Brainstorming day on polynomial optimization:

A hierarchy of spectral relaxations for polynomial optimization

Ngoc Hoang Anh Mai
PhD student
MAC Group

joint work with my supervisors:
Jean-Bernard Lasserre and Victor Magron

September 10, 2020
Scalability of moment relaxations for POP

Consider a general POP:

\[ [r] := \{1, \ldots, r\} \]

\[ f^* := \inf\{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \} . \] (1)
Scalability of moment relaxations for POP

Consider a general POP:

\[ [r] := \{1, \ldots, r\} \]

\[
  f^* := \inf \{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \}.
\] (1)

If a POP has sparsity ⇒ easy to scale thanks to:

▶ correlative sparsity: [Waki et al., 2006, Lasserre, 2006]

\[ f = f_1 + \cdots + f_p \]

and

\[ (g_i)_{i \in G_j}, \ (h_i)_{i \in H_j} \]

share the same few variables, e.g.,

\[ f = x_1 x_2 x_3 + x_2 x_3 x_4 + x_3 x_1 x_5. \]

▶ term sparsity: [Wang et al., 2019]

\[ f, g_i, h_j \]

have few nontrivial terms, e.g.,

\[ f = 1 + x_1 x_2 + x_2 x_3 + x_3 x_1. \]
Scalability of moment relaxations for POP

Consider a general POP: \[ [r] := \{1, \ldots, r\} \]

\[ f^* := \inf \{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \} . \] (1)

If a POP has sparsity \( \Rightarrow \) easy to scale thanks to:

- **correlative sparsity:** [Waki et al., 2006, Lasserre, 2006]

  \[ f = f_1 + \cdots + f_p \] and \( f_j, (g_i)_{i \in g_j}, (h_i)_{i \in h_j} \) share the same few variables,
Scalability of moment relaxations for POP

Consider a general POP:

\[ [r] := \{1, \ldots, r\} \]

\[ f^* := \inf \{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \} . \quad (1) \]

If a POP has sparsity \( \Rightarrow \) easy to scale thanks to:

- correlative sparsity: [Waki et al., 2006, Lasserre, 2006]
  \[ f = f_1 + \cdots + f_p \] and \( f_j, (g_i)_{i \in g_j}, (h_i)_{i \in h_j} \) share the same few variables, e.g., \( f = \underbrace{x_1x_2 + x_2x_3 + x_3x_1}_{f_1} + \underbrace{x_3x_4 + x_4x_5}_{f_2 + f_3} \).

- term sparsity: [Wang et al., 2019]
  \( f, g_i, h_j \) have few nontrivial terms, e.g., \( f = 1 + x_1x_2 + x_2x_3 + x_3x_1 \).
Scalability of moment relaxations for POP

Consider a general POP:

\[
[r] := \{1, \ldots, r\}
\]

\[
f^* := \inf \{ f(x) : g_i(x) \geq 0, i \in [m], h_j(x) = 0, j \in [l] \}.
\]

(1)

If a POP has sparsity \( \Rightarrow \) easy to scale thanks to:

- correlative sparsity: [Waki et al., 2006, Lasserre, 2006]
  \( f = f_1 + \cdots + f_p \) and \( f_j, (g_i)_{i \in g_j}, (h_i)_{i \in H_j} \) share the same few variables, e.g.,
  \[
  f = x_1 x_2 + x_2 x_3 + x_3 x_1 + x_3 x_4 + x_4 x_5.
  \]

- term sparsity: [Wang et al., 2019]
  \( f, g_i, h_j \) have few nontrivial terms, e.g.,
  \[
  f = 1 + x_1 x_2 + x_2 x_3 + x_3 x_1.
  \]
Scalability of moment relaxations for POP

Consider a general POP:

\[ r := \{1, \ldots, r\} \]

\[ f^* := \inf \{ f(x) : g_i(x) \geq 0, i \in [m], h_j(x) = 0, j \in [l] \} . \quad (1) \]

If a POP has sparsity \( \Rightarrow \) easy to scale thanks to:

- correlative sparsity: [Waki et al., 2006, Lasserre, 2006]

\[ f = f_1 + \cdots + f_p \text{ and } f_j, (g_i)_{i \in g_j}, (h_i)_{i \in H_j} \text{ share the same few variables}, \text{ e.g., } f = x_1x_2 + x_2x_3 + x_3x_1 + x_3x_4 + x_4x_5. \]

- term sparsity: [Wang et al., 2019] \( f, g_i, h_j \) have few nontrivial terms \( c_\alpha x^\alpha \),
Scalability of moment relaxations for POP

Consider a general POP:

\[ \begin{align*}
[r] & := \{1, \ldots, r\} \\
\end{align*} \]

\[ f^* := \inf \{ f(x) : g_i(x) \geq 0, \, i \in [m], \, h_j(x) = 0, \, j \in [l] \} . \quad (1) \]

If a POP has sparsity \( \Rightarrow \) easy to scale thanks to:

