BOUNDS FOR BINARY CODES RELATIVE TO PSEUDO-DISTANCES OF k POINTS

CHRISTINE BACHOC, GILLES ZÉMOR

ABSTRACT. We apply Schrijver's semidefinite programming method to obtain improved upper bounds on generalized distances and list decoding radii of binary codes.

1. INTRODUCTION

Let $H_n := \mathbb{F}_2^n$ denote the binary Hamming space, endowed with the Hamming distance. One of the longstanding problems of coding theory is to find estimates for the maximum cardinality $A(n,d)$ of a code $C \subset H_n$ with the constraint that the Hamming distance of any pair of distinct elements of C is at least equal to d. The best known upper bound for $A(n,d)$ is obtained with the so-called linear programming method, due to Philippe Delsarte, and is the optimal value of a linear program (LP for short) ([9], [8, Chapter 9]). Because linear programs come with efficient algorithms, this method yields good numerical bounds for given parameters (n,d). Moreover, close to optimal explicit feasible solutions have been found from which upper bounds in the form of explicit functions of n and d have been derived [12], as well as an upper bound in the asymptotic range [13]. After these significant achievements, the subject fell into a period of about twenty years during which nothing really new was discovered, until A. Schrijver in [15] obtained improved upper bounds on $A(n,d)$ for some small values of the parameters (n,d), using semidefinite programming. Although these improvements are numerically not all that impressive, the method behind them introduces genuinely new ideas. In order to explain them, it is good to go back to Delsarte's method. Let us recall that the variables of the Delsarte linear program represent the distribution of the Hamming distance in the constrained code. More precisely, let

$$x_i := \frac{1}{\mathrm{card}(C)} \mathrm{card}\{(x,y) \in C^2 : d(x,y) = i\}.$$

Then the main idea is to observe that these variables satisfy certain linear inequalities, the non trivial ones being related to the Krawtchouck polynomials $K_k^n(x)$, namely, for $0 \le k \le n$,

$$\sum_{i=0}^{n} K_k^n(i) x_i \ge 0.$$

Schrijver's new idea [15] is to exploit constraints on *triples* of points $(x,y,z) \in C^3$ rather than deal only with pairs. It turns out that the natural constraints are semidefinite positive (SDP) instead of linear. The variables of the program are

$$x_{a,b,c} := \frac{1}{\mathrm{card}(C)} \mathrm{card}\{(x,y,z) \in C^3 : d(y,z) = a, d(x,z) = b, d(x,y) = c\}$$

and the SDP constraints take the form

$$\sum_{a,b,c} x_{a,b,c} S(a,b,c) \succeq 0,$$

Date: February 15, 2010.

where $\succeq 0$ stands for "is positive semidefinite", for some symmetric matrices $S(a,b,c)$. These SDP constraints are closely related to the action of the symmetric group S_n on the functional space $\mathbb{R}^{H_n}$; more precisely, each S_n-irreducible module occuring in $\mathbb{R}^{H_n}$ gives rise to an SDP inequality with matrices of size equal to its multiplicity. It should be noted that the full group of automorphisms $\mathrm{Aut}(H_n)$ acts multiplicity free on the same space $\mathbb{R}^{H_n}$, and that it is the true reason why in the case of the Delsarte method, the constraints are linear.

The aim of the present paper is to show that the Schrijver method can be used not only to strengthen the LP bounds, but also to give bounds for other problems, to which the LP method does not apply. Indeed, in recent years several generalizations of the Hamming distance, in the form of functions (we will call them pseudo-distances) of $k \geq 3$ elements of H_n have attracted attention. We consider here three such functions $f(x_1,\ldots,x_k)$, namely the generalized Hamming distance $d(x_1,\ldots,x_k)$, the radial distance $r(x_1,\ldots,x_k)$ and the average radial distance $\overline{r}(x_1,\ldots,x_k)$. They share the crucial property of being invariant by the action of the automorphism group $\mathrm{Aut}(H_n)$ of the Hamming space.

The generalized Hamming *weights* of linear codes were introduced by Ozarow and Wyner [14] in view of cryptographic applications related to the so-called wire-tap channel. The concept was later made popular for its own sake by Wei [19]. The notion was extended to the non linear setting in [7] in order to derive bounds on generalized weights. The generalized Hamming *distance* $d(x_1,\ldots,x_k)$ of k points is the number of coordinates where the k points are not all equal. Thus $d(x_1,x_2)$ is the classical Hamming distance. In [7], the authors derive bounds for generalized distances, focusing on asymptotics, which are analogs of the classical Hamming, Plotkin and Elias-Bassalygo bounds. In the case of linear codes the best known asymptotic upper bounds were obtained in [1].

The radial distance and the average radial distance are related to the notion of list decoding. The *radial distance* or *radius* of k elements is the smallest radius of a Hamming ball that contains the k points. If a code C has the property that the radius of any k-tuple of pairwise distinct points is at least equal to some value r, then any ball of the Hamming space of radius $r-1$ intersects C in at most $k-1$ points. Thus a decoding procedure that outputs every codeword at distance at most $r-1$ of any given received vector yields a *list* of codewords of cardinality at most $k-1$. The search for large codes with given minimum k-radius is also studied in the litterature as the quest for dense *multiple packings*: indeed, a code of minimal k-radius r provides a packing of balls (centered at the codewords, of radius $r-1$) such that any element of H_n belongs to at most $k-1$ balls. These notions have a long standing history, going back to problems in Euclidean geometry and to early coding theory. They came back into the limelight some ten years ago when Sudan discovered his now famous algorithm for list decoding of Reed-Solomon codes [16]. Blinovskii [5] establishes asymptotic bounds on the maximal number of elements of a code with given minimal radius: in the process he defines an auxiliary quantity, the *average radius* of k elements that we will also investigate.

In general, we are given a function f from H_n^k into the set of non-negative integers, and we denote by $A_{k-1}(n,f,m)$ the maximal number of elements that a binary code C can have under the constraint that $f(x_1,\ldots,x_k) \geq m$ for every k-tuple of pairwise distinct codewords. Our goal is to show that the SDP method gives good upper bounds for $A_2(n,f,m)$ for modest values of n, when compared with the classical bounds. Our results provide strong motivation for the development of the SDP method, which is far from being at the same stage of achievement as the LP method.

The paper is organized as follows: Section 2 provides a description of the orbits of $\mathrm{Aut}(H_n)$ acting on H_n^k. This preliminary task is essential since the pseudo-distances we are dealing with only depend on these orbits. Section 3 recalls the definitions and basic properties of the three particular functions we consider. Section 4 defines the code invariants associated to these functions and recalls their significance for applications. Section 5 settles the "classical" bounds. These bounds already

appear in the litterature ([7], [5], [6]) but not in the precise form needed here: either they are settled only for linear codes, or the concern is in the asymptotic setting and they are not as tight as they can be for small parameters. Section 6 recalls the SDP method of [15] using the language of group representation, i.e. following [4], [17], [18]. Section 7 provides some numerical results.

