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

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

Scalability of moment relaxations for POP

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

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

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

Scalability of moment relaxations for POP

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

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

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, $(g_i)_{i \in \mathcal{G}_j}$, $(h_i)_{i \in \mathcal{H}_j}$ share the same few variables,

Scalability of moment relaxations for POP

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

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

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, $(g_i)_{i \in \mathcal{G}_j}$, $(h_i)_{i \in \mathcal{H}_j}$ share the same few variables,e.g., $f = \underbrace{x_1x_2 + x_2x_3 + x_3x_1}_{f_1} + \underbrace{x_3x_4}_{f_2} + \underbrace{x_4x_5}_{f_3}$.

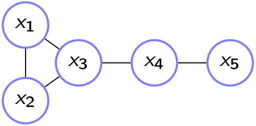
Scalability of moment relaxations for POP

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

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

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, $(g_i)_{i\in\mathcal{G}_j}$, $(h_i)_{i\in\mathcal{H}_j}$ share the same few variables,e.g., $f = \underbrace{x_1x_2 + x_2x_3 + x_3x_1}_{f_1} + \underbrace{x_3x_4}_{f_2} + \underbrace{x_4x_5}_{f_3}$.



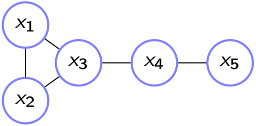
Scalability of moment relaxations for POP

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

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

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, $(g_i)_{i\in\mathcal{G}_j}$, $(h_i)_{i\in\mathcal{H}_j}$ share the same few variables,e.g., $f = \underbrace{x_1x_2 + x_2x_3 + x_3x_1}_{f_1} + \underbrace{x_3x_4}_{f_2} + \underbrace{x_4x_5}_{f_3}$.



- term sparsity: [Wang et al., 2019] f, g_i, h_j have few nontrivial terms $c_\alpha \mathbf{x}^\alpha$,

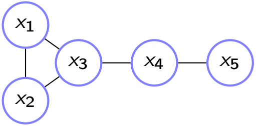
Scalability of moment relaxations for POP

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

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

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, $(g_i)_{i \in \mathcal{G}_j}$, $(h_i)_{i \in \mathcal{H}_j}$ share the same few variables,e.g., $f = \underbrace{x_1x_2 + x_2x_3 + x_3x_1}_{f_1} + \underbrace{x_3x_4}_{f_2} + \underbrace{x_4x_5}_{f_3}$.



- term sparsity: [Wang et al., 2019] f, g_i, h_j have few nontrivial terms $c_\alpha \mathbf{x}^\alpha$,e.g., $f = 1 + x_1x_2 + x_2x_3 + x_3$.

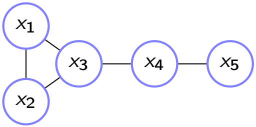
Scalability of moment relaxations for POP

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

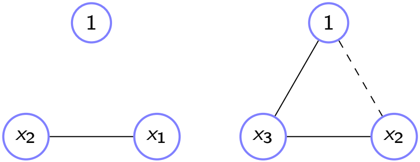
$$f^\star := \inf\{ f(\mathbf{x}) \ : \ g_i(\mathbf{x}) \geq 0 \,,\ i \in [m] \,,\ h_j(\mathbf{x}) \ = \ 0 \,,\ j \in [l] \,\} \,. \quad (1)$$

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, $(g_i)_{i \in \mathcal{G}_j}$, $(h_i)_{i \in \mathcal{H}_j}$ share the same few variables,e.g., $f = \underbrace{x_1x_2 + x_2x_3 + x_3x_1}_{f_1} + \underbrace{x_3x_4}_{f_2} + \underbrace{x_4x_5}_{f_3}$.



- term sparsity: [Wang et al., 2019] f, g_i, h_j have few nontrivial terms $c_\alpha \mathbf{x}^\alpha$,e.g., $f = 1 + x_1x_2 + x_2x_3 + x_3$.



Scalability of moment relaxations for POP

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

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

What if no sparsity (dense POP)?