- correlative sparsity: [Waki et al., 2006, Lasserre, 2006]

  \[ f = f_1 + \cdots + f_p \text{ and } f_j, (g_i)_{i \in g_j}, (h_i)_{i \in h_j} \text{ share the same few variables, e.g., } f = x_1 x_2 + x_2 x_3 + x_3 x_1 + x_3 x_4 + x_4 x_5. \]

- term sparsity: [Wang et al., 2019] \( f, g_i, h_j \) have few nontrivial terms \( c_\alpha x^\alpha \), e.g., \( f = 1 + x_1 x_2 + x_2 x_3 + x_3. \)
Scalability of moment relaxations for POP

Consider a general POP:
\[
[r] := \{1, \ldots, r\}
\]
\[
f^* := \inf \{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \}.
\]

If a POP has sparsity ⇒ easy to scale thanks to:

▶ correlative sparsity: [Waki et al., 2006, Lasserre, 2006]
  
  \[ f = f_1 + \cdots + f_p \text{ and } f_j, (g_i)_{i \in G_j}, (h_i)_{i \in H_j} \text{ share the same few variables, e.g., } f = x_1x_2 + x_2x_3 + x_3x_1 + x_3x_4 + x_4x_5. \]

▶ term sparsity: [Wang et al., 2019] \( f, g_i, h_j \) have few nontrivial terms \( c_\alpha x^\alpha \), e.g., \( f = 1 + x_1x_2 + x_2x_3 + x_3. \)
Scalability of moment relaxations for POP

Consider a general POP: 

\[ [r] := \{1, \ldots, r\} \]

\[ f^* := \inf \{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \} \] . (2)

What if no sparsity (dense POP)?
Scalability of moment relaxations for POP

Consider a general POP:

\[ \{r\} := \{1, \ldots, r\} \]

\[ f^* := \inf \{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \} . \quad (2) \]

What if no sparsity (dense POP)?
For instance, \( f = 1 + \sum_{i \in [6]} x_i + \sum_{(i,j) \in [6]^2} x_i x_j \)
Scalability of moment relaxations for POP

Consider a general POP:

$$[r] := \{1, \ldots, r\}$$

$$f^* := \inf \{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \} . \ (2)$$

What if no sparsity (dense POP)?

For instance, $f = 1 + \sum_{i \in [6]} x_i + \sum_{(i,j) \in [6]^2} x_i x_j$
Scalability of moment relaxations for POP

Consider a general POP: \[ [r] := \{1, \ldots, r\} \]

\[ f^* := \inf \{ f(x) : g_i(x) \geq 0, \ i \in [m], \ h_j(x) = 0, \ j \in [l] \} . \quad (2) \]

What if no sparsity (dense POP)?
For instance, \( f = 1 + \sum_{i\in[6]} x_i + \sum_{(i,j)\in[6]^2} x_i x_j \)

\[ \Rightarrow \text{use other information: constant trace property.} \]
Let $S$ be the set of real symmetric matrices of size $s$ with the inner product $\langle A, B \rangle = \text{trace}(AB)$. If constant trace property (CTP), i.e., $-\tau = \sup_{X \in S} \{ \text{trace}(CX) : A \preceq X = b, X \succeq 0, \text{trace}(X) = a \}$, e.g., SDP MAXCUT $\Rightarrow$ can be solved very efficiently thanks to:

1. Spectral Bundle Method (SBM): [Helmberg and Rendl, 2000] The dual of (3) is equivalent to $-\rho = \inf z a \lambda_{\max}(C - A^\top z) + b^\top z$, (4) where $\lambda_{\max}(A)$ is the largest eigenvalue of matrix $A$.

2. SketchyCGAL $\Rightarrow$ Sketchy + CG + AL: [Yurtsever et al., 2019] ▶ the augmented Lagrangian (AL) ▶ conditional gradient method (CG) ▶ randomized sketch (Sketchy)
SDP with constant trace property

Let $S$ be the set of real symmetric matrices of size $s$ with the inner product $\langle A, B \rangle = \text{trace}(AB)$.
SDP with constant trace property

Let $S$ be the set of real symmetric matrices of size $s$ with the inner product $\langle A, B \rangle = \text{trace}(AB)$.

If constant trace property (CTP), i.e.,

$$-\tau = \sup_{X \in S}\{\text{trace}(CX) : AX = b, X \succeq 0, \text{trace}(X) = a\},$$

(3)
SDP with constant trace property

Let $S$ be the set of real symmetric matrices of size $s$ with the inner product $\langle A, B \rangle = \text{trace}(AB)$.

If constant trace property (CTP), i.e.,

$$-\tau = \sup_{X \in S}\{\text{trace}(CX) : AX = b, X \succeq 0, \text{trace}(X) = a\},$$

(3)
e.g., SDP MAXCUT $\Rightarrow$ can be solved very efficiently thanks to:

1. **Spectral Bundle Method (SBM):**
   [Helmberg and Rendl, 2000] The dual of (3) is equivalent to

   $$-\rho = \inf_z a\lambda_{\text{max}}(C - A^Tz) + b^Tz,$$

   (4)
   where $\lambda_{\text{max}}(A)$ is the largest eigenvalue of matrix $A$. 

