness.

Proof: The lemma is proved in subsection 5.2 in the Appendix. ∎
From now on, let $\Lambda = \sum_{s=0}^{n} b_s K_s$ be an optimal solution of (6), implying $\Lambda(0) = A_{SLP}(n,d)$. We will show $\Lambda(0)$ is large. Note that the proof of lemma 4.1 shows that $\Lambda(i)$ is supported on even i.

Lemma 4.2: *([16] There is an even integer $0 \le \ell \le d-1$ such that*

$$\frac{\binom{n}{\ell}}{2^n}\Lambda(\ell) \ge \frac{1}{2^n}\sum_{i=0}^{n}\binom{n}{i}|\Lambda(i)| = \|\Lambda\|_1.$$

Note that $\|\Lambda\|_1 \ge b_0 = 1$.

The lemma essentially says that a feasible solution of (6) has to be large at some point of the interval $[0, d-1]$, even when normalized by the binomial measure. This, intuitively, should not be easy to achieve, and this at the final count is what gives us the lower bound on the value of the program.
Proof: Averaging Λ with respect to the binomial measure we get

$$\frac{1}{2^n}\sum_{i\ even}\binom{n}{i}\Lambda(i) = b_0 = 1 \ge 0.$$

The claim now is immediate from nonpositivity of $\Lambda(i)$, $i \ge d$. ∎
We will also be interested in the norm of a function on the even points in $\{0...n\}$. Recall $\|f\|_{1,\ even} = \frac{1}{2^n}\sum_{i\ even}\binom{n}{i}|f(i)|$.

We now define the two major players in the proof: functions f and g on $\{0...n\}$.

$$f(t) = \|\sum_{s=0}^{t} b_s K_s\|_{1,\ even}$$

and

$$g(t) = \frac{\binom{n}{\ell}}{2^n}\cdot\sum_{s=0}^{t} b_s K_s(\ell).$$

From now on we will assume, without loss of generality, that n is odd.
The following lemma shows that in the middle of the domain f and g are not too far apart.

Lemma 4.3:

$$f\left(\frac{n-1}{2}\right) \le \frac{\|\Lambda\|_1}{2}.$$
$$g\left(\frac{n-1}{2}\right) \ge \frac{\|\Lambda\|_1}{2(n+1)}.$$

Proof: By the symmetry of the coefficients b_s, and recalling $K_s(i) = K_{n-s}(i)$ for even i:

$$f\left(\frac{n-1}{2}\right) = \frac{1}{2}f(n) = \frac{\|\Lambda\|_{1,even}}{2} \le \frac{\|\Lambda\|_1}{2}.$$

And

$$g\left(\frac{n-1}{2}\right) = \frac{1}{2}g(n) = \frac{1}{2}\cdot\frac{\binom{n}{\ell}}{2^n}\Lambda(\ell) \ge \frac{\|\Lambda\|_1}{2(n+1)}.$$

∎
However, and the next lemma is the key one, when the argument is far from $\frac{n}{2}$ the two functions are far apart. First, some notation. Let $d = \delta n$, and let $\rho = \frac{1}{2} - \sqrt{\delta(1-\delta)}$. Let $\gamma = \min\left\{2\rho, 2H^{-1}\left(\frac{1}{2}\right)\right\}$.

Lemma 4.4: *(The Gap lemma) There exists a function $\epsilon(n) = o(n)$, such that in the interval $0 \le t \le \gamma n - \epsilon(n)$*

$$f(t) \ge \Omega\left(n^6\right) \cdot |g(t)|.$$

The 'tension' between f and g - far apart at start but have to come closer at the middle will lead to the lower bound.

Proof: The lemma is proved in subsection 5.2 in the Appendix. ∎

Remark 4.5: In fact it can be shown, following [15] and [10] (see also [13]) that for $t = \tau n$, $\tau < \gamma$ holds

$$f(t) \ge 2^{\frac{1}{2}\left[\int_{\delta}^{\chi(\tau)} \log_2 R_\tau(y) dy\right] \cdot n - o(n)} \cdot |g(t)|,$$

where $R_\tau = \frac{A}{C} \cdot \left(\frac{B+\sqrt{B^2-4AC}}{2A}\right)^2$, and $A(y) = (1-y)^2$, $B_\tau(y) = 1 - 2y(1-y) - \tau(2-\tau)$, $C(y) = y^2$. We have not been able, so far, to use this estimate in order to improve the lower bound on R_{LP}. ∎

Now, we introduce the discrete derivatives $(\Delta f)(t) = f(t) - f(t-1)$ and $(\Delta g)(t) = g(t) - g(t-1)$. It turns out that if one of these derivatives is non-negligible for large s, $\Lambda(0)$ has to be large. However, we will see that because of the forementioned 'tension' between f and g, one of these functions will have to change rapidly in the vicinity of γn.

