Semidefinite programming techniques in combinatorial optimization

Levent Tunçel

Dept. of Combinatorics and Optimization,
Faculty of Mathematics, University of Waterloo

July 22–26, 2016

We will start with the discussion of various forms of Semidefinite Programming (SDP) problems and some necessary background (no previous background on SDPs is required).
Then, we will formulate various problems in combinatorial optimization, graph theory, and discrete mathematics in general either as SDP problems or as nonconvex optimization problems with natural and useful SDP relaxations.
We will continue with some geometric representations of graphs as they relate to SDP, more recent work on Lovász theta body and its extensions, lift-and-project methods, and conclude with some of the more recent work in these research areas and the research area of lifted SDP-representations (or extended formulations) and some open research problems.

For some positive integer d, let $\mathcal{F} \subseteq \{0,1\}^d$. Suppose $c \in \mathbb{R}^d$ is given. Then, the problem

$$\max c^\top x \quad \text{subject to: } x \in \mathcal{F}$$

is a *combinatorial optimization problem*.
For the purposes of this short course, the above is a definition of combinatorial optimization problem.

How do we solve such problems?
Of course, if $\mathcal{F}$ is given as a list (as part of the input), this is trivial (we evaluate $c^\top x$ for each x in the list, ...).
What if $\mathcal{F}$ is given implicitly? I.e., let $G = (V, E)$ be a simple undirected graph, then $S \subseteq V$ is a *stable set* in G if for every $i, j \in S$, $\{i, j\} \notin E$. (No two elements in S are joined by an edge in G.) Then we define $\mathcal{F}$ in terms of the input graph G:

$$\mathcal{F} := \left\{ x \in \{0,1\}^V : x \text{ is the incidence vector of a stable set in } G \right\}.$$

How hard is this problem? Stability number of G

$$\alpha(G) := \max \{|S| : S \text{ is a stable set in } G\}$$

Computing $\alpha(G)$ is $\mathcal{NP}$-hard! So, our original problem is $\mathcal{NP}$-hard even if $c := \bar{e}$.

We can consider a convex optimization approach.
Line segment joining u and v in $\mathbb{R}^n$:

$$\{\lambda u + (1-\lambda)v : \lambda \in [0,1]\}.$$

Convex sets: $F \subseteq \mathbb{R}^n$ is convex if for every pair of points $u, v \in F$, the line segment joining u and v lies entirely in F.
Fact: Intersection of an arbitrary collection of convex sets is convex.

Externally: Convex Hulls:
Let $F \subseteq \mathbb{R}^n$. The intersection of all convex sets containing F is called the convex hull of F and is denoted by $\text{conv}(F)$.

Internally: Convex combinations of $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$:

$$\sum_{i=1}^{k} \lambda_i x^{(i)},$$

for some $\lambda \in \mathbb{R}^k_+$ such that $\sum_{i=1}^{k} \lambda_i = 1$.
Extreme points of convex sets: Let $F \subseteq \mathbb{R}^n$ be a convex set. $x \in F$ is called an extreme point of F if $\nexists u, v \in F \setminus \{x\}$ such that $x = \frac{1}{2}u + \frac{1}{2}v$.

Let $\mathcal{F} \subset \mathbb{R}^d$ be a compact set. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuous function. Then the optimization problem

$$\max \{f(x) : x \in \mathcal{F}\}$$

is equivalent to (by adding a new variable x_{d+1}),

$$\max \{x_{d+1} : f(x) \geq x_{d+1}, x \in \mathcal{F}, \ell_{d+1} \leq x_{d+1} \leq u_{d+1}\},$$

where $\ell_{d+1} \in \mathbb{R}$ is a lower bound on the minimum value of f over $\mathcal{F}$ and $u_{d+1} \in \mathbb{R}$ is an upper bound on the maximum value of f over $\mathcal{F}$.

We used:

Theorem

(Bolzano [1830], Weierstrass [1860]) Let $\mathcal{F} \subset \mathbb{R}^d$ be a nonempty compact set and $f : \mathcal{F} \to \mathbb{R}$ be a continuous function on $\mathcal{F}$. Then f attains its minimum and maximum values on $\mathcal{F}$.

So, without loss of generality, we may assume that the objective function is linear. By redefining, d, c and $\mathcal{F}$, our problem becomes

$$\max \left\{ c^\top x : x \in \mathcal{F} \right\},$$

for a given vector $c \in \mathbb{R}^d$, and for a given compact set $\mathcal{F} \subset \mathbb{R}^d$. Now, we have

$$\max \left\{ c^\top x : x \in \mathcal{F} \right\} = \max \left\{ c^\top x : x \in \operatorname{conv}(\mathcal{F}) \right\}.$$

We denote by H the d-dimensional unit hypersphere (in the context of the next theorem this is the set of all nontrivial objective functions).

Theorem

The set of all $c \in H$ for which the optimal solution of

$$\max \left\{ c^\top x : x \in \mathrm{conv}(\mathcal{F}) \right\}$$

is not unique has zero, $(d-1)$-dimensional Hausdorff measure.

An implication is that, if we were to pick the objective function c from the hypersphere H, each point in H having the "same" chance of being picked, probability of picking a c for which the optimal solution is not unique, would be zero.

In case $\mathcal{F}$ is a finite set (as in combinatorial optimization), conv($\mathcal{F}$) is a polytope (convex hull of finitely many points, equivalently, intersection of finitely many closed half spaces, that is bounded). Then, the situation is particularly nice.

Picture on the board...

If we are able to get our hands on a convex compact set F which is tractable (that is, we can optimize a linear function over it in polynomial time) and $F = \text{conv}(\mathcal{F})$, then for almost all objective function vectors c, the problem

$$\max\left\{c^\top x : x \in F\right\}$$

will have a unique minimizer $\hat{x}$. Thus, $\hat{x} \in \text{ext}(F) \subseteq \mathcal{F}$. Therefore, $\hat{x}$ will also be an optimal solution of our original (non-convex) problem

$$\max\left\{c^\top x : x \in \mathcal{F}\right\}.$$

How do we "compute" or approximate the convex hull?

Linear Optimization

Let $A \in \mathbb{R}^{m \times n}$, given.

$$
\begin{array}{lll} (LP_1) & \text{Maximize} & c^\top x \\ & & \begin{array}{rcl} Ax & \leq & b, \\ (x & \in & \mathbb{R}^n). \end{array} \end{array}
$$

All vectors are column vectors.

$$u, v \in \mathbb{R}^m, u \leq v \text{ means } u_i \leq v_i, \forall i \in \{1, 2, \ldots, m\}.$$

Let us use another form for our linear optimization problem:
Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ be given. Then, we have the LP problem

$$\begin{array}{llrcl} (LP) & \text{Max} & c^\top x & & \\ & & Ax & = & b, \\ & & x & \geq & 0. \end{array}$$

$x \geq 0$ or $x \in \mathbb{R}^n_+$ (x must lie in the nonnegative orthant—a convex cone).

$K \subseteq \mathbb{R}^n$ is a convex cone if $\forall x, v \in K$ and $\alpha \in \mathbb{R}_{++}$, we have $\alpha x \in K$ and $(x + v) \in K$.
Replace $\mathbb{R}^n_+$ by an arbitrary convex cone, to get a more general convex optimization problem.

Semidefinite Optimization

Let $\mathbb{S}^n$ denote the space of n-by-n symmetric matrices with entries in $\mathbb{R}$.

Definition

Let $X \in \mathbb{S}^n$.
X is *positive semidefinite* if

$$h^\top X h \geq 0, \quad \forall h \in \mathbb{R}^n.$$

X is *positive definite* if

$$h^\top X h > 0, \quad \forall h \in \mathbb{R}^n \setminus \{0\}.$$

The set of n-by-n, symmetric positive semidefinite matrices form a closed convex cone in $\mathbb{S}^n$. We denote it by $\mathbb{S}^n_+$.
The set of n-by-n, symmetric positive definite matrices form an open convex cone in $\mathbb{S}^n$. We denote it by $\mathbb{S}^n_{++}$.

For $A, B \in \mathbb{S}^n$, we use the trace inner-product:

$$\langle A, B \rangle := \operatorname{Tr}(A^\top B) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ij},$$

we write

$$A \succeq B$$

to mean $(A - B)$ is positive semidefinite;

$$A \succ B$$

to mean $(A - B)$ is positive definite.

Note that for $X \in \mathbb{S}^n$ all eigenvalues $\lambda_j(X)$ are real. Also,

$$X \succeq 0 \iff \lambda(X) \geq 0$$

and

$$X \succ 0 \iff \lambda(X) > 0.$$

Theorem

For every $X \in \mathbb{S}^n$, there exists $Q \in \mathbb{R}^{n \times n}$, orthogonal ($Q^\top Q = I$) such that

$$X = Q\mathrm{Diag}(\lambda(X))Q^\top.$$

Let us denote by $q^{(1)}, q^{(2)}, \ldots, q^{(n)}$ the columns of Q and denote by e_j the jth unit vector. Then,

$$Xq^{(j)} = Q\mathrm{Diag}(\lambda(X))e_j = \lambda_j(X)Qe_j = \lambda_j(X)q^{(j)},$$

that is, $q^{(j)}$ is an eigenvector of X associated with the eigenvalue $\lambda_j(X)$.