2. The orbits of $Aut(H_n)$ acting on H_n^k

The automorphism group of the binary Hamming space $H_n := \mathbb{F}_2^n$, denoted by $\mathrm{Aut}(H_n)$, is the semi-direct product of the group of translations by elements of H_n with the group of permutations on the n coordinates. The group $\mathrm{Aut}(H_n)$ acts two-point homogeneously on H_n, which means that the orbits of $\mathrm{Aut}(H_n)$ acting on H_n^2 are characterized by the Hamming distance. In other words

$$(x,y) \sim_{\mathrm{Aut}(H_n)} (x',y') \Leftrightarrow d(x,y) = d(x',y').$$

Here $(x,y) \sim_{\mathrm{Aut}(H_n)} (x',y')$ stands for: there exists $g \in \mathrm{Aut}(H_n)$ such that $g(x) = x'$ and $g(y) = y'$. We want to study the action of $\mathrm{Aut}(H_n)$ on k-tuples $(x_1,\dots,x_k) \in H_n^k$. We introduce:

Definition 2.1. *For $\underline{x} = (x_1,\dots,x_k) \in H_n^k$, and for $u \in \mathbb{F}_2^k$, let*

$$n_u(\underline{x}) := \mathrm{card}\{\, j,\ 1 \le j \le n : ((x_1)_j,\dots,(x_k)_j) = u\}$$

and let the "weight distribution" of $\underline{x}$ be defined by

$$\mathcal{W}(\underline{x}) := (n_u(\underline{x}))_{u\in\mathbb{F}_2^k}.$$

For $u \in \mathbb{F}_2^k$, the word obtained from u by flipping zeros and ones, will be denoted by $\overline{u}$. In other words $\overline{u} = u + 1^k$. One of $\{u,\overline{u}\}$ has the form $0w$ with $w \in \mathbb{F}_2^{k-1}$. Let

$$n_w(\underline{x}) := n_{0w}(\underline{x}) + n_{1\overline{w}}(\underline{x}).$$

The "symmetrized weight distribution" of $\underline{x}$ is defined by:

$$\overline{\mathcal{W}}(\underline{x}) := (n_w(\underline{x}))_{w\in\mathbb{F}_2^{k-1}}$$

Remarks:

(i) It is nice to identify $\underline{x}$ with the (k,n) matrix $M(\underline{x})$ whose i-th line equals x_i. Then $n_u(\underline{x})$ is the number of columns of $\underline{x}$ which are equal to u:

$$M(\underline{x}) = \begin{array}{lll} x_1 & = 000\dots0 & \dots\dots \\ x_2 & = 111\dots1 & \dots\dots \\ \vdots & = \quad\vdots \quad\quad \vdots & \\ x_k & = \underbrace{111\dots1}_{n_u(\underline{x})} & \dots\dots \end{array}$$

(ii) We have $\sum_{u\in\mathbb{F}_2^k} n_u(\underline{x}) = \sum_{w\in\mathbb{F}_2^{k-1}} n_w(\underline{x}) = n$.

Proposition 2.2.

$$\underline{x} \sim_{\mathrm{Aut}(H_n)} \underline{y} \Leftrightarrow \overline{\mathcal{W}}(\underline{x}) = \overline{\mathcal{W}}(\underline{y}).$$

Proof. It is clear that $\underline{x} \sim_{\mathrm{Aut}(H_n)} \underline{y}$ iff $\underline{x}' \sim_{S_n} \underline{y}'$ where $\underline{x}' = (0, x_2 - x_1,\dots,x_k - x_1)$ and $\underline{y}' = (0, y_2 - y_1,\dots,y_k - y_1)$. Then $\overline{\mathcal{W}}(\underline{x}') = \overline{\mathcal{W}}(\underline{y}')$ iff $\mathcal{W}(\underline{x}') = \mathcal{W}(\underline{y}')$ and is left unchanged if the coordinates are permuted. Conversely, for an appropriate permutation σ of the coordinates, $\sigma(\underline{x}')$ has its columns reordered in lexicographic order. Another permutation τ has the same effect on $\underline{y}'$; since $\mathcal{W}(\sigma(\underline{x}')) = \mathcal{W}(\tau(\underline{y}'))$, it means that $\sigma(\underline{x}') = \tau(\underline{y}')$. $\square$

Remark 2.3.

(i) *If $k=2$, we have of course $n_1(\underline{x}) = d(x_1,x_2)$ and $n_0(\underline{x}) = n - n_1(\underline{x})$. In the case $k=3$, we have*

$$\begin{aligned} n_{10} + n_{11} &= d(x_1,x_2) \\ n_{01} + n_{10} &= d(x_2,x_3) \\ n_{01} + n_{11} &= d(x_3,x_1) \end{aligned}$$

and the triple $(d(x_1,x_2), d(x_2,x_3), d(x_3,x_1))$ uniquely determines the orbit of (x_1,x_2,x_3).

(ii) *For arbitrary k, taking into account the relation $\sum_w n_w = n$, the orbits of $\mathrm{Aut}(H_n)$ on H_n^k are described by $2^{k-1}-1$ independent parameters. In contrast, the orbits of k-tuples of elements of the unit sphere of the Euclidean space S^{n-1} under the action of the orthogonal group $O(\mathbb{R}^n)$ need only $\binom{k}{2}$ real numbers in order to be uniquely determined, namely the pairwise inner products of the k vectors. The orbits of H_n^k under $\mathrm{Aut}(H_n)$ are determined by the pairwise distances $d(x_i,x_j)$ only if $k=2,3$.*

(iii) *In the next section we introduce several functions $f(x_1,\dots,x_k)$ such that*

$$f(\sigma(x_1),\dots,\sigma(x_k)) = f(x_1,\dots,x_k)$$

for all $\sigma \in \mathrm{Aut}(H_n)$. It follows from the above description of the orbits of H_n^k that such functions have an expression of the form $f(x_1,\dots,x_k) = \tilde{f}(\overline{\mathcal{W}}(\underline{x}))$.

3. $Aut(H_n)$-INVARIANT FUNCTIONS ON H_n^k.

3.1. The generalized Hamming distance.

Definition 3.1. *The generalized Hamming distance of k elements of H_n is defined by:*

$$\begin{aligned} d(x_1,\dots,x_k) &= \mathrm{card}\{j,\ 1\le j\le n:\ \mathrm{card}\{(x_1)_j,\dots,(x_k)_j\}\ge 2\} \\ &= \mathrm{card}\{j,\ 1\le j\le n:\ ((x_1)_j,\dots,(x_k)_j)\notin\{0^k,1^k\}\} \end{aligned}$$

Proposition 3.2. *The following properties hold for the generalized Hamming distance:*

(i) *$d(x_1,x_2)$ is the usual Hamming distance.*

(ii) *For all permutation τ of $\{1,\dots,k\}$, $d(x_1,\dots,x_k) = d(x_{\tau(1)},\dots,x_{\tau(k)})$.*

(iii) *For all $\sigma \in \mathrm{Aut}(H_n)$, $d(x_1,\dots,x_k) = d(\sigma(x_1),\dots,\sigma(x_k))$. The generalized distance $d(x_1,\dots,x_k)$ is related to the weight distribution by:*

$$d(x_1,\dots,x_k) = \sum_{w\neq 0^{k-1}} n_w(\underline{x}).$$

(iv) *$d(x_1,\dots,x_{k-1},x_k) = d(x_1,\dots,x_{k-1})$ if x_k belongs to the affine subspace generated by $x_1,\dots,x_{k-1}$.*

(v) *"Triangular" inequality: for all $y\in H_n$,*

$$d(x_1,\dots,x_k) \le \frac{1}{k-1}\sum_{i=1}^{k} d(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_k).$$

(vi) *The distance of three points can be expressed in terms of pairwise Hamming distances:*

$$d(x_1,x_2,x_3) = \frac{1}{2}(d(x_1,x_2)+d(x_2,x_3)+d(x_3,x_1)).$$

(vii) *For more than three points we have only the inequality:*

$$d(x_1, x_2, \ldots, x_k) \leq \frac{1}{k-1} \sum_{1 \leq i < j \leq k} d(x_i, x_j).$$

(viii) *We also have the inequalities:*

$$d(x_1, \ldots, x_k) \leq \frac{1}{k-1} \sum_{i=1}^{k} d(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$$

$$d(x_1, \ldots, x_k) \geq \frac{1}{k} \sum_{i=1}^{k} d(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k).$$

Proof. Properties (i), (ii), (iii) are obvious.

If x_k belongs to the affine subspace generated by $x_1, \ldots, x_{k-1}$, then we can write $x_k = \sum_{i=1}^{k-1} \alpha_i x_i$ with $\sum_{i=1}^{k-1} \alpha_i = 1$. Consequentely, if $((x_1)_j, \ldots, (x_{k-1})_j) = 0^{k-1}$, respectively 1^{k-1}, then we have $((x_1)_j, \ldots, (x_k)_j) = 0^k$, respectively 1^k. It follows (iv) that $d(x_1, \ldots, x_{k-1}, x_k) = d(x_1, \ldots, x_{k-1})$.