Lemma 4.6: *For all $1 \le t \le n$*

$$\Lambda(0) \ge \sqrt{\binom{n}{t}} \cdot |(\Delta f)(t)|,$$

and

$$\Lambda(0) \ge \sqrt{\frac{\binom{n}{t}2^n}{\binom{n}{\ell}}} \cdot |(\Delta g)(t)|.$$

Proof: Recall $\Lambda(0) = \sum_{s=0}^{n} b_s K_s(0) = \sum_{s=0}^{n} b_s \binom{n}{s}$. By the triangle inequality,

$$|(\Delta f)(t)| \le b_t \|K_t\|_{1,\ even} \le b_t \|K_t\|_1 \le b_t \|K_t\|_2 = b_t \sqrt{\binom{n}{t}}.$$

Therefore,

$$\Lambda(0) \ge b_t \binom{n}{t} \ge \sqrt{\binom{n}{t}} \cdot |(\Delta f)(t)|.$$

Similarly,

$$|(\Delta g)(t)| \le b_t \cdot \frac{\binom{n}{\ell}}{2^n} |K_t(\ell)|.$$

Since $\frac{\binom{n}{\ell}}{2^n} K_t^2(\ell) \le \|K_t\|_2^2$, we can upperbound $|K_t(\ell)| \le \sqrt{\frac{\binom{n}{t}2^n}{\binom{n}{\ell}}}$, and substituting

$$|(\Delta g)(t)| \le b_t \cdot \sqrt{\frac{\binom{n}{t}\binom{n}{\ell}}{2^n}}.$$

Therefore

$$\Lambda(0) \geq b_t \binom{n}{t} \geq \sqrt{\frac{\binom{n}{t}2^n}{\binom{n}{\ell}}} \cdot |(\Delta g)(t)|.$$

∎

Now we can prove the main claim of the section.

Theorem 4.7: *Let* $0 < \delta < \frac{1}{2}$. *Let* $\rho = \frac{1}{2} - \sqrt{\delta(1-\delta)}$. *Let* $\gamma = \min\left\{2\rho, 2H^{-1}\left(\frac{1}{2}\right)\right\}$. *Then*

$$R_{LP}(\delta) \geq \frac{1}{2}H(\gamma).$$

Proof: Let $\epsilon(n)$ be a function defined in lemma 4.4, and let $t = \gamma n - \epsilon(n)$. Then, by the Gap lemma $f(t) \geq \Omega\left(n^6\right) \cdot |g(t)|$. Now there are two cases to consider.

In the first case, there is $t \leq s \leq \frac{n-1}{2}$ for which the discrete derivative of g is large: $|(\Delta g)(s)| \geq \frac{1}{n^2}$. In such a case, by lemma 4.6

$$\Lambda(0) \geq \sqrt{\frac{\binom{n}{s}2^n}{\binom{n}{\ell}}} \cdot |(\Delta g)(s)| \geq \frac{1}{n^2} \cdot \sqrt{\frac{\binom{n}{t}2^n}{\binom{n}{\ell}}} \geq \frac{1}{n^2} \cdot \sqrt{\frac{\binom{n}{t}2^n}{\binom{n}{d}}}.$$

Therefore, recalling the well-known asymptotics of the binomial coefficients $\binom{n}{\alpha n} = 2^{H(\alpha)n - o(n)}$, we get

$$R_{LP}(\delta) \geq \frac{1}{2}\left(H(\gamma) + (1 - H(\delta))\right) \geq \frac{1}{2}H(\gamma).$$

If no such large derivative of g exists, we have $g(t) \geq g\left(\frac{n-1}{2}\right) - \sum_{s=t+1}^{\frac{n-1}{2}} |(\Delta g)(s)| \geq \frac{\|\Lambda\|_1}{4(n+1)}$. (Recall $\|\Lambda\|_1 \geq 1$.) Therefore $f(t) \geq \Omega\left(n^5\right) \cdot \|\Lambda\|_1$. On the other hand, $f\left(\frac{n-1}{2}\right) \leq \frac{\|\Lambda\|_1}{2}$.

