

# Biorthogonal codes with spherically-restricted positions and Krawtchouk polynomials<sup>1</sup>

ILYA DUMER

dumer@ee.ucr.edu

OLGA KAPRALOVA

okapralova@ee.ucr.edu

University of California at Riverside, USA

**Abstract.** Consider a biorthogonal code  $\text{RM}(1, m)$  defined on the hypercube  $E_2^m$  and let all its positions be restricted to  $m$ -tuples of a given Hamming weight  $b$ . In this paper, we define the parameters of this punctured code  $\text{RM}(1, m, b)$ . It is shown that the overall weight of any code vector is determined by the weight  $w$  of its information block and depends on the absolute values of the Krawtchouk polynomials  $K_b^m(w)$ . We then show that the minimum code weight of code  $\text{RM}(1, m, b)$  is achieved at the minimum input weight  $w = 1$  for any  $b < m/2$ . We further refine codes  $\text{RM}(1, m, b)$  by limiting the input weights  $w$  and show that some of the resulting codes meet the Griesmer bound.

## 1 Introduction

*Preliminaries.* In this paper we study punctured codes obtained by restricting biorthogonal Reed-Muller codes  $\text{RM}(1, m)$  to small subsets of their positions. Codes  $\text{RM}(1, m)$  have length  $n = 2^m$ , dimension  $m + 1$ , and distance  $2^{m-1}$ . Biorthogonal codes enjoy very simple and powerful decoding procedures. In particular, maximum likelihood decoding of these codes has complexity order of  $nm$  while a more limited recursive decoding reduces complexity to the order of  $n$ . For large  $n$ , both algorithms allow to correct most error patterns within decoding radius  $n(1 - \epsilon)/2$ , where  $\epsilon > 0$  is an arbitrarily small (or even vanishing) parameter. For a given  $\epsilon$ , full list decoding within this radius still has linear (in  $n$ ) complexity  $n\epsilon^{-2}$ . However, biorthogonal codes require very large block lengths to achieve good performance when used on the high-noise AWGN-channels. Therefore, in this paper we wish to modify these codes by restricting their huge length  $2^m$  to a much smaller order. More specifically, we will only use positions represented by the binary tuples  $(x_1, \dots, x_m)$  of a given Hamming weight  $b$ . Then we will derive the parameters of the obtained codes. We will also introduce a new precoding technique that drastically increases code distance and yields some code families that meet the Griesmer bound.

---

<sup>1</sup>This research was supported by NSF grant 1102074 and ARO grant W911NF-11-1-0027.

## 2 Punctured codes

Let  $F(1, m)$  be the set of  $m$ -variate affine functions. Then the binary code  $\text{RM}(1, m)$  is a set of truth-tables  $\mathbb{F}_2^m \xrightarrow{f} \mathbb{F}_2$ ,  $f \in F(1, m)$ . In other words, we form code vectors  $(\dots, f(x), \dots)$  on all  $2^m$  positions  $x = (x_1, \dots, x_m)$ . Now let  $S(b)$  be the Hamming sphere of radius  $b \in [1, m-1]$  in  $\mathbb{F}_2^m$ , which includes all vectors  $x$  of the Hamming weight  $\text{wt}(x) = b$ . In this paper we consider the punctured codes  $P(m, b) \equiv \text{RM}(1, m, b)$ , which are formed by the maps  $S(b) \xrightarrow{f} \mathbb{F}_2$  and have length  $n_b = \binom{m}{b}$ .

**Definition 1.** A code  $P(m, b)$  consists of the vectors  $(\dots, f(x), \dots)$  taken over positions  $x \in S_b$ , and affine functions  $f \in F(1, m)$ .

More generally, we can consider some subset of radii  $B \subseteq \{1, \dots, m-1\}$ . Then  $P(m, B)$  is obtained from  $\text{RM}(1, m)$  by restricting its positions to the set of spherical layers  $S(B) = \cup_{b \in B} S(b)$ . Obviously,  $P(m, B)$  is a linear code of length

$$n(B) = \sum_{b \in B} n_b.$$

Another generalization arises if we extend this construction to a general Reed-Muller code  $\text{RM}(r, m)$ . In this case we consider Boolean polynomials  $f \in F(r, m)$  of degree  $r$  or less. For any  $t \leq r$ , we then define the incidence matrix  $W_{t,b}(m)$  of  $t$ -subsets vs.  $b$ -subsets, which is an  $\binom{m}{t} \times \binom{m}{b}$  matrix whose rows represent the values of the Boolean monomials of degree  $t$  taken on a sphere  $S(b)$ . Then a punctured code  $P(r, m, b)$  is the linear span of the matrix<sup>2</sup>

