

New applications of semidefinite programming to coding theory

Christine Bachoc

Université Bordeaux I, IMB

DIAMANT seminar day, 13/03/2009

Outline

Joint research project with Gilles Zémor and Hervé Diet, work in progress..

- ▶ Review on the Hamming space
- ▶ Generalized Hamming distance
- ▶ List decoding radius
- ▶ SDP bounds

The Hamming space

- ▶ $H_n := \mathbb{F}_2^n = \{0, 1\}^n$.
- ▶ Hamming distance:

$$d(x, y) = \text{card}\{i, 1 \leq i \leq n : x_i \neq y_i\}.$$

- ▶ Automorphism group of H_n :

$$\text{Aut}(H_n) = T \rtimes S_n$$

where T is the group of translations $t_u : x \rightarrow x + u$, $T \simeq (\mathbb{F}_2^n, +)$ and S_n is the group of permutations acting on the n coordinates.

- ▶ $\text{Aut}(H_n)$ is *distance transitive* on H_n , i.e.

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

Binary error correcting codes

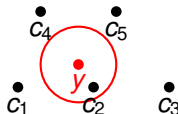
- ▶ A binary code of length n is a subset C of H_n .
- ▶ The minimal distance $d(C)$ of C :

$$d(C) = \min\{d(x, y) : x \neq y, (x, y) \in C^2\}.$$

is related to the error correction capacity of the code in the *binary symmetric channel*:



where the decoder outputs x' a codeword closest to y .

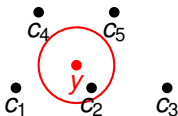


Binary error correcting codes

- ▶ Classical decoders return

$$B(y, \rho_1) \cap C$$

where $\rho_1 := \lfloor \frac{d(C)-1}{2} \rfloor$ which has at most one element.



- ▶ Given δ we are interested in the largest codes C such that $d(C) \geq \delta$ hence the classical optimization problem

$$A(n, \delta) := \max\{\text{card}(C) : d(C) \geq \delta\}.$$

Delsarte linear programming bound (1973), A. Schrijver semidefinite programming bound (2005).

Generalized Hamming weights of linear codes

- ▶ The Hamming weight $wt(x)$ of $x \in H_n$:

$$wt(x) := d(x, 0) = \text{card}\{i, 1 \leq i \leq n : x_i \neq 0\}.$$

Let $C \subset H_n$ be a linear code. Since $d(x, y) = wt(x - y)$ and $x - y \in C$ for all $(x, y) \in C^2$,

$$d(C) = \min\{wt(x) : x \neq 0, x \in C\}.$$

- ▶ Generalization: the Hamming weight of $D \subset H_n$:

$$wt(D) = \text{card}\{i, 1 \leq i \leq n : x_i \neq 0 \text{ for at least one } x \in D\}$$

The *s-th generalized Hamming weight* of a linear code was introduced independently by Ozarow and Wei (1991):

$$d_s(C) = \min\{wt(D) : D \subset C, D \text{ linear}, \dim(D) = s\}$$

- ▶ Clearly $d_1(C) = d(C)$.

Motivation: the wire-tap channel

- ▶ n bits are transmitted, carrying k bits of information. The channel is noiseless.
- ▶ An opponent can listen to any t bits of his choice out of the n . His goal is to recover the largest possible amount of information.
- ▶ Ozarow and Wyner, 1984, propose the following scheme:
 - ▶ Let C be a $[n, n - k]$ linear code. The messages are in one to one correspondance with its 2^k cosets. A word randomly chosen in the coset is transmitted.
 - ▶ Let H be a parity check matrix for C ($x \in C \Leftrightarrow xH^t = 0$). A given message $m \in \mathbb{F}_2^k$ is transmitted by x such that $xH^t = m$.
 - ▶ How to choose t bits of x such that knowledge of these t bits leaks the largest amount of information on m ?
 - ▶ How to choose C such that the information revealed on m cannot be too large?