It follows that for some $t \leq s \leq \frac{n-1}{2}$ the discrete derivative $|(\Delta f)(t)|$ is large, $|(\Delta f)(t)| \geq \|\Lambda\|_1 \geq 1$. This, by lemma 4.6, implies

$$\Lambda(0) \geq \sqrt{\binom{n}{t}} \cdot |(\Delta f)(t)| \geq \sqrt{\binom{n}{t}}.$$

The asymptotic form of this inequality is

$$R_{LP}(\delta) \geq \frac{1}{2}H(\gamma),$$

completing the proof. ∎

Theorem 1.1 now follows by observing that for $\rho \geq H^{-1}\left(\frac{1}{2}\right)$ the old lower bound (3) is already greater than the new one.

References

[1] N. Alon, L. Babai, and A. Itai *A fast and simple randomized parallel algorithm for the maximal independent set problem*, J. Algorithms, 7, vol. 4, 1986.

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

ignore

[3] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson, Randomness conductors and constant-degree expansion beyond the degree/2 barrier, in Proc. of 34st Symposium on Theory of Computing, ACM, pp. 659-668, 2002.

[4] L. Chihara and D. Stanton, *Zeroes of generalized Krawchouk polynomials*, J. Approx. Theory 60, no.1, 43-57, 1990.

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

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

ignore

[7] I. Dumer, D. Micciancio, and M. Sudan, *Hardness of approximating the minimum distance of a linear code*, IEEE Trans. Inform. Theory, vol. IT-49(1), 2003, 23-37.

[8] J. Friedman and J-P. Tillich, *Generalized Alon-Boppana Theorems and Error-Correcting Codes*, preprint, 2002.

ignore

[9] V. Guruswami and M. Sudan, *Improved decoding of Reed-Solomon codes and algebraic-geometry codes*, IEEE Trans. Inform. Theory, vol. IT-45(6), 1999, 1757-1767.

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

[11] G. Kalai and N. Linial, personal communication.

[12] N. Linial and A. Samorodnitsky, Linear codes and sums of characters, *Combinatorica*, vol. 22, 4, 2002.

[13] S. Litsyn, *New upper bounds on error exponents*, IEEE Trans. Inform. Theory.

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

[15] 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.

[16] A. Samorodnitsky, On the optimum of Delsarte's linear program, *Journal of Combinatorial Theory*, Ser. A, vol. 96, 2, 2001.

[17] A. Samorodnitsky, *Applications of Harmonic Analysis in Combinatorics and in Coding Theory*, PhD thesis, Department of Mathematics, Hebrew University, 1998.

[18] A. Samorodnitsky, Special classes of solutions for Delsarte's linear program, preprint (2000).

ignore

[19] M. Sudan, *Algorithmic Introduction to Coding Theory*, available from http://theory.lcs.mit.edu/ madhu/FT01

ignore

[20] L. Trevisan, Construction of Extractors Using Pseudorandom Generators, in Proc. of 31st Symposium on Theory of Computing, ACM, pp. 141-148, 1999.

5 Appendix

5.1 Properties of orthogonal polynomials

Properties of Krawchouk polynomials

Let $\mu(i) = \frac{\binom{n}{i}}{2^n}$, $0 \leq i \leq n$, be a discrete probability measure on $0...n$.

Definition:

$$K_s(x) = \sum_{k=0}^{s} (-1)^k \binom{x}{k}\binom{n-x}{s-k}. \tag{7}$$

Corollary 5.1:

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

Corollary 5.2: *$K_s(\frac{n}{2} - x)$ is an even (odd) function of x if s is even (odd).*

Orthogonality:

$$\sum_{i=0}^{n} \mu_K(i) K_s(i) K_t(i) = \begin{cases} \binom{n}{s} & \text{s = t} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Difference equation:

$$(n-i)K_s(i+1) - (n-2s)K_s(i) + iK_s(i-1) = 0. \tag{9}$$

First root behaviour, as n grows to infinity:

$$x_s = n \cdot \left(\frac{1}{2} - \sqrt{\frac{s}{n}(1 - \frac{s}{n})}\right) + o(n). \tag{10}$$

Properties of Hahn polynomials

Let $\eta(i) = \frac{\binom{n}{i}^2}{\binom{2n}{n}}$, $0 \leq i \leq n$, be a discrete probability measure on $0...n$. The orthogonal polynomials associated with μ are the Hahn polynomials $H_s(x)$, $s = 0...n$.