$$G(r, m, b) = \bigcup_{i=0}^r W_{i,b}(m),$$

that is a matrix constructed by stacking matrices  $W_{0,b}(m), \dots, W_{r,b}(m)$  one on top of the other. For a general order  $r \geq 2$ , finding parameters of the spherically restricted codes  $P(r, m, b)$  leads to some non-trivial problems associated with the weight distributions  $A(w)$  of the original RM codes  $\text{RM}(r, m)$ . Therefore, we will address the simplest case of codes  $P(m, B)$  obtained for  $r = 1$ .

## 3 Spherical restrictions of biorthogonal codes

Consider parameters of code  $P(m, b)$ . This code is the linear span of the matrix

$$G(m, b) = \bigcup_{i=0}^1 W_{i,b}(m), \quad (1)$$

Now let  $g = (g_0, g_1, \dots, g_m)$  be an information block of code  $P(m, b)$  and let  $w = \text{wt}(g_1, \dots, g_m)$  denote the Hamming weight of the last  $m$  information bits.

<sup>2</sup>We show in the sequel that this matrix can have linearly dependent rows even for  $r = 1$ .

In other words, given the input affine function

$$g(x) = g_0 + \sum_{i=1}^m g_i x_i \quad (2)$$

we count the number  $w$  of coefficients  $g_i = 1$  in the linear form  $g_{\text{lin}}(x) = g(x) - g_0$ . For integers  $m, b, w$ , we use the binary Krawtchouk polynomials [1]

$$K_b^m(w) = \sum_{j=0}^b (-1)^j \binom{w}{j} \binom{m-w}{b-j}. \quad (3)$$

The following lemma defines the weight of any vector  $y_g = gG(m, b)$ .

**Lemma 1.** *For any binary input vector  $g = (g_0, g_1, \dots, g_m)$ , the output vector  $y_g = gG(m, b)$  has weight*

$$\text{wt}(y_g) = n_b/2 - (-1)^{g_0} K_b^m(w)/2. \quad (4)$$

*Proof.* Consider a point  $x \in S_b$  and define the respective supports of  $g$  and  $x$  by

$$\mathbf{I} = \{i : g_i = 1, 1 \leq i \leq m\} \quad \text{and} \quad \mathbf{X} = \{i : x_i = 1, 1 \leq i \leq m\}.$$

Then the linear form

$$g_{\text{lin}}(x) = \sum_{i \in \mathbf{I}} g_i x_i = \sum_{i \in \mathbf{I} \cap \mathbf{X}} g_i x_i$$

is nonzero iff the set  $\mathbf{I} \cap \mathbf{X}$  has odd size  $j$ . The number of such points  $x \in S_b$  is

$$S_b^m(w) = \sum_{j=1, \text{odd}}^b \binom{w}{j} \binom{m-w}{b-j}.$$

So,

$$\text{wt}(y_g) = \begin{cases} S_b^m(w), & \text{for } g_0 = 0 \\ n_b - S_b^m(w), & \text{for } g_0 = 1 \end{cases} \quad (5)$$

which can be rewritten as (4).  $\square$

Next, consider the left null space  $\mathcal{N}(m, b) = \{g \in \mathbb{F}_2^{m+1} : gG(m, b) = 0\}$  of matrix  $G(m, b)$ . By definition of this matrix, the corresponding affine function (2) satisfies equality  $g(x) \equiv 0$  for all  $x \in S(b)$ . In this case, we also write  $g(x) \in \mathcal{N}(m, b)$ .

**Lemma 2.** *Matrix  $G(m, b)$  generates code  $P(m, b)$  of dimension  $m$ . Its null space  $\mathcal{N}(m, b)$  is generated by the single function*

$$g^\perp(x) = \begin{cases} 1 + \sum_{i=1}^m x_i & \text{for odd } b \\ \sum_{i=1}^m x_i & \text{for even } b. \end{cases} \quad (6)$$

*Proof.* It is easy to verify that function (6) belongs to  $\mathcal{N}(m, b)$ . Also, (5) shows that  $\text{wt}(y_g) = 0$  only if  $S_b^m(w) = 0, n_b$ . The latter implies that  $\text{wt}(g) = m$  and  $g_0 = m \pmod{2}$ . Thus,  $\mathcal{N}(m, b)$  includes one nonzero function (6) and code  $P(m, b)$  has dimension  $m$ .  $\square$

