Lower bounds on the size of SDP relaxations
James R. Lee, Prasad Raghavendra, David Steurer
Presented by Martin Böhm
Combinatorics PhD Seminar, 2014/15, MFF UK

Some notation
1. \( P_2(n) \) is a set of all subsets of \([n]\) of size at most 2t.
2. \( \mathbb{N}_{\geq}^n \) is a set of all monomials in \( n \) variables of degree at most \( t \).
3. \( ||\cdot||_p \) is the Frobenius norm, defined as \( \sqrt{\text{Tr}(A^t A)} \).
4. For a polynomial \( p(x) \) and a monomial \( \alpha \), \( ||p||_{\alpha} \) is the coefficient of the monomial \( \alpha \).
5. \( \text{vec}(p) \) is a vectorization of a polynomial \( p \) via \( \text{vec}(p) \) is therefore in \( \mathbb{R}^n \).

Semidefinite programming
\( \mathbb{D} \): A symmetric matrix \( M \in \mathbb{R}^{n \times n} \) is positive semidefinite if for all \( v \in \mathbb{R}^n \), \( v^t M v \geq 0 \). We define \( M \) to be positive semidefinite iff it has a square root, i.e., there exists \( U \) such that \( U^t U = M \). We write \( U = \sqrt{M} \).

\( \mathbb{D}(S_k) \): The set (cone) of all positive semidefinite matrices in \( \mathbb{R}^{k \times k} \) will be denoted \( S_k^+ \).

\( \mathbb{D} \): An linear operator \( \bullet \) : \( \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R} \) as defined as \( A \bullet B = \text{Tr}(A^t B) \).

\( \mathbb{D} \): A semidefinite program is a convex optimization program of the form:
\[ \begin{align*}
\text{max} \quad & C \bullet X \quad \text{subject to constraints} \quad A_i \bullet X = b_i \quad \text{and} \quad X \succeq 0.
\end{align*} \]

Extended formulations
\( \mathbb{D}(\text{Linear lift}) \): Consider some polytope \( P \). We say that a polytope \( Q \) is a linear lift of \( P \) if \( P \) is an image of \( Q \) under some linear map. We measure the size of the lift as the number of facets. The polytope \( Q \) is also called extension of \( P \).

\( \mathbb{D}(\text{PSD lift}) \): We say that a polytope \( P \) of dimension \( n \) admits a positive semidefinite lift if \( A \) is a positive semidefinite matrix such that there exists a finite set \( S \) of \( n \times n \) matrices such that \( P = \{ x : A x = b \} \) and \( x \in \mathbb{R}^n \).

Lasserre hierarchy

\( \mathbb{D}(\text{Moment matrix}) \): \( M_{t+1}(y) \) is the moment matrix of \( y \) of degree \( t+1 \), which is the set of all \( t \)-th moments of \( y \).

\( \mathbb{D}(\text{Moments of polynomials}) \): For some function \( f \) of degree \( \ell \), \( M_\ell(f)(y) \) is the \( \ell \)-th moment of \( f \). Then \( L_{\ell}(K) \) is the set of vectors \( y \in \mathbb{R}^{2d} \) that satisfy
\[ \begin{align*}
M_{\ell}(y) & \succeq 0; \\
M_{\ell}(y) & \succeq 0 \quad \forall \ell \in [m]; \\
\rho & = 1.
\end{align*} \]

Intuition: \( M_{t+1}(y) \geq 0 \) ensures consistency (y behaves locally as a distribution) while \( M_{t}(y) \geq 0 \) guarantees that \( y \) satisfies the \( t \)-th linear constraint.

We can solve any problem \( y \) in a pseudo-distribution of a pseudo-density of vertices of the polytope. In our \( \{0,1\}^n \) setting, it is a pseudo-density on \( \{0,1\}^n \).

Note: In our case, we only deal with problems of the form \( \max f(x) \) subject to \( x \in \{0,1\}^n \), and so we can simplify our Lasserre system to:
\[ \begin{align*}
\max f(y_1, y_2, \ldots, y_n) \quad \text{s.t.} \quad M_{t+1}(y) \geq 0; \\
y_0 = 1.
\end{align*} \]

Sum of Squares upper bounds

\( \mathbb{D} \): For a polynomial \( f : \{0,1\}^n \rightarrow \mathbb{R} \), a sum of squares program of degree \( d \) is a program of the form:
\[ \begin{align*}
\min & \quad \rho \\
\text{s.t.} & \quad \forall x \in \{0,1\}^n : \rho - f(x) = \sum_{i=1}^{k} g_i(x)^2; \\
& \quad \forall i \in [k] : \deg(g_i) \leq d/2.
\end{align*} \]

The number \( \rho \) is called the sum of squares upper bound of degree \( d \).

Original idea: verify that \( \forall x : \rho - f(x) \geq 0 \) using a sum of squares (which is always non-negative).

\( \mathbb{D} \): We can compute the sum of squares upper bound using a semidefinite program of size \( n^{O(d)} \).

The semidefinite program is as follows: the variable matrix \( X \) is indexed by a pair of monomials \( \alpha, \beta \) of degree at most \( d/2 \) each. The program itself is:
\[ \begin{align*}
\min & \quad \rho \\
\text{s.t.} & \quad \forall \gamma \in \mathbb{N}_{\geq}^n : \sum_{\alpha, \beta : |\alpha - \beta| = |\gamma|} X_{\alpha, \beta} = [\gamma](\rho - f); \\
& \quad X \geq 0.
\end{align*} \]

(L-eye-open lemma): A sum of squares semidefinite program of degree \( d \) is dual to the \( t/2 \)-th level of the Lasserre hierarchy.

Sum of squares vs. PSD rank

\( \mathbb{D} \): A degree of a function \( \{0,1\}^n \rightarrow \mathbb{R} \) will be the degree of the unique multilinear polynomial agreeing with \( f \) on every point of \( \{0,1\}^\ell \).

\( \mathbb{D}(\text{Sos degree}) \): Consider a non-negative function \( f : \{0,1\}^\ell \rightarrow \mathbb{R} \). We say that \( f \) has a sum-of-squares certificate of degree \( d \) if there exist functions \( g_1, \ldots, g_k : \{0,1\}^\ell \rightarrow \mathbb{R} \) such that \( \deg(g_i) \leq d/2 \) and \( f(x) = \sum_{i=1}^{k} g_i(x)^2 \) for all \( x \in \{0,1\}^\ell \).

We say that \( f \) has sos degree \( d' \) and write \( \text{deg}_{sos}(f) = d' \) if \( d' \) is the minimal degree of a sum-of-squares certificate for \( f \).

\( \mathbb{D}(\text{Main theorem}) \): For every \( m \geq 1 \) and \( f : \{0,1\}^m \rightarrow \mathbb{R} \), there exists a constant \( C > 0 \) such that the following holds: for \( n \gg 2m \), if \( \text{deg}_{sos}(f) = d + 2 \), then
\[ 1 + n^{d+4/2} \geq \text{rk}_{\text{psd}}(M_{d+1}(y)) \geq C \left( \frac{n}{\log n} \right)^{d/4}. \]
The main theorem gives us a good lower bound on \( f \) if \( \deg_{\text{sos}}(f) = d + 2 \). We finalize the proof of Theorem Bound on psd rank by applying the following:

\[
f : \{0, 1\}^m \to \mathbb{R}
\]

\[
f(x) = \left( \frac{m}{2} - \sum_{i=1}^{m} x_i \right)^2 - \frac{1}{4}.
\]

\( (\text{Grigoriev}) \): For every odd integer \( m \geq 1 \), the following function

**Further reading**