Applications of semidefinite programming to coding theory

Christine Bachoc

Université Bordeaux I, IMB

ALCOMA10, Thurnau, April 11-18 2010

Outline

- Delsarte linear programming method
- Lovász theta on graphs and SDP's
- Positive semidefinite functions and harmonic analysis
- Applications to bounds on codes in metric spaces
- Applications to constraints on several points

Based on joint work with : Dion Gisjwijt (Leiden), Cordian Riener (Frankfurt), Patrick Solé (Paris), Frank Vallentin (Delft), Gilles Zémor (Bordeaux).

Delsarte LP method on the binary Hamming space

- $H_n := \{0, 1\}^n$
- Hamming distance $d(x, y)$

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

- We want to estimate

$$A(n, d) := \max\{|C| : C \subset H_n,\ d(C) \geq d\}.$$

Delsarte LP method on the binary Hamming space

- Delsarte (1973): The Krawtchouk polynomials $K_k^n(t)$, $k = 0, 1, \ldots, n$ satisfy the positivity property:

$$\text{For all code } C \subset H_n, \sum_{(x,y)\in C^2} K_k^n(d(x,y)) \geq 0.$$

- The distance distribution of a code C

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

satisfies the following inequalities:

1. For all $0 \leq k \leq n$, $\sum_{i=0}^n K_k^n(i) x_i \geq 0$.
2. $x_i \geq 0$
3. $x_0 = 1$
4. If $d(C) \geq d$, $x_i = 0$ for $i = 1, \ldots, d-1$
5. $\sum_{i=0}^n x_i = |C|$

Delsarte LP method on the binary Hamming space

- We obtain the following linear program:

$$
\begin{aligned}
M(n,d) \quad := \max \Big\{ \textstyle\sum_{i=0}^{n} x_i : \quad & x_i \geq 0, \\
& x_0 = 1, \\
& x_i = 0 \ \text{ if } i = 1, \dots, d-1 \\
& \textstyle\sum_{i=0}^{n} K_k^n(i) x_i \geq 0 \text{ for all } 0 \leq k \leq n \Big\}
\end{aligned}
$$

which optimal value upper bounds $A(n,d)$:

$$A(n,d) \leq M(n,d)$$

Krawtchouk polynomials

- They are related to the irreducible decomposition of the space

$$\mathcal{C}(H_n) := \{f : H_n \to \mathbb{C}\}$$

under the action of the isometry group of H_n: $\mathrm{Aut}(H_n) = T \rtimes S_n$, $T \simeq (\mathbb{F}_2^n, +)$

- Let $\chi_z(x) := (-1)^{x \cdot z}$ denote the characters of $(\mathbb{F}_2^n, +)$.

$$\begin{aligned} \mathcal{C}(H_n) &= \oplus_{z \in H_n} \mathbb{C}\chi_z \\ &= \oplus_{k=0}^{n} P_k, \quad P_k := \oplus_{wt(z)=k} \mathbb{C}\chi_z \end{aligned}$$

$$\text{Then} \quad K_k^n(d(x,y)) = \sum_{wt(z)=k} \chi_z(x)\chi_z(y)$$

- $K_k^n(t) = \sum_{j=0}^{k} (-1)^j \binom{t}{j}\binom{n-t}{k-j}$.

Successful because..

- Delsarte LP method has lead to:
  - Excellent numerical bounds
  - Explicit bounds (Levenshtein)
  - Asymptotic bounds (MRRW)
- It has been generalized to:
  - The 2-point homogeneous spaces, finite (binary Johnson and q-Johnson, etc..), and also real compact (S^{n-1}, etc..). On each space a certain family of orthogonal polynomials plays the role of the Krawtchouk polynomials.
  - A few other symmetric spaces (non binary Johnson, permutation codes, ordered codes, real and complex Grassmannians, unitary codes, etc,..). Multivariate polynomials come into play.

But..