The announced "triangular" inequality (v) is easily checked in the case $n = 1$. The general case follows from the fact that

$$d(x_1, \ldots, x_k) = \sum_{j=1}^{n} d((x_1)_j, \ldots, (x_k)_j). \tag{1}$$

Again because of (1), it is enough to prove (vi) (vii) and (viii) for $n = 1$. In this case, let the Hamming weight of $(x_1, \ldots, x_k)$ be denoted by w, then

$$d(x_1, \ldots, x_k) = \begin{cases} 1 & \text{if } 1 \leq w \leq k-1 \\ 0 & \text{if } w = 0, k. \end{cases}$$

and

$$\sum_{1 \leq i < j \leq k} d(x_i, x_j) = w(k-w).$$

Obviously $w(k-w) \geq k-1$ if $w \neq 0, k$ and equals 0 otherwise. Inequality (vii) follows. In the case $k = 3$, $w(k-w)$ takes only the values 0 and 2 hence (vi).

To prove (viii), notice that we have

$$\sum_{i=1}^{k} d(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k) = \begin{cases} k & \text{if } 2 \leq w \leq k-2 \\ k-1 & \text{if } w = 1, k-1 \\ 0 & \text{if } w = 0, k \end{cases}$$

hence the announced inequalities. □

3.2. **The radial distance.**

Definition 3.3. *The radial distance or radius of k elements of H_n is defined by:*

$$\begin{aligned} r(x_1, \ldots, x_k) &= \min\{r : \textit{ there exists } y \in H_n \textit{ s.t. } \{x_1, \ldots, x_k\} \subset B(y, r)\} \\ &= \min_y \{\max_{1 \leq i \leq k} d(y, x_i)\}. \end{aligned}$$

Proposition 3.4. *The radial distance has the properties:*

(i) $r(x_1, x_2) = \lceil \frac{d(x_1, x_2)}{2} \rceil$.

(ii) *For all permutations τ of $\{1, \ldots, k\}$, $r(x_1, \ldots, x_k) = r(x_{\tau(1)}, \ldots, x_{\tau(k)})$.*

(iii) *For all $\sigma \in \mathrm{Aut}(H_n)$, $r(x_1, \ldots, x_k) = r(\sigma(x_1), \ldots, \sigma(x_k))$.*

(iv) *For all k,*

$$r(x_1,\dots,x_k) \geq \max_{1\leq i\leq k} r(x_1,\dots,x_{i-1},x_{i+1},\dots,x_k).$$

(v) *For $k=3$, we have*

$$r(x_1,x_2,x_3) = \max\{r(x_1,x_2), r(x_2,x_3), r(x_3,x_1)\}.$$

Proof. Properties (i) (ii) (iii) and (iv) are obvious.

Let $(x_1,x_2,x_3) \in H_n^3$. Without loss of generality we can assume that $d(x_2,x_3) \leq d(x_3,x_1) \leq d(x_1,x_2)$ and that $x_1 = 0$. With the notation of section 1 it amounts to assume that $n_{01} \leq n_{10} \leq n_{11}$. Let $y \in H_n$ be the center of a smallest ball containing the three words; clearly the coordinates of y at the positions corresponding to $w = 00$ in $M(\underline{x})$ must be equal to 0. Let y_w be the number of ones at the positions corresponding to w. We have:

$$\begin{aligned} d(y,x_1) &= y_{01} + y_{10} + y_{11} \\ d(y,x_2) &= y_{01} + n_{10} - y_{10} + n_{11} - y_{11} \\ d(y,x_3) &= n_{01} - y_{01} + y_{10} + n_{11} - y_{11} \end{aligned}$$

We choose y such that:

$$\begin{aligned} y_{01} &= 0 \\ y_{11} &= \lfloor \frac{n_{01}+n_{10}+2n_{11}}{4} \rceil \leq n_{11} \\ y_{10} &= \lfloor \frac{n_{10}-n_{01}}{4} \rceil \leq n_{10} \end{aligned}$$

Then one easily verifies that for $i = 1,2,3$, $d(y,x_i) \leq \lceil \frac{n_{10}+n_{11}}{2} \rceil$ thus the ball centered at y with radius $\lceil \frac{n_{10}+n_{11}}{2} \rceil$ contains the three words x_1,x_2,x_3. Since

$$n_{10} + n_{11} = d(x_1,x_2) = \max(d(x_2,x_3), d(x_3,x_1), d(x_1,x_2))$$

we have proved that

$$r(x_1,x_2,x_3) \leq \max\{r(x_1,x_2), r(x_2,x_3), r(x_3,x_1)\}.$$

□

Remark 3.5. *For $k \geq 4$ we cannot give a nice expression of $r(\underline{x})$ as an explicit function of $\overline{\mathcal{W}}(\underline{x})$. It should be noted that the determination of the center y and thus of $r(\underline{x})$ cannot be performed by a sequence of local decisions at each coordinate or even at each subset of coordinates corresponding to each u; in other words property* (1) *of $d(\quad)$ does not hold for r and it makes it more difficult to study. However for k randomly chosen points, the distances of each point to the center y of the smallest ball containing them are expected to have about the same value, in other words the points are expected to be close to the border of the ball. When it is the case, the radius of the k points is approximated by a much nicer function, called the average radius (or moment of inertia), introduced in* [5].

3.3. **The average radial distance.**

Definition 3.6. *The average radial distance or average radius (or moment distance or moment of inertia) of k elements of H_n is defined by:*

$$\overline{r}(x_1,\dots,x_k) = \min_y \frac{1}{k} \sum_{1\leq i\leq k} d(y,x_i).$$

Proposition 3.7. *The average radius has the properties:*

(i) $\overline{r}(x_1, x_2) = \frac{d(x_1,x_2)}{2}$.

(ii) *For all permutation τ of $\{1, \ldots, k\}$, $\overline{r}(x_1, \ldots, x_k) = \overline{r}(x_{\tau(1)}, \ldots, x_{\tau(k)})$.*

(iii) *For all $\sigma \in \mathrm{Aut}(H_n)$, $\overline{r}(x_1, \ldots, x_k) = \overline{r}(\sigma(x_1), \ldots, \sigma(x_k))$. In terms of the weight distribution $\overline{\mathcal{W}}(\underline{x}) = (n_w(\underline{x}))_{w \in \mathbb{F}_2^{k-1}}$,*

$$\overline{r}(x_1, \ldots, x_k) = \frac{1}{k} \sum_{w \in \mathbb{F}_2^{k-1}} \min(wt(w), k - wt(w)) n_w(\underline{x})$$

(iv) *For all k,*

$$\overline{r}(x_1, \ldots, x_k) \geq \frac{1}{k} \sum_{i=1}^{k} \overline{r}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$$

(v) *The above inequality is an equality for $k = 1 \mod 2$. In particular, we have*

$$\overline{r}(x_1, x_2, x_3) = \frac{\overline{r}(x_1, x_2) + \overline{r}(x_2, x_3) + \overline{r}(x_3, x_1)}{3}$$

(vi) *For all k,*

$$\overline{r}(x_1, \ldots, x_k) \leq \frac{2(k-1)}{k(k-2)} \sum_{i=1}^{k} \overline{r}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k).$$

(vii) *"Triangular" inequality: for all $y \in H_n$,*

$$\overline{r}(x_1, \ldots, x_k) \leq \frac{1}{k-1} \sum_{i=1}^{k} \overline{r}(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_k).$$

Proof. Properties (i) (ii) and the $\mathrm{Aut}(H_n)$-invariance are trivial.