By the above observations

$$\mathrm{Tr}\,(X) = \sum_{j=1}^{n} \lambda_j(X).$$

Some commonly used norms on $\mathbb{S}^n$.
For every $h \in \mathbb{R}^n$, $p \in [1, +\infty]$,

$$\|h\|_p := \left(\sum_{j=1}^{n} |h_j|^p \right)^{\frac{1}{p}}.$$

In the matrix space, we have Frobenius norm:

$$\|X\|_F := \langle X, X \rangle^{1/2} = \|\lambda(X)\|_2 = \sqrt{\sum_{j=1}^{n} (\lambda_j(X))^2}.$$

Operator p-norm (for every $p \in [1, +\infty]$):

$$\|X\|_p := \max \left\{ \|Xh\|_p : h \in \mathbb{R}^n, \|h\|_p = 1 \right\}.$$

Note that

$$\|X\|_2 = \max_j \{ |\lambda_j(X)| \}.$$

For $X \in \mathbb{S}^n_{++}$, we can define its unique symmetric positive semidefinite square root as

$$X^{1/2} := Q\,[\mathrm{Diag}(\lambda(X))]^{1/2}\,Q^\top.$$

We extend this definition to $X \in \mathbb{S}^n_+$.
Another notion of "square-root:"

Theorem

(Cholesky Decomposition) Let $X \in \mathbb{S}^n$. Then $X \in \mathbb{S}^n_+$ iff there exists $B \in \mathbb{R}^{n \times n}$ lower-triangular such that $X = BB^\top$.

Theorem

Let $X \in \mathbb{S}^n$. Then, TFAE

(a) *X is positive semidefinite;*

(b) *$\lambda_j(X) \geq 0$, for every $j \in \{1, 2, \ldots, n\}$;*

(c) *there exist $\mu \in \mathbb{R}^n_+$ and $h^{(i)} \in \mathbb{R}^n$, for every $i \in \{1, 2, \ldots, n\}$ such that*
$$X = \sum_{i=1}^{n} \mu_i h^{(i)} {h^{(i)}}^\top;$$

(d) *there exists $B \in \mathbb{R}^{n \times n}$ such that $X = BB^\top$ (here, B can be chosen as a lower triangular matrix—the Cholesky decomposition of X);*

(e) *for every nonempty $J \subseteq \{1, 2, \ldots, n\}$, $\det(X_J) \geq 0$, where $X_J := \{[X_{ij}] : i, j \in J\}$;*

(f) *for every $S \in \mathbb{S}^n_+$, $\langle X, S \rangle \geq 0$.*

Lemma

(Schur Complement Lemma) Let $X \in \mathbb{S}^n$ *and* $T \in \mathbb{S}^m_{++}$. *Then*

$$M := \begin{pmatrix} T & U^\top \\ U & X \end{pmatrix} \succeq 0 \iff X - UT^{-1}U^\top \succeq 0.$$

Moreover, $M \succ 0$ *iff* $X - UT^{-1}U^\top \succ 0$.

Proof.

$$\begin{pmatrix} I & 0 \\ UT^{-1} & I \end{pmatrix} \begin{pmatrix} T & 0 \\ 0 & X - UT^{-1}U^\top \end{pmatrix} \begin{pmatrix} I & T^{-1}U^\top \\ 0 & I \end{pmatrix} = M.$$

Note that we wrote

$$M = L \begin{pmatrix} T & 0 \\ 0 & X - UT^{-1}U^\top \end{pmatrix} L^\top,$$

where L is nonsingular (in our case $\det(L) = 1$). Therefore,

$$M \succeq 0 \iff X - UT^{-1}U^\top \succeq 0.$$

Also,

$$M \succ 0 \iff X - UT^{-1}U^\top \succ 0.$$

$\square$

One interesting special case is $\begin{pmatrix} 1 & x^\top \\ x & X \end{pmatrix}$.

We have

$$\begin{pmatrix} 1 & x^\top \\ x & X \end{pmatrix} \succeq 0 \iff X - xx^\top \succeq 0$$

and

$$\begin{pmatrix} 1 & x^\top \\ x & X \end{pmatrix} \succ 0 \iff X - xx^\top \succ 0.$$

Theorem

(1) $\mathbb{S}^n_{++} = \text{int}\left(\mathbb{S}^n_+\right).$

(2) *Let $X \in \mathbb{S}^n$. Then, TFAE*

(a) *X is positive definite;*

(b) *$\lambda_j(X) > 0$, for every $j \in \{1, 2, \ldots, n\}$;*

(c) *there exist $\mu \in \mathbb{R}^n_{++}$ and $h^{(i)} \in \mathbb{R}^n$, for every $i \in \{1, 2, \ldots, n\}$ linearly independent such that*

$$X = \sum_{i=1}^{n} \mu_i h^{(i)} {h^{(i)}}^\top;$$

(d) *there exists $B \in \mathbb{R}^{n \times n}$ nonsingular such that $X = BB^\top$ (here, B can be chosen as a lower triangular matrix—the Cholesky decomposition of X);*

(e) *for every $J_k := \{1, 2, \ldots, k\}$, $k \in \{1, 2, \ldots, n\}$, $\det(X_{J_k}) > 0$;*

(f) *for every $S \in \mathbb{S}^n_+ \setminus \{0\}$, $\langle X, S \rangle > 0$;*

(g) *$X \succeq 0$ and $\text{rank}(X) = n$.*

Given $C, A_1, A_2, \ldots, A_m \in \mathbb{S}^n$, $b \in \mathbb{R}^m$ we have

$$
\begin{array}{lrcl}
(P) \quad \inf & \langle C, X \rangle & & \\
 & \langle A_i, X \rangle & = & b_i, \quad \forall i \in \{1, 2, \ldots, m\} \\
 & X & \succeq & 0,
\end{array}
$$

$$
\begin{array}{lrcl}
(D) \quad \sup & b^\top y & & \\
 & \sum_{i=1}^{m} y_i A_i & \preceq & C.
\end{array}
$$

For SDP: Slater point $\bar{y} \in \mathbb{R}^m$ such that

$$\sum_{i=1}^{m} \bar{y}_i A_i \prec C.$$

Theorem

(A Strong Duality Theorem) Suppose (D) has a Slater point. If the objective value of (D) is bounded from above then (P) attains its optimum value and the optimum values of (P) and (D) coincide.

Why do we need extra assumptions for the duality theorem of SDP?
Consider the examples (with parameter $\gamma > 0$):

$$C := \begin{pmatrix} \gamma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, b := \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$A_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then for every feasible solution of (D), we have $y_2 = 0$ and $y_1 \leq 0$. So, the set of feasible solutions of (D) (in the S-space, $S := C - \sum_{i=1}^{m} A_i y_i$) is

$$\left\{ \begin{pmatrix} \gamma & 0 & 0 \\ 0 & S_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} : S_{22} \geq 0 \right\}.$$

Moreover, every feasible solution of (D) is optimal in (D).

For the primal, it is also true that the set of feasible solutions and the set of optimal solutions coincide, which is (in the X-space):

$$\left\{ \begin{pmatrix} 1 & 0 & X_{31} \\ 0 & 0 & 0 \\ X_{31} & 0 & X_{33} \end{pmatrix} : X_{33} \geq X_{31}^2 \right\}.$$

Even though both (P) and (D) have finite optimal objective values, both of these values are attained, there is a duality gap of γ.

In the domain of (LP): Let v^* denote the optimal objective value... Suppose we have a feasible solution $\bar{x}$ for (P) and a feasible solution $\bar{y}$ of (D). Then

$$c^\top \bar{x} \leq v^* \leq b^\top \bar{y}.$$

Moreover, using an optimal $\bar{x}$ for (P) and an optimal $\bar{y}$ of (D), we have

$$c^\top \bar{x} = v^* = b^\top \bar{y}.$$

In the domain of (SDP): Let v^* denote the optimal objective value...
Suppose we have a feasible solution $\bar{X}$ for (P) and a feasible solution $\bar{y}$ of (D). Then

$$\langle C, \bar{X} \rangle \geq v^* \geq b^\top \bar{y}.$$

Moreover, using an optimal $\bar{X}$ for (P) and an optimal $\bar{y}$ of (D) (for a suitably defined (D), or under a Slater type assumption), we have

$$\langle C, \bar{X} \rangle = v^* = b^\top \bar{y}.$$

In the domain of Combinatorial Optimization problems (e.g., 0,1 Integer Programming Problems): Suppose we have an integer feasible solution $\bar{x}$ of the LP or SDP relaxation, and a feasible solution $\bar{y}$ of the dual of the convex relaxation C_k. Then

$$c^\top \bar{x} \leq v^* \leq b^\top \bar{y}.$$

Moreover, using an optimal $\bar{x}$ for the original Integer Programming Problem and an optimal solution $\bar{y}$ of the dual of the convex relaxation $C_d = \mathrm{conv}(\mathcal{F})$, we have

$$c^\top \bar{x} = v^* = b^\top \bar{y}.$$

We can solve such semidefinite optimization problems *efficiently*
both in terms of

- computational complexity theory and
- practical computation.

However, in terms of both computational complexity theory and practical computation the situation in SDP is much worse than that of LP. (That is, there are very many deep, open problems.)