Note also that removing the last row in  $G_m(m, b)$  makes it a full-rank generator matrix for code  $P(m, b)$ . Our next goal is to estimate the minimum distance

$$d(m, b) = n_b/2 - \max_w |K_b^m(w)|/2.$$

First, note that we have two trivial cases

$$d(m, 1) = d(m, m-1) = 1 \quad \text{for all } m \geq 2. \quad (7)$$

In the sequel, we consider the nontrivial cases  $b \in [2, m-2]$  and replace matrix  $G(m, b)$  with two matrices

$$\bar{\mathcal{G}}(m, b) = \left[ \begin{array}{c|c} G(m-1, b) & G(m-1, b-1) \\ \hline 0 & 11 \dots 1 \end{array} \right] \quad (8a)$$

$$\mathcal{G}(m, b) = \left[ \begin{array}{c|c} G(m-1, b) & G(m-1, b-1) \end{array} \right]. \quad (8b)$$

To obtain matrix  $\bar{\mathcal{G}}(m, b)$ , we re-order the columns of  $G(m, b)$ , by taking  $\binom{m-1}{b}$  positions  $x \in S(b)$  with  $x_m = 0$  and then using  $\binom{m-1}{b-1}$  positions with  $x_m = 1$ . Then we remove the last (linearly dependent) row in  $\bar{\mathcal{G}}(m, b)$  and obtain the generator matrix  $\mathcal{G}(m, b)$  of code  $P(m, b)$ .

We will now use recursive representation of the matrix  $\mathcal{G}(m, b)$ . Given an information block  $g$ , we consider two cases:

(A)  $g$  belongs to one of the null-spaces:

$$g \in \mathcal{N}(m-1, b) \text{ or } g \in \mathcal{N}(m-1, b-1).$$

(B)  $g$  belongs to neither of the null-spaces:

$$g \notin \mathcal{N}(m-1, b) \text{ and } g \notin \mathcal{N}(m, b-1).$$

The following Lemma 3 concerns case A.

**Lemma 3.** *Let  $g(x) \in \mathcal{N}(m, b_1)$ , then  $g(x) = 1$  for all  $x \in S(b_2)$  if  $b_2 \not\equiv b_1 \pmod{2}$ .*

*Proof.* For  $g(x)$  in (6), lemma follows by comparing the parity of  $b_1$  and  $b_2$ .  $\square$

Thus, given a vector  $g$  such that  $gG = 0$  on one sub-matrix  $G = G(m-1, b), G(m-1, b-1)$ , we obtain an all-one vector on the other. Correspondingly, case A gives minimum weight

$$\text{wt}_{\min}(gG) = \min \left\{ \binom{m-1}{b-1}, \binom{m-1}{b} \right\}.$$

In case B we will use induction to bound the minimum distance  $d(m, b)$ . This induction will suffice for all  $m$  except the boundary case  $m = 2b \pm 1$ , which we handle separately in Lemma 5. Consider the sum of distances

$$\delta(m, b) = d(m-1, b) + d(m-1, b-1) \quad (9)$$

of two subcodes generated by the recursive representation (8b). The following Lemma allows us to analyze the minimum distance  $d(m, b)$  by induction.

**Lemma 4.** For any  $b \in [2, m - 2]$ ,

$$d(m, b) \geq \min \left\{ \delta(m, b), \binom{m-1}{b-1}, \binom{m-1}{b} \right\} \quad (10)$$

*Proof.* The two latter estimates come directly from case A. For case B, we simply use (8b), which shows that vector  $g\mathcal{G}(m, b)$  has weight  $\text{wt}_g(m, b) = \text{wt}_g(m-1, b) + \text{wt}_g(m-1, b-1)$ .  $\square$

Note that case A gives the last two estimates in (10), which are tight. The first estimate uses case B and is tight only if both minimum distances  $d(m-1, b)$  and  $d(m-1, b-1)$  are reached on the same input function  $g(x)$ . Below, we first consider two special cases, for which the estimate  $\delta(m, b)$  of case B is not tight.