If the j-th column of the matrix $M(\underline{x})$ equals $u \in \mathbb{F}_2^k$, the contribution of this column in $\sum_i d(y, x_i)$ is equal to $wt(u)$ if $y_j = 0$ and to $wt(\overline{u})$ if $y_j = 1$. So the minimum of this sum over all y equals

$$\sum_u \min(wt(u), wt(\overline{u})) n_u(\underline{x}).$$

which leads to the formula announced in (iii). It also shows that

$$\overline{r}(x_1, \ldots, x_k) = \sum_{j=1}^{n} \overline{r}((x_1)_j, \ldots, (x_k)_j).$$

Consequently, in order to prove the remaining assertions, we can assume $n = 1$. Let the weight of $\underline{x}$ be denoted by w. Without loss of generality we assume that either $w < k/2$ or $w = k/2$. This last case can only happen if $k = 0 \mod 2$. We prove (v) and (vi): in the case $w < k/2$, removing $x_i = 1$ makes $k\overline{r}(\underline{x})$ drop by 1 while removing $x_i = 0$ does not change $k\overline{r}(\underline{x})$. In the case $w = k/2$, $k\overline{r}(\underline{x})$ always drops by 1. In other words,

$$(k-1)\overline{r}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k) = \begin{cases} w - 1 & \text{if } x_i = 1 \text{ and } w < \frac{k}{2} \\ w & \text{if } x_i = 0 \text{ and } w < \frac{k}{2} \\ w - 1 & \text{if } w = \frac{k}{2} \end{cases}$$

and

$$(k-1) \sum_{i=1}^{k} \overline{r}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k) = \begin{cases} (k-1)w & \text{if } w < \frac{k}{2} \\ k(k-2)/2 & \text{if } w = \frac{k}{2}. \end{cases}$$

We obtain

$$\frac{1}{k}\sum_{i=1}^{k}\overline{r}(x_1,\dots,x_{i-1},x_{i+1},\dots,x_k)=\begin{cases}\overline{r}(\underline{x}) & \text{if } w\neq\frac{k}{2}\\ \frac{(k-2)}{(2k-2)}\overline{r}(\underline{x}) & \text{if } w=\frac{k}{2}\end{cases}$$

hence the inequalities

$$\frac{(k-2)}{(2k-2)}\overline{r}(\underline{x})\leq\frac{1}{k}\sum_{i=1}^{k}\overline{r}(x_1,\dots,x_{i-1},x_{i+1},\dots,x_k)\leq\overline{r}(\underline{x}).$$

If $k=1 \mod 2$, the case $w=k/2$ never happens so the second inequality is always an equality.

For the triangular inequality (vii), we find

$$k\sum_{i=1}^{k}\overline{r}(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_k)=$$

$$\begin{cases}(k-1)w+k & \text{if } w<\lfloor\frac{k}{2}\rfloor \text{ and } y=1\\ (k-1)w & \text{if } w=\frac{k}{2} \text{ and } y=1\\ kw & \text{if } w=\frac{k-1}{2} \text{ and } y=1\\ (k-1)w & \text{if } y=0\end{cases}$$

hence

$$k\sum_{i=1}^{k}\overline{r}(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_k)\geq(k-1)w=k(k-1)\overline{r}(\underline{x}).$$

□

3.4. **Relationships between d, r, $\overline{r}$.**

Proposition 3.8. *The following hold:*

(i) *For all $\underline{x}=(x_1,\dots,x_k)\in H_n^k$,*

$$\frac{1}{k}d(\underline{x})\leq\overline{r}(\underline{x})\leq r(\underline{x})\leq d(\underline{x})$$

and

$$\overline{r}(\underline{x})\leq\frac{1}{2}d(\underline{x}).$$

(ii) *For $k=2,3$, $d(\underline{x})=k\overline{r}(\underline{x})$.*

(iii) *If $\overline{r}(\underline{x})=r(\underline{x})$ then the center of any of the balls of minimal radius $r(\underline{x})$ containing the points $(x_1,\dots,x_k)$ is equidistant to these points. The converse is false, in the sense that the points may be equidistant to some y while $\overline{r}(\underline{x})<r(\underline{x})$.*

Proof. Since

$$\frac{1}{k}\sum_{i=1}^{k}d(y,x_i)\leq\max_i d(y,x_i),$$

we obviously have $\overline{r}(\underline{x})\leq r(\underline{x})$. From

$$1\leq\min(wt(w),k-wt(w))\leq k/2$$

for $w\neq 0^{k-1}$ and from the expressions given in Proposition 3.7 (iii) for $\overline{r}(\underline{x})$ and in Proposition 3.2 (iii) for $d(\underline{x})$ we have

$$\frac{1}{k}d(\underline{x})\leq\overline{r}(\underline{x})\leq\frac{1}{2}d(\underline{x}).$$

Let J be the set of coordinates where $\underline{x}_j \in \{0^k, 1^k\}$. If we choose y such that y_j agrees with $(x_i)_j$ for $j \in J$, then $d(y, x_i) \leq n - |J| = d(\underline{x})$. Thus $r(\underline{x}) \leq d(\underline{x})$. This concludes point (i).

(ii) is obvious from previous formulas.

Let us assume that $\overline{r}(\underline{x}) = r(\underline{x}) = r$ and let y be the center of a ball of radius r containing all x_i. Then we have the inequalities

$$r = \overline{r}(\underline{x}) \leq \frac{1}{k} \sum_i d(y, x_i) \leq \max_i d(y, x_i) = r$$

thus $\frac{1}{k} \sum_i d(y, x_i) = \max_i d(y, x_i)$ which means that all $d(y, x_i)$ are equal to r.

We build a counterexample with $k = 3$. If n_{01}, n_{10}, n_{11} are even numbers, the points will be equidistant to some point y with $y_w = n_w/2$. We assume moreover that $n_{01} \leq n_{10} \leq n_{11}$. From Proposition 3.4, $r(\underline{x}) = (n_{10} + n_{11})/2$ and from Proposition 3.7, $\overline{r}(\underline{x}) = (n_{01} + n_{10} + n_{11})/3$ so if $2n_{01} < n_{10} + n_{11}$ we are done. □

4. Code invariants and their significance

4.1. **Code invariants.**

Definition 4.1. *For any $C \subset H_n$, and for $f = d, r, \overline{r}$, we define*

$$f_{k-1}(C) = \min f(x_1, \ldots, x_k)$$

where the minimum is taken over all k-tuples of pairwise distinct elements of C. Moreover we define

$$d^{\text{aff}}_{k-1}(C) = \min d(x_1, \ldots, x_k)$$

where the minimum is taken over all k-tuples of affinely independent elements of C. Following standard notation in coding theory, we let $A_{k-1}(n, f, m)$ be the maximal number of elements that a code $C \subset H_n$ can have under the condition $f_{k-1}(C) \geq m$.

Proposition 4.2. *The following hold:*

(i) $d_1(C) = d^{\text{aff}}_1(C)$ *is the Hamming distance of the code C.*
(ii) $d_2(C) = d^{\text{aff}}_2(C)$.
(iii) *If C is a linear code, and $2^{t-1} < k \leq 2^t$, $d_{k-1}(C) = d^{\text{aff}}_t(C)$.*
(iv) *If C is a linear code, $d^{\text{aff}}_{k-1}(C)$ coincides with the minimum $(k-1)$-th generalized weight as defined in* [19], *namely:*

$$d^{\text{aff}}_{k-1}(C) = \min\{w(D) : D \subset C, D \; linear, \dim(D) = k - 1\}.$$

where $w(D)$ is the set of coordinates i at which at least one element of D is non zero.

Proof. Obvious. □

Remark. The quantity $d_k(C)$ is more natural and easier to deal with than the more intricate $d^{\text{aff}}_k(C)$. Unfortunately, $d_k(C)$ only coincides with the minimum k-th generalized weight of a linear code for $k = 1, 2$, hence the definition of $d^{\text{aff}}_k(C)$, originally stated in [7]. In [2] yet another generalisation of the minimum k-th generalized weight to non-linear codes is introduced that does not consider affinely independent sets of vectors. We will not dwell on the differences here and our study will mostly focus on the quantity $d_k(C)$ itself, of interest in its own right since it has a natural interpretation in terms of list decoding "radius" for lists of size k when decoding from erasures (see section 4.2 below).