We can consider complexity analyses based on the ellipsoid method. Which typically assumes the existence balls with radii R and r (the former for a ball containing some optimal solutions and the latter for a ball contained in the set of feasible solutions intersected with the first ball). It is also possible to analyze the algorithms for convex optimization in the case of unknown R, r, and to express the computational complexity of the algorithms in terms of R and r. For example, Freund and Vera [2009] prove that if a convex set $F \subseteq \mathbb{R}^d$ is given by a separation oracle, the problem of computing $\bar{x} \in F$ can be solved in

$$O\left(d\left(\ln(d) + d\ln\left(\frac{1}{r} + \frac{R}{r}\right) + \ln(R+1)\right)\right) \text{ iterations,}$$

where each iteration makes one call to the separation oracle (and all the constants hidden by $O(\cdot)$ are at most 2).

Many interior-point algorithms lead to similar polynomial iteration complexity bounds, under similar (but not the same) assumptions. See for instance, Nesterov and Nemirovskii [1994], Alizadeh [1995], Nesterov and Todd [1997,1998], Nesterov, Todd and Ye [1999], Nemirovksii and T. [2005,2010].
For a bit complexity analysis of the SDP feasibility problem (which admits an exponential bound in the size of the input), see Porkolab and Khachiyan [1997].

Given $A_1, A_2, \ldots, A_m \in \mathbb{S}^n \cap \mathbb{Z}^{n \times n}$ and $b \in \mathbb{Z}^m$, the *SDP-Feasibility Problem* is to decide whether there exists $X \in \mathbb{S}^n_+$ such that $\langle A_i, X \rangle = b_i$, for every $i \in \{1, 2, \ldots, m\}$.

Open Problem: *Is SDP-feasibility in $\mathcal{P}$?*

In fact,

Open Problem: *Is SDP-feasibility in $\mathcal{NP}$?*

Theorem

(Ramana [1997]) In the real number computational model, the problem of deciding SDP feasibility is in $\mathcal{NP} \cap$ co-$\mathcal{NP}$.

Theorem

(Ramana [1997]) In the Turing machine model, if SDP feasibility is in $\mathcal{NP}$ then it is in $\mathcal{NP} \cap$ co-$\mathcal{NP}$.

We can start with a possibly nonlinear algebraic representation of the constraint

$$x \in \mathcal{F}.$$

(I) Homogeneous Lifting to the matrix space $\mathbb{S}^{d+1}$: To enforce the constraint $x \in \{0,1\}^d$, it suffices to write

$$x_j^2 - x_j = 0 \quad \forall j \in \{1, 2, \cdots, d\}.$$

Consider representing x in the space $\mathbb{S}^{d+1}$ by the matrix

$$\begin{pmatrix} 1 & x^\top \\ x & xx^\top \end{pmatrix} =: Y.$$

Then, the linear constraints on Y,

$$Y_{00} = 1, diag(Y) = Ye_0,$$

together with the requirements that $Y \succeq 0$ and

$$\text{rank}(Y) = 1,$$

allows us to transform any 0,1 integer programming formulation to an SDP formulation with an additional rank constraint. Removing the rank constraint, leads to an SDP relaxation.

(II) Non-homogeneous Lifting to the matrix space $\mathbb{S}^d$: To enforce the constraint $x \in \{-1, 1\}^d$, it suffices to write

$$x_j^2 = 1 \;\; \forall j \in \{1, 2, \cdots, d\}.$$

Consider representing x in the space $\mathbb{S}^d$ by the matrix

$$xx^\top =: X.$$

Then, the linear constraints on X,

$$diag(X) = \bar{e}, X \succeq 0,$$

together with the requirement that

$$\text{rank}(X) = 1,$$

allows us to transform any -1,1 pure quadratic programming formulation to an SDP formulation with an additional rank constraint.
Removing the rank constraint, leads to an SDP relaxation.

Goemans and Williamson [1995] used the following SDP relaxation of MaxCut

$$\begin{array}{lrcl} \max & -\frac{1}{4}\langle W, X\rangle & & \qquad \left(+\frac{1}{4}\langle W, \bar{e}\bar{e}^\top\rangle\right) \\ \text{subject to:} & \operatorname{diag}(X) & = & \bar{e}, \\ & X & \succeq & 0, \end{array}$$

where W_{ij} is either zero (if $\{i,j\} \notin E$) or is equal to the given nonnegative weight of the edge $\{i,j\}$, in proving their theorem:

Theorem

(Goemans and Williamson [1995]) Let $G = (V, E)$ with $w \in \mathbb{Q}_+^E$ be given. Then a cut in G of value at least 0.87856.*(weight of the max. cut) can be computed in polynomial-time.*

(III) Vector representation of variables: Another approach is to start with a formulation that is based on polynomial equations and inequalities. Suppose we are using $x \in \{-1, 1\}^d$ formulation.
Beat all the higher degree monomials down to quadratics by adding new equations and new variables.
E.g., consider the system

$$\begin{aligned} x_1^4 x_2^2 + x_2^3 x_3 + x_1^5 - 1 &\geq 0, \\ 2x_1^3 - x_2^4 &\geq 0 \end{aligned}$$

$$\begin{aligned} x_1^4 x_2^2 + x_2^3 x_3 + x_1^5 - 1 &\geq 0, \\ 2x_1^3 - x_2^4 &\geq 0 \end{aligned}$$

which is equivalent to the quadratic system:

$$\begin{aligned} x_5 x_6 + x_6 x_7 + x_1 x_5 - 1 &\geq 0, \\ 2x_1 x_4 - x_6^2 &\geq 0, \\ x_4 &= x_1^2, \\ x_5 &= x_4^2, \\ x_6 &= x_2^2, \\ x_7 &= x_2 x_3, \end{aligned}$$

...

Say we have a quadratic inequality

$$x_1^2 + 12x_1x_2 + 23x_2x_3 + x_4^2 \leq 32.$$

Let us represent each variable x_j, by a vector $v^{(j)} \in \mathbb{R}^d$. Then our quadratic inequality becomes

$$\left\langle v^{(1)}, v^{(1)} \right\rangle + 12 \left\langle v^{(1)}, v^{(2)} \right\rangle + 23 \left\langle v^{(2)}, v^{(3)} \right\rangle + \left\langle v^{(4)}, v^{(4)} \right\rangle \leq 32.$$

The quadratic equations $x_j^2 = 1 \ \ \forall j$ become

$$\left\langle v^{(j)}, v^{(j)} \right\rangle = 1 \ \ \forall j.$$

Recall that

$$X \in \mathbb{S}_+^d \text{ iff } X = BB^\top$$

for some $B \in \mathbb{R}^{d \times d}$.
Define

$$B^\top =: \left[v^{(1)} v^{(2)} \dots v^{(d)} \right].$$

Then

$$X_{ij} = \langle v^{(i)}, v^{(j)} \rangle \quad \forall i, j.$$

I.e., Semidefinite Optimization allows us to express a problem by enforcing linear constraints on the variables: $\langle v^{(i)}, v^{(j)} \rangle$.

For example,

$$\operatorname{diag}(X) = \bar{e}$$

is equivalent to:

$$\left\langle v^{(j)}, v^{(j)} \right\rangle = 1 \quad \forall j.$$

Similarly,

$$X_{11} + 12X_{12} + 23X_{23} + X_{44} \leq 32$$

is equivalent to:

$$\left\langle v^{(1)}, v^{(1)} \right\rangle + 12 \left\langle v^{(1)}, v^{(2)} \right\rangle + 23 \left\langle v^{(2)}, v^{(3)} \right\rangle + \left\langle v^{(4)}, v^{(4)} \right\rangle \leq 32.$$

If there are no additional restrictions on the dimension of $v^{(j)}$, and all the constraints are as above, then the SDP formulation is an exact formulation (we will see and example of this next); otherwise (if there is a restriction that say $v^{(j)} \in \mathbb{R}^3$), this leads to an SDP relaxation (for $d \geq 4$).

For example, Arora, Rao and Vazirani [2004] use an approach as above to derive an SDP relaxation for the sparsest cut problem. de Carli Silva, Harvey and Sato [2016] utilize SDP techniques to derive polynomial time algorithms that approximate sums of symmetric positive semidefinite matrices. Such algorithms have applications in finding sparsifiers of hypergraphs, sparse solutions to semidefinite programs, sparsifiers of unique games etc.

(IV) SoS Relaxations of prescribed degree:
Since any system of polynomial inequalities can be reformulated as a system of quadratic inequalities, the above results can be translated to the setting of polynomial optimization problems (POP):

$$\begin{array}{rrcl} \min & p_0(x) & & \\ & p_i(x) & \geq & 0, \quad i \in \{1, 2, \ldots, m\}, \end{array}$$

where $p_0, p_1, \ldots, p_m : \mathbb{R}^n \to \mathbb{R}$ are polynomials.
Before going forward, let's take a break and consider a question about a single multivariate polynomial.