**Lemma 5.**  $d(2b+1, b) = d(2b+1, b+1) = \binom{2b}{b-1}$ .

*Proof.* We use the original equality (4) and estimate the values of the Krawtchouk polynomials. These polynomials satisfy the following relations [1]:

$$K_b^m(w+1) = \frac{m-2b}{m-w} K_b^m(w) - \frac{w}{m-w} K_b^m(w-1) \quad (11a)$$

$$K_{b+1}^m(w) = K_{b+1}^m(w-1) - K_b^m(w) - K_b^m(w-1) \quad (11b)$$

$$K_b^m(w) = (-1)^w K_{m-b}^m(w) \quad (11c)$$

First, (11c) shows we can consider only one case, for example  $d(2b+1, b)$ , since

$$|K_b^{2b+1}(w)| = |K_{b+1}^{2b+1}(w)|.$$

Second, (11c) with (11b) show that we can consider only even values of  $w$  to find  $\max |K_b^{2b+1}(w)|$ , since for  $w > 1$

$$K_b^{2b+1}(w) = -K_b^{2b+1}(w-1).$$

Then relation (11a) shows that for even  $w$

$$\left| \frac{K_b^{2b+1}(w+1)}{K_b^{2b+1}(w)} \right| = \frac{w+1}{2b+1-w} < 1, \text{ if } w < b.$$

So, for  $1 \leq w < b$ , function  $|K_b^{2b+1}(w)|$  decreases in  $w$ . It has maximum at  $w = 1$ , which gives minimum-distance codewords on the weight-1 input functions  $g(x) = x_i$ ,  $1 \leq i \leq 2b$ . To extend our proof for  $w \geq b$ , we use the relation [1]

$$K_b^m(w) = (-1)^b K_b^m(m-w),$$

which shows that  $|K_b^{2b+1}(1)|$  is still the maximum value for all  $w \in [1, 2b]$ . Finally,  $d(2b+1, b)$  is derived by calculating

$$K_b^{2b+1}(1) = \binom{2b}{b} - \binom{2b}{b-1}.$$

$\square$

**Theorem 1.** Code  $P(m, b)$  has minimum distance

$$d(m, b) = \begin{cases} \binom{m-1}{b-1}, & \text{if } m > 2b \\ 2\binom{m-2}{b}, & \text{if } m = 2b \\ \binom{m-1}{b}, & \text{if } m < 2b \end{cases} \quad (12)$$

*Proof.* We first bound  $d(m, b)$  from above. Then equalities (12) are immediately satisfied if we take

$$\begin{aligned} g(x) &= x_i \text{ for any } i \in [1, m-1] \text{ and } m > 2b; \\ g(x) &= 1 + x_i + x_j \text{ for any } 1 \leq i < j \leq m-1, \quad m = 2b; \\ g(x) &= 1 + x_i \text{ for any } i \in [1, m-1] \text{ and } m < 2b. \end{aligned}$$

Next we use equalities (12) to prove the lower bounds for  $d(m, b)$  in case B. Our base case includes equalities (7). By induction, we assume that (12) holds for some  $m$  and any  $b \in [2, m-1]$ . Then we consider the inductive step  $m \rightarrow m+1$  and recalculate function (9):

$$\delta(m+1, b) = d(m, b) + d(m, b-1)$$

using (12). Direct substitution shows that

$$\delta(m+1, b) = \begin{cases} \binom{m}{b-1}, & \text{if } m > 2b \\ 2\binom{m-1}{b}, & \text{if } m = 2b-1 \\ \binom{m}{b}, & \text{if } m < 2b-2 \end{cases}$$

To complete our inductive step for case B, we add the two special cases  $m = 2b-2, 2b$  (in both cases  $\delta(m+1, b)$  gives a low estimate and the inductive step  $m \rightarrow m+1$  fails). Then Lemma 5 gives

$$d(m+1, b) = \begin{cases} d(2b+1, b) = \binom{2b}{b-1}, & \text{if } m = 2b \\ d(2b-1, b) = \binom{2b-2}{b}, & \text{if } m = 2b-2 \end{cases}$$

Finally, to minimize the bounds in (10), we compare the obtained estimates of  $d(m, b)$  for both cases A and B. Comparing (12) with (10), we see that both estimates are identical for the first and the last line of (12). The second line of (12) also gives the lowest estimate, since for  $m = 2b$

$$2\binom{m-2}{b} < \binom{m-1}{b-1} = \binom{m-1}{b},$$

which completes the proof of the theorem.  $\square$

Theorem 1 can also be reformulated for the Krawtchouk polynomials. First, note [1–3] that the Krawtchouk polynomial  $K_b^m(w)$  has  $b$  simple (non-repeating) roots  $0 < r_1 < \dots < r_b < m$ , which are symmetric with respect to  $m/2$ . These roots are also interlaced with  $b - 1$  roots of the derivative of  $K_b^m(w)$ . Due to this,  $K_b^m(w)$  decreases in the subinterval  $I_1 = [1, r_1]$  and oscillates in the second subinterval  $I_2 = [r_1, m/2]$ . Here  $r_1 > 1$  for all  $b < m/2$  and  $r_1 = 1$  for  $b = m/2$ . Now we reformulate Theorem 1 as follows.