Definition:

$$H_s(i) = \frac{w_s}{\binom{n}{i}^2} \text{coef}_{y^i z^i}(1 - yz)^s (1+y)^{n-s}(1+z)^{n-s}, \tag{11}$$

where $w_s = \binom{2n}{s} - \binom{2n}{s-1}$.

Corollary 5.3: $H_0 \equiv 1,\ H_s(0) = w_s$

Orthogonality:

$$\sum_{i=0}^{n} \eta(i) H_s(i) H_t(i) = \begin{cases} w_s & \text{s = t} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

The first root behaviour:

$$x_{s,1} = 2n \cdot \frac{\frac{1}{4} - \frac{s}{2n}(1 - \frac{s}{2n})}{1 + 2\sqrt{\frac{s}{2n}(1 - \frac{s}{2n})}} + o(n). \tag{13}$$

5.2 Proofs of lemmas from the body of the paper

Proof: Of lemma 3.4. We will assume that f is positive on the interval $I = \left[0, \frac{1-\sqrt{1-\lambda^2}}{2}(1+\epsilon)\cdot n\right]$ and obtain a contradiction. Set $t = \frac{1-\sqrt{1-\lambda^2}}{2}$.

By the definition of f

$$f(r+1) = \frac{\lambda n f(r) - r f(r-1)}{n-r}.$$

Set $\theta(r) = \frac{f(r)}{f(r-1)}$. Since f is positive on I, for any r in I must hold $\theta(r) \geq \frac{r}{\lambda n} \geq \frac{(1+\epsilon)t}{\lambda}$. On the other hand, we will show that for any $r > (1+\epsilon/2)tn$ holds $\theta(r) < \theta(r-1)\cdot(1-\delta)$ for an absolute positive constant δ depending on ϵ. These two facts evidently cannot coexist, giving the desired contradiction.

It remains to see the exponential decrease of $\theta(r)$ for $r > (1+\epsilon/2)tn$. For any $r \geq 1$ holds $(n-r)\theta(r) = \lambda n - \frac{r}{\theta(r-1)}$, or $\theta(r) = \frac{\lambda n}{n-r} - \frac{r}{(n-r)\theta(r-1)}$. We claim that for $r > (1+\epsilon/2)tn$, and for any $x > 0$ holds

$$\frac{\lambda n}{n-r} - \frac{r}{(n-r)x} < (1-\delta)x,$$

for some $\delta > 0$. Indeed, this is a quadratic inequality which is easily seen to hold (since, by the definition of t, for $r > (1+\epsilon/2)tn$ and for a sufficiently small δ holds $(1-\delta)r(n-r) > \lambda^2 n^2$). ∎

Proof: Of lemma 4.1. One direction is immediate. $A_{LP}(n,d) \leq A_{SLP}(n,d)$ because $A_{LP}(n,d)$ is a minimum over a larger set.

For the other direction, let $\Lambda = \sum_{s=0}^{n} b_s K_s$ be an optimal solution of (1). Define $\Lambda_S = \sum_{s=0}^{n} c_s K_s$ by $c_s = \frac{b_s + b_{n-s}}{b_0 + b_n}$. We claim Λ_S is a feasible solution of (6). Indeed, by definition $c_s \geq 0, c_s = c_{n-s}$, and $c_0 = 1$. It remains to check the non-positivity of $\Lambda_S(i)$ for $d \leq i \leq n$. Here we use a property of Krawchouk polynomials: $K_s(i) = (-1)^i K_{n-s}(i)$. Using this it is easy to see that

$$\Lambda_S(i) = \sum_{s=0}^{n} c_s K_s(i) = \frac{1}{b_0+b_n}\sum_{s=0}^{n}(b_s + b_{n-s})K_s(i) = \frac{1}{b_0+b_n}\sum_{s=0}^{n}\left(1+(-1)^i\right)K_s(i).$$

This is zero if i is odd, and equals $\frac{2}{b_0+b_n}\Lambda(i)$ if i is even. Therefore, for $d \leq i \leq n$, $\Lambda_S(i)$ is nonnegative, and Λ_S is a feasible solution of (6). This implies

$$A_{SLP}(n,d) \leq \Lambda_S(0) = \frac{2}{b_0+b_n}\Lambda(0) = \frac{2}{b_0+b_n}A_{LP}(n,d) \leq 2A_{LP}(n,d).$$

∎

Proof: Of lemma 4.4. It is known ([17], [18] prop. 1.1) that a polynomial of low degree cannot attain its l_1 norm at points far from $n/2$.