- Some spaces of interest in coding theory cannot be treated, examples:
  - The projective space over $\mathbb{F}_q$:

$$\mathcal{P}_{q,n} := \{x \subset \mathbb{F}_q^n : x \text{ is a linear subspace } \}$$

    with the distance $d_S(x, y) := \dim(x) + \dim(y) - 2\dim(x \cap y)$ (Koetter, Kschichang 2007) or may be the injection distance $d_i(x, y) := \max(\dim(x), \dim(y)) - \dim(x \cap y)$ (da Silva, Kschichang, 2009).
  - The balls of H_n:

$$B_n(w) := \{x \in H_n : wt(x) \leq w\}$$

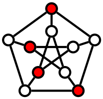
  - Their isometry group (resp. $\mathrm{Gl}_n(\mathbb{F}_q)$ and S_n) is not transitive.

But..

- When available, Delsarte LP bound is not optimal. It only exploits constraints on pairs of points. Can it be strengthened, exploiting triples of points, or more generally k-tuples of points ?
- Schrijver (2005): exploits SDP constraints on triples of points and improves the known bounds for $A(n, d)$ for some values of (n, d). Gijswijt, Schrijver, Tanaka: non binary Hamming space.
- Pseudo-distances $f(x_1, \ldots, x_k)$ involving k-tuples of points have been introduced (generalized Hamming distance, radial distance,..). How can we bound the size of codes subject to a constraint on k-tuples e.g. with given minimal pseudo-distance ?

Graphs

- $G = (V, E)$ a finite graph.
- An independence set S of G is a subset of V such that $S^2 \cap E = \emptyset$



- The independence number of G:

$$\alpha(G) = \max_{S \text{ independent}} |S|$$

- $V = H_n$, $E = \{(x, y) : d(x, y) < d\}$. The independence sets of $G = (V, E)$ are exactly the codes with minimal distance at least equal to d and

$$A(n, d) = \alpha(G).$$

Lovász ϑ

- 1978, L. Lovász, *On the Shannon capacity of a graph* introduces the theta number $\vartheta(G)$ and proves the Sandwich Theorem:

Theorem

$$\alpha(G) \leq \vartheta(G) \leq \chi(\overline{G})$$

- $\vartheta(G)$ is the optimal value of a semidefinite program (SDP).
- With ϑ, he proves that the capacity of the pentagon equals $\sqrt{5}$ (a conjecture of Shannon).

Lovász ϑ

- $G = (V, E)$ and $V = \{1, \ldots, v\}$

$$\vartheta(G) = \max \Big\{ \sum_{i,j} B_{i,j} : \begin{array}{l} B = (B_{i,j})_{1 \le i,j \le v},\ B \succeq 0 \\ \sum_i B_{i,i} = 1, \\ B_{i,j} = 0 \quad (i,j) \in E \end{array} \Big\}$$

- Proof of $\alpha(G) \le \vartheta(G)$:
  - If S is an independence set, the matrix B defined by

$$B_{i,j} = \frac{1}{|S|} \mathbf{1}_S(i) \mathbf{1}_S(j)$$

    satisfies the above conditions.
  - Moreover $\sum_{i,j} B_{i,j} = |S|$.
  - Thus $|S| \le \vartheta(G)$.

SDP's

- Primal program:

$$\gamma := \min\{ \quad c_1 x_1 + \cdots + c_m x_m : \\ -A_0 + x_1 A_1 + \cdots + x_m A_m \succeq 0\}$$

where A_i are real symmetric matrices of size r.

- Dual program:

$$\gamma^* := \max\{ \quad \mathrm{Trace}(A_0 Z) : \\ Z \succeq 0, \quad \mathrm{Trace}(A_i Z) = c_i, \quad i = 1, \ldots, m\}$$

- Linear programs (LP) occur when the matrices A_i are diagonal.
- In general, $\gamma \geq \gamma^*$. Under some mild conditions, $\gamma = \gamma^*$.
- In this case, interior point methods lead to algorithms that allow to approximate γ to an arbitrary precision in polynomial time. Good free solvers are available (NEOS)!