2. SketchyCGAL $\Rightarrow$ Sketchy + CG + AL: [Yurtsever et al., 2019]
SDP with constant trace property

Let $S$ be the set of real symmetric matrices of size $s$ with the inner product $\langle A, B \rangle = \text{trace}(AB)$.

If constant trace property (CTP), i.e.,

$$-\tau = \sup_{X \in S} \{ \text{trace}(CX) : AX = b, X \succeq 0, \text{trace}(X) = a \},$$

(3)

e.g., SDP MAXCUT $\Rightarrow$ can be solved very efficiently thanks to:

1. **Spectral Bundle Method (SBM):**
   [Helmberg and Rendl, 2000] The dual of (3) is equivalent to

   $$-\rho = \inf_z a\lambda_{\max}(C - A^Tz) + b^Tz,$$

   (4)

   where $\lambda_{\max}(A)$ is the largest eigenvalue of matrix $A$.

2. **SketchyCGAL** = Sketchy+CG+AL: [Yurtsever et al., 2019]
   - the augmented Lagrangian (AL)
   - conditional gradient method (CG)
   - randomized sketch (Sketchy)
Contribution

Exploit constant trace property (CTP) for moment relaxations of POP.

Convert moment relaxations with CTP to spectral relaxations.

Build an algorithm called SpectralPOP to solve spectral relaxations and extract the optimal solutions of POP.

Display several numerical results of SpectralPOP for a sample of random dense POPs.
Contribution

▶ Exploit constant trace property (CTP) for moment relaxations of POP.
▶ Convert moment relaxations with CTP to spectral relaxations.
▶ Build an algorithm called SpectralPOP to solve spectral relaxations and extract the optimal solutions of POP.
▶ Display several numerical results of SpectralPOP for a sample of random dense POPs.
Contribution

▶ Exploit constant trace property (CTP) for moment relaxations of POP.
▶ Convert moment relaxations with CTP to spectral relaxations.
Contribution

- Exploit constant trace property (CTP) for moment relaxations of POP.
- Convert moment relaxations with CTP to spectral relaxations.
- Build an algorithm called \textit{SpectralPOP} to solve spectral relaxations and extract the optimal solutions of POP.
Contribution

▶ Exploit constant trace property (CTP) for moment relaxations of POP.
▶ Convert moment relaxations with CTP to spectral relaxations.
▶ Build an algorithm called **SpectralPOP** to solve spectral relaxations and extract the optimal solutions of POP.
▶ Display several numerical results of **SpectralPOP** for a sample of random dense POPs.
Consider an equality constrained POP on the unit sphere:

\[ f^{\star} := \inf \left\{ f(x) : h_j(x) = 0, j \in [l] \right\} \]

with

\[ h_1 = 1 - \|x\|_2^2. \]

For every \( k \in \mathbb{N} \), consider SDP [Lasserre, 2001]:

\[ \tau_k = \inf y \left\{ L y(f) : M_k(y) \succeq 0, M_k - w_j(h_j y) = 0, y_0 = 1 \right\}. \]

Strong duality holds due to [Josz and Henrion, 2016].

Convergence rate: [Schweighofer, 2004] \( \tau_k \uparrow f^{\star} \) with \( O\left( k^{-1/c} \right) \).

Moreover \( \tau_k = f^{\star} \) under some assumption (see [Lasserre, 2015]).
Consider an equality constrained POP on the unit sphere:

\[ f^* := \inf \{ f(x) : h_j(x) = 0, \ j \in [l] \} \quad \text{with} \quad h_1 = 1 - \|x\|_2^2. \quad (5) \]
Moment-SOS hierarchy

Consider an equality constrained POP on the unit sphere:

\[ f^* := \inf \{ f(x) : h_j(x) = 0, j \in [l] \} \quad \text{with} \quad h_1 = 1 - \|x\|_2^2. \quad (5) \]

For every \( k \in \mathbb{N} \), consider SDP [Lasserre, 2001]:

\[ \tau_k = \inf_y \{ L_y(f) : M_k(y) \succeq 0, M_{k-w_j}(h_j y) = 0, y_0 = 1 \}. \quad (6) \]
Consider an equality constrained POP on the unit sphere:

$$f^* := \inf \{ f(x) : h_j(x) = 0, \; j \in [l] \} \quad \text{with} \quad h_1 = 1 - \|x\|^2_2. \quad (5)$$

For every $k \in \mathbb{N}$, consider SDP [Lasserre, 2001]:

$$\tau_k = \inf_y \{ L_y(f) : M_k(y) \succeq 0, \; M_{k-w_j}(h_j y) = 0, \; y_0 = 1 \}. \quad (6)$$

Strong duality holds due to [Josz and Henrion, 2016].
Consider an equality constrained POP on the unit sphere:

\[ f^* := \inf \{ f(x) : h_j(x) = 0, j \in [l] \} \quad \text{with} \quad h_1 = 1 - \|x\|_2^2. \quad (5) \]

For every \( k \in \mathbb{N} \), consider SDP [Lasserre, 2001]:

\[ \tau_k = \inf_y \{ L_y(f) : M_k(y) \succeq 0, M_{k-w_j}(h_j y) = 0, y_0 = 1 \}. \quad (6) \]

Strong duality holds due to [Josz and Henrion, 2016].