To be more precise, let $\chi(\tau) = \frac{1}{2} - \frac{1}{2}\sqrt{\tau(2-\tau)}$ for $0 < \tau < \frac{1}{2}$. There exists a function $\alpha(n) = o(n)$ such that if P is a polynomial of degree $t = \tau n$, and $i \leq \chi(\tau)\cdot n - \alpha(n)$ then $\frac{\binom{n}{i}}{2^n}|P(i)| \leq O\left(\frac{1}{n^7}\right)\cdot\|P\|_1$. [6]

Let $P = \sum_{s=0}^{t} b_s K_s$. Then P is a polynomial of degree (at most) t. Observe that χ is a decreasing function. Therefore

$$\chi\left(\frac{t}{n}\right) \geq \chi\left(\gamma - \frac{\epsilon(n)}{n}\right) \geq \chi(\gamma) + \frac{\epsilon(n)}{n}\cdot|\chi'(\gamma)| - o\left(\frac{\epsilon(n)}{n}\right) \geq$$

[6] For completeness, the lemma is proved in Appendix (lemma 5.4).

$$\chi(2\rho) + \frac{\epsilon(n)}{n} \cdot |\chi\prime(\gamma)| - o\left(\frac{\epsilon(n)}{n}\right) = \delta + \frac{\epsilon(n)}{n} \cdot |\chi\prime(\gamma)| - o\left(\frac{\epsilon(n)}{n}\right).$$

This means that

$$\chi\left(\frac{t}{n}\right) \cdot n \geq d + \epsilon(n) \cdot |\chi\prime(\gamma)| - o\left(\epsilon(n)\right) \geq \ell + \epsilon(n) \cdot |\chi\prime(\gamma)| - o\left(\epsilon(n)\right).$$

We can choose $\epsilon(n)$ appropriately, so that $\epsilon(n) \cdot |\chi\prime(\gamma)| - o\left(\epsilon(n)\right) > \alpha(n)$. Now,

$$g(t) = \frac{\binom{n}{\ell}}{2^n} \cdot \sum_{s=0}^{t} b_s K_s(\ell) \leq O\left(\frac{1}{n^7}\right) \cdot \| \sum_{s=0}^{t} b_s K_s \|_1.$$

To complete the proof of the lemma we need to replace $\| \cdot \|_1$ with $\| \cdot \|_{1,even}$ in this inequality. For this purpose we use corollary 5.5 in the Appendix. By the corollary, there is a function $\phi(n) = o(n)$ such that if P is a polynomial of degree t, and $t < 2H^{-1}\left(\frac{1}{2}\right) \cdot n - \phi(n)$, then $\|P\|_1 \leq O\left(\sqrt{n} \cdot \|P\|_{1,even}\right)$. Choosing $\epsilon(n) \geq \phi(n)$:

$$g(t) \leq O\left(\frac{1}{n^7}\right) \cdot \| \sum_{s=0}^{t} b_s K_s \|_1 \leq O\left(\frac{1}{n^6}\right) \cdot \| \sum_{s=0}^{t} b_s K_s \|_{1,even} = O\left(\frac{1}{n^6}\right) \cdot f(t).$$

∎

5.3 Even norms

Let $\{0...n\}$ be endowed with the binomial probability measure $\mu(i) = \frac{\binom{n}{i}}{2^n}$. We consider two norms on the space of functions on $\{0...n\}$. The first is the usual l_1 norm given by $\|f\|_1 = \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} |f(i)|$. The second is the same, but restricting the functions to a subset of the domain - to all the even points $0 \leq i \leq n$. Namely, it is given by $\|f\|_{1,\ even} = \frac{1}{2^n} \sum_{i\ even} \binom{n}{i} |f(i)|$. The goal in this subsection is to show that on the subspace of polynomials of a sufficiently low degree, the new l_1 norm is not much smaller than the usual one.