Motivation: the wire-tap channel

- ▶ Let $J \subset \{1, \dots, n\}$, $\text{card}(J) = t$, J^c its complementary set and let

$$H_J := \{x \in H_n : x_i = 0 \text{ for all } i \in J^c\}$$

and $x = x_J + x_{J^c}$ the decomposition of x according to $H_n = H_J \oplus H_{J^c}$.

- ▶ Since

$$xH^t = x_J H^t + x_{J^c} H^t,$$

if x_J is known, the uncertainty left on $m = xH^t$ is that of $x_{J^c} H^t$ and has dimension $\dim((H_{J^c})H^t)$.

- ▶ We have

$$\dim((H_{J^c})H^t) = k - \dim(C^\perp \cap H_J)$$

- ▶ There exists J , $\text{card}(J) = t$ such that $\dim(C^\perp \cap H_J) \geq s$ iff

$$d_s(C^\perp) \leq t.$$

Generalized Hamming distance of non linear codes

- ▶ Introduced in 1994 (Cohen, Litsyn, Zémor):

$$d(x_0, x_1, \dots, x_s) = \text{card}\{i, 1 \leq i \leq n : ((x_0)_i, \dots, (x_s)_i) \notin \{0^{s+1}, 1^{s+1}\}\}$$

$$x_0 = 0 \dots 01 \dots 1100 \dots 0$$

$$x_1 = 0 \dots 01 \dots 1011 \dots 0$$

⋮

$$x_s = 0 \dots 01 \dots 1 \underbrace{001 \dots 1}_{d(x_0, \dots, x_s)}$$

Generalized Hamming distance of non linear codes

► Properties:

1. $d(x_0, x_1)$ is the usual Hamming distance.
2. For all permutation τ of $\{0, \dots, s\}$, $d(x_0, \dots, x_s) = d(x_{\tau(0)}, \dots, x_{\tau(s)})$.
3. For all $\sigma \in \text{Aut}(H_n)$, $d(x_0, \dots, x_s) = d(\sigma(x_0), \dots, \sigma(x_s))$.

► If $C \subset H_n$ is a binary code, let

$$d_s(C) := \min\{d(x_0, \dots, x_s) : (x_0, \dots, x_s) \in C^{s+1}, \dim_{\text{aff}}(x_0, \dots, x_s) = s\}.$$

► Coincides with Wei definition if C is linear.

List-decoding

- ▶ Classical decoders return $B(y, \rho_1(C)) \cap C$ where $\rho_1(C) = \lfloor \frac{d(C)-1}{2} \rfloor$ hence at most one codeword.
- ▶ A *list-decoding procedure* returns a list (c_1, c_2, \dots, c_s) of codewords found in a larger Hamming ball $B(y, \rho)$. The size s of the list should be “small” (i.e. polynomial in n).
- ▶ This notion was introduced by Elias (1957) and is a hot topic since Madhu Sudan gave a nice algorithm for list decoding of Reed-Solomon codes (≈ 2000).
- ▶ The *s-list decoding radius* $\rho_s(C)$ of a code C is defined by:

$$\rho_s(C) := \max\{r : \text{for all } y \in H_n, \text{ card}(B(y, r) \cap C) \leq s\}.$$

- ▶ We want to define an invariant attached to tuples of codewords.

List decoding

- ▶ We introduce the *radius of* $(x_0, \dots, x_s) \in H_n^{s+1}$:

$$r(x_0, \dots, x_s) := \min\{r : \text{there exists } y \in H_n \text{ s.t. } \{x_0, \dots, x_s\} \subset B(y, r)\}$$

and the *s-radius of a code C*:

$$r_s(C) := \min\{r(x_0, \dots, x_s) : (x_0, \dots, x_s) \in C^{s+1}, x_i \neq x_j \text{ for all } i < j\}.$$

Then

$$\rho_s(C) = r_s(C) - 1$$

- ▶ Properties:

1. $r(x_0, x_1) = \lfloor \frac{d(x_0, x_1) + 1}{2} \rfloor$.
2. For all permutation τ of $\{0, \dots, s\}$, $r(x_0, \dots, x_s) = r(x_{\tau(0)}, \dots, x_{\tau(s)})$.
3. For all $\sigma \in \text{Aut}(H_n)$, $r(x_0, \dots, x_s) = r(\sigma(x_0), \dots, \sigma(x_s))$.