**Convergence rate:** [Schweighofer, 2004] \( \tau_k \uparrow f^* \) with \( O(k^{-1/c}) \).
Consider an equality constrained POP on the unit sphere:

\[ f^* := \inf \{ f(x) : h_j(x) = 0, j \in [l] \} \quad \text{with} \quad h_1 = 1 - \|x\|_2^2. \]  

(5)

For every \( k \in \mathbb{N} \), consider SDP [Lasserre, 2001]:

\[ \tau_k = \inf_y \{ L_y(f) : M_k(y) \succeq 0, M_{k-w_j}(h_j y) = 0, y_0 = 1 \} . \]  

(6)

Strong duality holds due to [Josz and Henrion, 2016].

**Convergence rate:** [Schweighofer, 2004] \( \tau_k \uparrow f^* \) with \( \mathcal{O}(k^{-1/c}) \). Moreover \( \tau_k = f^* \) under some assumption (see [Lasserre, 2015]).
Example

Consider a simple example of POP (5) with $n = 1$:

$-1 = \inf\{ x : x_1 - x_2 = 0 \}$.

Then the second order moment relaxation ($k = 2$) has the form:

$\tau^2 = \inf y_1 y_2 \text{s.t.} \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \succeq 0,$

obviously, the psd matrix of this form has trace 3.
Example

Consider a simple example of POP (5) with $n = 1$:

$$-1 = \inf\{x : 1 - x^2 = 0\}.$$
Example

Consider a simple example of POP (5) with $n = 1$:

$$-1 = \inf \{ x : 1 - x^2 = 0 \}.$$ 

Then the second order moment relaxation ($k = 2$) has the form:

$$\tau_2 = \inf_{y} y_1$$

$$\begin{bmatrix}
1 & x & x^2 \\
1 & y_0 & y_1 & y_2 \\
x & y_1 & y_2 & y_3 \\
x^2 & y_2 & y_3 & y_4 \\
\end{bmatrix} \succeq 0,$$

$$\begin{bmatrix}
1 & x \\
y_0 - y_2 & y_1 - y_3 \\
y_1 - y_3 & y_2 - y_4 \\
\end{bmatrix} = 0, \ y_0 = 1.$$
Example

Consider a simple example of POP (5) with $n = 1$:

$$-1 = \inf\{x : 1 - x^2 = 0\}.$$  

Then the second order moment relaxation ($k = 2$) has the form:

$$\tau_2 = \inf_y y_1$$

$$\begin{bmatrix} 1 & x & x^2 \\ 1 & y_0 & y_1 & y_2 \\ x & y_1 & y_2 & y_3 \\ x^2 & y_2 & y_3 & y_4 \end{bmatrix} \succeq 0,$$

$$\begin{bmatrix} 1 & x \\ 1 & y_0 - y_2 & y_1 - y_3 \\ x & y_1 - y_3 & y_2 - y_4 \end{bmatrix} = 0, \; y_0 = 1.$$

By removing equality constraints, it can be rewritten as

$$\tau_2 = \inf_y y_1$$

$$\begin{bmatrix} 1 & y_1 & 1 \\ y_1 & 1 & y_1 \\ 1 & y_1 & 1 \end{bmatrix} \succeq 0,$$
Example

Consider a simple example of POP (5) with \( n = 1 \):

\[-1 = \inf \{ x : 1 - x^2 = 0 \}.\]

Then the second order moment relaxation \((k = 2)\) has the form:

\[
\tau_2 = \inf_y y_1
\]

\[
\begin{bmatrix}
1 & x & x^2 \\
1 & y_0 & y_1 & y_2 \\
x & y_1 & y_2 & y_3 \\
x^2 & y_2 & y_3 & y_4
\end{bmatrix} \succeq 0, \quad 1 \begin{bmatrix}
y_0 - y_2 & y_1 - y_3 \\
y_1 - y_3 & y_2 - y_4
\end{bmatrix} = 0, \quad y_0 = 1.
\]

By removing equality constraints, it can be rewritten as

\[
\tau_2 = \inf_y y_1
\]

\[
\begin{bmatrix}
1 & y_1 & 1 \\
y_1 & 1 & y_1 \\
1 & y_1 & 1
\end{bmatrix} \succeq 0,
\]

Obviously, the psd matrix of this form has trace 3.
Exploiting constant trace property

Fix $k \in \mathbb{N}$. Let $P_k$ be a diagonal matrix with nonzero entries being the square roots of the coefficients of polynomial $(1 + \|x\|_2^2)^k$.