Scalability of moment relaxations for POP

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

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

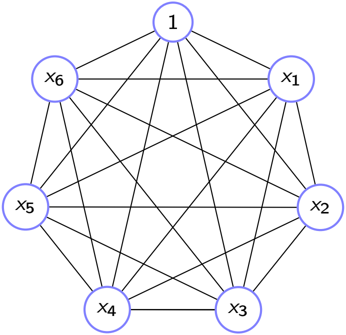
What if no sparsity (dense POP)?
For instance, $f = 1 + \sum_{i \in [6]} x_i + \sum_{(i,j) \in [6]^2} x_i x_j$

Scalability of moment relaxations for POP

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

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

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

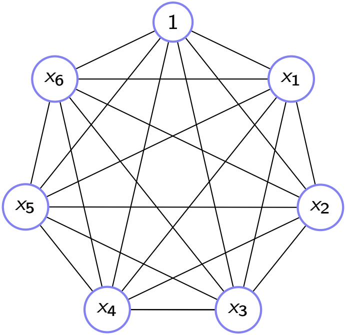


Scalability of moment relaxations for POP

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

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

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



$\Rightarrow$ use other information: constant trace property.

SDP with constant trace property

SDP with constant trace property

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

SDP with constant trace property

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

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

$$-\tau = \sup_{X \in \mathcal{S}} \{\text{trace}(\mathbf{CX}) \,:\, \mathcal{A}\mathbf{X} = \mathbf{b}\,,\, \mathbf{X} \succeq 0\,,\, \text{trace}(\mathbf{X}) = a\}\,, \tag{3}$$

SDP with constant trace property

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

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

$$-\tau = \sup_{X \in \mathcal{S}} \{\text{trace}(\mathbf{CX}) \, : \, \mathcal{A}\mathbf{X} = \mathbf{b} \, , \, \mathbf{X} \succeq 0 \, , \, \text{trace}(\mathbf{X}) = a\} \, , \tag{3}$$

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

$$-\rho = \inf_{\mathbf{z}} a\lambda_{\max}(\mathbf{C} - \mathcal{A}^\top \mathbf{z}) + \mathbf{b}^\top \mathbf{z} \, , \tag{4}$$

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

SDP with constant trace property

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

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

$$-\tau = \sup_{X \in \mathcal{S}} \{ \text{trace}(\mathbf{CX}) \, : \, \mathcal{A}\mathbf{X} = \mathbf{b} \, , \, \mathbf{X} \succeq 0 \, , \, \text{trace}(\mathbf{X}) = a \} \, , \tag{3}$$

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

   $$-\rho = \inf_{\mathbf{z}} a \lambda_{\max}(\mathbf{C} - \mathcal{A}^\top \mathbf{z}) + \mathbf{b}^\top \mathbf{z} \, , \tag{4}$$

   where $\lambda_{\max}(\mathbf{A})$ is the largest eigenvalue of matrix $\mathbf{A}$.
2. **SketchyCGAL**=Sketchy+CG+AL: [Yurtsever et al., 2019]
   - the augmented Lagrangian (AL)
   - conditional gradient method (CG)
   - randomized sketch (Sketchy)

Contribution

Contribution

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

Contribution

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

Contribution

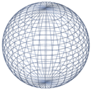
- Exploit constant trace property (CTP) for moment relaxations of POP.
- Convert moment relaxations with CTP to spectral relaxations.
- Build an algorithm called **SpectralPOP** to solve spectral relaxations and extract the optimal solutions of POP.

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.

Moment-SOS hierarchy

Moment-SOS hierarchy

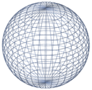


Consider an equality constrained POP on the unit sphere:

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

Moment-SOS hierarchy

Consider an equality constrained POP on the unit sphere:

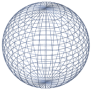


$$f^\star := \inf\{f(\mathbf{x}) \,:\, h_j(\mathbf{x}) \,=\, 0\,, j \in [l]\,\} \quad \text{with} \quad h_1 = 1 - \|\mathbf{x}\|_2^2\,. \quad (5)$$

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

$$\tau_k = \inf_{\mathbf{y}}\{\,L_{\mathbf{y}}(f) \,:\, \mathbf{M}_k(\mathbf{y}) \succeq 0\,,\, \mathbf{M}_{k-w_j}(h_j\;\mathbf{y}) = 0\,,\, y_0 = 1\}\,. \quad (6)$$

