Semidefinite Programming, Combinatorial Optimization and Real Algebraic Geometry

assoc. prof. Janez Povh, Ph.D.¹

¹UNM - Faculty of information studies

Edinburgh, 16. September 2014
Outline

Application of SDP to comb. optim.
  Max Cut
  Graph partitioning problem

Application of SDP to RAG
  Commutative RAG
  Non-commutative RAG
Some notation

- $I$ - unit matrix,
- $X \in \mathbb{R}^{m \times n} \Rightarrow x = \text{vec}(X) \in \mathbb{R}^{mn}$
- $\langle x, y \rangle = x^T y = \sum_i x_i y_i$
- $\langle A, B \rangle = \text{tr}(A^T B) = \sum_{i,j} a_{ij} b_{ij}$
- $e$ - vector of all ones.
The Max Cut problem

Given weighted graph $G = (V, E)$ with edge weights $W$:
The Max Cut problem

- Given weighted graph $G = (V, E)$ with edge weights $W$.
- Find $S \subset V$ with maximum cut edges.
MCP - formally

Motivation: network design, cluster analysis.
MCP - formally

- **Motivation**: network design, cluster analysis.

- **Def. MCP**

  Given a weighted graph \( G = (V, E) \) with edge weights \( W \), find a subset \( S \subset V \) such that

  \[
  \text{cut}(S) := \sum_{i \in S, j \in V \setminus S} w_{ij}
  \]

  is maximum.
MCP - int. prog. formulation

Def. MCP

Given weighted graph $G = (V, E)$ with edge weights $W$, solve

$$\max \quad \frac{1}{4} \sum_{i,j} w_{ij}(1 - x_i x_j)$$

$x \in \{-1, 1\}^n$
Def. **MCP**

**Given** weighted graph $G = (V, E)$ with edge weights $W$, solve

$$\max \frac{1}{4} \sum_{i,j} w_{ij} (1 - x_i x_j)$$

$x \in \{-1, 1\}^n$

**Solution** for MCP: $S = \{i : x_i = 1\}$.

**Thm.:** (Karp, 1972) MCP is an NP-complete problem.

**Thm.:** MCP is polynomial if graph is **planar** (Orlova, Dorfman), **weakly** bipartite (Grötchal, Pulleyblank), graphs **without long odd cycles** (Grötchal and Nemhauser), **line** graphs (Arbib), graphs with **bounded tree width** (Bodlaender, Jansen) and some others.
MCP - relaxation

**Def. MCP**

**Given** weighted graph $G = (V, E)$ with edge weights $W$, solve

$$\max \quad \frac{1}{4} \sum_{i,j} w_{ij} (1 - x_i x_j)$$

$$x \in \{-1, 1\}^n$$
MCP - relaxation

Def. MCP

Given weighted graph $G = (V, E)$ with edge weights $W$, solve

$$\max \frac{1}{4} \sum_{i,j} w_{ij} (1 - x_i x_j)$$

$x \in \{-1, 1\}^n$

Def. MCP-SDP

Given weighted graph $G = (V, E)$ with edge weights $W$, solve

$$\max \frac{1}{4} \sum_{i,j} w_{ij} (1 - \langle v_i, v_j \rangle)$$

$v_i \in S_{n-1}$
MCP - relaxation

**Def. MCP**

Given weighted graph $G = (V, E)$ with edge weights $W$, solve

$$\max \frac{1}{4} \sum_{i,j} w_{ij} (1 - x_i x_j)$$

$x \in \{-1, 1\}^n$

**Def. MCP-SDP**

Given weighted graph $G = (V, E)$ with edge weights $W$, solve

$$\max \frac{1}{4} \sum_{i,j} w_{ij} (1 - \langle v_i, v_j \rangle)$$

$v_i \in S_{n-1}$

**Idea:** Introduce $Y = [y_{ij}], \ y_{ij} = \langle v_i, v_j \rangle$. 

Janez Povh
MCP - relaxation

**Def. MCP-SDP**

**Given** weighted graph $G = (V, E)$ with edge weights $W$, solve

$$\max \frac{1}{4} \sum_{i,j} w_{ij} (1 - \langle v_i, v_j \rangle)$$

$$v_i \in S_{n-1}$$
MCP - relaxation

Def. MCP-SDP

Given weighted graph \( G = (V, E) \) with edge weights \( W \), solve

\[
\max \frac{1}{4} \sum_{i,j} w_{ij} (1 - \langle v_i, v_j \rangle)
\]
\[v_i \in S_{n-1}\]

Def. MCP-SDP

Given weighted graph \( G = (V, E) \) with edge weights \( W \), solve

\[
\max \frac{1}{4} \sum_{i,j} w_{ij} (1 - y_{ij})
\]
\[Y \succeq 0, \quad y_{ii} = 1 \ \forall i.\]
MCP - relaxation

Def. MCP-SDP

**Given** weighted graph \( G = (V, E) \) with edge weights \( W \), solve

\[
\max \quad \frac{1}{4} \sum_{i,j} w_{ij} (1 - \langle v_i, v_j \rangle)
\]

\[
v_i \in S_{n-1}
\]

Def. MCP-SDP

**Given** weighted graph \( G = (V, E) \) with edge weights \( W \), solve

\[
\max \quad \frac{1}{4} \sum_{i,j} w_{ij} (1 - y_{ij})
\]

\[
Y \succeq 0, \quad y_{ii} = 1 \ \forall i.
\]

Back: Compute \( v_i \in \mathbb{R}^n \) such that \( \langle v_i, v_j \rangle = y_{ij} \).
How to obtain a good cut?

1. **Input:** \( \{v_i\} \) obtained from SDP relaxation MCP-SDP.
2. Generate random \( r \in S_{n-1} \).
3. **Define** \( S = \{i : \langle r, v_i \rangle \geq 0\} \).