Proposition 4.3. *For $f = d, d^{\text{aff}}, r, \overline{r}$ and for any code C,*

$$f_{k-1}(C) \leq f_k(C)$$

Proof. It follows from Propositions 3.2 (viii), 3.4 (iv) and 3.7 (iv) that for pairwise distinct $\underline{x}$

$$f(x_1, \ldots, x_{k+1}) \geq f_{k-1}(C)$$

respectively for affinely independent $\underline{x}$

$$d(x_1, \ldots, x_{k+1}) \geq d^{\text{aff}}_{k-1}(C).$$

Hence $f_{k-1}(C) \leq f_k(C)$. □

4.2. **List decoding.** A list decoding procedure is a decoding procedure that outputs a list of codewords. The length L of the list is determined in advance. This list is usually obtained by the enumeration of all codewords in a ball $B(y, r)$. For a given code C, the associated value of r is known as the L-list decoding radius of C:

Definition 4.4. *The L-list decoding radius $R_L(C)$ is the largest value of r such that, for all $y \in H_n$,*

$$\text{card}(B(y, r) \cap C) \leq L.$$

In the case $L = 1$, we recover the notion of the (unique) decoding radius of a code, $R_1(C) = \lfloor (d(C) - 1)/2 \rfloor$. This number is also the largest value of r such that the balls of radius r centered at the codewords have the property that any $L+1$ of them have an empty intersection. A set of balls with this property is called a L-multiple packing. Thus a classical packing of balls is a 1-multiple packing.

Proposition 4.5.

$$R_L(C) = r_L(C) - 1.$$

Proof. There exists $(x_1, \ldots, x_{L+1}) \in C^{L+1}$ and $y \in H_n$ such that for all $1 \leq i \leq L+1$, $x_i \in B(y, r_L(C))$ and $x_i \neq x_j$ thus $\text{card}(B(y, r_L(C)) \cap C) = L + 1$ and $R_L(C) < r_L(C)$. Moreover, if $r < r_L(C)$, $L + 1$ codewords cannot be elements of the same ball of radius r thus $R_L(C) = r_L(C) - 1$. □

The notion of list decoding can also be investigated in the framework of erasure decoding, see [11].

Definition 4.6. *The L-list decoding radius for erasures $R^{\text{er}}_L(C)$ is the largest value of r such that, for all $E \subset \{1, \ldots, n\}$, $\text{card}(E) \leq r$, and for any $y = (y_i)_{i \notin E} \in \{0, 1\}^{n - \text{card}(E)}$*

$$\text{card}(\{x \in C : (x_i)_{i \notin E} = y\}) \leq L.$$

The following proposition, which is a straightforward consequence of the definition of d_L, makes generalized distances relevant to erasure decoding [11, 20, 21].

Proposition 4.7.

$$R^{\text{er}}_L(C) = d_L(C) - 1.$$

5. Upper bounds for d_k, r_k, $\overline{r}_k$.

In this section we gather the analogs of the Singleton, Hamming, Plotkin and Elias bounds for $f = d, r, \overline{r}$. The methods are well-known and some of the bounds may be found explicitely in the litterature, but not always in form precise enough for numerical computation (in particular only asymptotic versions of the Elias bounds can be found) which we need to compare them to the new SDP bounds.

5.1. **The Singleton bound.** This bound for d is the most elementary and is a natural generalisation of the classical Singleton bound for the ordinary Hamming distance.

Proposition 5.1. *Let $C \subset H_n$. Then, if $d_{k-1}(C) \geq d_{k-1}$*

$$|C| \leq (k-1)2^{n-d_{k-1}+1}.$$

Proof. Consider the restriction of the codewords on a fixed set of $(n-d_{k-1}+1)$ indices. The number of possible images is of course $2^{n-d_{k-1}+1}$. If $|C| > (k-1)2^{n-d_{k-1}+1}$, there is a subset of k codewords having the same image. Thus they have a generalized Hamming distance at most equal to $d_{k-1}+1$ and we have a contradiction. □

It is worth noticing that the Singleton bound for $k=3$ is tight for $d=3$ and for $d=n$.

5.2. **Hamming type bound.** This volume type bound is established in [7][Prop II.I] for d_{k-1} and for linear codes and generalized to the non-linear case in [2]. We take the following notations: the number of elements of a ball of radius r in H_n is denoted b_r^n or b_r if n is clear from the context. We recall the formula

$$b_r^n = \sum_{k=0}^{r} \binom{n}{k}.$$

Proposition 5.2. *Let $C \subset H_n$. Then*

(i) *If $r_{k-1}(C) \geq r_{k-1}$ or $\overline{r}_{k-1}(C) \geq r_{k-1}$ then*

$$|C| \leq \frac{(k-1)2^n}{b^n_{r_{k-1}-1}}$$

(ii) *If $d_{k-1}(C) \geq d_{k-1}$ then*

$$|C| \leq \frac{(k-1)2^n}{b^n_{\lceil d_{k-1}/k \rceil - 1}}$$

(iii) *If $d^{\text{aff}}_{k-1}(C) \geq d_{k-1}$ then*

$$|C| \leq \frac{2^{n+k-2}}{b^n_{\lceil d_{k-1}/k \rceil - 1}}$$

Proof. (i) If $r_{k-1}(C) \geq r_{k-1}$ or $\overline{r}_{k-1}(C) \geq r_{k-1}$, from Proposition 3.8 (i) and Proposition 4.5 we have, for all $y \in H_n$, $\text{card}(B(y, r_{k-1}-1) \cap C) \leq k-1$. In order to establish the announced inequality, we count in two ways the elements of

$$E := \{(c,y), c \in C, y \in H_n : d(c,y) \leq r_{k-1}-1\}.$$

We have

$$\begin{aligned}
\text{card}(E) &= \sum_{c \in C} \text{card}\{y \in H_n : d(y,c) \leq r_{k-1}-1\} \\
&= |C| b^n_{r_{k-1}-1} \\
&= \sum_{y \in H_n} \text{card}\{c \in C : d(y,c) \leq r_{k-1}-1\} \\
&\leq \text{card}(H_n)(k-1) = (k-1)2^n.
\end{aligned}$$

(ii) If $d_{k-1}(C) \geq d_{k-1}$, from Proposition 3.8 (i) we have $r_{k-1}(C) \geq \lceil \frac{d_{k-1}}{k} \rceil$ thus we can apply the previous result.

(iii) Let $(x_1, \ldots, x_k) \in C^k$ be affinely independent and let $y \in H_n$. We have

$$\begin{aligned} d_{k-1} \leq d(x_1, \ldots, x_k) &\leq \frac{1}{k-1} \sum_{1 \leq i < j \leq k} d(x_i, x_j) \\ &\leq \frac{1}{k-1} \sum_{1 \leq i < j \leq k} (d(x_i, y) + d(y, x_j)) \\ &\leq \sum_{1 \leq i \leq k} d(x_i, y). \end{aligned}$$

Thus for some i, $1 \leq i \leq k$, $d(x_i, y) \geq \lceil \frac{d_{k-1}}{k} \rceil$. Since any subset of H_n with at least $2^{k-2} + 1$ elements contains k affinely independent ones, we have for all $y \in H_n$,

$$\mathrm{card}(B(y, \lceil \frac{d_{k-1}}{k} \rceil - 1) \cap C) \leq 2^{k-2}.$$

and we follow the same line as in (i). □

5.3. **Plotkin type bound.** This type of bound is usually derived from the estimate of the average value of f among C^k. This average value can be estimated when f can be calculated from its value at each coordinate, which is the case for $f = d, \overline{r}$.

We take the following notations: let C be a binary code with M elements; let w_j be the number of ones in the j-th column of the $M \times n$ matrix whose rows are the M elements of C. Let $J_k(C)$, respectively $J_k^{\mathrm{aff}}(C)$ be the set of k-tuples of pairwise distinct, respectively affinely independent codewords. We moreover define

$$j_k(x) := \begin{cases} 0 & \text{if } x \leq k-1 \\ \prod_{t=0}^{k-1}(x-t) & \text{if } x \geq k-1 \end{cases}$$

and

$$j_k^{\mathrm{aff}}(x) := \begin{cases} 0 & \text{if } x \leq 2^{k-2} \\ x \prod_{t=0}^{k-2}(x-2^t) & \text{if } x \geq 2^{k-2}. \end{cases}$$

We have obviously $|J_k(C)| = j_k(M)$ and $|J_k^{\mathrm{aff}}(C)| \geq j_k^{\mathrm{aff}}(M)$, this last inequality being an equality if C is linear. For $x \in \mathbb{R}$, we also denote as is usual $\binom{x}{k} := j_k(x)/k!$.