Moment-SOS hierarchy



Consider an equality constrained POP on the unit sphere:

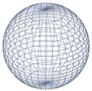
$$f^\star := \inf\{ f(\mathbf{x}) \ : \ h_j(\mathbf{x}) \ = \ 0\,, j \in [l]\,\} \quad \text{with} \quad h_1 = 1 - \|\mathbf{x}\|_2^2\,. \quad (5)$$

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

$$\tau_k = \inf_{\mathbf{y}} \{ \, L_{\mathbf{y}}(f) \ : \ \mathbf{M}_k(\mathbf{y}) \succeq 0\,, \ \mathbf{M}_{k-w_j}(h_j \ \mathbf{y}) = 0\,, \ y_0 = 1 \}\,. \quad (6)$$

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

Moment-SOS hierarchy



Consider an equality constrained POP on the unit sphere:

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

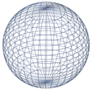
For every $k \in \mathbb{N}$, consider SDP [Lasserre, 2001]:

$$\tau_k = \inf_{\mathbf{y}}\{\, L_{\mathbf{y}}(f) \,:\, \mathbf{M}_k(\mathbf{y}) \succeq 0\,,\, \mathbf{M}_{k-w_j}(h_j\ \mathbf{y}) = 0\,,\, y_0 = 1\}\,. \quad (6)$$

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

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

Moment-SOS hierarchy



Consider an equality constrained POP on the unit sphere:

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

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

$$\tau_k = \inf_{\mathbf{y}}\{\, L_{\mathbf{y}}(f) \,:\, \mathbf{M}_k(\mathbf{y}) \succeq 0\,,\, \mathbf{M}_{k-w_j}(h_j\ \mathbf{y}) = 0\,,\, y_0 = 1\}\,. \quad (6)$$

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

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

Example

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:

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

$$\text{s.t.} \quad \begin{array}{c} \\ 1 \\ x \\ x^2 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & x & x^2 \end{array} \\ \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \end{array} \succeq 0\,, \begin{array}{c} \\ 1 \\ x \end{array} \begin{array}{c} \begin{array}{cc} 1 & x \end{array} \\ \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} \end{array} = 0\,,\ y_0 = 1\,.$$

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:

$$\begin{aligned}
\tau_2 = \inf_{\mathbf{y}} \quad & y_1 \\
\text{s.t.} \quad & \begin{array}{c} \\ 1 \\ x \\ x^2 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 1 & x & x^2 \end{array} \\ \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \end{array} \succeq 0\,,\ \begin{array}{c} \\ 1 \\ x \end{array}\!\!\begin{array}{c} \begin{array}{cc} 1 & x \end{array} \\ \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} \end{array} = 0\,,\ y_0 = 1\,.
\end{aligned}$$

By removing equality constraints, it can be rewritten as

$$\begin{aligned}
\tau_2 = \inf_{\mathbf{y}} \quad & y_1 \\
\text{s.t.} \quad & \begin{bmatrix} 1 & y_1 & 1 \\ y_1 & 1 & y_1 \\ 1 & y_1 & 1 \end{bmatrix} \succeq 0\,,
\end{aligned}$$

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:

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

$$\text{s.t.} \quad \begin{array}{c} \\ 1 \\ x \\ x^2 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 1 & x & x^2 \end{array} \\ \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \end{array} \succeq 0\,,\; \begin{array}{c} \\ 1 \\ x \end{array}\!\!\begin{array}{c} \begin{array}{cc} 1 & x \end{array} \\ \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} \end{array} = 0\,,\; y_0 = 1\,.$$

By removing equality constraints, it can be rewritten as

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

$$\text{s.t.} \quad \begin{bmatrix} 1 & y_1 & 1 \\ y_1 & 1 & y_1 \\ 1 & y_1 & 1 \end{bmatrix} \succeq 0\,,$$

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

Exploiting constant trace property

Exploiting constant trace property

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

Exploiting constant trace property

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

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

$$\mathbf{P}_k = \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{matrix} 1 \\ x \\ x^2 \end{matrix} \\ \begin{matrix} 1 & x & x^2 \end{matrix} \end{array}. \tag{7}$$