**Thm.:** (Goemans, Williamson, 1995) Let \( r \in S_{n-1} \) be obtained by uniform distribution yielding \( S \). Then

\[
\mathbb{E}(\text{cut}(S)) > 0.87856 \cdot \text{cut}(S^{opt}).
\]
Solving to optimality

- Common approach: Branch and Bound.
- Good lower bound and upper bounds are needed.
- Good upper bounds obtained by SDP relaxations (improved by further constraints - see Rendl, Rinaldi, Wiegele 2010).
- Biq Mac (Wiegele) - web based solver for MCP (up to 300 vertices).
- Solving MCP and MkCP in practise: heuristics.
The graph partitioning problem - a picture

- Partition the nodes of a graph into sets with prescribed sizes such that the number of edges between different sets is minimal.

\[
\text{Slika: } \text{cut}(S_1, S_2, S_3) = 4
\]
Motivations and complexity

- GP problem appears in floor planning, analysis of networks etc.
- GP is connected with vertex separator problem and bandwidth problem.
Motivations and complexity

- GP problem appears in floor planning, analysis of networks etc.
- GP is connected with vertex separator problem and bandwidth problem.
- If $k = 2$, $m_1 = \lceil \frac{n}{2} \rceil$, $m_2 = \lfloor \frac{n}{2} \rfloor$, we get an NP-complete graph bisection problem as a special case.
Application to the graph partitioning problem

- **INPUT** Graph $G = (V, E)$, $k \in \mathbb{N}$, $m = (m_1, \ldots, m_k) \in \mathbb{N}^k$. 

\[ \text{Problem: find a partition of graph nodes into sets } S_1, \ldots, S_k \text{ with } |S_i| = m_i, \text{ which gives minimum cut edges.} \]

\[ \text{OPT GPP} = \min \frac{1}{2} \langle X, A X B \rangle \quad \text{s.t.} \quad X \in \mathbb{R}^{n \times k}, \quad X^T X = \mathbf{M} := \text{Diag}(m) \cdot \text{diag}(XX^T) = u A \ldots \text{adjacency matrix of } G, \quad B = J - I. \]
Application to the graph partitioning problem

- **INPUT** Graph $G = (V, E)$, $k \in \mathbb{N}$, $\mathbf{m} = (m_1, \ldots, m_k) \in \mathbb{N}^k$.
- **Problem**: find a partition of graph nodes into sets $S_1, \ldots, S_k$ with $|S_i| = m_i$, which gives minimum cut edges.
Application to the graph partitioning problem

- **INPUT** Graph $G = (V, E)$, $k \in \mathbb{N}$, $\mathbf{m} = (m_1, \ldots, m_k) \in \mathbb{N}^k$.
- **Problem:** find a partition of graph nodes into sets $S_1, \ldots S_k$ with $|S_i| = m_i$, which gives minimum cut edges.

\[
\text{OPT}_{\text{GPP}} = \min \frac{1}{2} \langle \mathbf{X}, \mathbf{AXB} \rangle \\
\text{s. t. } \mathbf{X} \in \mathbb{R}^{n \times k}, \\
\quad \mathbf{X}^T \mathbf{X} = \mathbf{M} := \text{Diag}(\mathbf{m}) \\
\quad \text{diag}(\mathbf{XX}^T) = u_n
\]

Application to the graph partitioning problem

- **INPUT** Graph \( G = (V, E) \), \( k \in \mathbb{N} \), \( m = (m_1, \ldots, m_k) \in \mathbb{N}^k \).

- **Problem**: find a partition of graph nodes into sets \( S_1, \ldots, S_k \) with \( |S_i| = m_i \), which gives minimum cut edges.

\[
\text{OPT}_{GPP} = \min \frac{1}{2} \langle X, AXB \rangle
\]

\[
\text{s. t. } \quad X \in \mathbb{R}_{+}^{n \times k},
X^T X = M := \text{Diag}(m)
\]

\[
\text{diag}(XX^T) = u_n
\]

\( A \ldots \) adjacency matrix of \( G \), \( B = J - I \).
Adding redundant constraints

Every partition matrix $X$ satisfies $XM^{-1}X^T \preceq I$, $e_n^TXe_k = n$.

$$OPT_{GPP} = \min \frac{1}{2} \langle X, AXB \rangle$$\hspace{1cm}s.t. \hspace{0.5cm}X \in \mathbb{R}^{n \times k}_+, \hspace{1cm}X^TX = M := \text{Diag}(m) \hspace{1cm}\text{diag}(XX^T) = u_n \hspace{1cm}XM^{-1}X^T \preceq I$$

We use $x = \text{vec}(X)$ and $\langle X, AXB \rangle = \langle B \otimes A, xx^T \rangle$. 
We introduce $V = xx^T \in S^{+}_{kn}$. 

$$OPT_{GPP} \geq OPT_{DH} = \min \langle B \otimes A, V \rangle$$

$V \in S^{+}_{kn}$, $W \in S^{+}_n$

$$\sum_{i=1}^{k} \frac{1}{m_i} V^{ii} + W = I, \quad \langle I, V^{ij} \rangle = m_i \delta_{ij}, \forall i, j$$

$$\langle I \otimes E_{ii}, V \rangle = 1, \quad 1 \leq i \leq n$$
We introduce $V = xx^T \in S_{kn}^+$. 