**Corollary 2.** *For all integers  $b \in [1, m - 1]$  and  $w \in [1, m - 1]$ , polynomial  $K_b^m(w)$  has maximum absolute value at  $w = 1$ :*

$$|K_b^m(1)| \geq |K_b^m(w)| \quad (13)$$

except for  $b = m/2$ , in which case the maximum is achieved at  $w = 2$ .

*Proof.* Equality (4) shows that the minimum weight in  $P(m, b)$  is obtained at the maximum  $|K_b^m(w)|$  of the Krawtchouk polynomial. This maximum, according to the proof of Theorem 1 is achieved for  $b \neq m/2$  at functions  $g(x)$  that have weight  $w = 1$  (recall that we do not count coefficient  $g_0$ ) or weight  $w = 2$  for  $b = m/2$ .  $\square$

We note that the result similar to Corollary 2 is known in the asymptotic setting [4], where it is shown that condition (13) holds for all  $1 \leq w \leq m/2$  given parameters  $\eta \in (0, 1)$ ,  $b = (1 + \eta)m/2$ , and sufficiently large  $m > m_0(\eta)$ . Our Corollary 2 extends this result for all  $b$  and  $m$ .

Since the minimum weight  $d_{\min}$  in codes  $P(m, b)$  is achieved at small input weights  $w = 1, 2$ , our next goal is to increase  $d_{\min}$  by limiting the range of input weights  $w$ . We describe this technique in the following section for the second layer  $b = 2$ .

## 4 Precoding on the second layer

Consider a code  $P(m, \{1, 2\})$  of length  $m(m+1)/2$ . By Lemma 1, an information word  $g$  of weight  $w$  generates a codeword of weight

$$w + w(m - w) = w(m - w + 1). \quad (14)$$

Thus, the minimum distance of  $P(m, \{1, 2\})$  keeps growing as long as the entire weight range  $[w_0, w_1]$  of nonzero information blocks gets closer to the midpoint  $(m + 1)/2$ . Therefore we will now restrict the set  $\mathbf{G}$  of possible information blocks and consider this set as a code  $\mathbf{G}[m, k, \delta]$  of dimension  $k$  and distance  $\delta$ . Our encoding now becomes a two-step procedure. First, the information word  $u \in \mathbb{F}_2^k$  is encoded into a vector  $g \in \mathbf{G}$  of length  $m$ , which in turn is encoded into  $y \in P(m, B)$ . This procedure yields the smaller code  $P_{\mathbf{G}}(m, B)$  and can be depicted as follows

$$u \in \mathbb{F}_2^k \xrightarrow{\text{Precoding}} g \in \mathbf{G} \xrightarrow{P(m, B)} y \in P_{\mathbf{G}}(m, B). \quad (15)$$

For example, the parity-check code  $\mathbf{G}[m, m-1, 2]$  increases the distance (14) of  $P(m, \{1, 2\})$  from  $m$  to  $2m-2$ , while reducing its dimension  $m$  to  $m-1$ . Some of the resulting codes can have good parameters. For example, code  $\mathbf{G}[7, 6, 2]$  generates a  $[28, 6, 12]$  linear code that meets the Griesmer bound. Similarly, the (extended) Hamming code  $\mathbf{G}[m, m - \lceil \log_2 m \rceil - 1, 4]$  increases the distance of  $P(m, \{1, 2\})$  almost fourfold, while insignificantly reducing the dimension of this code.

As another example, we take odd  $s$  and consider the code  $\mathbf{G}[m = 2^s - 1, 2s, 2^{s-1} - 2^{(s-1)/2}]$  which is the dual of the double-error-correcting BCH code [7]. Then the two-layer code  $P_{\mathbf{G}}(m, \{1, 2\})$  has length  $m(m+1)/2$  and distance  $(m^2 - 1)/4$ , which is very close to half the code length.