Exploiting constant trace property

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

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

$$\mathbf{P}_k = \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{array}{l} 1 \\ x \\ x^2 \end{array} \\ \begin{array}{ccc} 1 & x & x^2 \end{array} \end{array} . \tag{7}$$

$$\left.\begin{array}{l} \mathbf{M}_{k-1}((1 - \|\mathbf{x}\|_2^2)\,\mathbf{y}) = 0\,, \\ y_0 = 1 \end{array}\right\} \Rightarrow \operatorname{trace}(2^{-k}\mathbf{P}_k\mathbf{M}_k(y)\mathbf{P}_k) = 1\,. \tag{8}$$

Exploiting constant trace property

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

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

$$\mathbf{P}_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{matrix} 1 \\ x \\ x^2 \end{matrix} . \quad (7)$$
$$\begin{matrix} 1 & x & x^2 \end{matrix}$$

$$\left.\begin{matrix} \mathbf{M}_{k-1}((1 - \|\mathbf{x}\|_2^2)\,\mathbf{y}) = 0\,, \\ y_0 = 1 \end{matrix}\right\} \Rightarrow \mathrm{trace}(2^{-k}\mathbf{P}_k\mathbf{M}_k(y)\mathbf{P}_k) = 1\,. \quad (8)$$

Since $\mathbf{P}_k \succ 0$, (6) is equivalent to SDP:

$$\begin{aligned} \tau_k = \inf_{\mathbf{y}} \quad & L_{\mathbf{y}}(f) \\ \text{s.t.} \quad & 2^{-k}\mathbf{P}_k\mathbf{M}_k(y)\mathbf{P}_k \succeq 0\,, \\ & \mathbf{M}_{k-w_j}(h_j\mathbf{y}) = 0\,,\ y_0 = 1\,, \end{aligned} \quad (9)$$

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:

$$
\begin{aligned}
-\tau_k = \sup_{\mathbf{X}\in\mathcal{S}_k} \quad & \mathrm{trace}(\mathbf{C}_k\mathbf{X}) \\
\text{s.t.} \quad & \mathcal{A}_k\mathbf{X} = \mathbf{b}_k\,, \\
& \mathrm{trace}(\mathbf{X}) = 1\,, \\
& \mathbf{X} \succeq 0\,,
\end{aligned}
\tag{10}
$$

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:

$$\begin{aligned} -\tau_k = \sup_{\mathbf{X}\in\mathcal{S}_k} \quad & \mathrm{trace}(\mathbf{C}_k\mathbf{X}) \\ \text{s.t.} \quad & \mathcal{A}_k\mathbf{X} = \mathbf{b}_k\,, \\ & \mathrm{trace}(\mathbf{X}) = 1\,, \\ & \mathbf{X} \succeq 0\,, \end{aligned} \tag{10}$$

where $\mathcal{A}_k : \mathcal{S}_k \to \mathbb{R}^{m_k}$ is a linear operator of the form

$$\mathcal{A}_k\mathbf{X} = \left[\mathrm{trace}(\mathbf{A}_k^{(1)}\mathbf{X}), \dots, \mathrm{trace}(\mathbf{A}_k^{(m_k)}\mathbf{X})\right]\,.$$

Spectral relaxations

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

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z}\,. \tag{11}$$

Spectral relaxations

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

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z}\,. \tag{11}$$

Here:

$$\mathcal{A}_k^\top z = \sum_{i \in [m_k]} z_i \mathbf{A}_k^{(i)}\,.$$

Spectral relaxations

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

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z}\,. \tag{11}$$

Here:

$$\mathcal{A}_k^\top z = \sum_{i\in[m_k]} z_i \mathbf{A}_k^{(i)}\,.$$

Notice that:

Spectral relaxations

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

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z}\,. \tag{11}$$

Here:

$$\mathcal{A}_k^\top z = \textstyle\sum_{i\in[m_k]} z_i \mathbf{A}_k^{(i)}\,.$$

Notice that:

- $\mathbf{A}_k^{(i)}$ and $\mathbf{C}_k$ are very sparse and have size $s_k = \binom{n+k}{n}$;

Spectral relaxations

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

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z}\,. \tag{11}$$

Here:

$$\mathcal{A}_k^\top z = \sum_{i\in[m_k]} z_i \mathbf{A}_k^{(i)}\,.$$

Notice that:

- $\mathbf{A}_k^{(i)}$ and $\mathbf{C}_k$ 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:

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z} \,. \tag{11}$$

Here:

$$\mathcal{A}_k^\top z = \sum_{i \in [m_k]} z_i \mathbf{A}_k^{(i)} \,.$$

Notice that:

- $\mathbf{A}_k^{(i)}$ and $\mathbf{C}_k$ are very sparse and have size $s_k = \binom{n+k}{n}$;
- The number of equality constraints: $m_k = \mathcal{O}\left(\binom{n+k}{n}^2\right)$
- $\mathbf{b}_k = [0, \ldots, 0, 1]^\top$.

Spectral relaxations

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

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z}\,. \tag{11}$$

Here:

$$\mathcal{A}_k^\top z = \textstyle\sum_{i\in[m_k]} z_i \mathbf{A}_k^{(i)}\,.$$

Notice that:

- $\mathbf{A}_k^{(i)}$ and $\mathbf{C}_k$ are very sparse and have size $s_k = \binom{n+k}{n}$;
- The number of equality constraints: $m_k = \mathcal{O}\left(\binom{n+k}{n}^2\right)$
- $\mathbf{b}_k = [0, \ldots, 0, 1]^\top$.

$m_k \geq 200 \times s_k$ when $n \geq 30$ and $k \geq 2$
$\Rightarrow$ not suitable for spectral bundle method [Helmberg et al., 2014];

SpectralPOP Algorithm

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:** τ_k and $\mathbf{x}^\star$

SpectralPOP Algorithm

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

1. Compute the optimal value and an optimal solution $\mathbf{z}^\star$ of

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z} \tag{12}$$

$\triangleright$ Limited Memory Bundle Method [Haarala et al., 2007];

SpectralPOP Algorithm

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

1. Compute the optimal value and an optimal solution $\mathbf{z}^\star$ of

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z} \qquad (12)$$

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

2. Compute uniform eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$ corresponding to the largest eigenvalue of $\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}^\star$;

SpectralPOP Algorithm

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

1. Compute the optimal value and an optimal solution $\mathbf{z}^\star$ of

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z} \tag{12}$$

   $\triangleright$ Limited Memory Bundle Method [Haarala et al., 2007];

2. Compute uniform eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_r$ corresponding to the largest eigenvalue of $\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}^\star$;
3. Compute $\xi^\star \in \operatorname{argmin}_{\xi \in \Delta^{(r-1)}} \left\| \mathbf{b}_k - \mathcal{A}_k \left(\sum_{j\in[r]} \xi_j \mathbf{u}_j \mathbf{u}_j^\top \right) \right\|_2^2$, $\Delta^{(r-1)}$: the standard $(r-1)$-simplex;

SpectralPOP Algorithm

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

1. Compute the optimal value and an optimal solution $\mathbf{z}^\star$ of

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z} \qquad (12)$$

$\triangleright$ Limited Memory Bundle Method [Haarala et al., 2007];

2. Compute uniform eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$ corresponding to the largest eigenvalue of $\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}^\star$;
3. Compute $\xi^\star \in \operatorname{argmin}_{\xi \in \Delta^{(r-1)}} \left\| \mathbf{b}_k - \mathcal{A}_k \left(\sum_{j\in[r]} \xi_j \mathbf{u}_j \mathbf{u}_j^\top \right) \right\|_2^2$, $\Delta^{(r-1)}$: the standard $(r-1)$-simplex;
4. $\mathbf{X}^\star \leftarrow \sum_{j\in[r]} \xi_j^\star \mathbf{u}_j \mathbf{u}_j^\top$ and $\mathbf{M}_k(\mathbf{y}^\star) \leftarrow 2^k \mathbf{P}_k^{-1} \mathbf{X}^\star \mathbf{P}_k^{-1}$;

SpectralPOP Algorithm

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

1. Compute the optimal value and an optimal solution $\mathbf{z}^\star$ of

$$-\tau_k = \inf_{\mathbf{z}} \lambda_{\max}(\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}) + \mathbf{b}_k^\top \mathbf{z} \tag{12}$$