Graphs and codes

- $V = H_n$. **Bad news:** the number of vertices 2^n is exponential in n, thus also the complexity of the computation of ϑ!
- **Good news:** the group $\Gamma := \mathrm{Aut}(H_n)$ acts on the SDP thus it has the same optimal value as its symmetrization

$$\vartheta^\Gamma = \vartheta$$

where ϑ^Γ is restricted to the Γ-invariant matrices B: $B_{\gamma i, \gamma j} = B_{i,j}$.

$$\vartheta = \vartheta^\Gamma \quad = \max\Big\{ \sum_{i,j} B_{i,j} : \quad \begin{array}{l} B = (B_{i,j})_{1 \le i,j \le v}, \\ B \succeq 0, \ \gamma B = B \text{ for all } \gamma \in \Gamma, \\ \sum_i B_{i,i} = 1, \\ B_{i,j} = 0 \quad (i,j) \in E \Big\} \end{array}$$

Graphs and codes

- ${\vartheta'}^{\Gamma}$ (where in ϑ' the condition $B \geq 0$ is added) is exactly equal to Delsarte LP (Mc Eliece, Rodemich, Rumsey ; independently Schrijver, 79) and has polynomial complexity. Proof:
- $B(\gamma x, \gamma y) = B(x, y)$ for all $\gamma \in \mathrm{Aut}(H_n)$ and $B \succeq 0$ iff

$$B(x, y) = \sum_{k=0}^{n} a_k K_k^n(d(x, y)) \quad \text{with } a_k \geq 0$$

- Thus

$$\vartheta' = {\vartheta'}^{\Gamma} \quad = \max \Big\{ 2^{2n} a_0 : \quad \begin{array}{l} a_0, \ldots, a_n \geq 0 \\ \sum_{k=0}^{n} \binom{n}{k} 2^n a_k = 1, \\ \sum_{k=0}^{n} K_k^n(i) a_k = 0 \text{ if } i = 1, \ldots, n \\ \sum_{k=0}^{n} K_k^n(i) a_k \geq 0 \text{ for all i } \Big\} \end{array}$$

An SDP bound for codes

- It is a general fact that the spaces of interest in coding theory are huge (if not infinite!) and have a large group of isometries.
- We want to follow the same line for a metric space (X, d) with isometry group Γ. Main task: we need a decription of the Γ-invariant positive semidefinite functions $F : X^2 \mapsto \mathbb{R}$ ($F \succeq 0$, meaning the matrix $\big(F(x, y)\big)_{x,y \in X^2}$ is psd).
- There is a recipe using harmonic analysis (the study of the space of functions on X as a Γ-module).

The recipe

- Let $\mathcal{C}(X) := \{f : X \mapsto \mathbb{C}\}$ with the action of Γ: $(\gamma f)(x) = f(\gamma^{-1}x)$.
- Decompose the space $\mathcal{C}(X)$ under the action of Γ

$$\mathcal{C}(X) = R_0^{m_0} \perp R_1^{m_1} \perp \cdots \perp R_s^{m_s}$$

- For $k = 0, \ldots, s$, compute a certain Γ-invariant matrix $E_k(x, y)$, of size m_k, associated to $R_k^{m_k}$
- Then, $F \succeq 0$ and F is Γ-invariant iff

$$F(x, y) = \sum_{k=0}^{s} \mathrm{Trace}(F_k E_k(x, y)) \text{ with } F_k \succeq 0.$$

- Then ϑ'^{Γ} transforms into an SDP with variables the F_k.

The end of the recipe

- To compute $E_k(x,y)$: take an explicit decomposition

$$R_k^{m_k} = R_{k,1} \perp \cdots \perp R_{k,m_k}.$$

Let $(e_{k,i,1}, \ldots, e_{k,i,d_k})$ be compatible basis of $R_{k,i}$. Then

$$E_{k,i,j}(x,y) = \sum_{s=1}^{d_k} e_{k,i,s}(x)\overline{e_{k,j,s}(y)}.$$

- $E_{k,i,j}(x,y)$ is Γ-invariant thus is a function of the orbits of X^2 under Γ. Compute $Y_{k,i,j}$ such that

$$E_{k,i,j}(x,y) = Y_{k,i,j}(O_\Gamma(x,y)).$$

An example

X	Projective space $\mathcal{P}_{q,n}$	Hamming space H_n
q	p^t	1
Γ	$\mathrm{Gl}_n(\mathbb{F}_q)$	S_n
$\lvert x\rvert$	$\dim(x)$	$wt(x)$
Orbits of X	$X_k := \{x \in X : \lvert x\rvert = k\}$	
Orbits of X^2	$X_{a,b,c} := \{(x,y) \in X^2 : \lvert x\rvert = a, \lvert y\rvert = b, \lvert x \cap y\rvert = c\}$	