For instance, with $n = 1$ and $k = 2$, $(1 + x^2)^2 = 1 + 2x^2 + x^4$ yields $P_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

$(7)$

$M_k - 1((1 - \|x\|_2^2)y) = 0$, $y_0 = 1$ $\Rightarrow \text{trace} \left( 2 - kP_k M_k(y)P_k \right) = 1$.

$(8)$

Since $P_k \succ 0$, (6) is equivalent to SDP:

$\tau_k = \inf y \text{ } L y(f) \text{ s.t. } 2 - kP_k M_k(y)P_k \succeq 0$, $M_k - w_j(h_jy) = 0$, $y_0 = 1$.

$(9)$
Exploiting constant trace property

Fix $k \in \mathbb{N}$. Let $P_k$ be a diagonal matrix with nonzero entries being the square roots of the coefficients of polynomial $(1 + \|x\|_2^2)^k$. 

\[ M_{k-1}(1 - \|x\|_2^2) = 0, \]
\[ y_0 = 1 \implies \text{trace}(2 - k P_k M_k y P_k) = 1. \]

Since $P_k \succ 0$, (6) is equivalent to SDP:
\[ \tau_k = \inf_y L(y)(f) \text{ s.t. } 2 - k P_k M_k y P_k \succeq 0, M_k - w_j(h_j y) = 0, y_0 = 1. \]
Exploiting constant trace property

Fix $k \in \mathbb{N}$. Let $P_k$ be a diagonal matrix with nonzero entries being the square roots of the coefficients of polynomial $(1 + \|x\|_2^2)^k$.

For instance, with $n = 1$ and $k = 2$, $(1 + x^2)^k = 1 + 2x^2 + x^4$ yields

$$P_k = \begin{bmatrix}
1 & 0 & 0 \\
0 & \sqrt{2} & 0 \\
0 & 0 & 1
\end{bmatrix} x .$$

$$P_k = \begin{bmatrix}
1 & 0 & 0 \\
0 & \sqrt{2} & 0 \\
0 & 0 & 1
\end{bmatrix} x .$$

$$P_k = \begin{bmatrix}
1 & 0 & 0 \\
0 & \sqrt{2} & 0 \\
0 & 0 & 1
\end{bmatrix} x .$$

$(7)$
Exploiting constant trace property

Fix $k \in \mathbb{N}$. Let $P_k$ be a diagonal matrix with nonzero entries being the square roots of the coefficients of polynomial $(1 + \|x\|_2^2)^k$.

For instance, with $n = 1$ and $k = 2$, $(1 + x^2)^2 = 1 + 2x^2 + x^4$ yields

$$P_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} x.$$  \hspace{1cm} (7)

$$M_{k-1}((1 - \|x\|_2^2) y) = 0, \quad \begin{array}{c} y_0 = 1 \\ y_0 = 1 \end{array} \Rightarrow \text{trace}(2^{-k} P_k M_k(y) P_k) = 1. \hspace{1cm} (8)$$
Exploiting constant trace property

Fix $k \in \mathbb{N}$. Let $P_k$ be a diagonal matrix with nonzero entries being the square roots of the coefficients of polynomial $(1 + \|x\|_2^2)^k$.

For instance, with $n = 1$ and $k = 2$, $(1 + x^2)^k = 1 + 2x^2 + x^4$ yields

$$P_k = \begin{bmatrix}
1 & 0 & 0 \\
0 & \sqrt{2} & 0 \\
0 & 0 & 1
\end{bmatrix} x \cdot
\begin{bmatrix}
1 \\
\sqrt{2} \\
0
\end{bmatrix} x^2
1 \times \ x^2
$$

$$M_{k-1}((1 - \|x\|_2^2) y) = 0, \quad \Rightarrow \quad \text{trace}(2^{-k}P_k M_k(y)P_k) = 1. \quad (8)$$

Since $P_k \succ 0$, (6) is equivalent to SDP:

$$\tau_k = \inf_y \quad L_y(f)$$

s.t. \quad $2^{-k}P_k M_k(y)P_k \succeq 0,$

$M_{k-w_j}(h_j y) = 0, \quad y_0 = 1,$

$$\quad (9)$$
Exploiting constant trace property
Exploiting constant trace property

With \( \mathbf{X} := 2^{-k} \mathbf{P}_k \mathbf{M}_k(\mathbf{y}) \mathbf{P}_k \), SDP (9) can be written in the form:

\[
-\tau_k = \sup_{\mathbf{X} \in \mathcal{S}_k} \text{trace}(\mathbf{C}_k \mathbf{X}) \\
\text{s.t.} \quad \mathcal{A}_k \mathbf{X} = \mathbf{b}_k, \\
\text{trace}(\mathbf{X}) = 1, \\
\mathbf{X} \succeq 0, 
\]