Proposition 5.3. *With the above notations:*

(i) *If $d_{k-1}(C) \geq d_{k-1}$ then*

$$\delta_{k-1} := \frac{d_{k-1}}{n} \leq 1 - 2\frac{\binom{M/2}{k}}{\binom{M}{k}}.$$

(ii) *If C is linear or if $k = 3$, and if $d_{k-1}^{\mathrm{aff}}(C) \geq d_{k-1}$, we have*

$$\delta_{k-1} := \frac{d_{k-1}}{n} \leq \big(1 - \frac{1}{2^{k-1}}\big)\frac{M}{M-1}.$$

(iii) *If $\overline{r}_{k-1}(C) \geq r_{k-1}$ then*

$$\rho_{k-1} := \frac{r_{k-1}}{n} \leq \frac{\sum_{i=1}^{k-1} \frac{1}{k}\binom{M/2}{i}\binom{M/2}{k-i}\min(i, k-i)}{\binom{M}{k}}$$

Proof. (i) For the generalized Hamming distance, we have

$$\begin{aligned}\sum_{\underline{x}\in J_k(C)} d(\underline{x}) &= \sum_{\underline{x}\in J_k(C)} \big(\sum_{j=1}^{n} d((x_1)_j,\dots,(x_k)_j)\big)\\ &= \sum_{j=1}^{n}\big(\sum_{\underline{x}\in J_k(C)} d((x_1)_j,\dots,(x_k)_j)\big)\\ &= \sum_{j=1}^{n}\sum_{i=1}^{k-1} k!\binom{w_j}{i}\binom{M-w_j}{k-i}.\end{aligned}$$

The function $x \to \binom{x}{i}\binom{M-x}{k-i} + \binom{x}{k-i}\binom{M-x}{i}$ is concave and invariant by $x \to M-x$ thus it takes its maximum at $x = M/2$. We derive the inequalities:

$$j_k(M)d_{k-1} \le d(\underline{x}) \le n\sum_{i=1}^{k-1} k!\binom{M/2}{i}\binom{M/2}{k-i} = nk!\Big(\binom{M}{k} - 2\binom{M/2}{k}\Big).$$

(ii) In the special case $k = 3$, we obtain from (i) the desired inequality. In the case C linear, we observe that $w_j = 0, M, M/2$ and that $d((x_1)_j,\dots,(x_k)_j)$ is non zero only if $w_j = M/2$ and $x_1,\dots,x_k$ do not all belong to $\{x \in C : x_j = 0\}$ or to $\{x \in C : x_j = 1\}$, which have $M/2$ elements. Thus

$$j_k^{\mathrm{aff}}(M)d_{k-1} \le n\big(j_k^{\mathrm{aff}}(M) - 2j_k^{\mathrm{aff}}(M/2)\big)$$

hence the announced inequality.

(iii) The result for $\overline{r}$ is derived similarly to the result (i) in the d case. □

Remark 5.4. *The upper bounds established in Proposition 5.3 can easily be turned into upper bounds for $M = |C|$. Indeed, if $\phi_{k-1} := f_{k-1}/n \le A(M)/B(M)$ where A and B are polynomials of the same degree, with respective leading coefficients α and β, with $B(M) > 0$, then, if $\phi_{k-1} \ge \alpha/\beta$, M is upper bounded by the largest zero of the polynomial $\phi_{k-1}B(M) - A(M)$. The bound obtained this way holds for $\delta_{k-1} \ge 1 - 1/2^{k-1}$ and $\rho_{k-1} \ge 1/2 - \binom{k-1}{\lfloor (k-1)/2\rfloor}/2^k$.*

5.4. **The Elias-Bassalygo technique and constant weight codes.** We recall that $A_{k-1}(n,f,m)$ denotes the maximal number of elements that a code $C \subset H_n$ can have under the condition $f_{k-1}(C) \ge m$; analogously let $A_{k-1}(n,w,f,m)$ be the maximum among the codes with constant weight w. With a standard argument, the following inequality holds:

$$\frac{A_{k-1}(n,f,m)}{\mathrm{card}(H_n)} \le \frac{A_{k-1}(n,w,f,m)}{\mathrm{card}(J_n^w)} \tag{2}$$

where J_n^w is the set of the $\binom{n}{w}$ binary words of length n and weight w. This so-called *Elias Bassalygo technique* is expected to improve the bounds on H_n, if similar bounds are established on the Johnson spaces J_n^w. Note that the value of w on the right hand side can be chosen freely. This line was followed in [7] for the generalized Hamming distance, and required moreover to extend the methods to non linear codes. In view of (2), we work out Plotkin type bounds for constant weight codes:

Proposition 5.5. *Let $C \subset J_n^w$ have M elements and let $\omega := w/n$.*

(i) *If $d_{k-1}(C) \ge d_{k-1}$ then*

$$\delta_{k-1} := \frac{d_{k-1}}{n} \le 1 - \frac{\binom{M\omega}{k} + \binom{M(1-\omega)}{k}}{\binom{M}{k}}.$$

(ii) *If* $d^{\mathrm{aff}}_{k-1}(C) \geq d_{k-1}$, *we have*

$$\delta_{k-1} := \frac{d_{k-1}}{n} \leq \frac{j_k(M)}{j^{\mathrm{aff}}_k(M)}\Big(1 - \frac{\binom{M\omega}{k} + \binom{M(1-\omega)}{k}}{\binom{M}{k}}\Big).$$

(iii) *If* $\overline{r}_{k-1}(C) \geq r_{k-1}$ *then*

$$\rho_{k-1} := \frac{r_{k-1}}{n} \leq \frac{\sum_{i=1}^{k-1} \frac{1}{k}\binom{M\omega}{i}\binom{M(1-\omega)}{k-i}\min(i, k-i)}{\binom{M}{k}}.$$

Proof. For d and $\overline{r}$, we follow the same line as for the proof of Proposition 5.3. There we applied the inequality $\sum_{j=1}^{n} g(w_j) \leq ng(M/2)$ for relevant functions g, being concave and invariant by $x \to M - x$. Since $C \subset J_n^w$, we have $\sum_{j=1}^{n} w_j = Mw$, so we can instead use the stronger inequality $\sum_{j=1}^{n} g(w_j) \leq ng(Mw/n)$. □

6. THE SDP BOUND FOR d_2, r_2, $\overline{r}_2$.

The method developed in [15] can be used to derive upper bounds for the cardinality of a binary code C with given $d_2(C)$ (respectively $r_2(C)$, $\overline{r}_2(C)$). Recall that $d_2(C) \geq d$ if and only if, for all $(x, y, z) \in C^3$ such that $x \neq y$, $y \neq z$, $d(x, y) + d(y, z) + d(z, x) \geq 2d$, $z \neq x$ (respectively $r_2(C) \geq r$ iff $\max(\lceil \frac{d(x,y)}{2} \rceil, \lceil \frac{d(y,z)}{2} \rceil, \lceil \frac{d(z,x)}{2} \rceil) \geq r$ and $\overline{r}_2(C) \geq \overline{r}$ iff $d(x, y)+d(y, z)+d(z, x) \geq 6\overline{r}$ for all $(x, y, z) \in C^3$).

The SDP constraints at work in [15] are exactly SDP constraints on triples of points. In order to describe these constraints we adopt the group theoretic point of view of [4], [17], [18].