We have (Delsarte, 78):

$$\begin{array}{llllllll}
\mathcal{C}(X) = & \mathcal{C}(X_0) \perp & \mathcal{C}(X_1) \perp & \dots & \perp \mathcal{C}(X_{\lfloor \frac{n}{2} \rfloor}) \perp & \dots & \perp \mathcal{C}(X_{n-1}) & \perp \mathcal{C}(X_n) \\
= & H_{0,0} \perp & H_{0,1} \perp & \dots & \perp H_{0,\lfloor \frac{n}{2} \rfloor} \perp & \dots & \perp H_{0,n-1} & \perp H_{0,n} \\
 & & H_{1,1} \perp & \dots & & & \perp H_{1,n-1} & \\
 & & & \ddots & \vdots & & & \\
 & & & & H_{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor} & & &
\end{array}$$

An example

Theorem (B., Vallentin, 2007)

For all $0 \leq k \leq \lfloor n/2 \rfloor$, $H_{k,k} \perp \cdots \perp H_{k,n-k} \simeq R_k^{n-2k+1}$, *and the coefficients of the associated matrix* $E_k(x,y)$ *are equal to:*

$$E_{k,i,j}(x,y) = |X| h_k \frac{\begin{bmatrix} j-k \\ i-k \end{bmatrix} \begin{bmatrix} n-2k \\ j-k \end{bmatrix}}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} q^{k(j-k)} Q_k(n,i,j; i - |x \cap y|)$$

if $k \leq i \leq j \leq n-k$, $|x| = i$, $|y| = j$, *and* $E_{k,i,j}(x,y) = 0$ *if* $|x| \neq i$ *or* $|y| \neq j$, *and* $Q_k(n,i,j;t)$ *are q-Hahn polynomials with parameters* n, i, j.

Numerical applications

- $X = B_n(w)$ the ball of radius w, center 0 in H_n, min distance 8.

$n \backslash w$	8	9	10	11	12	13	14	15	16	$A(n,8) \leq$
18	67									72
19	100	123	137							142
20	154	222	253							256
21	245	359	465							512
22	349	598	759	870	967	990	1023			1024
23	507	831	1112	1541	1800	1843	1936	2047		2048
24	760	1161	1641	2419	3336	3439	3711	3933	4095	4096

- $X = \mathcal{P}_{2,8}$ the projective space

d_S	constructions in $\mathcal{G}_{2,4,8}$	LP bound for $\mathcal{G}_{2,4,8}$	SDP bound
8	17	17	17
7			18
6	256	308	308
5			364
4	4098	6477	6479
3			9273
2	65536	200787	222378

Pseudo-distances on the binary Hamming space

The relative position of two binary words is measured by their Hamming distance. What about $k \geq 3$ words ?

Let us call pseudo distance any fonction $f(x_1, \ldots, x_k)$ such that:

- $f(x_1, \ldots, x_k) \in \mathbb{R}^+$
- $f(x_1, \ldots, x_k) = f(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$ for all permutation σ of $\{1, \ldots, k\}$
- $f(x_1, \ldots, x_k) = f(\gamma x_1, \ldots, \gamma x_k)$ for all $\gamma \in \mathrm{Aut}(H_n) = T \rtimes S_n$.