(10)
With $X := 2^{-k}P_k M_k(y)P_k$, SDP (9) can be written in the form:

$$-\tau_k = \sup_{X \in S_k} \text{trace}(C_k X)$$

s.t. $A_k X = b_k$,  
$\text{trace}(X) = 1$,  
$X \succeq 0$,  

where $A_k : S_k \to \mathbb{R}^{m_k}$ is a linear operator of the form

$$A_k X = \begin{bmatrix} \text{trace}(A_k^{(1)} X), \ldots, \text{trace}(A_k^{(m_k)} X) \end{bmatrix}.$$
Spectral relaxations

Following [Helmberg and Rendl, 2000], we obtain:

$$-\tau_k = \inf_{z} \lambda_{\text{max}}(C_k - A_k^T z) + b_k^T z.$$  \hspace{1cm} (11)
Spectral relaxations

Following [Helmberg and Rendl, 2000], we obtain:

\[- \tau_k = \inf_z \lambda_{\text{max}}(C_k - A_k^T z) + b_k^T z.\] (11)

Here:

\[A_k^T z = \sum_{i \in [m_k]} z_i A_k^{(i)}.\]
Spectral relaxations

Following [Helmberg and Rendl, 2000], we obtain:

\[-\tau_k = \inf_{z} \lambda_{\text{max}}(C_k - A_k^T z) + b_k^T z.\] (11)

Here:

\[A_k^T z = \sum_{i \in [m_k]} z_i A_k^{(i)}.\]

Notice that:

▶ \(A_k^{(i)}\) and \(C_k\) are very sparse and have size \(s_k = (n + k)n\);
▶ The number of equality constraints: \(m_k = O((n + kn)^2)\);
▶ \(b_k = [0, \ldots, 0, 1]^T\).

\(m_k \geq 200 \times s_k\) when \(n \geq 30\) and \(k \geq 2\) ⇒ not suitable for spectral bundle method [Helmberg et al., 2014].
Spectral relaxations

Following [Helmberg and Rendl, 2000], we obtain:

$$- \tau_k = \inf_z \lambda_{\text{max}}(C_k - A_k^T z) + b_k^T z. \quad (11)$$

Here:

$$A_k^T z = \sum_{i \in [m_k]} z_i A_k^{(i)}.$$

Notice that:

- $A_k^{(i)}$ and $C_k$ are very sparse and have size $s_k = \binom{n+k}{n}$;
Spectral relaxations

Following [Helmberg and Rendl, 2000], we obtain:

\[-\tau_k = \inf_z \lambda_{\max}(C_k - A_k^T z) + b_k^T z.\]  \hspace{1cm} (11)

Here:

\[A_k^T z = \sum_{i \in [m_k]} z_i A_k^{(i)}.\]

Notice that:

- \(A_k^{(i)}\) and \(C_k\) are very sparse and have size \(s_k = \binom{n+k}{n}\);
- The number of equality constraints: \(m_k = \mathcal{O}\left(\binom{n+k}{n}^2\right)\)
Spectral relaxations

Following [Helmberg and Rendl, 2000], we obtain:

\[ -\tau_k = \inf_z \lambda_{\max}(C_k - A_k^T z) + b_k^T z. \]  \hspace{1cm} (11)

Here:

\[ A_k^T z = \sum_{i \in [m_k]} z_i A^{(i)}_k. \]

Notice that:

- \( A^{(i)}_k \) and \( C_k \) are very sparse and have size \( s_k = \binom{n+k}{n} \);
- The number of equality constraints: \( m_k = O \left( \binom{n+k}{n}^2 \right) \);
- \( b_k = [0, \ldots, 0, 1]^T. \)
Spectral relaxations

Following [Helmberg and Rendl, 2000], we obtain:

$$-\tau_k = \inf_z \lambda_{\text{max}}(C_k - A_k^T z) + b_k^T z. \quad (11)$$

Here:

$$A_k^T z = \sum_{i \in [m_k]} z_i A_k^{(i)}.$$

Notice that:

- $A_k^{(i)}$ and $C_k$ are very sparse and have size $s_k = \binom{n+k}{n}$;
- The number of equality constraints: $m_k = O\left(\binom{n+k}{n}^2\right)$;
- $b_k = [0, \ldots, 0, 1]^T$.