Let $X := H_n$ and, for all $k := 0 \ldots n$, the so-called Johnson spaces $X_k := \{x, x \in X : wt(x) = k\}$. We consider the action of the symmetric group S_n on H_n. The Johnson spaces X_k are exactly the orbits of this action. Now we consider the decomposition of the functional space $L^2(X) = \mathbb{R}^X$ of real valued functions on X under the action of S_n. The space $\mathbb{R}^X$ is endowed with the S_n-invariant scalar product

$$(f, g) = \frac{1}{|X|}\sum_{x \in X} f(x)g(x).$$

We have the obvious decomposition into pairwise orthogonal S_n-invariant subspaces:

$$\mathbb{R}^X = \mathbb{R}^{X_1} \perp \mathbb{R}^{X_1} \perp \cdots \perp \mathbb{R}^{X_n}.$$

The decomposition of $\mathbb{R}^{X_k}$ into S_n-irreducible subspaces is described in [10]. We have

$$\mathbb{R}^{X_k} = H_{0,k} \perp H_{1,k} \perp \cdots \perp H_{\min(k,n-k),k}$$

where the $H_{i,k}$ are pairwise isomorphic for fixed i and pairwise non isomorphic for fixed k. The picture looks like:

$$\begin{array}{rcccccccc}
\mathbb{R}^X = & \mathbb{R}^{X_1} \perp & \mathbb{R}^{X_1} \perp & \ldots & \perp \mathbb{R}^{X_{\lfloor \frac{n}{2} \rfloor}} \perp & \ldots & \perp \mathbb{R}^{X_{n-1}} & \perp \mathbb{R}^{X_n} \\
 & H_{0,0} \perp & H_{0,1} \perp & \ldots & \perp H_{0,\lfloor \frac{n}{2} \rfloor} \perp & \ldots & \perp H_{0,n-1} & \perp H_{0,n} \\
 & & H_{1,1} \perp & \ldots & & & \perp H_{1,n-1} & \\
 & & & \ddots & \vdots & & & \\
 & & & & H_{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor} & & &
\end{array}$$

where the columns represent the decomposition of $\mathbb{R}^{X_k}$ and the rows the isotypic components of $\mathbb{R}^X$, with multiplicity $n-2k+1$, i.e. we have for $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$,

$$H_{k,k} \perp H_{k,k+1} \perp \cdots \perp H_{k,n-k} \simeq H_{k,k}^{n-2k+1}.$$

To each of these isotypic components, indexed by k, for $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$, we associate a matrix E_k of size $n-2k+1$ as explained in [3], indexed with s,t subject to $k \leq s,t \leq n-k$, in the following way: Let $(e_{k,k,1}, e_{k,k,2}, \ldots, e_{k,k,h_k})$ be an orthonormal basis of $H_{k,k}$ and let $e_{k,s,j} = \psi_{k,s}(e_{k,k,j})$. The application $\psi_{k,s}$ is defined by:

$$\begin{array}{rccl} \psi_{k,s}: & \mathbb{R}^{X_k} & \to & \mathbb{R}^{X_k} \\ & f & \mapsto & \psi_{k,s}(f) : \psi_{k,s}(f)(y) = \sum_{\substack{wt(x)=k \\ x \subset y}} f(x) \end{array}$$

and has the property to send and orthonormal basis of $H_{k,k}$ to an orthogonal basis of $H_{k,s}$, the elements of this basis having constant square norm equal to $\binom{n-2k}{s-k}$. The (s,t) coefficient of E_k is defined by:

$$E_{k,s,t}(x,y) = \frac{1}{h_k} \sum_{j=1}^{h_k} e_{k,s,j}(x) e_{k,t,j}(y).$$

From [3], $E_{k,s,t}(x,y) = E_{k,s,t}(gx,gy)$ for all $g \in S_n$. Thus for $k \leq s \leq t \leq n-k$, we can define $P_{k,s,t}$ by $E_{k,s,t}(x,y) = P_{k,s,t}(s - |x \cap y|)$. It turns out that these $P_{k,s,t}$ express in terms of Hahn polynomials.

The Hahn polynomials associated to the parameters n,s,t with $0 \leq s \leq t \leq n$ are the polynomials $Q_k(n,s,t;x)$ with $0 \leq k \leq \min(s, n-t)$ uniquely determined by the properties:

(i) Q_k has degree k in the variable x

(ii) They are orthogonal polynomials for the weights

$$0 \leq i \leq s \quad w(n,s,t;i) = \binom{s}{i}\binom{n-s}{t-s+i}$$

(iii) $Q_k(0) = 1$

The combinatorial meaning of the above weights is the following:

Lemma 6.1. *Given $x \in X_k$, the number of elements $y \in X_t$ such that $|x \cap y| = s - i$ is equal to $w(n,s,t;i)$.*

Finaly we have:

Proposition 6.2. *If $k \leq s \leq t \leq n-k$, $wt(x) = s$, $wt(y) = t$,*

$$E_{k,s,t}(x,y) = |X| \frac{\binom{t-k}{s-k}\binom{n-2k}{t-k}}{\binom{n}{t}\binom{t}{s}} Q_k(n,s,t;s-|x \cap y|)$$

If $wt(x) \neq s$ or $wt(y) \neq t$, $E_{k,s,t}(x,y) = 0$.

By the construction, the matrices E_k satisfy the semidefinite positivity properties:

Theorem 6.3. *For all k, $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$, for all $C \subset H_n$,*

$$\sum_{(c,c') \in C^2} E_k(c,c') \succeq 0.$$

These constraints are not interesting for pairs because they are not stronger than the positivity properties from Delsarte method. They are only interesting if triples of points are involved: namely we associate to $(x,y,z) \in H_n^3$ the matrices

$$F_k(x,y,z) := E_k(x-z, y-z).$$

We have for all $C \subset H_n$, and for all $z \in H_n$,

$$\sum_{(c,c') \in C^2} F_k(c,c',z) \succeq 0$$

which leads to the two positive semidefinite conditions:

$$\left\{ \begin{array}{l} \sum_{(c,c',c") \in C^3} F_k(c,c',c") \succeq 0 \\ \sum_{(c,c') \in C^2, c" \notin C} F_k(c,c',c") \succeq 0 \end{array} \right. \tag{3}$$

From Proposition 6.2, $E_k(x-z, y-z)$ only depends on the values of $wt(x-z)$, $wt(y-z)$, $wt(x-y)$; so with $a := d(y,z)$, $b := d(x,z)$, $c := d(x,y)$, we have for some matrices $T_k(a,b,c)$,

$$F_k(x,y,z) = T_k(a,b,c).$$

We introduce the unknowns $x_{a,b,c}$ of the SDP. Let, for

$$(a,b,c) \in \Omega := \left\{ (a,b,c) \in [0 \dots n]^3 : \quad \begin{array}{l} a+b+c \equiv 0 \mod 2 \\ a+b+c \leq 2n \\ c \leq a+b \\ b \leq a+c \\ a \leq b+c \end{array} \right\}$$

$$x_{a,b,c} := \frac{1}{|C|} \operatorname{card}\{(x,y,z) \in C^3 : d(y,z) = a, d(x,z) = b, d(x,y) = c\}.$$

Note that

$$x_{0,c,c} = \frac{1}{|C|} \operatorname{card}\{(x,y) \in C^3 : d(x,y) = c\}.$$

With the definition

$$\begin{array}{ll} t(a,b,c) & := \operatorname{card}\{z \in H_n : d(x,z) = b \text{ and } d(y,z) = a\} \text{ for } d(x,y) = c \\ & = \binom{c}{i}\binom{n-c}{a-i} \text{ where } a-b+c = 2i \end{array}$$

the following inequalities hold for $x_{a,b,c}$:

(i) $x_{0,0,0} = 1$
(ii) $x_{a,b,c} \geq 0$
(iii) $x_{a,b,c} = x_{\tau(a),\tau(b),\tau(c)}$ for every permutation τ of $\{a,b,c\}$
(iv) $x_{a,b,c} \leq t(a,b,c) x_{0,c,c}$
(v) $x_{a,b,c} \leq t(b,c,a) x_{0,a,a}$
(vi) $x_{a,b,c} \leq t(c,a,b) x_{0,b,b}$
(vii) $\sum_{a,b,c} T_k(a,b,c) x_{a,b,c} \succeq 0$ for all $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$
(viii) $\sum_{a,b,c} T_k(a,b,c)(t(a,b,c) x_{0,c,c} - x_{a,b,c}) \succeq 0$ for all $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$

where conditions (vii) and (viii) are equivalent to (3). To the above semidefinite constraints we add the extra condition (ix) that translates the assumption that $d_2(C) \geq d$ for some given value d (respectively $r_2(C) \geq r$, $\overline{r}_2(C) \geq \overline{r}$), namely

(ix)^d $x_{a,b,c} = 0$ if $abc \neq 0$ and $a + b + c \leq 2(d-1)$

respectively

(ix)$^{\overline{r}}$ $x_{a,b,c} = 0$ if $abc \neq 0$ and $a + b + c < 6\overline{r}$

or

(ix)^r $x_{a,b,c} = 0$ if $abc \neq 0$ and $\max(\lceil \frac{a}{2} \rceil, \lceil \frac{b}{2} \rceil, \lceil \frac{c}{2} \rceil) \leq r - 1$.