$$OPT_{GPP} \geq OPT_{DH} = \min \langle B \otimes A, V \rangle$$

$V \in S_{kn}^+, W \in S_n^+$

$$\sum_{i=1}^{k} \frac{1}{m_i} V^{ii} + W = I, \quad \langle I, V^{ij} \rangle = m_i \delta_{ij}, \quad \forall i, j$$

$$\langle I \otimes E_{ii}, V \rangle = 1, \quad 1 \leq i \leq n$$

**Thm.** (P. 2009) The semidefinite bound $OPT_{DH}$ is exactly the Donath-Hoffman eigenvalue bound for $OPT_{GPP}$

$$OPT_{DH} = \max \left\{ \frac{1}{2} \sum_{i=1}^{k} m_{k-i+1} \lambda_i (L + D) : D = \text{Diag}(d), \; e^T d = 0 \right\}$$
Improved SDP bounds for GPP

- If we add $e_n^T X e_k = n$ (actually: $\langle X, J_n X J_k \rangle = n^2$), we obtain:

$$OPT_{GPP} \geq OPT_{new1} = \min \langle B \otimes A, V \rangle$$

$$V \in S^{+}_{kn}, \quad W \in S^{+}_n$$

$$\sum_{i=1}^{k} \frac{1}{m_i} V^{ii} + W = I, \quad \langle I, V^{ij} \rangle = m_i \delta_{ij}, \quad \forall i, j$$

$$\langle I \otimes E_{ii}, V \rangle = 1, \quad 1 \leq i \leq n$$

$$\langle J_{kn}, V \rangle = n^2.$$
If we further add $x_{ij}x_{i\ell} = 0 \ \forall i, j, \ell, j \neq \ell$, we obtain:

$$OPT_{GPP} \geq OPT_{\text{new2}} = \min \langle B \otimes A, V \rangle$$

$s \in S_{kn}^+, \ W \in S_n^+$

$$\sum_{i=1}^{k} \frac{1}{m_i} V^{ii} + W = I, \ \langle I, V^{ij} \rangle = m_i \delta_{ij}, \ \forall i, j$$

$$\langle I \otimes E_{ii}, V \rangle = 1, \ \ 1 \leq i \leq n$$

$$\langle J_{kn}, V \rangle = n^2$$

$$\langle E_{j\ell} \otimes E_{ii}, V \rangle = 0 \ \forall i, j, \ell, j \neq \ell.$$
### Numerical results

| name   | $n$ | $|E|$ | $OPT_{DH}$ | $OPT_{new1}$ | $OPT_{new2}$ | PRGP [WoZh 99] |
|--------|-----|------|------------|--------------|--------------|----------------|
| g50.01 | 50  | 111  | 17.922     | 22.762       | 23.570       | 23.549         |
| g50.02 | 50  | 256  | 81.956     | 95.920       | 99.983       | 99.423         |
| g50.03 | 50  | 342  | 124.718    | 148.701      | 152.231      | 151.225        |
| g50.04 | 50  | 478  | 204.303    | 236.697      | 242.578      | 242.063        |
| g50.05 | 50  | 611  | 287.204    | 332.791      | 338.494      | 338.529        |
| g50.06 | 50  | 759  | 378.250    | 440.780      | 443.184      | 442.966        |
| g50.07 | 50  | 897  | 470.157    | 544.238      | 550.335      | 549.934        |
| g50.08 | 50  | 984  | 530.486    | 615.035      | 620.326      | 620.168        |
| g50.09 | 50  | 1098 | 618.867    | 719.456      | 722.990      | 722.270        |

**Tabela:** Semidefinite lower bounds for Graph partitioning problem, where $m = (5, 10, 15, 20)$
Stronger relaxations

- **Further strengthening:** adding new constraints, redundant from original constraints:
  - Triangle constraints for 0-1 programs.
  - Row sum/column sum constraints.
  - **Completely positive constraint:**

**Def.** $X$ is *completely positive* iff $X = \sum_{i=1}^{r} x_i x_i^T$ for some $r \in \mathbb{N}$ and $x_i \geq 0$.

**Def.** The cone of *completely positive* matrices $\mathcal{CP}$.

**Def.** The cone of *copositive* matrices $\mathcal{COP} = \{ A \in S : x^T A x \geq 0 \ \forall x \geq 0 \}$. 
Second main result

**Theorem 2** (P., 2009) Adding completely positive constraint $V \in CP_{kn}$ to $OPT_{new1}$ we obtain the **exact value** for GPP.

$$OPT_{GPP} = \min \langle B \otimes A, V \rangle$$

$V \in CP_{kn}, W \in S_n^+$

$$\sum_i \frac{1}{m_i} V^{ii} + W = I, \quad \langle I, V^{ij} \rangle = m_i \delta_{ij}, \ \forall i, j$$

$$\langle I \otimes E_{ii}, V \rangle = 1, \ 1 \leq i \leq n$$

$$\langle J_{kn}, V \rangle = n^2.$$

**Proof technique:**

- If $V$ is feasible solution of completely positive program, then

$$V = \sum_i \lambda_i x_i x_i^T,$$

where $x_i = \text{vec}(X_i)$ and $X_i$ feasible for GPP.

- We explore the structure of the equations.
The contribution of copositive formulation

- The GPP problem remains NP-hard.
- GPP can be rewritten as completely positive program using other techniques (Burer 2008, P. 2007, P. 2009).
- We can approximate $OPT_{GPP}$ using semidefinite approximations of cone $\mathcal{CP}_{kn}$ (De Klerk, Pasechnik, 2002) or direct heuristics (Bomze, Jarre, Rendl, 2009; Duer, Bundfuss, 2008, 2009).
Real algebraic geometry

**Problem:** Let $f \in \mathbb{R}[x]$. Is $f(x) \geq 0$ for all $x \in \mathbb{R}^n$?
Real algebraic geometry

Problem: Let $f \in \mathbb{R}[x]$. Is $f(x) \geq 0$ for all $x \in \mathbb{R}^n$?