Just a question... How can I convince you that

$$\begin{aligned} f(x) \quad := \quad & 83 - 108x_1 + 216x_2 + x_3^2 x_1^2 - 2x_3^3 x_1 - 32x_1 x_2^3 \\ & +24x_1^2 x_2^2 - 8x_1^3 x_2 - 144x_1 x_2^2 + 72x_1^2 x_2 \\ & -216x_1 x_2 + x_3^2 + 54x_1^2 + 216x_2^2 + 432x_1^5 x_2^2 x_4 \\ & -432x_1^4 x_2^3 x_4 + 144x_1^3 x_2^4 x_4 - 576x_1^3 x_2^3 x_4^2 \\ & +256x_1^2 x_2^2 x_4^4 + 864x_1^4 x_2^2 x_4^2 + 768x_1^3 x_2^2 x_4^3 \\ & -16x_1^2 x_2^5 x_4 - 12x_1^3 + 96x_2^3 + x_1^4 + 16x_2^4 + x_3^4 \\ & +96x_1^2 x_2^4 x_4^2 - 256x_1^2 x_2^3 x_4^3 - 12x_1^3 x_2^5 \\ & +54x_1^4 x_2^4 - 108x_1^5 x_2^3 + 81x_1^6 x_2^2 + x_1^2 x_2^6 \\ & +2x_3^2 x_1 - 2x_3^3 \qquad \geq 2, \quad \forall x \in \mathbb{R}^4? \end{aligned}$$

What if I claim ...

$$\begin{aligned} f(x) &= (x_1 - 2x_2 - 3)^4 + x_3^2 (x_3 - x_1 - 1)^2 \\ &\quad + x_1^2 x_2^2 (3x_1 - x_2 + 4x_4)^4 + 2 \\ &\geq 2, \quad \forall x \in \mathbb{R}^4 ? \end{aligned}$$

That is, $f(x) = \text{Sum-of-Squares} + 2, \ \forall x \in \mathbb{R}^4$.

There is a lot of work in the area of solving (POP) by utilizing Linear Optimization, Semidefinite Optimization and Convex Optimization techniques. See Lasserre [2001-...], Parrilo [2003-...], Laurent [2003-...], Gouveia, Parrilo, Thomas [2009].

Lasserre uses the connections to Putinar's Theorem [1993]:

Theorem

Suppose $\mathcal{F} := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i \in \{1, 2, \ldots, m\}\}$ *is compact, the polynomials* p_i *have even degree, and their highest degree homogeneous parts do not have common zeroes in* $\mathbb{R}^n$ *except 0. Then every polynomial that is positive on* $\mathcal{F}$ *can be written as a nonnegative combination of polynomials of the form*

$$[h_0(x)]^2 + \sum_{i=1}^{m} [h_i(x)]^2 \, p_i(x).$$

Perhaps one of the most fundamental problems here is the *K-moment problem* which is, given $K \subset \mathbb{R}^n$ to decide when a real valued function f of set of monomials in n variables is a moment function $\int_K x^m d\mu$ for some nonnegative Borel measure μ on K. Schmüdgen [1991] characterized the solutions to the K-moment problem (called *K-moment sequences*) for all compact semi-algebraic sets K in terms of the positive definiteness of matrices arising from the moment functions. Schmüdgen's proof utilizes Positivstellensatz in proving the above-mentioned algebraic fact. The result also generalizes many other preexisting beautiful results such as Handelman's Theorem [1988]; some of these connections are old and they generalize results some of which go all the way back to Minkowski in late 1800's.

Positivstellensatz (Stengle [1974]):
$\mathcal{F} := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i \in \{1, 2, \ldots, m\}\} = \emptyset$ iff that there exists $g \in \text{cone}\,(p_1, p_2, \ldots, p_m)$ such that $g(x) = -1$.
That is, iff there exist $s_0, s_J, \ldots \in \text{SoS}(n, *)$ such that

$$g = \sum_{J \subseteq \{1,2,\ldots,m\}} s_J \prod_{i \in J} p_i = -1.$$

Positivstellensatz is a "common" generalization of Farkas' Lemma

Lemma

(Farkas' Lemma) Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ be given. Then exactly one of the following systems has a solution:

(I) $Ax = b, x \geq 0;$

(II) $A^\top y \geq 0, b^\top y < 0.$

and Hilbert's Nullstellensatz [1901] (characterizing when a system of polynomial equations has no solution over $\mathbb{C}^n$): Only for this slide, let p_i be polynomials in complex variables ($x \in \mathbb{C}^n$). Then exactly one of the following systems has a solution:

(I) $p_i(x) = 0, \forall i \in \{1, 2, \ldots, m\};$

(II) $\exists$ polynomials h_i such that $\sum_{i=1}^m h_i(x) p_i(x) = -1.$

Parrilo uses the Positivstellensatz and gets SoS certificates!
Given $x \in \mathbb{R}^n$ and polynomial $f : \mathbb{R}^n \to \mathbb{R}$ of degree $2d$, let

$$h(x) := [1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, x_1x_3, \ldots, x_2^2, \ldots, x_n^d]^\top \in \mathbb{R}^N,$$

where $N := \binom{n+d}{d}$. We are interested in

$$F(f) := \left\{ X \in \mathbb{S}^N : [h(x)]^\top X h(x) = f(x) \right\}.$$

The following well-known fact connects SoS and semidefinite optimization.

Theorem

Let $\bar{z} \in \mathbb{R}$. Then $[f(x) - \bar{z}]$ is SoS iff

$$\left\{ X \in F(f) : X \succeq \bar{z} e_1 e_1^\top \right\} \neq \emptyset.$$

Let $G = (V, E)$ be given. *A unit distance representation of G* is $v : V \to \mathbb{R}^d$ for some $d \geq 1$ such that

$$\|v^{(i)} - v^{(j)}\|_2 = 1, \quad \forall \{i, j\} \in E.$$

Theorem

(Lovász [2003]) Let $G = (V, E)$ be a given graph. Then, the optimal objective value of the following semidefinite optimization problem is finite, it is attained and is equal to the square of the radius of the smallest ball containing a unit distance representation of G:

$$\begin{array}{rlrcll} t(G) := & \min & t & & & \\ & \text{subject to:} & \operatorname{diag}(X) - t\bar{e} & \leq & 0, & \\ & & X_{ii} - 2X_{ij} + X_{jj} & = & 1, & \forall \{i, j\} \in E, \\ & & X & \in & \mathbb{S}_{+}^{V}. & \end{array}$$

Theorem

(Lovász [2003]) Let $G = (V, E)$ be a given graph. Then, the optimal objective value of the following semidefinite optimization problem is finite, it is attained and is equal to the square of the radius of the smallest ball containing a smallest hypersphere representation of G:

$$\begin{array}{lrcll} \min & t & & & \\ \text{subject to:} & \operatorname{diag}(X) - t\bar{e} & = & 0, & \\ & X_{ii} - 2X_{ij} + X_{jj} & = & 1, & \forall \{i,j\} \in E, \\ & X & \in & \mathbb{S}^V_+. & \end{array}$$

We can consider generalization of such representations to ellipsoids. Find a unit distance representation of a given graph G which lies in the "smallest" ellipsoid. See, de Carli Silva [2013], de Carli Silva and T. [2013]. Many of these problems are hard, and there are many open problems related to such ellipsoidal representation problems.

Unit representations of graphs in a ball and unit representations of graphs on hyperspheres are closely connected to *orthonormal representations of graphs*:

$$\{u^{(i)} \in \mathbb{R}^d : i \in V\}$$

is called an *orthonormal representation of G* if

- $\|u^{(i)}\|_2 = 1, \ \forall i \in V$, and
- $\langle u^{(i)}, u^{(j)} \rangle = 0, \ \forall \{i,j\} \in \bar{E}$.

We will cover the connections of such representations to the stable set problem in the next section.

We have A and b given, describing a nonempty polytope

$$P := \left\{ x \in \mathbb{R}^d : Ax \le b,\ 0 \le x \le \bar{e} \right\}.$$

We are interested in $0,1$ vectors in P:

$$P_I := \operatorname{conv}\left(P \cap \{0,1\}^d \right).$$

We will consider operators Γ that take a compact convex set $C_k \subseteq [0,1]^d$ and return a compact convex set C_{k+1} such that

- $C_k \cap \{0,1\}^d \subseteq C_{k+1} := \Gamma(C_k) \subseteq C_k$,
- $C_{k+1} \neq C_k$ unless $C_k = \text{conv}\left(C_k \cap \{0,1\}^d\right)$.

Note that the given system of equations and inequalities is:

$$Ax \le b, \quad 0 \le x \le \bar{e},$$

$$x_j^2 - x_j = 0, \quad \forall j \in \{1, 2, \ldots, d\}.$$

$$P := \left\{x \in \mathbb{R}^d : Ax \leq b,\ 0 \leq x \leq \bar{e}\right\}.$$

We are interested in $0,1$ vectors in P:

$$P_I := \operatorname{conv}\left(P \cap \{0,1\}^d\right).$$

$$BCC_{(j)}(P) := \operatorname{conv}\left\{(P \cap \{x : x_j = 0\}) \cup (P \cap \{x : x_j = 1\})\right\},$$

where $j \in \{1, 2, \ldots, d\}$. Since the inclusions

$$P_I \subseteq BCC_{(j)}(P) \subseteq P$$

are clear for every j, it makes sense to consider applying this operator iteratively, each time for a new index j.

Let us define

$$J := \{j_1, j_2, \ldots, j_k\} \subseteq \{1, 2, \ldots, d\}.$$

Let us denote

$$BCC_{(J)}(P) := BCC_{(j_k)}\left(BCC_{(j_{k-1})}\left(\cdots BCC_{(j_1)}(P)\cdots\right)\right).$$

It is easy to check that in the above context, the operators $BCC_{(j)}$ commute with each other.
Therefore, the notation $BCC_{(J)}(\cdot)$ is justified.