List decoding from erasures

$$\begin{array}{rcl} y & * * * * * & 1010 \dots \\ x_1 & \dots \dots \dots & 1010 \dots \\ \vdots & & \\ x_s & \underbrace{\dots \dots \dots}_e & 1010 \dots \end{array}$$

We have $d(x_1, \dots, x_s) \leq e$.

- ▶ The *s-erasure list decoding radius* $\rho_s^{er}(C)$ is defined by (Guruswami 2003):

$$\rho_s^{er}(C) = \max\{e : \text{card}(\{x \in H_n : x_J = z, |J| = n - e\} \cap C) \leq s\}$$

- ▶ We have $\rho_s^{er}(C) = d'_s(C) - 1$ where

$$d'_s(C) := \min\{d(x_0, \dots, x_s) : (x_0, \dots, x_s) \in C^{s+1}, \text{ for all } i < j, x_i \neq x_j\}.$$

Optimization problems

- ▶ A bunch of optimization problems:

$$A_s(n, \delta) := \max\{\text{card}(C) : d_s(C) \geq \delta\}$$

$$A'_s(n, \delta) := \max\{\text{card}(C) : d'_s(C) \geq \delta\}$$

$$B_s(n, \rho) := \max\{\text{card}(C) : r_s(C) \geq \rho\}$$

- ▶ The functions d and r on tuples of elements of H_n are invariant under the action of $\text{Aut}(H_n)$.

Orbits of $\text{Aut}(H_n)$ acting on H_n^{s+1}

- ▶ Recall $\text{Aut}(H_n) = T \rtimes S_n$ where $T \simeq (\mathbb{F}_2^n, +)$ is the group of translations and S_n is the group of permutations acting on the n coordinates.
- ▶ Let $\underline{x} := (x_0, \dots, x_s) \in H_n^{s+1}$.

$$\begin{array}{rcl} x_0 & = & 000 \dots 0 \quad \dots \\ x_1 & = & 111 \dots 1 \quad \dots \\ \vdots & = & \vdots \quad \vdots \\ x_s & = & \underbrace{111 \dots 1}_{n_u(\underline{x})} \quad \dots \end{array}$$

For all $u \in \mathbb{F}_2^{s+1}$, let

$$n_u(\underline{x}) := \text{card}\{i, \quad 1 \leq i \leq n : \begin{array}{c} (x_0)_i \\ \vdots \\ (x_s)_i \end{array} = u\}.$$

i.e. the number of occurrences of u as a column in the above matrix.

Orbits of $\text{Aut}(H_n)$ acting on H_n^{s+1}

- ▶ The orbits of S_n acting on H_n^{s+1} are characterized by

$$\left(n_u(\underline{x}) \right)_{u \in \mathbb{F}_2^{s+1}}.$$

- ▶ Let $\bar{u} = \mathbf{1} + u$; the sets $\{u, \bar{u}\}$ are in correspondance with elements of \mathbb{F}_2^s (set $u_1 = 0$).
- ▶ The orbits of $\text{Aut}(H_n)$ acting on H_n^{s+1} are characterized by

$$\left(n_u(\underline{x}) + n_{\bar{u}}(\underline{x}) \right)_{u = \begin{smallmatrix} 0 \\ w \end{smallmatrix}, w \in \mathbb{F}_2^s}$$

- ▶ Notation: $n_w(\underline{x}) := n_u(\underline{x}) + n_{\bar{u}}(\underline{x})$ for $u = \begin{smallmatrix} 0 \\ w \end{smallmatrix}$. Moreover

$$\sum_{w \in \mathbb{F}_2^s} n_w(\underline{x}) = n.$$

Orbits of $\text{Aut}(H_n)$ acting on H_n^3

- ▶ Example: $s = 2$. The orbits of triples $(\underline{x}) = (x_0, x_1, x_2)$ are characterized by

$$w \in \left\{ \begin{matrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{matrix} \right\}$$

with

$$n_0^0(\underline{x}) + n_1^0(\underline{x}) + n_0^1(\underline{x}) + n_1^1(\underline{x}) = n.$$