It remains to notice that

(x) $|C| = \sum_c x_{0,c,c}$.

Thus an upper bound on $|C|$ is obtained with the optimal value of the program that maximizes $\sum_c x_{0,c,c}$ under the constraints (i) to (ix).

It is worth noticing that the conditions (ix) can be replaced by any other conditions of the type

(ix*) $x_{a,b,c} = 0$ if $(a, b, c) \in I$

where I is a set of forbidden values in C related to some other situation. In the classical case treated in [15], $d_1(C) \geq \delta$, $I = \{(a, b, c) : a$ or b or $c \in [1 \ldots (\delta - 1)]^3\}$.

7. Numerical results

In this section we compare the SDP bounds obtained for $A_2(n, d, m)$ and for $A_2(n, r, m)$ with the previously known bounds, stated in Section 5. We recall the obvious values $A_2(n, d, 3) = 2^{n-1}$, $A_2(n, d, n) = 4$, $A_2(n, r, 1) = 2^n$, $A_2(n, r, \lfloor n/2 \rfloor) = 4$.

Table 1 gives two upper bounds for $A_2(n, d, m)$: one is the tightest of the combinatorial bounds of section 5, with a superscript $1, 2, 3, 4$ denoting which of the four methods, Singleton, Hamming, Plotkin, Elias (respectively) achieves this best, and the other is the bound obtained by the SDP method of Section 6. As we can see, in the non-trivial cases the SDP bound gives a substantial improvement almost all the time.

For the radius r, we can restrict ourselves to codes in which the pairwise distances are even. Let us denote $A_2^+(n, r, m)$ the maximal number of elements of such a code with minimal radius at least equal to m; then one easily sees that $A_2(n, r, m) = A_2^+(n + 1, r, m)$, with the standard extension of an optimal code to an even code with an extra coordinate. Table 2 compares the best bound for $A_2^+(n, r, m)$ (in italics) given by the combinatorial methods of Section 5 to the SDP bound. Again we have improvements in almost every instance.

References

[1] A. Ashikhmin, A. Barg, S. Litsyn, *New upper bounds on generalized weights*, IEEE Trans. Inform. Theory IT-45, (1999) 1258-1263.

[2] L. A. Bassalygo, *Supports of a code*, in proc. AAECC 11, (1995) LNCS 948 pp. 1–3.

[3] C. Bachoc, F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. 21 (2008), 909-924.

[4] C. Bachoc, *Semidefinite programming, harmonic analysis and coding theory*, arXiv:0909.4767 .

[5] V. M. Blinovskii, *Bounds for codes in the case of list decoding of finite volume*, Problems of Information Transmission **22(1)** (1986), 7-19

[6] V. M. Blinovskii, *Generalization of Plotkin bound to the case of multiple packing*, ISIT 2009

[7] G. Cohen, S. Litsyn, G. Zémor, *Upper bounds on generalized Hamming distances*, IEEE Trans. Inform. Theory **40** (1994), 2090-2092.

[8] J.H. Conway, N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, Springer-Verlag, 1988.

[9] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl. (1973), vi+97.

[10] P. Delsarte, *Hahn polynomials, discrete harmonics and t-designs*, SIAM J. Appl. Math. **34**-1 (1978)

$n\backslash m$	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
10	170 186^2	85 128^1	42 64^1	24 32^1	12 16^1	6 6^3										
11	290 341^2	170 256^1	85 128^1	35 61^2	24 32^1	12 12^3	5 5^3									
12	554 630^2	277 512^1	170 256^1	68 103^2	33 64^1	24 32^1	8 10^3	5 5^3								
13	1042 1170^2	521 1024^1	266 512^1	130 178^2	64 128^1	32 64^1	16 32^1	8 8^3	5 5^3							
14	2048 2184^2	1024 2048^1	512 1024^1	257 309^2	128 256^1	64 128^1	32 64^1	16 22^3	8 8^3	5 5^3						
15	3616 4096^2	2048 4096^1	1024 2048^1	414 541^2	256 512^1	128 256^1	43 113^2	32 64^1	16 16^3	6 7^3	5 5^3					
16	6963 7710^2	3489 7710^2	2048 4096^1	766 956^2	382 956^2	256 512^1	83 188^2	41 128^1	32 64^1	10 13^3	6 7^3	5 5^3				
17	13296 14563^2	6696 14563^2	3407 7710^4	1395 1702^2	708 1702^2	359 963^4	151 314^2	80 256^1	41 128^1	20 52^3	10 11^3	6 6^3	4 4^3			
18	26214 27594^2	13107 27594^2	6555 15420^4	2559 3048^2	1313 3048^2	682 1927^4	288 530^2	142 512^1	80 256^1	40 128^1	20 28^3	10 10^3	6 6^3	4 4^3		
19	47337 52428^2	26214 52428^2	13107 27594^4	4531 5489^2	2431 5489^2	1284 3246^4	513 903^2	276 903^2	142 512^1	51 208^2	40 128^1	20 20^3	8 9^3	6 6^3	4 4^3	
20	91750 99864^2	46113 99864^2	26214 55188^4	8133 9939^2	4342 9939^2	2373 5518^4	1024 1552^2	512 1514^4	274 1024^1	94 338^2	50 256^1	40 128^1	12 16^3	8 8^3	6 6^3	4 4^3

TABLE 1. Bounds on $A_2(n, d, m)$.

	m=2	3	4	5	6	7	8
n=10	96 *102*	16 *22*					
11	174 *186*	26 *36*	5 *11*				
12	341 *341*	48 *61*	10 *17*				
13	582 *630*	89 *103*	14 *27*	5 *10*			
14	1109 *1170*	161 *178*	22 *43*	5 *14*			
15	2085 *2184*	283 *309*	36 *69*	9 *22*	5 *9*		
16	4096 *4096*	526 *541*	64 *113*	13 *33*	5 *13*		
17	7235 *7710*	848 *956*	123 *188*	18 *52*	5 *19*	4 *8*	
18	13926 *14563*	1550 *1702*	216 *314*	30 *81*	10 *27*	5 *12*	
19	21883 *27594*	2852 *3048*	379 *530*	48 *129*	12 *41*	5 *16*	4 *8*

TABLE 2. Bounds on $A_2^+(n, r, m)$.

[11] V. Guruswami, *List Decoding from Erasures: Bounds and Code Constructions*, IEEE Trans. Inform. Theory IT-49, 2003, 2826–2833.
[12] V. I. Levenshtein, ”Universal bounds for codes and designs”, in *Handbook of Coding Theory*, eds V. Pless and W. C. Huffmann, Amsterdam: Elsevier, 1998, 499-648.
[13] R. J. McEliece, E. R. Rodemich, H. Rumsey, L. Welch, ”New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities”, *IEEE Trans. Inform. Theory* IT-23, 1977, 157-166.
[14] L. H. Ozarow and A. D. Wyner, *Wire Tap Channel II*, Bell System Technical Journal 63, Dec. 1984, 2135-2157.
[15] A. Schrijver, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Trans. Inform. Theory IT-51 (2005), 2859–2866.
[16] M. Sudan, *Decoding of Reed Solomon codes beyond the error-correction bound*, Journal of Complexity 13(1) (1997), 180–193.
[17] F. Vallentin, *Lecture notes: Semidefinite programs and harmonic analysis*, arXiv.math:0809.2017 .
[18] F. Vallentin, *Symmetry in semidefinite programs*, Linear Algebra and Appl. 430 (2009), 360-369.
[19] V.K. Wei, *Generalized Hamming weights for linear codes*, IEEE Trans. Inform. Theory IT-37 (1991), 1412-1418.
[20] G. Zémor, *Threshold effects in codes*, in First French-Israeli workshop on algebraic coding, Paris, july 1993 LNCS 781, Springer-Verlag, pp. 278–286.
[21] G. Zémor and G. Cohen, *The threshold probability of a code*, IEEE Trans. on Information theory, IT-41 (1995) pp. 469–477.