Ex. Motzkin polynomial:

$$M(x, y) = x^2y^4 + x^4y^2 + 1 - 3x^2y^2$$
Real algebraic geometry

**Problem:** Let $f \in \mathbb{R}[x]$. Is $f(x) \geq 0$ for all $x \in \mathbb{R}^n$?

**Ex.** Motzkin polynomial:

$$M(x, y) = x^2 y^4 + x^4 y^2 + 1 - 3x^2 y^2$$

$$(x^2 + y^2 + 1) M(x, y) = (x^2 y - y)^2 + (x y^2 - x)^2 + (x^2 y^2 - 1)^2 +$$

$$+ \frac{1}{4}(xy^3 - x^3 y)^2 + \frac{3}{4}(xy^3 + x^3 y - 2xy)^2$$
**Problem:** Let \( f \in \mathbb{R}[x] \). Is \( f(x) \geq 0 \) for all \( x \in \mathbb{R}^n \)?
**Problem:** Let \( f \in \mathbb{R}[x] \). Is \( f(x) \geq 0 \) for all \( x \in \mathbb{R}^n \)?

**Ex.** \( f_A(x) = \sum_{i,j} a_{ij}x_i^2x_j^2 \geq 0 \) for all \( x \in \mathbb{R}^n \) \( \text{IFF} \) \( A \) is copositive.

**Rem.** The strong membership problems for \( CP \) and \( COP \) are NP-hard (Dickinson and Gijben, 2014).

- The Hilbert **17th problem** (1900): Is every non-negative polynomial with real coefficients a sum of squares of rational functions?
  - Positive answer by Emil Artin in 1927.
  - **Additional question:** Is every non-negative polynomial with real coefficients a sum of squares of real polynomials?
  - The answer: **NO** (known already by Hilbert).
Def. Let $PSD_{n,d}$ be the set (cone) of all non-negative polynomials in $n$ variables of degree $\leq d$. 
Def. Let $PSD_{n,d}$ be the set (cone) of all non-negative polynomials in $n$ variables of degree $\leq d$.

Def. $f \in \mathbb{R}[x]$ is sum of squares (SOS) iff $f = \sum_i p_i^2$ for some $p_i \in \mathbb{R}[x]$. 
Def. Let $PSD_{n,d}$ be the set (cone) of all non-negative polynomials in $n$ variables of degree $\leq d$.

Def. $f \in \mathbb{R}[x]$ is sum of squares (SOS) iff $f = \sum_i p_i^2$ for some $p_i \in \mathbb{R}[x]$.

Def. Let $SOS_{n,2d}$ be the set (cone) of polynomials in $n$ vars with degree $\leq 2d$, which are SOS.
Positivity of polynomials

**Def.** Let $PSD_{n,d}$ be the set (cone) of all non-negative polynomials in $n$ variables of degree $\leq d$.

**Def.** $f \in \mathbb{R}[x]$ is sum of squares (SOS) iff $f = \sum_i p_i^2$ for some $p_i \in \mathbb{R}[x]$.

**Def.** Let $SOS_{n,2d}$ be the set (cone) of polynomials in $n$ vars with degree $\leq 2d$, which are SOS.

**Thm.** $PSD_{n,2d} = SOS_{n,2d}$ iff

- $2d = 2$;
- $n = 1$;
- $n = 2, \ 2d = 4$. 
Positivity of polynomials

**Def.** Let $PSD_{n,d}$ be the set (cone) of all non-negative polynomials in $n$ variables of degree $\leq d$.

**Def.** $f \in \mathbb{R}[x]$ is sum of squares (SOS) iff $f = \sum_i p_i^2$ for some $p_i \in \mathbb{R}[x]$.

**Def.** Let $SOS_{n,2d}$ be the set (cone) of polynomials in $n$ vars with degree $\leq 2d$, which are SOS.

**Thm.** $PSD_{n,2d} = SOS_{n,2d}$ iff

- $2d = 2$;
- $n = 1$;
- $n = 2$, $2d = 4$.

**Ex.** Motzkin polynomial:

$$M(x_1, x_2) = x_1^2 x_2^4 + x_1^4 x_2^2 + 1 - 3x_1^2 x_2^2 \in PSD_{2,6} \setminus SOS_{2,6}.$$
SOS polynomials

**Lem.** If $f \in SOS_{n,2d}$, then $f(x) \in PSD$. 

**Quest.** How to figure out whether $f \in SOS_{n,2d}$?

**Answ.** with SDP.
**Lem.** If \( f \in SOS_{n,2d} \), then \( f(x) \in PSD \).

**Lem.** Let \( f = \sum_i p_i q_i^2 \) with \( p_i \in PSD \). Then \( f(x) \in PSD \).
SOS polynomials

**Lem.** If \( f \in SOS_{n,2d} \), then \( f(x) \in PSD \).

**Lem.** Let \( f = \sum_i p_i q_i^2 \) with \( p_i \in PSD \). Then \( f(x) \in PSD \).

**Lem.** Let \((1 + p)f = \sum_i p_i q_i^2 \) with \( p, p_i \in PSD \). Then \( f(x) \in PSD \).
SOS polynomials

\textbf{Lem.} If $f \in \text{SOS}_{n,2d}$, then $f(x) \in \text{PSD}$.

\textbf{Lem.} Let $f = \sum_i p_i q_i^2$ with $p_i \in \text{PSD}$. Then $f(x) \in \text{PSD}$.

\textbf{Lem.} Let $(1 + p)f = \sum_i p_i q_i^2$ with $p, p_i \in \text{PSD}$. Then $f(x) \in \text{PSD}$.