   ▷ Limited Memory Bundle Method [Haarala et al., 2007];
2. Compute uniform eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$ corresponding to the largest eigenvalue of $\mathbf{C}_k - \mathcal{A}_k^\top \mathbf{z}^\star$;
3. Compute $\xi^\star \in \operatorname{argmin}_{\xi \in \Delta^{(r-1)}} \left\| \mathbf{b}_k - \mathcal{A}_k \left(\sum_{j \in [r]} \xi_j \mathbf{u}_j \mathbf{u}_j^\top \right) \right\|_2^2$, $\Delta^{(r-1)}$: the standard $(r-1)$-simplex;
4. $\mathbf{X}^\star \leftarrow \sum_{j \in [r]} \xi_j^\star \mathbf{u}_j \mathbf{u}_j^\top$ and $\mathbf{M}_k(\mathbf{y}^\star) \leftarrow 2^k \mathbf{P}_k^{-1} \mathbf{X}^\star \mathbf{P}_k^{-1}$;
5. Extract $\mathbf{x}^\star$ from $\mathbf{y}^\star$ ▷ [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

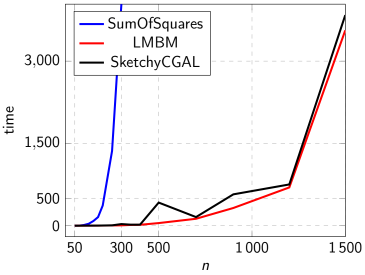
LMBM: Limited Memory Bundle Method

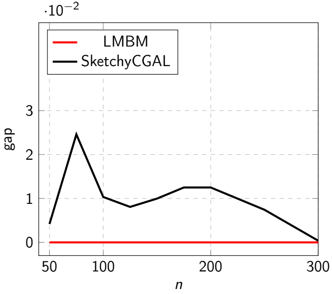
Spectral relaxations for quadratic forms ($k = 1$)

matrix size $s_k = n + 1$ 2 equality constraints

gap: the relative optimality gap w.r.t. SumOfSquares

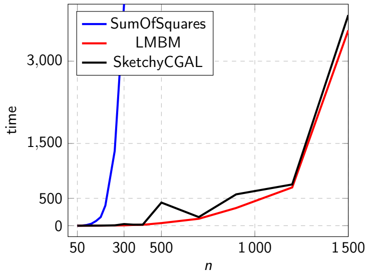
LMBM: Limited Memory Bundle Method

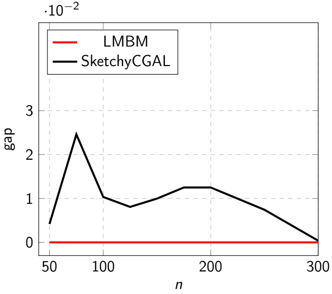




Spectral relaxations for quadratic forms ($k = 1$)

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





- **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 ($k = 2$)

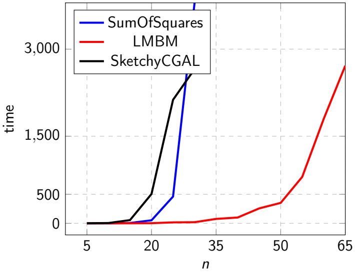
size of matrix: $s_k = \binom{n+k}{n}$

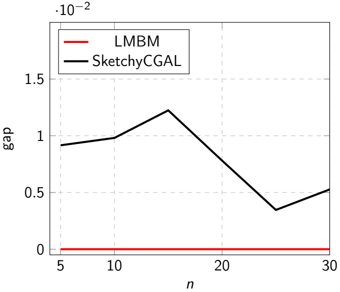
number of equality constraints: $m_k = \mathcal{O}\left(\binom{n+k}{n}^2\right) \geq 200 \times s_k$

Spectral relaxations for QCQP ($k = 2$)

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

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

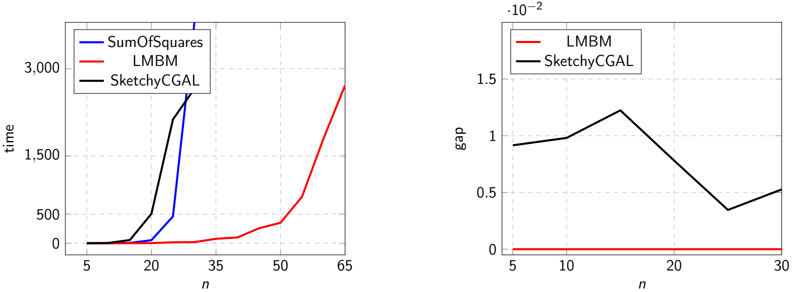




Spectral relaxations for QCQP ($k = 2$)