A beautiful, fundamental property of these operators is:

Lemma

For every $J \subseteq \{1, 2, \ldots, d\}$, *we have*

$$BCC_{(J)}(P) = \operatorname{conv}\left(P \cap \{x : x_j \in \{0, 1\}, \forall j \in J\}\right).$$

The lemma directly leads to the convergence theorem.

Theorem

(Balas [1974]) Let P be as above. Then

$$BCC_{(\{1,2,\ldots,d\})}(P) = P_I.$$

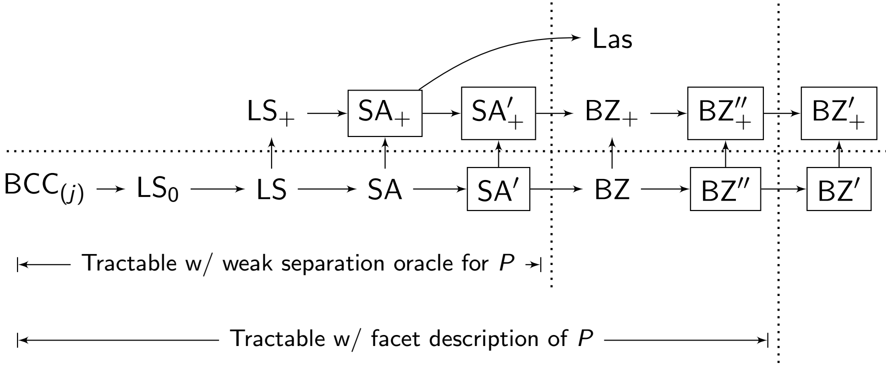


Figure: Various properties of lift-and-project operators (Au and T. [2011, 2015, 2016]).

Lovász and Schrijver [1991] proposed:

$$M_0(K) := \Big\{ Y \in \mathbb{R}^{(d+1)\times(d+1)} : \quad Ye_0 = Y^\top e_0 = \mathrm{diag}(Y), \; Ye_i \in K, \; Y(e_0 - e_i) \in K, \; \forall i \in \{1, 2, \ldots, d\} \Big\}$$

$$\mathrm{LS}_0(K) := \{ Ye_0 : Y \in M_0(K) \}.$$

Tighter,

$$M(K) := M_0(K) \cap \mathbb{S}^{d+1},$$

$$\mathrm{LS}(K) := \{Ye_0 : Y \in M(K)\}.$$

and tighter,

$$M_+(K) := M_0(K) \cap \mathbb{S}_+^{d+1},$$

$$\mathrm{LS}_+(K) := \{ Ye_0 : Y \in M_+(K) \}.$$

Lemma

Let K be as above. Then

$$K_I \subseteq \mathrm{LS}_+(K) \subseteq \mathrm{LS}(K) \subseteq \mathrm{LS}_0(K) \subseteq K.$$

Theorem

(Lovász and Schrijver [1991]) Let P be as above. Then

$$P \supseteq \mathrm{LS}_0(P) \supseteq \mathrm{LS}_0^2(P) \supseteq \cdots \supseteq LS_0^d(P) = P_I.$$

Similarly for LS as well as LS_+.

Moreover, the relaxations obtained after a few iterations are still tractable if the original relaxation P is.

Theorem

(Lovász and Schrijver [1991]) Let P be as above. If we have a polynomial time weak separation oracle for P then we can optimize any linear function over any of $LS_0^k(P)$, $LS^k(P)$, $LS_+^k(P)$ in polynomial time, provided $k = O(1)$.

There is a wide spectrum of lift-and-project type operators: Balas [1974], Sherali and Adams [1990], Lovász and Schrijver [1991], Balas, Ceria and Cornuéjols [1993], Kojima and T. [2000], Lasserre [2001], de Klerk and Pasechnik [2002], Parrilo [2003], Bienstock and Zuckerberg [2004].

Such a general method (it applies to every combinatorial optimization problem)...
Can it be really good on any problem?

Let $G = (V, E)$ be an undirected graph.
We define the *fractional stable set polytope* as

$$\mathrm{FRAC}(G) := \left\{ x \in [0,1]^V : x_i + x_j \leq 1 \text{ for all } \{i,j\} \in E \right\}.$$

This polytope is used as the initial approximation to the convex hull of incidence vectors of the *stable sets* of G, which is called the *stable set polytope*:

$$\mathrm{STAB}(G) := \mathrm{conv}\left(\mathrm{FRAC}(G) \cap \{0,1\}^V \right).$$

Let us define the class of *odd-cycle inequalities*. Let $\mathcal{H}$ be the node set of an odd-cycle in G then the inequality

$$\sum_{i \in \mathcal{H}} x_i \leq \frac{|\mathcal{H}| - 1}{2}$$

is valid for $\mathrm{STAB}(G)$. We define

$$\mathrm{OC}(G) := \{x \in \mathrm{FRAC}(G) : x \text{ satisfies all odd-cycle constraints for } G\}.$$

If $\mathcal{H}$ is an odd-anti-hole then the inequality

$$\sum_{i \in \mathcal{H}} x_i \leq 2$$

is valid for STAB(G).

If we have an odd-wheel in G with *hub* node represented by x_{2k+2} and the *rim* nodes represented by $x_1, x_2, \ldots, x_{2k+1}$, then the *odd-wheel inequality*

$$kx_{2k+2} + \sum_{i=1}^{2k+1} x_i \leq k$$

is valid for $\mathrm{STAB}(G)$.

Based on these classes of inequalities we define the polytopes

$$\mathrm{OC}(G), \mathrm{ANTI\text{-}HOLE}(G), \mathrm{WHEEL}(G).$$

Theorem

(Lovász and Schrijver [1991]) For every graph G,

$$\mathrm{LS}_0(G) = \mathrm{LS}(G) = \mathrm{OC}(G).$$

Note that this theorem provides a compact lifted representations of the odd-cycle polytope of G (in the spaces $\mathbb{R}^{(\{0\}\cup V)\times(\{0\}\cup V)}$ and $\mathbb{S}^{\{0\}\cup V}$). This polytope can have exponentially many facets in the worst case.

However, $M(G)$ is represented by

$$|V|(|V|+1)/2 \text{ variables and } O\left(|V|^3\right) \text{ linear inequalities.}$$

What about $\mathrm{LS}_0^2(G), \mathrm{LS}^2(G)$?

What about $\mathrm{LS}_0^2(G), \mathrm{LS}^2(G)$?
There exist graphs G for which

$\mathrm{LS}_0^2(G) \neq \mathrm{LS}^2(G)$ Au and T. [2009].

Open Problem: Give good combinatorial characterizations for $\mathrm{LS}_0^2(G)$ and $\mathrm{LS}^2(G)$.

Some partial results by Lipták [1999] and by Lipták and T. [2003].

A *clique* in G is a subset of nodes in G so that every pair of them are joined by an edge. The *clique polytope* of G is defined by

$$\mathrm{CLQ}(G) := \left\{ x \in \mathbb{R}^V_+ : \sum_{j \in C} x_j \leq 1 \text{ for every clique } C \text{ in } G \right\}.$$

Optimizing a linear function over $\mathrm{FRAC}(G)$ is easy!
Linear optimization over $\mathrm{CLQ}(G)$ (and $\mathrm{STAB}(G)$) is $\mathcal{NP}$-hard!

Orthonormal Representations of Graphs and the Theta Body of $G := (V, E)$

$u^{(1)}, u^{(2)}, \ldots, u^{(|V|)} \in \mathbb{R}^d$ such that

$$\langle u^{(i)}, u^{(j)} \rangle = 0, \text{ for all } i \neq j, \{i, j\} \notin E,$$

and

$$\langle u^{(i)}, u^{(i)} \rangle = 1, \quad \text{for all } i \in V.$$

$$\mathrm{TH}(G) := \left\{ x \in \mathbb{R}^V_+ : \sum_{i \in V} \left(c^\top u^{(i)} \right)^2 x_i \leq 1, \right.$$
$$\left. \forall \text{ ortho. representations and } c \in \mathbb{R}^d \text{ s.t. } \|c\|_2 = 1 \right\}$$

$$\mathrm{TH}(G) \supseteq \mathrm{STAB}(G)$$

but infinitely many linear inequalities!

Let $\chi(G)$ denote the *chromatic number* of G.
Clique number:

$$\omega(G) := \max\{|C| : C \text{ is a clique in } G\}$$

G is *perfect* if for every node-induced subgraph H of G, we have

$$\chi(H) = \omega(H).$$

Theorem

Let $G = (V, E)$. Then TFAE

(i) *G is perfect*

(ii) $\mathrm{TH}(G)$ *is a polytope*

(iii) $\mathrm{TH}(G) = \mathrm{CLQ}(G)$

(iv) $\mathrm{STAB}(G) = \mathrm{CLQ}(G)$

(v) *G does not contain an odd-hole or odd anti-hole*

(vi) *the ideal generated by* $\left\{(x_v^2 - x_v), \forall v \in V; x_u x_v, \forall \{u, v\} \in E\right\}$ *is* $(1, 1)$*-SoS.*

Follows from the work of Chudnovsky, Chvátal, Fulkerson, Gouveia, Grötschel, Lovász, Parrilo, Robertson, Schrijver, Seymour, Rekha Thomas, Robin Thomas.