Proposition 5.4: *Let $\beta < 2H^{-1}(\frac{1}{2})$. If f is a polynomial of degree at most βn then:*

$$\|f\|_{1,\ even} \geq \Omega\left(\frac{1}{\sqrt{n}}\right) \cdot \|f\|_1.$$

In fact we will need a following slightly more precise formulation which easily follows from the proof of the proposition:

Corollary 5.5: *There exists a function $\phi(n) = o(n)$ such that if f is a polynomial of degree at most $2H^{-1}(\frac{1}{2}) \cdot n - \phi(n)$ then:*

$$\|f\|_{1,\ even} \geq \Omega\left(\frac{1}{\sqrt{n}}\right) \cdot \|f\|_1.$$

To prove the proposition we will endow the space $\{0...n\}$ with the squared binomial probability measure $\eta(i) = \frac{\binom{n}{i}^2}{\binom{2n}{n}}$ and work with the family of Hahn polynomials $\{H_s\}_{i=0}^{n}$ orthogonal w.r.t. the measure η.

For notational convenience we define a new inner product on the space of functions on $\{0...n\}$. $\langle f, g \rangle = \frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 f(i) g(i)$. Recall also $\|\|f\|\|_2^2 = \frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 f^2(i)$.

We start the proof with several technical lemmas. Let $w_s = \|\|H_s\|\|_2^2 = \binom{2n}{s} - \binom{2n}{s-1}$.

Lemma 5.6: *For $s > 0$*

$$\left| \frac{1}{\binom{2n}{n}} \sum_{i \ even} \binom{n}{i}^2 H_s(i) \right| \leq O\left(\frac{w_s}{2^n}\right).$$

For $s = 0$

$$\frac{1}{\binom{2n}{n}} \sum_{i \ even} \binom{n}{i}^2 H_0(i) = \frac{1}{\binom{2n}{n}} \sum_{i \ even} \binom{n}{i}^2 \geq \frac{1}{2} - o(1).$$

Proof: First we compute, for $m \geq 0$, the alternating sum

$$\sum_{i=0}^{m} (-1)^i \binom{m}{i}^2 = (-1)^{\frac{m}{2}} \binom{m}{m/2}.$$

(When m is odd the expression on the left is to be interpreted as zero.) To see this observe that the left hand side is the coefficient at zero of

$$\left(1 - \frac{1}{y}\right)^m (1+y)^m = (-1)^m \mathrm{coef}_{y^m} \left(1 - y^2\right)^m.$$

Now, recall

$$H_s(i) = \frac{w_s}{\binom{n}{i}^2} \mathrm{coef}_{y^i z^i} (1 - yz)^s (1+y)^{n-s} (1+z)^{n-s}.$$

Therefore

$$\frac{1}{\binom{2n}{n}} \sum_{i \ even} \binom{n}{i}^2 H_s(i) = \frac{w_s}{\binom{2n}{n}} \cdot \sum_{i \ even} \mathrm{coef}_{y^i z^i} (1 - yz)^s (1+y)^{n-s} (1+z)^{n-s} =$$

$$\frac{w_s}{\binom{2n}{n}} \cdot \sum_{i \ even} \sum_{k=0}^{i} (-1)^k \binom{s}{k} \binom{n-s}{i-k}^2 = \frac{w_s}{\binom{2n}{n}} \cdot \sum_{k=0}^{s} (-1)^k \binom{s}{k} \sum_{i \geq k, \ i \ even} \binom{n-s}{i-k}^2 =$$

$$\frac{w_s}{\binom{2n}{n}} \cdot \sum_{k=0}^{s} (-1)^k \binom{s}{k} \sigma_k,$$

where $\sigma_k = \sum_{even \ j}^{n-s} \binom{n-s}{j}^2$ if k is even, and $\sigma_k = \sum_{odd \ j}^{n-s} \binom{n-s}{j}^2$ if k is odd. Therefore for $s > 0$ the last expression is

$$\frac{w_s}{\binom{2n}{n}} \cdot 2^s \sum_{i=0}^{n-s} (-1)^i \binom{n-s}{i}^2 = \frac{w_s \cdot 2^s}{\binom{2n}{n}} \cdot (-1)^{n-s} \binom{n-s}{(n-s)/2},$$

The absolute value of which is at most $O\left(\frac{w_s}{2^n}\right)$, proving the second statement of the lemma.