\textbf{Rem.} SOS is a cone in $\mathbb{R}[x]$. It is convex, pointed, closed, full dimensional cone.
SOS polynomials

**Lem.** If \( f \in \text{SOS}_{n,2d} \), then \( f(x) \in \text{PSD} \).

**Lem.** Let \( f = \sum_i p_i q_i^2 \) with \( p_i \in \text{PSD} \). Then \( f(x) \in \text{PSD} \).

**Lem.** Let \((1 + p)f = \sum_i p_i q_i^2 \) with \( p, p_i \in \text{PSD} \). Then \( f(x) \in \text{PSD} \).

**Rem.** \( \text{SOS} \) is a cone in \( \mathbb{R}[x] \). It is convex, pointed, closed, full dimensional cone.

**Quest.:** How to figure out whether \( f \in \text{SOS}_{n,2d} \)?
**SOS polynomials**

**Lem.** If $f \in \text{SOS}_{n,2d}$, then $f(x) \in PSD$.

**Lem.** Let $f = \sum_i p_i q_i^2$ with $p_i \in PSD$. Then $f(x) \in PSD$.

**Lem.** Let $(1 + p)f = \sum_i p_i q_i^2$ with $p, p_i \in PSD$. Then $f(x) \in PSD$.

**Rem.** \text{SOS} is a cone in $\mathbb{R}[x]$. It is convex, pointed, closed, full dimensional cone.

**Quest.:** How to figure out whether $f \in \text{SOS}_{n,2d}$? **Answ.:** with SDP.
Let \( p(x_1, x_2) = 2x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4 \).\n
\( p \) is SOS since

\[
p(x_1, x_2) = \frac{1}{2}((2x_1^2 - 3x_2^2 + x_1x_2)^2 + (x_2^2 + 3x_1x_2)^2).
\]
Let \( p(x_1, x_2) = 2x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4. \)

\( p \) is SOS since

\[
p(x_1, x_2) = \frac{1}{2}((2x_1^2 - 3x_2^2 + x_1x_2)^2 + (x_2^2 + 3x_1x_2)^2).
\]

We can obtain sos decomp. by **Gram matrix method**: 

**Thm.** Let \( f \in \mathbb{R}[x] \) with degree \( 2d \). \( f \) is SOS IFF there exists \( Q \succeq 0 \) such that

\[
f(x) = V_d^T Q V_d,
\]

where \( V_d \) is the vector of all monomials of degree \( \leq d \).
Let \( p(x_1, x_2) = 2x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4 \).

Corresponding SDP:

\[
\begin{align*}
\text{min} & \quad 0 \\
p. p. & \quad Q = \begin{bmatrix} 2 & a & 1 \\ a & 5 & 0 \\ 1 & 0 & b \end{bmatrix} \succeq 0, \\
2a + b & = -1.
\end{align*}
\]

where \( V_2 = (x_1^2, x_2^2, x_1x_2)^T \).
If $b = 5$ we get

$$Q = \begin{bmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{bmatrix}^T \cdot \begin{bmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{bmatrix} \succeq 0.$$ 

Therefore:

$$p(x_1, x_2) = \frac{1}{2}((2x_1^2 - 3x_2^2 + x_1 x_2)^2 + (x_2^2 + 3x_1 x_2)^2).$$
Complexity of SOS SDP

- There are $\binom{n+d-1}{d}$ monomials of degree $d$. This is the order of $Q$. 
Complexity of SOS SDP

- There are \( \binom{n+d-1}{d} \) monomials of degree \( d \). This is the order of \( Q \).

**Thm.** It is enough to consider only the monomials from one half of the **Newton polytope**:

\[
\begin{align*}
\text{Newton polytope for } f &= 1 + 2y^2 - 4x^5 \\
&= y = x, \quad y = 2, \quad x = 5
\end{align*}
\]
Thm. (Pólya, 1929, Hardy, Littlewood, Pólya, 1988, Powers, Reznick, 2001) Let $f \in \mathbb{R}[x]$ be a homogeneous polynomial on $\mathbb{R}^n$ such that $f(x) > 0$ for all $x \in \mathbb{R}_+^n \setminus \{0\}$. Then for some $r \in \mathbb{N}$, we have that all the coefficients of $(e^T x)^r f(x)$ are non-negative (positive).

Application: In copositive programming - LP ali SDP certificates for copositivity: if $(\sum_i x_i^2)^r \sum_{i,j} A_{i,j} x_i^2 x_j^2$ has non-negative coefficients then $A$ is copositive (LP problem).
**Thm.** (Putinar, 1993) Let $m \in \mathbb{N}$ and $f, g_1 = 1, g_2, \ldots, g_m \in \mathbb{R}[x]$. If $f(x) > 0$ for all $x \in K := \{x \in \mathbb{R}^n : g_i(x) \geq 0, \text{ for } i = 1, \ldots, m\} \setminus \{0\}$,

then there exists $s_1, \ldots, s_m \in \text{SOS}$ such that

$$ f(x) = \sum_{i=1}^{m} s_i g_i(x), \text{ provided e.g. } K \text{ compact.} $$
Application in optimization

- \( f_{\inf} = \inf \{ f(x) : x \in K \} = \sup \{ \varepsilon : f(x) - \varepsilon \geq 0 \text{ for all } x \in K \} \)

- \( f_{\inf} \geq f_{\operatorname{sos}} = \sup \{ \varepsilon : f(x) - \varepsilon = \sum_i s_i g_i, \ s_i \in \operatorname{SOS} \} \).
Application in optimization