$m_k \geq 200 \times s_k$ when $n \geq 30$ and $k \geq 2$

$\Rightarrow$ not suitable for spectral bundle method [Helmberg et al., 2014];
SpectralPOP Algorithm

1. Compute the optimal value and an optimal solution \( z^\star \) of
   \[
   -\tau_k = \inf_{\lambda} \lambda \max (C_k - A_k^\top z) + b_k^\top z
   \]

2. Compute uniform eigenvectors \( u_1, \ldots, u_r \) corresponding to the largest eigenvalue of
   \( C_k - A_k^\top z^\star \)

3. Compute \( \xi^\star \in \arg\min_{\xi \in \Delta} (r - 1) \| b_k - A_k \|_2 \sum_{j \in \{r\}} \xi_j u_j u_j^\top \),
   \( \Delta \): the standard \( (r - 1) \)-simplex

4. \( X^\star \leftarrow \sum_{j \in \{r\}} \xi_j^\star u_j u_j^\top \) and \( M_k(y^\star) \leftarrow 2^k P - 1^k X^\star P - 1^k \)

5. Extract \( x^\star \) from \( y^\star \)

[Henrion and Lasserre, 2005]

Limited Memory Bundle Method [Haarala et al., 2007]
SpectralPOP Algorithm

Input: \( f, h_j \) of POP (5) and \( k \in \mathbb{N} \)
SpectralPOP Algorithm

**Input:** \( f, h_j \) of POP (5) and \( k \in \mathbb{N} \)  

**Output:** \( \tau_k \) and \( x^* \)
SpectralPOP Algorithm

Input: \( f, h_j \) of POP (5) and \( k \in \mathbb{N} \)  
Output: \( \tau_k \) and \( x^* \)

1. Compute the optimal value and an optimal solution \( z^* \) of

\[
-\tau_k = \inf_z \lambda_{\text{max}}(C_k - A_k^T z) + b_k^T z
\]  

\( \triangleright \) Limited Memory Bundle Method [Haarala et al., 2007];
SpectralPOP Algorithm

Input: \( f, h_j \) of POP (5) and \( k \in \mathbb{N} \)  
Output: \( \tau_k \) and \( x^* \)

1. Compute the optimal value and an optimal solution \( z^* \) of

\[
\tau_k = \inf_z \lambda_{\text{max}}(C_k - A_k^T z) + b_k^T z
\]

\( \triangleq \) Limited Memory Bundle Method [Haarala et al., 2007];

2. Compute uniform eigenvectors \( u_1, \ldots, u_r \) corresponding to the largest eigenvalue of \( C_k - A_k^T z^* \);
SpectralPOP Algorithm

**Input:** $f, h_j$ of POP (5) and $k \in \mathbb{N}$

**Output:** $\tau_k$ and $x^*$

1. Compute the optimal value and an optimal solution $z^*$ of

   $$-\tau_k = \inf_z \lambda_{\text{max}}(C_k - A_k^T z) + b_k^T z$$

   (12)

   ▶ Limited Memory Bundle Method [Haarala et al., 2007];

2. Compute uniform eigenvectors $u_1, \ldots, u_r$ corresponding to the largest eigenvalue of $C_k - A_k^T z^*$;

3. Compute $\xi^* \in \text{argmin}_{\xi \in \Delta^{(r-1)}} \left\| b_k - A_k \left( \sum_{j \in [r]} \xi_j u_j u_j^T \right) \right\|_2^2$,

   $\Delta^{(r-1)}$: the standard $(r - 1)$-simplex;
SpectralPOP Algorithm

**Input:** \( f, h_j \) of POP (5) and \( k \in \mathbb{N} \)  

**Output:** \( \tau_k \) and \( x^* \)

1. Compute the optimal value and an optimal solution \( z^* \) of

\[
-\tau_k = \inf_z \lambda_{\max}(C_k - A_k^T z) + b_k^T z
\]

▷ Limited Memory Bundle Method [Haarala et al., 2007];

2. Compute uniform eigenvectors \( u_1, \ldots, u_r \) corresponding to the largest eigenvalue of \( C_k - A_k^T z^* \);

3. Compute \( \xi^* \in \text{argmin}_{\xi \in \Delta^{(r-1)}} \| b_k - A_k \left( \sum_{j \in [r]} \xi_j u_j u_j^T \right) \|_2^2 \), \( \Delta^{(r-1)} \): the standard \((r-1)\)-simplex;

4. \( X^* \leftarrow \sum_{j \in [r]} \xi_j^* u_j u_j^T \) and \( M_k(y^*) \leftarrow 2^k P_k^{-1} X^* P_k^{-1} \);
SpectralPOP Algorithm

**Input:** \( f, h_j \) of POP (5) and \( k \in \mathbb{N} \)  

**Output:** \( \tau_k \) and \( x^* \)

1. Compute the optimal value and an optimal solution \( z^* \) of

\[
-\tau_k = \inf_z \lambda_{\max}(C_k - A_k^T z) + b_k^T z
\]

\( (12) \)

▷ Limited Memory Bundle Method [Haarala et al., 2007];

2. Compute uniform eigenvectors \( u_1, \ldots, u_r \) corresponding to the largest eigenvalue of \( C_k - A_k^T z^* \);

3. Compute \( \xi^* \in \arg\min_{\xi \in \Delta^{(r-1)}} \left\| b_k - A_k \left( \sum_{j \in [r]} \xi_j u_j u_j^T \right) \right\|_2^2 \),

\( \Delta^{(r-1)} \): the standard \((r-1)\)-simplex;

4. \( X^* \leftarrow \sum_{j \in [r]} \xi_j^* u_j u_j^T \) and \( M_k(y^*) \leftarrow 2^k P_k^{-1} X^* P_k^{-1} \);