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

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

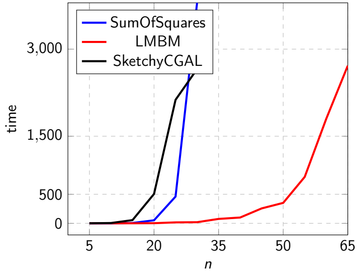


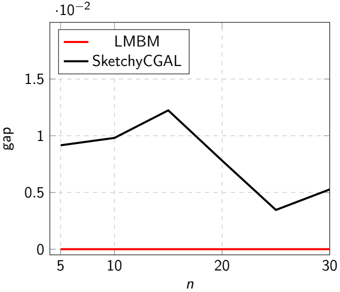
- If $n = 65$, then $s_k = 2211$ and $m_k = 1618453$.

Spectral relaxations for QCQP ($k = 2$)

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

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





- If $n = 65$, then $s_k = 2211$ and $m_k = 1618453$.
- **SpectralPOP** (LMBM):
  - provide the optimal value with high accuracy (gap $< 10^{-5}$)
  - up to more than 150 times faster than **SumOfSquares** (Mosek) when $n \geq 25$.

Summary and Future works

Summary and Future works

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

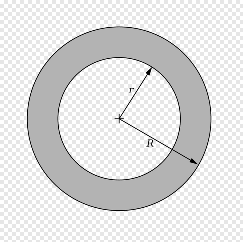
  SpectralPOP (LMBM) is *cheaper, faster, but maintains the same accuracy* as **SumOfSquares**.

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: $r \le \|\mathbf{x}(T)\|_2^2 \le R$

  - exploiting CTP for sparse POPs

End

Thank you for your attention!

SpectralPOP: https://github.com/maihoanganh/SpectralPOP
Homepage: https://sites.google.com/view/hoanganhmai

Haarala, N., Miettinen, K., and Mäkelä, M. M. (2007).
Globally convergent limited memory bundle method for large-scale nonsmooth optimization.
Mathematical Programming, 109(1):181–205.

Helmberg, C., Overton, M. L., and Rendl, F. (2014).
The spectral bundle method with second-order information.
Optimization Methods and Software, 29(4):855–876.

Helmberg, C. and Rendl, F. (2000).
A spectral bundle method for semidefinite programming.
SIAM Journal on Optimization, 10(3):673–696.

Henrion, D. and Lasserre, J.-B. (2005).
Detecting global optimality and extracting solutions in gloptipoly.
In *Positive polynomials in control*, pages 293–310. Springer.

Josz, C. and Henrion, D. (2016).
Strong duality in Lasserre's hierarchy for polynomial optimization.
Optimization Letters, 10(1):3–10.

Lasserre, J. B. (2001).
Global optimization with polynomials and the problem of moments.
SIAM Journal on optimization, 11(3):796–817.

Lasserre, J.-B. (2006).
Convergent SDP-Relaxations in Polynomial Optimization with Sparsity.
SIAM Journal on Optimization, 17(3):822–843.

Lasserre, J. B. (2015).
An introduction to polynomial and semi-algebraic optimization, volume 52.
Cambridge University Press.

Schweighofer, M. (2004).
On the complexity of Schmüdgen's positivstellensatz.
Journal of Complexity, 20(4):529–543.

Waki, H., Kim, S., Kojima, M., and Muramatsu, M. (2006).
Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity.
SIAM Journal on Optimization, 17(1):218–242.

Wang, J., Magron, V., and Lasserre, J.-B. (2019).
TSSOS: A Moment-SOS hierarchy that exploits term sparsity.
arXiv preprint arXiv:1912.08899.

Yurtsever, A., Tropp, J. A., Fercoq, O., Udell, M., and Cevher, V. (2019).
Scalable semidefinite programming.
arXiv preprint arXiv:1912.02949.