- $f_{\inf} = \inf \{ f(x) : x \in K \} = \sup \{ \varepsilon : f(x) - \varepsilon \geq 0 \text{ for all } x \in K \}$
- $f_{\inf} \geq f_{sos} = \sup \{ \varepsilon : f(x) - \varepsilon = \sum_i s_i g_i, \ s_i \in SOS \}.$
- $f_{sos}^{(t)} = \sup \{ \varepsilon : f(x) - \varepsilon = \sum_i s_i g_i, \ s_i \in SOS, \ \deg(s_i g_i) \leq 2t \}$
Application in optimization

\[ f_{\inf} = \inf \{ f(x) : x \in K \} = \sup \{ \varepsilon : f(x) - \varepsilon \geq 0 \text{ for all } x \in K \} \]

\[ f_{\inf} \geq f_{sos} = \sup \{ \varepsilon : f(x) - \varepsilon = \sum_i s_i g_i, \ s_i \in \text{SOS} \} \]

\[ f_{sos}^{(t)} = \sup \{ \varepsilon : f(x) - \varepsilon = \sum_i s_i g_i, \ s_i \in \text{SOS}, \ \deg(s_i g_i) \leq 2t \} \]

**Thm.**\( \lim_{t \to \infty} f_{sos}^{(t)} = f_{sos} = f_{\inf} \) provided e.g. \( K \) compact.
Read more in...


How do we get NC polynomials?

Formal **construction:**

1. Start with NC letters $\mathbf{x} := (x_1, \ldots, x_n)$ and operation "multiplication".
2. Consider $\langle \mathbf{x} \rangle$ - monoid freely generated by $\mathbf{x}$ (empty word is 1).
3. Free algebra $\mathbb{R}\langle \mathbf{x} \rangle$: noncommutative (nc) polynomials.
4. Add **involution** $\ast$ which:
   - fixes $\mathbb{R} \cup \{\mathbf{x}\}$ pointwise
   - and reverses words, e.g. $(x_1x_2^2x_3 - 2x_3^3)^* = x_3x_2^2x_1 - 2x_3^3$.
5. $\text{Sym} \mathbb{R}\langle \mathbf{x} \rangle$ - the set of all symmetric elements:

$$\text{Sym} \mathbb{R}\langle \mathbf{x} \rangle = \{ f \in \mathbb{R}\langle \mathbf{x} \rangle \mid f = f^* \}.$$
Why NC polys are relevant?

1. Lots of applications in **control theory, systems engineering** and **optimization** (see Helton, McCullough, Oliveira, Putinar, 2008),
2. Applications to **quantum physics** (Pironio, Navascués, Acín, 2010)
3. Applications in **quantum information science** (Pál and T. Vértesi, 2009),
4. **Quantum chemistry** (e.g. to compute the ground-state electronic energy of atoms or molecules) - see cf. Mazziotti, 2004.
5. Certificates of positivity via sums of squares are used to get general bounds on **quantum correlations** (cf. Glauber, 1963).
6. The Bessis-Moussa-Villani conjecture (BMV) from **quantum statistical mechanics** is tackled by NC polynomials (Klep, Schweighofer, 2009; Cafuta, Klep, Povh, 2011)
The BMV conjecture

1. **Bessis - Moussa - Villani (BMV) conjecture (1975):**
   - For symmetric matrices $A, B$ with $B$ positive semidefinite, the function
     $$\Phi^{A,B} : \mathbb{R} \to \mathbb{R}, t \mapsto \text{tr}(e^{A-tB})$$
     is the Laplace transform of a positive measure $\mu^{A,B}$ on $[0, \infty)$:
     $$\text{tr}(e^{A-tB}) = \int_0^\infty e^{-tx} d\mu^{A,B}(x).$$
The BMV conjecture

1. **Bessis - Moussa - Villani (BMV) conjecture (1975):**
   - For symmetric matrices $A, B$ with $B$ positive semidefinite, the function
     $$\Phi^{A,B} : \mathbb{R} \rightarrow \mathbb{R}, \ t \mapsto \text{tr}(e^{A-tB})$$
     is the Laplace transform of a positive measure $\mu^{A,B}$ on $[0, \infty)$:
     $$\text{tr}(e^{A-tB}) = \int_0^{\infty} e^{-tx} d\mu^{A,B}(x).$$
   - **BMV equivalently** (Lieb–Seiringer, 2004): The polynomial
     $$\text{tr}((A + tB)^m) \in \mathbb{R}[t] = \sum_{k=0}^{m} t^k \text{tr}(S_{m,k}(A, B))$$
     has only nonnegative coefficients whenever $A, B$ are PSD of order $s$, for all $m$. 

2. The conjecture was recently proved by H.R. Stahl: Proof of the BMV conjecture, 2011.
The BMV conjecture

   - For symmetric matrices $A, B$ with $B$ positive semidefinite, the function
     \[ \Phi^{A,B} : \mathbb{R} \rightarrow \mathbb{R}, \ t \mapsto \text{tr}(e^{A-tB}) \]
     is the Laplace transform of a positive measure $\mu^{A,B}$ on $[0, \infty)$:
     \[ \text{tr}(e^{A-tB}) = \int_0^\infty e^{-tx} d\mu^{A,B}(x). \]
   - **BMV equivalently** (Lieb–Seiringer, 2004): The polynomial
     \[ \text{tr}((A + tB)^m) \in \mathbb{R}[t] = \sum_{k=0}^{m} t^k \text{tr}(S_{m,k}(A, B)) \]
     has only nonnegative coefficients whenever $A, B$ are PSD of order $s$, for all $m$.

2. The conjecture was recently proved by H.R. Stahl: Proof of the BMV conjecture, 2011.
Positivity of NC polynomials

1. \( f(x) \geq 0? \)
Positivity of NC polynomials

1. \( f(x) \geq 0? \)

2. Is \( f(x) \geq 0 \) component-wise?
**Positivity of NC polynomials**

1. $f(x) \geq 0$?