For $s = 0$, we get

$$\frac{1}{\binom{2n}{n}} \sum_{i \ even} \binom{n}{i}^2 H_0(i) = \frac{1}{\binom{2n}{n}} \sum_{i \ even} \binom{n}{i}^2 = \frac{\binom{2n}{n} + (-1)^n \binom{n}{n/2}}{2 \cdot \binom{2n}{n}} \geq \frac{1}{2} - o(1),$$

completing the proof. ∎

Lemma 5.7: *Let $\beta < 2H^{-1}(\frac{1}{2})$. If $s, t \leq \beta n$ then:*

$$\| |H_s| \|_{2, \ even}^2 \geq \left(\frac{1}{2} - o(1)\right) \cdot \| |H_s| \|_2^2.$$

And for $s \neq t$

$$| \langle H_s, H_t \rangle | \leq O\left(\frac{w_s w_t}{2^n}\right).$$

Proof: Expand

$$H_s H_t = \sum_{r=0}^{s+t} \alpha_r H_r,$$

where the coefficients α_r are known to be nonnegative [5]. First, consider α_0. By the orthogonality of Hahn polynomials, for $s \neq t$:

$$0 = \frac{1}{\binom{2n}{n}} \sum_{i=0}^{n} \binom{n}{i}^2 H_s(i) H_t(i) = \sum_{r=0}^{s+t} \alpha_r \cdot \frac{1}{\binom{2n}{n}} \sum_{i=0}^{n} \binom{n}{i}^2 H_r(i) = \alpha_0.$$

Therefore, by lemma 5.6

$$\left| \frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 H_s(i) H_t(i) \right| = \left| \sum_{r=1}^{s+t} \alpha_r \cdot \frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 H_r(i) \right| \leq$$

$$\sum_{r=1}^{s+t} \alpha_r \cdot \left| \frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 H_r(i) \right| \leq O\left(\frac{1}{2^n} \sum_{r=1}^{s+t} \alpha_r w_r \right) =$$

$$O\left(\frac{1}{2^n} \sum_{r=1}^{s+t} \alpha_r H_r(0) \right) = O\left(\frac{1}{2^n} (H_s H_t)(0) \right) = O\left(\frac{w_s w_t}{2^n} \right),$$

proving the claim of the lemma for the inner product. For the norms, we have similarly to the above,

$$H_s^2 = \sum_{r=0}^{s+t} \alpha_r H_r,$$

with $\alpha_0 = w_s$, and therefore

$$\frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 H_s^2(i) = \sum_{r=0}^{s+t} \alpha_r \cdot \frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 H_r(i) \geq$$

$$\alpha_0 \cdot \frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 H_0(i) - \left| \sum_{r=1}^{s+t} \alpha_r \cdot \frac{1}{\binom{2n}{n}} \sum_{i\ even} \binom{n}{i}^2 H_r(i) \right| \geq$$

$$\left(\frac{1}{2} - o(1) \right) \cdot w_s - \frac{w_s^2}{2^n}.$$

For $s \leq \beta n$, $w_s \leq 2^{2H\left(\frac{s}{2n}\right) \cdot n} \leq 2^{2H\left(\frac{\beta}{2}\right) \cdot n}$ is exponentially smaller than 2^n and therefore

$$|||H_s||_{2,\ even}^2 \geq \left(\frac{1}{2} - o(1) \right) \cdot w_s.$$

∎

Corollary 5.8: *Let f be a polynomial of degree $s \leq \beta n$. Then*

$$|||f|||_{2,\ even}^2 \geq \left(\frac{1}{2} - o(1) \right) \cdot |||f|||_2^2.$$

Proof: Expand $f = \sum_{r=0}^{s} \alpha_r H_r$. Then by the orthogonality of Hahn polynomials:

$$|||f|||_2^2 = \sum_{r=0}^{s} \alpha_r^2 |||H_r|||_2^2 = \sum_{r=0}^{s} \alpha_r^2 w_r.$$