Note also that any linear code  $\mathbf{G}$  can be replaced with a nonlinear code, such as the Kerdock code with an even parameter  $s$  in the above example. Indeed, it is easy to see from Lemma 1 that a non-linear precoding of vectors  $g'$  and  $g''$  separated by some distance  $w$  gives vectors

$$y(g') = g'W_{1,b}(m), \quad y(g'') = g''W_{1,b}(m) : d(y(g'), y(g'')) = S_b^m(w).$$

Thus, the parameters of codes  $P_{\mathbf{G}}(m, B)$  can be improved by using the precoded sets  $\mathbf{G}$  with good distance distributions.

Finally, let us describe some infinite families of codes that achieve the Griesmer bound. Let  $\mathbf{G}(s) = \mathbf{G}[2^s - 1, s, 2^{s-1}]$  denote the shortened code RM(1,  $s$ ). We then use encoding (15) to consider code  $P_{\mathbf{G}(s)}(2^s - 1, B)$  on the spherical layers  $B = \{1, 2\}$ .

**Lemma 6.** *Code  $P_{\mathbf{G}(s)}(2^s - 1, B)$  meets the Griesmer bound for  $B = \{1, 2\}$ .*

*Proof.*  $\mathbf{G}(s)$  is a constant-weight equidistant code that has length  $m = 2^s - 1$  and nonzero weight  $w = 2^{s-1}$ . Then code  $P_{\mathbf{G}(s)}$  retains the dimension  $s$  of code  $\mathbf{G}(s)$ , has length and distance

$$m + \binom{m}{2} = 2^{2s-1} - 2^{s-1}, \quad d = w + w(m - w) = 2^{2s-2}.$$

Thus,  $P_{\mathbf{G}(s)}(2^s - 1, B)$  meets the Griesmer bound  $n \geq \sum_{i=0}^{s-1} \lceil \frac{d}{2^i} \rceil = 2^{2s-1} - 2^{s-1}$ .  $\square$

Note also that  $P_{\mathbf{G}(s)}$  is a constant-weight equidistant code with all weights and pairwise distances equal to  $2^{2s-2}$ . It is also easy to verify that the extended set  $\{1, 2, 2^s - 2, 2^s - 3\}$  and all of its subsets  $B$  also give codes  $P_{\mathbf{G}(s)}(2^s - 1, B)$  that meet the Griesmer bound. Both families of codes were previously designed using different techniques (see, for example, [5]). Our input-weight-limiting technique can also be extended to the higher layers  $b > m/2$  (some of the preliminary results will be reported in [6]) and also for codes of moderate length  $n$  (see [8]). In summary, this technique holds substantial promise and can be investigated further for general RM codes, including efficient decoding algorithms.



## 5 Open problems

Code design described above combines single-layer-restricted biorthogonal codes with the precoded information blocks. Some of these codes have high distance and meet or approach the Griesmer bound. To improve this design, we wish to use the following directions. First, our design can be extended using multi-layer constructions, which bring code distance closer to the Griesmer bound. Another important direction is to extend our spherically-restricted design to the general codes  $RM(r, m)$ . Finally, our preliminary observations indicate that the spherically-punctured codes can enable efficient decoding procedures that can correct high-noise errors and operate close to channel capacity.

## 6 Acknowledgment

The authors thank I. Krasikov for helpful remarks.

## References

- [1] I. Krasikov and S. Litsyn, “Survey of Krawtchouk Polynomials,” *DIMACS: Discrete Math and Theor. Comp. Sci.*, Vol. 56, pp. 199-2011, 2001.
- [2] V. Levenshtein, “Krawtchouk polynomials and universal bounds for Hamming spaces,” *IEEE Trans. Inform. Theory*, Vol. 41:5, pp. 1303-1321, 1995.
- [3] G. Kalai and N. Linial, “On the distance distribution of codes,” *IEEE Trans. Inform. Theory*, Vol. 41:5, pp. 1467—1472, 1995.
- [4] N. Alon and B. Sudakov, “Bipartite Subgraphs and the Smallest Eigenvalue,” *Combinatorics, Probability and Computing*, Vol. 9, pp. 1-12, 2000.
- [5] T. Helleseth, “Further classifications of codes meeting the Griesmer bound,” *IEEE Trans. Inform. Theory*, vol. 30:2, pp. 395-403, 1984.
- [6] I. Dumer and O. Kapralova, “Spherically punctured biorthogonal codes,” ISIT 2012, Boston, MA, July 1-6, 2012 (accepted).
- [7] F.J. MacWilliams and N.J.A. Sloane, “The Theory of Error-Correcting Codes,” *North-Holland, Amsterdam*, 1981.
- [8] Available online at <http://www.ee.ucr.edu/~okapralova/codes.html>