TH($\cdot$) behaves well under duality and graph complementation:

Theorem

For every graph $G = (V, E)$,

$$[\mathrm{TH}(G)]^o \cap \mathbb{R}_+^V = \mathrm{TH}(\overline{G}).$$

That is, the polar of TH(G) when restricted to the nonnegative orthant, coincides with the TH($\cdot$) set of the complement of G. (Polar of a set intersected with the nonnegative orthant is called *the antiblocker* of the set— another notion of duality.)
Convex sets that satisfy such a beautiful property have been characterized (under mild assumptions). See, de Carli Silva [2013], de Carli Silva and T. [2016]. This generalization covers a large class of relaxations of the stable set polytope of graphs. It also leads to various variants of the Lovász theta number.

There is a strong connection between $\mathrm{LS}_+(G)$ and $\mathrm{TH}(G)$:

Theorem

(Lovász and Schrijver [1991]) Let $G = (V, E)$. Then

$$\mathrm{TH}(G) = \left\{ x \in \mathbb{R}^V : \begin{pmatrix} 1 \\ x \end{pmatrix} = Ye_0;\ Y_{ij} = 0, \forall \{i,j\} \in E; \right. \\ \left. Ye_0 = \mathrm{diag}(Y);\ Y \succeq 0 \right\}.$$

Theorem

(Lovász and Schrijver [1991]) For every graph G,

$$LS_+(G) \subseteq \mathrm{OC}(G) \cap \text{ANTI-HOLE}(G) \cap \mathrm{WHEEL}(G) \cap \mathrm{CLQ}(G) \cap \mathrm{TH}(G).$$

Open Problem: Give full, elegant, combinatorial characterizations for $\mathrm{LS}_+(G)$.

Is $\mathrm{LS}_+(G)$ polyhedral for every G?

Let $G_{\alpha\beta}$ be the graph in the following figure:

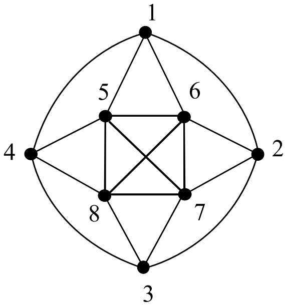


Figure: The graph $G_{\alpha\beta}$.

A two dimensional cross-section of the compact convex relaxation $\mathrm{LS}_+(G_{\alpha\beta})$ has a nonpolyhedral piece on its boundary.
We say that $z \in \mathbb{R}^8$ is an *$\alpha\beta$-point*, if α and β are both nonnegative and $z_i := \begin{cases} \alpha & \text{if} \quad i \in \{1,2,3,4\}, \\ \beta & \text{if} \quad i \in \{5,6,7,8\}. \end{cases}$

Theorem

(Bianchi, Escalante, Nasini, T. [2014]) An $\alpha\beta$-point with

$$\frac{1}{4} \leq \alpha \leq \frac{1}{2}$$

belongs to $\mathrm{LS}_+(G_{\alpha\beta})$ *if and only if*

$$\beta \leq \frac{3 - \sqrt{1 + 8(-1 + 4\alpha)^2}}{8}.$$

Moreover, if an $\alpha\beta$-point is in the set with the maximum allowed value for β, then it is an extreme point of the set.

The SDP relaxation $\mathrm{LS}_+(G)$ of $\mathrm{STAB}(G)$ is stronger than $\mathrm{TH}(G)$. By following the same line of reasoning used for perfect graphs, MWSSP can be solved in polynomial time for the class of graphs for which $\mathrm{LS}_+(G) = \mathrm{STAB}(G)$.

We call these LS_+*-perfect graphs*.

If G' is a node-induced subgraph of G ($G' \subseteq G$), we consider every point in $\mathrm{STAB}(G')$ as a set of points in $\mathrm{STAB}(G)$, although they do not belong to the same space
(for the missing nodes, we take direct sums with the interval $[0,1]$, since originally $\mathrm{STAB}(G) \subseteq \mathrm{STAB}(G') \oplus [0,1]^{V(G)\setminus V(G')}$).
With this notation, given any family of graphs $\mathcal{F}$ and a graph G, we denote by $\mathcal{F}(G)$ the relaxation of $\mathrm{STAB}(G)$ defined by

$$\mathcal{F}(G) := \bigcap_{G' \subseteq G;\, G' \in \mathcal{F}} \mathrm{STAB}(G').$$

A graph is called *near-bipartite* if after deleting the closed neighborhood of *any* node, the resulting graph is bipartite. Let us denote by NB the class of all near-bipartite graphs.
For every graph G,

- $\mathrm{LS}_+(G) \subseteq \mathrm{NB}(G)$ and
- $\mathrm{NB}(G) \subseteq \mathrm{CLQ}(G) \cap \mathrm{OC}(G) \cap \mathrm{ANTI\text{-}HOLE}(G) \cap \mathrm{WHEEL}(G)$.

If G' is a node-induced subgraph of G ($G' \subseteq G$), we consider every point in $\text{STAB}(G')$ as a point in $\text{STAB}(G)$, although they do not belong to the same space (for the missing nodes, we take direct sums with the interval $[0, 1]$, since originally $\text{STAB}(G) \subseteq \text{STAB}(G') \oplus [0, 1]^{V(G) \setminus V(G')}$). With this notation, given any family of graphs $\mathcal{F}$ and a graph G, we denote by $\mathcal{F}(G)$ the relaxation of $\text{STAB}(G)$ defined by

$$\mathcal{F}(G) := \bigcap_{G' \subseteq G;\, G' \in \mathcal{F}} \text{STAB}(G').$$

Open Problem: Find a combinatorial characterization of LS_+-perfect graphs.

Current best characterization (Bianchi, Escalante, Nasini, T. [2014])

$$\mathrm{LS}_+(G) \subseteq \mathrm{NB}(G) \cap \hat{\mathrm{TH}}(G).$$

What is the smallest graph which is LS_+-imperfect?

In a related context, Knuth (1993) asked what is the smallest graph for which $\mathrm{STAB}(G) \neq \mathrm{LS}_+(G)$?

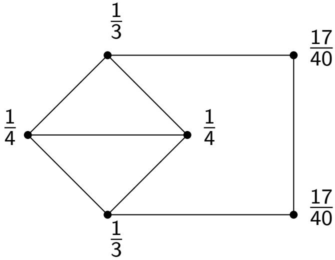


Figure: Little graph that could! G_2 with corresponding weights

Proposition

(Lipták, T., 2003) G_2 is the smallest graph for which $\mathrm{LS}_+(G) \neq \mathrm{STAB}(G)$.

The LS-*rank of P* is the smallest k for which $\mathrm{LS}^k(P) = P_I$.
Analogously, LS_0-*rank of P*, LS_+-*rank of* P_I *relative to P* ...
We denote these ranks by $r(G)$, $r_0(G)$, and $r_+(G)$, respectively.

Theorem

(Lipták, T. [2003]) For every graph $G = (V, E)$, $r_+(G) \leq \left\lfloor \frac{|V|}{3} \right\rfloor$.

$$n_+(k) := \min\{|V(G)| : r_+(G) = k\}.$$

Open Problem: What are the values of $n_+(k)$ for every $k \in \mathbb{Z}_+$? In particular,
Conjecture (Lipták, T. [2003]): Is it true that $n_+(k) = 3k$ for all $k \in \mathbb{Z}_+$?

$$n_+(k) := \min\{|V(G)| : r_+(G) = k\}.$$

Open Problem: What are the values of $n_+(k)$ for every $k \in \mathbb{Z}_+$? In particular,

Conjecture (Lipták, T. [2003]): Is it true that $n_+(k) = 3k$ for all $k \in \mathbb{Z}_+$?

$k = 1$ (triangle is the answer);

$k = 2$ the above graph G_2 is the answer;

$k = 3$, Escalante, Montelar, Nasini (2006);

$k = 4$?

What about the polyhedral graph ranks?

Conjecture (Lipták, T. [2003]): $r_0(G) = r(G) \quad \forall \text{ graphs } G.$

True for:

bipartite graphs, series-parallel graphs, perfect graphs and odd-star-subdivisions of graphs in $\mathcal{B}$ (which contains cliques and wheels, among many other graphs), antiholes and graphs that have LS_0-rank ≤ 2. Also true for all 8-node graphs, and for 9-node graphs that contain a 7-hole or a 7-antihole as an induced subgraph Au [2008].

Other lower bound results: Stephen and T. [1999], Cook and Dash [2000], Goemans and T. [2001], Laurent [2002], Laurent [2003], Aguilera, Bianchi and Nasini [2004], Escalante, Montelar and Nasini [2006], Arora, Bollobás, Lovász and Tourlakis [2006], Cheung [2007], Georgiou, Magen, Pitassi, Tourlakis [2007], Schoeneback, Trevisan and Tulsiani [2007], Charikar, Makarychev and Makarychev [2009], Mathieu and Sinclair [2009], Raghavendra and Steurer [2009], Benabbas and Magen [2010], Karlin, Mathieu and Thach Nguyen [2011], Chan, Lee, Raghavendra and Steurer [2013], Thapper and Zivny [2016].

Many of the lower bound proofs have been unified/generalized: Hong and T. [2008].

Other work on convex relaxation methods on the stable set problem using cone of copositive matrices (and the cone of completely positive matrices): de Klerk and Pasechnik [2002], Peña, Vera and Zuluaga [2008], ...

Stronger “lower bound” results via study of extended complexity.