2. Is $f(x) \geq 0$ component-wise?

3. Is $f(x) \succeq 0$?
Positivity of NC polynomials

1. \( f(x) \geq 0? \)
2. Is \( f(x) \geq 0 \) component-wise?
3. Is \( f(x) \succeq 0? \)
4. Is \( \text{tr } f(x) \geq 0? \)
Positivity of NC polynomials

1. \( f(x) \geq 0? \)

2. Is \( f(x) \geq 0 \) component-wise?

3. Is \( f(x) \succeq 0? \)

4. Is \( \text{tr} f(x) \geq 0? \)

5. We stick to: “Is \( f(x) \succeq 0? \)”
Main questions

**Given:** real polynomial $f$ in non-commuting variables $\mathbf{x} = (x_1, \ldots, x_n)$.

**Q 1:** Is $f(A) \succeq 0$ for all $n$-tuples of symmetric matrices $A = (A_1, \ldots, A_n)$ of the same size?
Main questions

**Given:** real polynomial $f$ in non-commuting variables $\mathbf{x} = (x_1, \ldots, x_n)$.

**Q 1:** Is $f(A) \succeq 0$ for all $n$-tuples of symmetric matrices $A = (A_1, \ldots, A_n)$ of the same size?

**Q 2:** Find the smallest eigenvalue of $f$, i.e. compute

$$\lambda_{\min}(f) = \inf \langle f(A)v, v \rangle$$

$A$ an $n$-tuple of symmetric matrices,

$v$ a unit vector.
Sums of hermitian squares (SOHS)

\[ \text{SOHS} = \{ \sum_i g_i^* g_i : g_i \in \mathbb{R}\langle x \rangle \} \subset Sym \mathbb{R}\langle x \rangle. \]

\[ \text{SOHS}_d = \{ \sum_i g_i^* g_i : g_i \in \mathbb{R}\langle x \rangle, \ deg(g_i^* g_i) \leq d \} \subset Sym \mathbb{R}\langle x \rangle. \]
Sums of hermitian squares (SOHS)

- \( \text{SOHS} = \{ \sum_i g_i^* g_i : g_i \in \mathbb{R}\langle x \rangle \} \subset \text{Sym} \mathbb{R}\langle x \rangle. \)
- \( \text{SOHS}_d = \{ \sum_i g_i^* g_i : g_i \in \mathbb{R}\langle x \rangle, \text{deg}(g_i^* g_i) \leq d \} \subset \text{Sym} \mathbb{R}\langle x \rangle \)
Examples

1. Let $f = x_1 x_2 + x_2 x_1 + 4 - x_1^2 - x_2^2$. It is not non-negative (take $x_1 = 0$, $x_2 = 3i$).
Examples

1. Let \( f = x_1x_2 + x_2x_1 + 4 - x_1^2 - x_2^2 \). It is not non-negative (take \( x_1 = 0, x_2 = 3I \)).
2. Let \( f = (x_1 \ast x_2 + x_2)(x_1 \ast x_2 + x_2) + 2 \). It is non-negative.
Examples

1. Let $f = x_1 x_2 + x_2 x_1 + 4 - x_1^2 - x_2^2$. It is not non-negative (take $x_1 = 0$, $x_2 = 3i$).

2. Let $f = (x_1 * x_2 + x_2) * (x_1 * x_2 + x_2) + 2$. It is non-negative.
Question 1

Recall: Question 1.

Is \( f(A) \succeq 0 \) for all \( n \)-tuples of symmetric matrices \( A = (A_1, \ldots, A_n) \) of the same size?
Recall: Question 1.

\[
f(A) \succeq 0
\]
for all n-tuples of symmetric matrices \( A = (A_1, \ldots, A_n) \) of the same size?

Thm.: (Helton, Annals of math., 2002)

\[
f \in \mathbb{R}\langle x \rangle \text{ is SOHS } \iff f(x) \succeq 0 \text{ whenever we replace } x_i \text{ by symmetric matrices } A_i \text{ of dimension } k \times k, \forall k \geq 1.
\]
Problem 1: Given: \( f \in \mathbb{R}\langle x \rangle \), is \( f \in \text{SOHS} \)?

If YES: Provide (SDP) certificate!
If NO: Provide (SDP) certificate!
**Prop.:** Suppose \( f \in \mathbb{R}\langle x \rangle \) is of degree \( \leq 2d \). Then \( f \in SOHS \) if and only if there exists a positive semidefinite (PSD) matrix \( G \) satisfying

\[
f = W_d^* GW_d,
\]

where

\[
W_d = \{ p \in \langle x \rangle : \deg(p) \leq d \}.
\]
SDP certificate for SOHS

**Prop.:** Suppose $f \in \mathbb{R}\langle x \rangle$ is of degree $\leq 2d$. Then $f \in SOHS$ if and only if there exists a positive semidefinite (PSD) matrix $G$ satisfying

$$f = W_d^* GW_d,$$

where

$$W_d = \{ p \in \langle x \rangle : \deg(p) \leq d \}.$$ 

**Rem.:** Given such a PSD matrix $G$ with rank $r$, the SOHS decomposition is

$$f = \sum_{i=1}^{r} g_i^* g_i,$$

where $g_i = H(i,:) W_d$, $G = H^T H$. 