5. Extract \( x^* \) from \( y^* \) [Henrion and Lasserre, 2005].
Spectral relaxations for quadratic forms \((k = 1)\)

matrix size \(s_k = n + 1\)  
2 equality constraints  
gap: the relative optimality gap w.r.t. SumOfSquares  
LMBM: Limited Memory Bundle Method
Spectral relaxations for quadratic forms ($k = 1$)

matrix size $s_k = n + 1$  
2 equality constraints  
gap: the relative optimality gap w.r.t. SumOfSquares  
LMBM: Limited Memory Bundle Method  

![Graph showing time vs. n for different methods.](image1)  
![Graph showing gap vs. n for different methods.](image2)
Spectral relaxations for quadratic forms \((k = 1)\)

matrix size \(s_k = n + 1\)  
2 equality constraints  
gap: the relative optimality gap w.r.t. SumOfSquares  
LMBM: Limited Memory Bundle Method  

\[50 \quad 300 \quad 500 \quad 1000 \quad 1500\]

\[0 \quad 500 \quad 1000 \quad 1500 \quad 3000\]

\[0 \quad 1 \quad 2 \quad 3 \quad 10^{-2}\]

▶ **SpectralPOP** (LMBM):  
▶ can solve POPs with \(n = 300\) variables in 5 seconds while it takes 1 hour in **SumOfSquares** (Mosek).  
▶ provide optimal value (gap < \(10^{-5}\)) and optimal solution with high accuracy for \(n = 1500\) in 1 hour.
Spectral relaxations for QCQP ($k = 2$)

size of matrix: $s_k = \binom{n+k}{n}$

number of equality constraints: $m_k = \mathcal{O} \left( \left( \binom{n+k}{n} \right)^2 \right) \geq 200 \times s_k$
Spectral relaxations for QCQP ($k = 2$)

size of matrix: $s_k = \binom{n+k}{n}$

number of equality constraints: $m_k = \mathcal{O}\left(\left(\frac{n+k}{n}\right)^2\right) \geq 200 \times s_k$
Spectral relaxations for QCQP \((k = 2)\)

size of matrix: \(s_k = \binom{n+k}{n}\)

number of equality constraints: \(m_k = \mathcal{O}\left(\left(\binom{n+k}{n}\right)^2\right) \geq 200 \times s_k\)

If \(n = 65\), then \(s_k = 2211\) and \(m_k = 1618453\).
Spectral relaxations for QCQP ($k = 2$)

size of matrix: $s_k = \binom{n+k}{n}$

number of equality constraints: $m_k = \mathcal{O}\left(\left(\frac{n+k}{n}\right)^2\right) \geq 200 \times s_k$

If $n = 65$, then $s_k = 2211$ and $m_k = 1618453$.

**SpectralPOP (LMBM):**

▶ provide the optimal value with high accuracy ($\text{gap} < 10^{-5}$)
▶ up to more than 150 times faster than **SumOfSquares** (Mosek) when $n \geq 25$. 
Summary and Future works

Comparison with SumOfSquares (Mosek) and SketchyCGAL on the last tested sample of POPs: SpectralPOP (LMBM) is cheaper, faster, but maintains the same accuracy as SumOfSquares.

Future works:
- exploiting CTP for POPs with annulus constraints: $r \leq \|x(T)\|_2^2 \leq R$
- exploiting CTP for sparse POPs
Summary and Future works

- Comparison with **SumOfSquares** (Mosek) and **SketchyCGAL** on the last tested sample of POPs:

  **SpectralPOP** (LMBM) is *cheaper, faster, but maintains the same accuracy* as **SumOfSquares**.
Summary and Future works

▶ Comparison with **SumOfSquares** (Mosek) and **SketchyCGAL** on the last tested sample of POPs:

**SpectralPOP** (LMBM) is *cheaper, faster, but maintains the same accuracy* as **SumOfSquares**.

▶ Future works:
  ▶ exploiting CTP for POPs with annulus constraints:
    \[ r \leq \|x(T)\|_2^2 \leq R \]
  ▶ exploiting CTP for sparse POPs
Thank you for your attention!

**SpectralPOP:** https://github.com/maihoanganh/SpectralPOP

**Homepage:** https://sites.google.com/view/hoanganhmai
Globally convergent limited memory bundle method for large-scale nonsmooth optimization.

The spectral bundle method with second-order information.

A spectral bundle method for semidefinite programming.

Detecting global optimality and extracting solutions in gloptipoly.
In *Positive polynomials in control*, pages 293–310. Springer.

Strong duality in Lasserre’s hierarchy for polynomial optimization.

Global optimization with polynomials and the problem of moments.

Convergent SDP-Relaxations in Polynomial Optimization with Sparsity.

*An introduction to polynomial and semi-algebraic optimization*, volume 52.
Cambridge University Press.

On the complexity of Schmûdgen’s positivstellensatz.