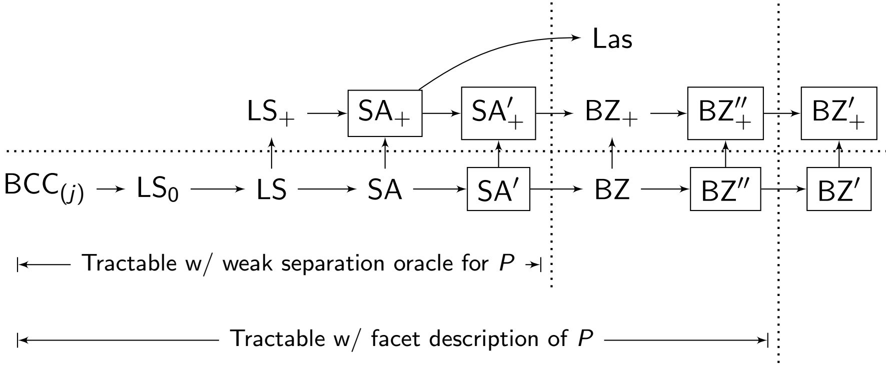


Figure: Various properties of lift-and-project operators (Au and T. [2011, 2015, 2016]).

Denote $\{0,1\}^d$ by $\mathcal{F}$. Define $\mathcal{A} := 2^{\mathcal{F}}$. For each $x \in \mathcal{F}$, we define the vector $x^{\mathcal{A}} \in \mathbb{R}^{\mathcal{A}}$ such that

$$x^{\mathcal{A}}_{\alpha} = \begin{cases} 1, & \text{if } x \in \alpha; \\ 0, & \text{otherwise.} \end{cases}$$

For any given $x \in \mathcal{F}$, if we define $Y^x_{\mathcal{A}} := x^{\mathcal{A}}(x^{\mathcal{A}})^\top$, then, the following must hold:

- $Y^x_{\mathcal{A}} e_0 = (Y^x_{\mathcal{A}})^\top e_0 = \mathrm{diag}\,(Y^x_{\mathcal{A}}) = x^{\mathcal{A}}$;
- $Y^x_{\mathcal{A}} e_\alpha \in \left\{0, x^{\mathcal{A}}\right\},\ \forall \alpha \in \mathcal{A}$;
- $Y^x_{\mathcal{A}} \in \mathbb{S}^{\mathcal{A}}_+$;
- $Y^x_{\mathcal{A}}[\alpha, \beta] = 1 \iff x \in \alpha \cap \beta$;
- If $\alpha_1 \cap \beta_1 = \alpha_2 \cap \beta_2$, then $Y^x_{\mathcal{A}}[\alpha_1, \beta_1] = Y^x_{\mathcal{A}}[\alpha_2, \beta_2]$.

Given $S \subseteq [d]$ and $t \in \{0, 1\}$, we define

$$S|_t := \{x \in \mathcal{F} : x_i = t, \ \forall i \in S\}.$$

For any integer $i \in [0, d]$, define

$$\mathcal{A}_i := \{S|_1 \cap T|_0 : S, T \subseteq [n], S \cap T = \emptyset, |S| + |T| \leq i\}$$

and

$$\mathcal{A}_i^+ := \{S|_1 : S \subseteq [d], |S| \leq i\}.$$

- Let $\tilde{\mathrm{SA}}^k(P)$ denote the set of matrices $Y \in \mathbb{R}^{\mathcal{A}_1^+ \times \mathcal{A}_k}$ that satisfy all of the following conditions:

(SA 1) $Y[\mathcal{F}, \mathcal{F}] = 1$;

(SA 2) $\hat{x}(Ye_\alpha) \in K(P)$ for every $\alpha \in \mathcal{A}_k$;

(SA 3) For each $S|_1 \cap T|_0 \in \mathcal{A}_{k-1}$, impose

$$Ye_{S|_1 \cap T|_0} = Ye_{S|_1 \cap T|_0 \cap j|_1} + Ye_{S|_1 \cap T|_0 \cap j|_0}, \ \forall j \in [n] \setminus (S \cup T).$$

(SA 4) For each $\alpha \in \mathcal{A}_1^+, \beta \in \mathcal{A}_k$ such that $\alpha \cap \beta = \emptyset$, impose $Y[\alpha, \beta] = 0$;

(SA 5) For every $\alpha_1, \alpha_2 \in \mathcal{A}_1^+, \beta_1, \beta_2 \in \mathcal{A}_k$ such that $\alpha_1 \cap \beta_1 = \alpha_2 \cap \beta_2$, impose $Y[\alpha_1, \beta_1] = Y[\alpha_2, \beta_2]$.

- Let $\tilde{\mathrm{SA}}_+^k(P)$ denote the set of matrices $Y \in \mathbb{S}_+^{\mathcal{A}_k}$ that satisfies all of the following conditions:

(SA$_+$ 1) (SA 1), (SA 2) and (SA 3);

(SA$_+$ 2) For each $\alpha, \beta \in \mathcal{A}_k$ such that $\operatorname{conv}(\alpha) \cap \operatorname{conv}(\beta) \cap P = \emptyset$, impose $Y[\alpha, \beta] = 0$;

(SA$_+$ 3) For any $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathcal{A}_k$ such that $\alpha_1 \cap \beta_1 = \alpha_2 \cap \beta_2$, impose $Y[\alpha_1, \beta_1] = Y[\alpha_2, \beta_2]$.

- Define

$$\mathrm{SA}^k(P) := \left\{ x \in \mathbb{R}^d : \exists Y \in \tilde{\mathrm{SA}}^k(P) : Y e_{\mathcal{F}} = \hat{x} \right\}$$

and

$$\mathrm{SA}_+^k(P) := \left\{ x \in \mathbb{R}^d : \exists Y \in \tilde{\mathrm{SA}}_+^k(P) : \hat{x}(Y e_{\mathcal{F}}) = \hat{x} \right\}.$$

The SA_+^k operator extends the lifted space of the SA^k operator to a set of square matrices, and imposes an additional positive semidefiniteness constraint. Moreover, SA_+^k refines the operator LS_+^k.

Given $P := \left\{x \in [0,1]^d : Ax \leq b\right\}$, and an integer $k \in [d]$,

1. Let $\tilde{\text{Las}}^k(P)$ denote the set of matrices $Y \in \mathbb{S}_+^{\mathcal{A}_{k+1}^+}$ that satisfy all of the following conditions:

(Las 1) $Y[\mathcal{F}, \mathcal{F}] = 1$;

(Las 2) For each $j \in [m]$, let A^j be the j^{th} row of A. Define the matrix $Y^j \in \mathbb{S}^{\mathcal{A}_k^+}$ such that

$$Y^j[S|_1, S'|_1] := b_j Y[S|_1, S'|_1] - \sum_{i=1}^{n} A_i^j Y[(S \cup \{i\})|_1, (S' \cup \{i\})|_1]$$

and impose $Y^j \succeq 0$.

(Las 3) For every $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathcal{A}_k^+$ such that $\alpha_1 \cap \beta_1 = \alpha_2 \cap \beta_2$, impose $Y[\alpha_1, \beta_1] = Y[\alpha_2, \beta_2]$.

2. Define

$$\text{Las}^k(P) := \left\{x \in \mathbb{R}^d : \exists Y \in \tilde{\text{Las}}^k(P) : \hat{x}(Ye_{\mathcal{F}}) = \hat{x}\right\}.$$

In our setting, the Las-rank of a polytope P (the smallest k such that $\text{Las}^k(P) = P_I$) is equal to the Theta-rank, defined by Gouveia, Parrilo, Thomas [2010].

Consider the set

$$P_{d,\alpha} := \left\{ x \in [0,1]^d : \sum_{i=1}^{d} x_i \leq d - \alpha \right\}.$$

Theorem

(Au and T. [2015]) Suppose an integer $d \geq 5$ is not a perfect square. Then there exists

$$\alpha \in \left(\left\lfloor \sqrt{d} \right\rfloor, \left\lceil \sqrt{d} \right\rceil \right)$$

such that the BZ'_+-rank of $P_{d,\alpha}$ is at least $\left\lfloor \frac{\sqrt{d}+1}{2} \right\rfloor$.

Theorem

(Au and T. [2015]) For every $d \geq 2$, the SA_+*-rank of* $P_{d,\alpha}$ *is d for all* $\alpha \in (0,1)$.

Theorem

(Cheung [2007])

1. *For every even integer $d \geq 4$, the* Las-rank *of $P_{d,\alpha}$ is at most $d-1$ for all $\alpha \geq \frac{1}{d}$;*
2. *For every integer $d \geq 2$, there exists $\alpha \in \left(0, \frac{1}{d}\right)$ such that the* Las-rank *of $P_{d,\alpha}$ is d.*

Theorem

(Au and T. [2015]) Suppose $d \geq 2$, and

$$0 < \alpha \leq d \left(\frac{3-\sqrt{5}}{4d-4} \right)^{d}.$$

Then $P_{d,\alpha}$ has Las-rank *d.*

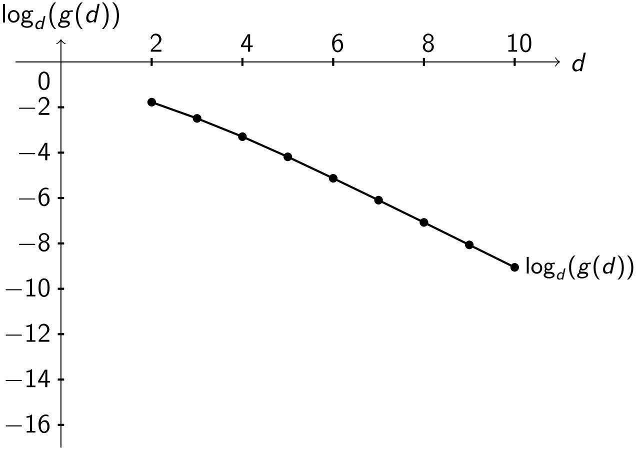


Figure: Computational results and upper bounds for $g(d) := \max\left\{\alpha : \mathrm{Las}^{d-1}(P_{d,\alpha}) \neq (P_{d,\alpha})_I\right\}$ (Au and T. [2015]).