Prop.: Given $f = \sum_{w \in \langle x \rangle} a_w w$ of degree $2d$, then $f \in SOHS$ iff exists $G \succeq 0$ such that:

$$\sum_{p,q \in W_d, p^* q = w} G_{p,q} = a_w, \quad \forall w \in W_{2d}$$
SDP certificate for SOHS

Prop.: Given \( f = \sum_{w \in \langle x \rangle} a_w w \) of degree 2\( d \), then \( f \in SOHS \) iff exists \( G \succeq 0 \) such that:

\[
\sum_{\substack{p,q \in W_d \\text{such that} \quad p^* q = w}} G_{p,q} = a_w, \quad \forall w \in W_{2d}
\]

Note: "Is \( f \) in \( SOHS? \)" is SDP feasibility problem.
**Prop.**: Given $f = \sum_{w \in \langle x \rangle} a_w w$ of degree $2d$, then $f \in SOHS$ iff exists $G \succeq 0$ such that:

$$\sum_{p, q \in W_d, p^* q = w} G_{p, q} = a_w, \quad \forall w \in W_{2d}$$

**Note**: "Is $f$ in SOHS?" is SDP feasibility problem.

**SDP**: 

$$\inf \langle I, G \rangle \quad (\text{SDP}_{\text{SOHS}}) \quad \text{s.t.} \quad \langle A_w, G \rangle = a_w + a_{w^*} \quad \forall w \in W_{2d}$$

$$G \succeq 0.$$ 

where

$$(A_w)_{u, v} = \begin{cases} 2; & \text{if } u^* v \in \{w, w^*\}, \ w^* = w, \\ 1; & \text{if } u^* v \in \{w, w^*\}, \ w^* \neq w, \\ 0; & \text{otherwise.} \end{cases}$$
Optimization of NC polys.

Problem: Given $f \in \text{Sym} \mathbb{R}\langle x \rangle$

find smallest eigenvalue of $f$:

$$\lambda_{\text{min}}(f) = \inf \langle f(A)v, v \rangle$$

$A$ an $n$-tuple of symmetric matrices,
$v$ a unit vector.
Problem: Given $f \in \text{Sym} \mathbb{R}\langle x \rangle$

find smallest eigenvalue of $f$:

$$\lambda_{\text{min}}(f) = \inf \langle f(A)v, v \rangle$$

$A$ an $n$-tuple of symmetric matrices,
$v$ a unit vector.

By Helton-McCullough SOHS theorem:

$$\lambda_{\text{min}}(f) = \sup \lambda$$

s. t. $f - \lambda \in \text{SOHS}_{d}$. (SDP_{eig-min})
Prop: The dual to \((SDP_{\text{eig-min}})\) is

\[
L_{\text{sohs}} = \inf \langle G_f, H \rangle
\]

s.t.

\[
H \in S^+, \quad H_{1,1} = 1
\]

\[
H_{p,q} = H_{r,s} \quad \text{for all } p, q, r, s, \quad p^* q = r^* s.
\]

\[(DSDP_{\text{eig-min}})_d\]
Prop: The dual to \((\text{SDP}_{\text{eig-min}})\) is

\[
L_{\text{sohs}} = \inf \langle G_f, H \rangle \\
\text{s.t.} \quad H \in S^+ \\
H_{1,1} = 1 \\
H_{p,q} = H_{r,s} \quad \text{for all } p, q, r, s, \quad p^* q = r^* s.
\]

\(\text{(DSDP}_{\text{eig-min}})_{d}\)

Prop.: 

\[
L_{\text{sohs}} = \lambda_{\min}(f).
\]
Thm.: (Cafuta, Klep, P., 2010) Let $f \in \mathbb{R}\langle x\rangle_{\leq 2d}$.

(a) Then $\lambda_{\text{min}}(f)$ is attained if and only if there is a feasible point $L$ for $(\text{DSDP}_{\text{eig-min}})_{d+1}$ satisfying $L(f) = L_{\text{sohs}} = \lambda_{\text{min}}(f)$.

(b) If $L_{\text{sohs}}$ is attained, we can find symmetric $s \times s$ matrices $A_1, \ldots, A_n$ and unit vector $v$ such that

$$
\lambda_{\text{min}}(f) = \langle f(A)v, v \rangle.
$$

Proof:

- Gelfand-Naimark-Segal (GNS) construction
- use flat extensions
Example

\[
\begin{align*}
\text{>> } f &= (1 - 3*x*y + y*x)'*(1 - 3*x*y + y*x) + (-1 + x^2)^2 + (-y+y^2)^2; \\
\text{>> } \text{NCmin}(f) \\
\lambda_{\text{min}}(f) &= 0. \\
\text{>> } [X,fX,eig\_val,eig\_vec]=\text{NCopt}(f)
\end{align*}
\]
Example - cnt.

\[ A = \begin{bmatrix}
0.9644 & -0.0379 & -0.1276 & 0.0879 \\
-0.0379 & -0.9828 & 0.1588 & 0.0235 \\
-0.1276 & 0.1588 & 0.4923 & 0.2253 \\
0.0879 & 0.0235 & 0.2253 & -0.9790 \\
\end{bmatrix} \]

\[ B = \begin{bmatrix}
0.8367 & 0.1790 & 0.3326 & 0.0832 \\
0.1790 & 0.0215 & 0.1388 & 0.5320 \\
0.3326 & 0.1388 & -0.0227 & -0.6871 \\
0.0832 & 0.5320 & -0.6871 & -0.1778 \\
\end{bmatrix} \]

\[ f(A, B) = \begin{bmatrix}
0.7978 & 1.2130 & 0.8094 & 0.6920 \\
1.2130 & 3.3989 & -2.6498 & -0.0064 \\
0.8094 & -2.6498 & 10.5185 & 3.0781 \\
0.6920 & -0.0064 & 3.0781 & 7.9733 \\
\end{bmatrix} \]

\[ \lambda_{\min}(f) = \lambda_{\min}(f(A, B)) = 0.0000, \]

\[ v = \begin{bmatrix}
-0.8741 \\
0.4515 \\
0.1789 \\
0.0072 \\
\end{bmatrix}^t. \]

Note: Commutative min. of \( f \) is 0.0625 (for \( x = 1, \ y = 1/2 \)).
1. \( f \in SOHS_d \) has low complexity (size of matrix is \( O(kd/2) \)) (Newton chip method from P., Klep, 2010)