Examples:

- The generalized Hamming distance
- The radial and average radial distances

The generalized Hamming distance

- Introduced by Ozarow and Wei (1991) for the linear codes in view of cryptographic applications; extended to non linear codes in 1994 (Cohen, Litsyn, Zémor):

$$d(x_1, \ldots, x_k) = \text{card}\left\{j,\ 1 \leq j \leq n:\ ((x_1)_j, \ldots, (x_k)_j) \notin \{0^k, 1^k\}\right\}.$$

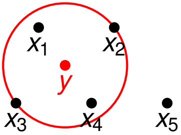
$$\begin{aligned} x_1 &= 0\ldots01\ldots1100\ldots0 \\ x_2 &= 0\ldots01\ldots1011\ldots0 \\ &\vdots \\ x_k &= 0\ldots01\ldots1\underbrace{001\ldots1}_{d(x_1,\ldots,x_k)} \end{aligned}$$

- When $k = 2$ it is the usual Hamming distance.

The radial distance

- Related to the notion of list decoding (Elias, $\approx$ 1950)

$$r(x_1, \dots, x_k) = \min \{ r : \text{ there exists } y \in H_n \text{ s.t. } \{x_1, \dots, x_k\} \subset B(y, r) \}$$
$$= \min_y \{ \max_{1 \le i \le k} d(y, x_i) \}.$$



- Difficult to analyse, often replaced by the average radial distance (Blinovski, Ashwelde).

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

Generalizations of $A(n,d)$

- Let $C \subset H_n$ and f a pseudo-distance f, we define

$$f_{k-1}(C) = \min\left\{ f(x_1,\ldots,x_k) \ : \ (x_1,\ldots,x_k) \in C^k, x_i \neq x_j \right\}$$

- Let

$$A_{k-1}(n,f,m) = \max\left\{ |C| \ : \ C \subset H_n,\ f_{k-1}(C) \geq m \right\}.$$

- $A_1(n,d,m) = A(n,m)$.
- Problem: upper bound $A_{k-1}(n,f,m)$.

An SDP upper bound for $A_{k-1}(n, f, m)$

- Joint work with G. Zémor ($k = 3$) and with Cordian Riener ($k \geq 4$). Goal: generalize Lovász ϑ.

$$\chi(z_1, \dots, z_k) := \frac{1}{|C|} \mathbf{1}_C(z_1) \dots \mathbf{1}_C(z_k)$$

- χ satisfies
  (1) $\chi(z_1, \dots, z_k) = \chi(\{z_1, \dots, z_k\})$
  (2) For all $(z_1, \dots, z_{k-2}) \in H_n^{k-2}$, for all $I \subset \{1, \dots, k-2\}$,

$$(x, y) \mapsto \sum_{J \,:\, I \subset J \subset \{1,\dots,k-2\}} (-1)^{|J|-|I|} \chi(z_{j_{(j \in J)}}, x, y) \succeq 0 \text{ and } \geq 0$$

  (3) $\chi(z_1, \dots, z_k) = 0$ if $f(z_1, \dots, z_k) \leq m - 1$ and $z_i \neq z_j$ (under the assumption $f_{k-1}(C) \geq m$)
  (4) $\sum_{x \in H_n} \chi(x) = 1$
  (5) $\sum_{(x,y) \in H_n^2} \chi(x, y) = |C|$

An SDP upper bound for $A_{k-1}(n, f, m)$

Theorem

The optimal value of the following SDP is an upper bound of $A_{k-1}(n, f, m)$:

$$\max \left\{ \sum_{(x,y)\in H_n^2} F(x,y) : \quad \begin{array}{l} F : H_n^k \to \mathbb{R}, \\ F \textit{ satisfies } (1)-(4) \end{array} \right\} \qquad (P_k)$$

(1) $F(z_1, \ldots, z_k) = F(\{z_1, \ldots, z_k\})$

(2) $(x, y) \mapsto \sum_{J\,:\, I\subset J\subset\{1,\ldots,k-2\}} (-1)^{|J|-|I|} F(z_{j_{(j\in J)}}, x, y) \succeq 0$ and ≥ 0

(3) $F(z_1, \ldots, z_k) = 0$ if $f(z_1, \ldots, z_k) \leq m-1$ et $z_i \neq z_j$