- ▶ We have

$$d(x_0, x_1) = n_1^0(\underline{x}) + n_1^1(\underline{x})$$

$$d(x_0, x_2) = n_0^1(\underline{x}) + n_1^1(\underline{x})$$

$$d(x_1, x_2) = n_0^1(\underline{x}) + n_0^0(\underline{x})$$

hence equivalently by

$$(d(x_1, x_2), d(x_2, x_0), d(x_0, x_1)).$$

The functions $d(x_0, \dots, x_s)$ and $r(x_0, \dots, x_s)$

- ▶ Remember for all $\sigma \in \text{Aut}(H_n)$, $d(x_0, \dots, x_s) = d(\sigma(x_0), \dots, \sigma(x_s))$ and same for r . Hence $d(\underline{x}) = \tilde{d}(n_w(\underline{x}))$ and $r(\underline{x}) = \tilde{r}(n_w(\underline{x}))$.
- ▶ For d it is easy:

$$d(\underline{x}) = \sum_{w \neq 0^s} n_w(\underline{x}).$$

- ▶ For $s = 2$:

$$d(x_0, x_1, x_2) = \frac{1}{2} (d(x_0, x_1) + d(x_1, x_2) + d(x_2, x_0)).$$

- ▶ For r no nice formula so far.. although given ϕ , $\tilde{r}(\phi)$ can be easily calculated.

Semidefinite programs (SDP)

► **Primal problem:**

$$m = \min\{ \begin{array}{l} b_1 x_1 + \dots + b_k x_k : \\ -A_0 + x_1 A_1 + \dots + x_k A_k \succeq 0 \end{array} \}$$

where A_i are symmetric $r \times r$ matrices .

► **Dual problem:**

$$m^* = \max\{ \begin{array}{l} \text{Trace}(A_0 Z) : \\ Z \succeq 0, \quad \text{Trace}(A_i Z) = b_i, \quad i = 1, \dots, k \end{array} \}$$

- Linear programs (LP) correspond to the case where the matrices A_i are diagonal.
- In general, $m \geq m^*$. Under certain conditions, $m = m^*$.
- In this case, interior point methods lead to algorithms that approximate m with an arbitrary precision, with polynomial complexity.

SDP relaxations (road map)

- ▶ Let $\Phi := \{(n_w(\underline{x}))_{w \in \mathbb{F}_2^s} : \underline{x} \in H_n^{s+1}\}$.
- ▶ The variables are:

$$X_\phi := \frac{1}{\text{card}(C)} \text{card}\{\underline{x} \in C^{s+1} : (n_w(\underline{x}))_{w \in \mathbb{F}_2^{s+1}} = \phi\}.$$

- ▶ The condition e.g. $r_s(C) \geq \rho$ or $d_s(C) \geq \delta$ translate to

$$X_\phi = 0 \text{ for some } \phi \in \Phi.$$

- ▶ The objective function is

$$\sum_{\phi: \phi(w)=0 \text{ if } w \neq 0^s, 0^{s-1}1} X_\phi = \text{card}(C).$$

- ▶ *We need (non trivial) semidefinite constraints on the X_ϕ .*

SDP bounds for $s = 2$

- ▶ Notations: $(a, b, c) := (d(x_1, x_2), d(x_0, x_2), d(x_0, x_1))$ and

$$X_{a,b,c} := \frac{1}{\text{card}(C)} \text{card}\{\underline{x} \in C^3 : (d(x_1, x_2), d(x_0, x_2), d(x_0, x_1)) = (a, b, c)\}.$$