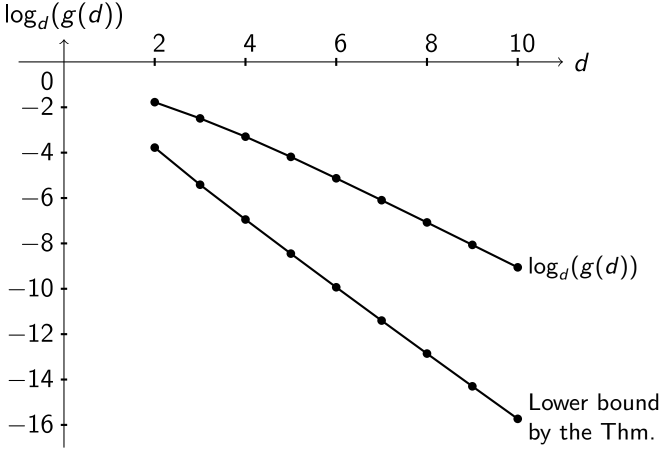


Figure: Computational results and upper bounds for
$g(d) := \max\left\{\alpha : \mathrm{Las}^{d-1}(P_{d,\alpha}) \neq (P_{d,\alpha})_I\right\}$ (Au and T. [2015]).

Given $\alpha > 0$, we define the set

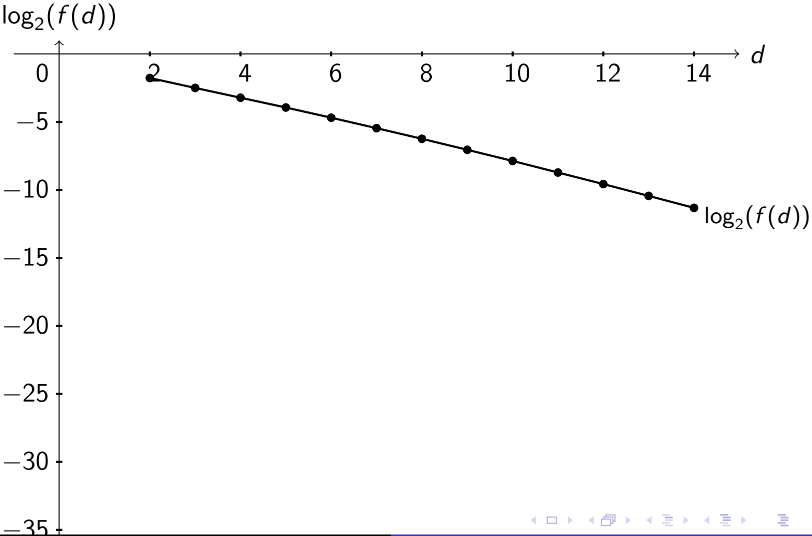
$$Q_{d,\alpha} := \left\{ x \in [0,1]^d : \sum_{i \in S} (1 - x_i) + \sum_{i \notin S} x_i \geq \alpha, \ \forall S \subseteq [d] \right\}.$$

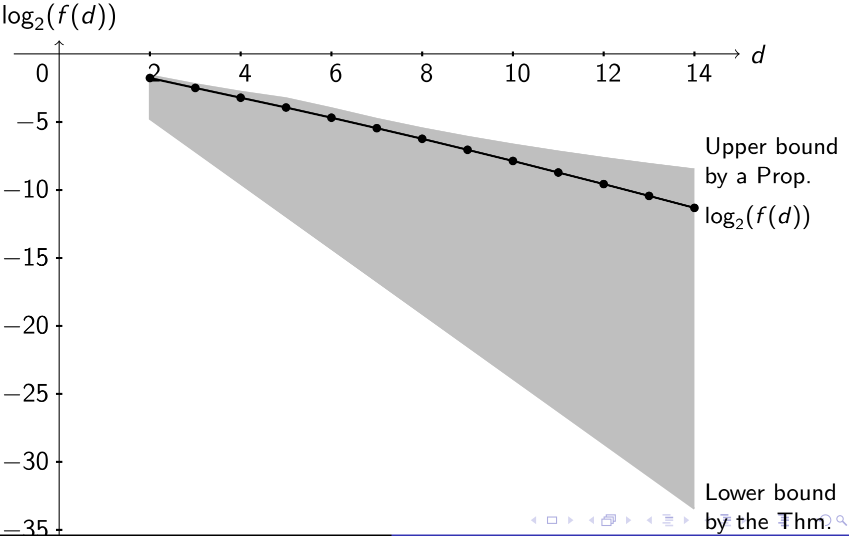
Theorem

(Au and T. [2015]) Suppose $d \geq 2$, and

$$0 < \alpha \leq \left(\frac{3 - \sqrt{5}}{4} \right)^d.$$

Then $Q_{d,\alpha}$ has Las*-rank d.*


$\log_2(f(d))$
d
0
2
4
6
8
10
12
14
-5
-10
-15
-20
-25
-30
-35
$\log_2(f(d))$


$\log_2(f(d))$
d
0
2
4
6
8
10
12
14
-5
-10
-15
-20
-25
-30
-35
Upper bound by a Prop.
$\log_2(f(d))$
Lower bound by the Thm.

For the complete graph $G := K_d$, FRAC(G) has rank 1 with respect to LS_+, SA_+ and Las operators. However, the rank is known to be $\Theta(d)$ for all other operators that yield only polyhedral relaxations, such as SA and Lovász and Schrijver's LS operator.

For the complete graph $G := K_d$, FRAC(G) has rank 1 with respect to LS$_+$, SA$_+$ and Las operators. However, the rank is known to be $\Theta(d)$ for all other operators that yield only polyhedral relaxations, such as SA and Lovász and Schrijver's LS operator.

Theorem

(Au [2014], Au and T. [2016]) Suppose G is the complete graph on $d \geq 3$ vertices. Then the BZ$'$*-rank (and the* BZ*-rank) of* FRAC(G) *is between* $\left\lceil \frac{d}{2} \right\rceil - 2$ *or* $\left\lceil \frac{d+1}{2} \right\rceil$.

Stronger "lower bound" results via study of extended complexity.
Let $C \subset \mathbb{R}^d$ be a compact convex set with nonempty interior. We may assume, $0 \in \operatorname{int}(C)$.
Let $\mathcal{U} \subseteq \mathbb{R}^N$ be an affine subspace and $K \subset \mathbb{R}^N$ be a pointed closed convex cone with nonempty interior. Let $\mathcal{L} : \mathbb{R}^N \to \mathbb{R}^d$ be a linear map.
If

$$C = \mathcal{L}(K \cap \mathcal{U}),$$

then $K \cap \mathcal{U}$ is a *lifted K-representation of C*.

If in addition, $\operatorname{int}(K) \cap \mathcal{U} \neq \emptyset$, then $K \cap \mathcal{U}$ is a *proper lifted K-representation of C*.

Define the *slack function of* C:

$$S_C : \text{ext}(C) \oplus \text{ext}(C^o) \to \mathbb{R}, \quad S_C(x, s) := 1 - \langle s, x \rangle .$$

A slack function S_C is *factorizable* if there exist maps $A : \text{ext}(C) \to K$ and $\tilde{A} : \text{ext}(C^o) \to K^*$ such that

$$S_C(x, s) = \left\langle \tilde{A}(s), A(x) \right\rangle \quad \forall (x, s) \in \text{ext}(C) \oplus \text{ext}(C^o).$$

Theorem

(Gouveia, Parrilo and Thomas [2013])

- *If C has a proper lifted K-representation then S_C is K-factorizable.*
- *If S_C is K-factorizable then C has a lifted K-representation.*

Let us focus on Linear Programming representations, next.

Extension complexity of a polytope P, is the minimum number of linear inequalities required to describe K in a lifted K-representation of P.

Theorem

(Rothvoß [2014]) The matching polytope has exponential extended complexity. In particular, the extension complexity of the perfect matching polytope of a complete n-node graph is $2^{\Omega(n)}$.

A corollary of the above theorem is that the extension complexity of the TSP polytope is $2^{\Omega(n)}$.

There are many other exciting results in this direction, and many open problems.
See, for instance, Fiorini, Massar, Pokutta, Tiwary and de Wolf [2012], Goemans [2015], Kaibel [2011], Fiorini, Kaibel, Pashkovich and Theis [2013], Kaibel and Weltke [2015], Braun and Pokutta [2015], ...

- The convex relaxation methods I discussed can all be phrased so that they are based on polynomial systems of inequalities. This area which is a meeting place for combinatorial optimization, convex optimization and real algebraic geometry continues to be very exciting and vibrant.
- There are many interesting open problems!