On the other hand, by lemma 5.7

$$|||f|||_{2,\ even}^2 = \sum_{r=0}^{s} \alpha_r^2 \cdot |||H_r|||_{2,\ even}^2 + \sum_{r \neq t}^{s} \alpha_r \alpha_t \cdot \langle H_r, H_t \rangle \geq$$

$$\sum_{r=0}^{s} \alpha_r^2 \cdot |||H_r|||_{2,\ even}^2 - \sum_{r \neq t}^{s} |\alpha_r||\alpha_t| \cdot |\langle H_r, H_t \rangle| \geq$$

$$\left(\frac{1}{2} - o(1)\right) \cdot \sum_{r=0}^{s} \alpha_r^2 w_r - O\left(\sum_{r \neq t}^{s} |\alpha_r||\alpha_t| \cdot \frac{w_r w_t}{2^n}\right) \geq \left(\frac{1}{2} - o(1)\right) \cdot \sum_{r=0}^{s} \alpha_r^2 w_r.$$

∎

Now we can prove proposition 5.4

Proof:

$$\|f\|_1^2 \leq O\left(\sqrt{n} \cdot |||f|||_2^2\right).$$

On the other hand

$$\|f\|_{1,\ even}^2 \geq \Omega\left(\frac{1}{\sqrt{n}} \cdot |||f|||_{2,\ even}^2\right),$$

which, by preceding corollary is at least

$$\Omega\left(\frac{1}{\sqrt{n}} \cdot |||f|||_2^2\right).$$

∎

5.4 Gap lemma for the l_1 norm

Lemma 5.9: *Let $\chi(\tau) = \frac{1}{2} - \frac{1}{2}\sqrt{\tau(2-\tau)}$ for $0 < \tau < \frac{1}{2}$. There exists a function $\alpha(n) = o(n)$ such that if P is a polynomial of degree $t = \tau n$, and $i \leq \chi(\tau) \cdot n - \alpha(n)$ then $\frac{\binom{n}{i}}{2^n}|P(i)| \leq O\left(\frac{1}{n^7}\right) \cdot \|P\|_1$.*

Proof: Expand P as a linear combination of the Hahn polynomials $\{H_s\}$: $P = \sum_{s=0}^{t} c_s H_s$. Recall that these are the orthogonal polynomials w.r.t. the measure $\eta(i) = \frac{\binom{n}{i}^2}{\binom{2n}{n}}$ on $\{0...n\}$. Then

$$\frac{\binom{n}{i}}{2^n}|P(i)| = \frac{\sqrt{\binom{2n}{n}}}{2^n}\sqrt{\eta(i)} \cdot |P(i)| \leq O\left(\frac{1}{n^{1/4}}\right) \cdot \sqrt{\eta(i)P^2(i)} =$$

$$O\left(\frac{1}{n^{1/4}}\right) \cdot \sqrt{\eta(i)(\sum_{s=0}^{k} c_s H_s(i))^2} \leq O\left(n^{1/4}\right) \cdot \sqrt{\sum_{s=0}^{k} c_s^2 \eta(i) H_s^2(i)}.$$

By the choice of t, for any $s \leq t$ we have $i \leq \chi(\tau) \cdot n - \alpha(n) \leq \chi\left(\frac{s}{n}\right) \cdot n - \alpha(n)$.

We need to know two facts about the Hahn polynomials (see subsection 2. The location of the first root x_s of H_s: $x_s \geq \chi\left(\frac{s}{n}\right) \cdot n - o(n)$. The second fact ([16]) is that there is a function $\beta(n) = o(n)$ such

that for any s and $j \le x_s - \beta(n)$ holds $\eta(j)H_s^2(j) \le \frac{1}{n^{15}} \cdot |||H_s|||_2^2 = \sum_{j=0}^{n} \eta(j)H_s^2(j)$. Therefore, for the appropriate choice of $\alpha(n)$, we have for any $s \le t$

$$\eta(i)H_s^2(i) \le \frac{1}{n^{15}} \cdot |||H_s|||_2^2.$$

Therefore, by orthogonality of the Hahn polynomials,

$$\frac{\binom{n}{i}}{2^n}|P(i)| \le \frac{1}{n^{\frac{29}{4}}}\sqrt{\sum_{s=0}^{t} c_s^2 |||H_s|||_2^2} = \frac{1}{n^{\frac{29}{4}}}|||P|||_2.$$

Note, that for any function f:

$$|||f|||_2 \le O(n^{\frac{1}{4}}) \cdot \|f\|_1.$$

And so, at last (for n big enough):

$$\frac{\binom{n}{i}}{2^n}|P(i)| \le O\left(\frac{1}{n^7}\right) \cdot \|P\|_1,$$

proving the lemma. ∎