- ▶ Some SDP constraints are given by A. Schrijver (*New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE IT, 2005) where the aim is to bound the size of codes with $d_1(C) \geq \delta$.
- ▶ *Only the variables set to zero will change*. For d_2 :

$$X_{a,b,c} = 0 \text{ for all } (a, b, c) : abc \neq 0 \text{ and } a + b + c < 2\delta.$$

SDP bounds for $s = 2$

For all k , $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$, there are $(n - 2k + 1) \times (n - 2k + 1)$ matrices $F_k(x, y, z)$ such that

$$\text{for all } z \in H_n, \quad (x, y) \rightarrow F_k(x, y, z) \succeq 0$$

leading to:

$$\sum_{(x_0, x_1, x_2) \in \mathcal{C}^3} F_k(x_0, x_1, x_2) \succeq 0$$
$$\sum_{(x_0, x_1) \in \mathcal{C}^2, x_2 \notin \mathcal{C}} F_k(x_0, x_1, x_2) \succeq 0.$$

Moreover for all $\sigma \in \text{Aut}(H_n)$, $F_k(\sigma(x), \sigma(y), \sigma(z)) = F_k(x, y, z)$, hence there are matrices T_k such that

$$F_k(x_0, x_1, x_2) = T_k(a, b, c)$$

SDP bounds for $s = 2$

- ▶ The SDP conditions become:

$$\sum_{(a,b,c)} \chi_{a,b,c} T_k(a, b, c) \succeq 0$$

$$\sum_c \chi_{0,c,c} \left(\sum_{a,b} T_k(a, b, c) t(a, b, c) \right) - \sum_{(a,b,c)} \chi_{a,b,c} T_k(a, b, c) \succeq 0$$

where $t(a, b, c) = \binom{c}{i} \binom{n-c}{i}$, $i = (a - b + c)/2$.

- ▶ The matrix coefficients of T_k are computed by A. Schrijver, and by F. Vallentin in terms of Hahn polynomials.

A group theoretic construction of F_k

- ▶ The group S_n acts on H_n hence on \mathbb{R}^{H_n} . We have a decomposition into S_n -irreducible subspaces:

$$\mathbb{R}^{H_n} = H_0 \perp H_1^2 \perp \dots \perp H_k^{n-2k+1} \perp \dots \perp H_{\lfloor \frac{n}{2} \rfloor}^{n-2\lfloor \frac{n}{2} \rfloor+1}$$

where $H_k \simeq [n-k, k]$.

- ▶ For all k , we construct a matrix $E_k(x, y)$ from a decomposition:

$$H_k^{n-2k+1} = H_{k,1} \oplus \dots \oplus H_{k,n-2k+1}$$

with

$$(E_k)_{i,j}(x, y) := \frac{1}{h_k} \sum_{s=1}^{h_k} e_{k,i,s}(x) e_{k,j,s}(y)$$

where $(e_{k,i,s})_s$ is a suitably chosen orthonormal basis of $H_{k,i}$.

- ▶ Then $(x, y) \rightarrow E_k(x, y) \succeq 0$, $E_k(\sigma(x), \sigma(y)) = E_k(x, y)$ for $\sigma \in S_n$ and $F_k(x, y, z) = E_k(x-z, y-z)$.

Some numerical results: upper bounds for $A_2(n, \delta)$.

n	δ	SDP bound	Hamming/Plotkin bound
7	5	16*	36
7	6	8*	8
8	6	16*	64
16	4	6963	8192
16	6	2048*	8192
16	8	382	963
16	10	83	188
16	12	32*	188
16	14	6	7
23	11	4096*	8196
23	12	2048*	8196
24	12	4096*	14438

For the next future

- ▶ Bounds for $B_2(n, \rho)$.
- ▶ Compute explicitly the sdp constraints for 4-tuples, etc.. using the action of $\text{Stab}(x_0) \cap \text{Stab}(x_1)$, etc..
- ▶ Derive explicit bounds i.e. explicit functions of n and δ .
- ▶ Derive asymptotic bounds.
- ▶ Consider similar problems on other spaces like e.g. the unit sphere of Euclidean space, apparently of interest in signal processing, compressed sensing, etc..