(4) $\sum_{x\in H_n} F(x) = 1$

Symmetrization

- The SDP (P_k) is invariant under $\Gamma = \mathrm{Aut}(H_n)$. Thus one can restrict in (P_k) to the functions F which are Γ-invariant.
- It is enough to have an expression for the $F \succeq 0$ and $\Gamma_{\underline{z}}$-invariant where

$$\Gamma_{\underline{z}} := \mathrm{Stab}(\mathrm{Aut}(H_n), z_1, \ldots, z_{k-2})$$

Invariante positive semidefinite functions

- $k = 2$, $\Gamma_{\underline{z}} = \Gamma$, we have: $F(x,y) = \sum_{i=0}^{n} a_i K_i^n(d(x,y))$, with $a_k \geq 0$ and K_k^n are the Krawtchouk polynomials.
- $k = 3$, $z_1 = 0^n$, $\Gamma_{\underline{z}} = S_n$,

$$F(x,y) = \sum_{i=0}^{\lfloor n/2 \rfloor} \mathrm{Trace}(A_i E_i^n(x,y)),\ A_i \succeq 0$$

- $k \geq 4$, we obtain matrices $E_k(x,y)$ built from tensor powers of the previous ones.

Numerical results

- The variables of the resulting SDP are indexed by the orbits of the subsets of H_n with at most k elements, their number is of order $n^{2^{k-1}-1}$.
- For $k = 3$ one recovers the SDP given by A. Schrijver in order to strengthen the known upper bounds for $A(n, d)$ with a different condition (3).
- These SDP constraints are in fact part of Lasserre hierarchy (Dion Gijswijt, 2009: algorithmic method to symmetrize the full Lasserre hierarchy).
- Implemented for $k = 3, 4$.
- The numerical results show that the SDP bound improves the previous ones.

Bounds for $A_2(n, d, m)$

	m=4	5	6	7	8	9	10	11	12	13	14	15	16	17
n=10	170	85	42	24	12	6								
	186	*128*	*64*	*32*	*16*	*6*								
11	*288*	170	85	*33*	24	12	5							
	341	*256*	*128*	*61*	*32*	*12*	*5*							
12	554	*270*	170	*64*	*32*	24	8	5						
	630	*512*	*256*	*103*	*64*	*32*	*10*	*5*						
13	1042	521	266	130	64	32	16	8	5					
	1170	*1024*	*512*	*178*	*128*	*64*	*32*	*8*	*5*					
14	2048	1024	512	257	128	64	32	16	8	5				
	2184	*2048*	*1024*	*309*	*256*	*128*	*64*	*22*	*8*	*5*				
15	3616	2048	1024	414	256	128	43	32	16	6	5			
	4096	*4096*	*2048*	*541*	*512*	*256*	*113*	*64*	*16*	*7*	*5*			
16	6963	3489	2048	766	382	256	83	41	32	10	6	5		
	7710	*7710*	*4096*	*956*	*956*	*512*	*188*	*128*	*64*	*13*	*7*	*5*		
17	13296	6696	3407	1395	708	359	151	80	41	20	10	6	4	
	14563	*14563*	*7710*	*1702*	*1702*	*963*	*314*	*256*	*128*	*52*	*11*	*6*	*4*	
18	26214	13107	6555	2559	1313	682	288	142	80	40	20	10	6	4
	27594	*27594*	*15420*	*3048*	*3048*	*1927*	*530*	*512*	*256*	*128*	*28*	*10*	*6*	*4*
19	47337	26214	13107	4531	2431	1284	513	276	142	51	40	20	8	6
	52428	*52428*	*27594*	*5489*	*5489*	*3246*	*903*	*903*	*512*	*208*	*128*	*20*	*9*	*6*
20	91750	46113	26214	8133	4342	2373	1024	512	274	94	50	40	12	8
	99864	*99864*	*55188*	*9939*	*9939*	*5518*	*1552*	*1514*	*1024*	*338*	*256*	*128*	*16*	*8*

Bounds for $A_3(n, d, m)$

	m=3	4	5	6	7	8	9	10	11
n=8	172	90	45	24	12				
9	344	179	89	44	24	12			
10	687	355	177	87	43	24	12		
11	1373	706	342	169	84	41	24	12	
12	2744	1402	665	307	167	79	40	24	12