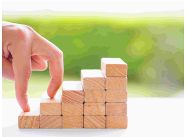
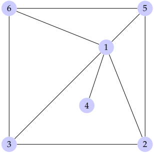
Exploiting sparsity & symmetries in polynomial optimization

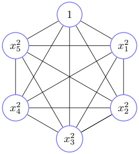
Victor Magron
LAAS CNRS

Lectures on polynomial optimization

University of Murcia 19 May 2022







What is a sparse/symmetric POP?

Looks like a regular polynomial optimization problem (POP):

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

What is a sparse/symmetric POP?

Looks like a regular polynomial optimization problem (POP):

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

But the input data f, g_j are "SPARSE" or "SYMMETRIC"!

What is a sparse/symmetric POP?

Looks like a regular polynomial optimization problem (POP):

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

But the input data f, g_j are "SPARSE" or "SYMMETRIC"!

Correlative sparsity: few products between each variable and the others in f, g_j

What is a sparse/symmetric POP?

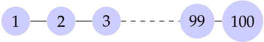
Looks like a regular polynomial optimization problem (POP):

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

But the input data f, g_j are "SPARSE" or "SYMMETRIC"!

Correlative sparsity: few products between each variable and the others in f, g_j

$\rightsquigarrow f(\mathbf{x}) = x_1 x_2 + x_2 x_3 + \ldots x_{99} x_{100}$



What is a sparse/symmetric POP?

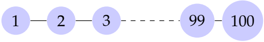
Looks like a regular polynomial optimization problem (POP):

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

But the input data f, g_j are “SPARSE” or “SYMMETRIC”!

Correlative sparsity: few products between each variable and the others in f, g_j

$\rightsquigarrow f(\mathbf{x}) = x_1 x_2 + x_2 x_3 + \ldots x_{99} x_{100}$



Term sparsity: few terms in f, g_j

What is a sparse/symmetric POP?

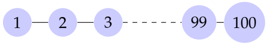
Looks like a regular polynomial optimization problem (POP):

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

But the input data f, g_j are “SPARSE” or “SYMMETRIC”!

Correlative sparsity: few products between each variable and the others in f, g_j

$\rightsquigarrow f(\mathbf{x}) = x_1x_2 + x_2x_3 + \ldots x_{99}x_{100}$



Term sparsity: few terms in f, g_j

$\rightsquigarrow f(\mathbf{x}) = x_1^{99}x_2 + x_1x_2^{100}$

What is a sparse/symmetric POP?

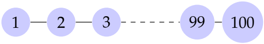
Looks like a regular polynomial optimization problem (POP):

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

💡 But the input data f, g_j are "SPARSE" or "SYMMETRIC"!

Correlative sparsity: few products between each variable and the others in f, g_j

$\rightsquigarrow f(\mathbf{x}) = x_1 x_2 + x_2 x_3 + \ldots x_{99} x_{100}$



Term sparsity: few terms in f, g_j

$\rightsquigarrow f(\mathbf{x}) = x_1^{99} x_2 + x_1 x_2^{100}$

Symmetry under a subgroup of $\mathrm{GL}(n)$:

What is a sparse/symmetric POP?

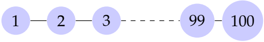
Looks like a regular polynomial optimization problem (POP):

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

💡 But the input data f, g_j are "SPARSE" or "SYMMETRIC"!

Correlative sparsity: few products between each variable and the others in f, g_j

$\rightsquigarrow f(\mathbf{x}) = x_1 x_2 + x_2 x_3 + \ldots x_{99} x_{100}$

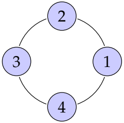


Term sparsity: few terms in f, g_j

$\rightsquigarrow f(\mathbf{x}) = x_1^{99} x_2 + x_1 x_2^{100}$

Symmetry under a subgroup of GL(n):

$\rightsquigarrow f(\mathbf{x}) = x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_1$



Where do we find sparse/symmetric POPs?

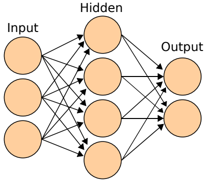
Where do we find sparse/symmetric POPs?

Everywhere (almost)!

Where do we find sparse/symmetric POPs?

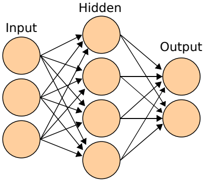
Everywhere (almost)!

Deep learning
$\rightsquigarrow$ robustness, computer vision

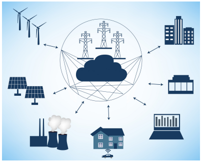


Where do we find sparse/symmetric POPs?

Everywhere (almost)!



Deep learning
⇝ robustness, computer vision



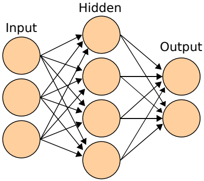
Power systems
⇝ AC optimal power-flow, stability

Where do we find sparse/symmetric POPs?

Everywhere (almost)!

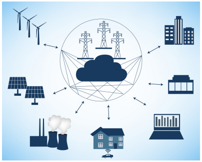
Deep learning

⇝ robustness, computer vision



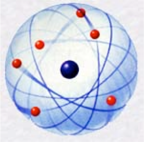
Power systems

⇝ AC optimal power-flow, stability



Quantum Systems

⇝ condensed matter, entanglement



The Moment-SOS Hierarchy for POP

Correlative sparsity

Term sparsity

Symmetries

The Moment-SOS Hierarchy for POP

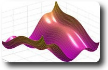
NP-hard NON CONVEX Problem $f_{\min} = \inf f(\mathbf{x})$

Theory

(Primal)

$\inf \quad \int f \, d\mu$

with μ proba $\Rightarrow$



INFINITE LP

(Dual)

$\sup \quad \lambda$

$\Leftarrow$ with $f - \lambda \geqslant 0$

The Moment-SOS Hierarchy for POP

NP-hard NON CONVEX Problem $f_{\min} = \inf f(\mathbf{x})$

Practice

(Primal **Relaxation**)

moments $\int \mathbf{x}^{\alpha} \, d\mu$

finite number $\Rightarrow$

SDP

(Dual **Strengthening**)

$f - \lambda =$ **sum of squares**

$\Leftarrow$ **fixed** degree

LASSERRE'S HIERARCHY of **CONVEX PROBLEMS** $\uparrow f^*$
[Lasserre '01]

degree d & n vars $\Longrightarrow \binom{n+2d}{n}$ **SDP** VARIABLES

The Moment-SOS Hierarchy for POP

NP-hard NON CONVEX Problem $f_{\min} = \inf f(\mathbf{x})$

Practice

(Primal **Relaxation**)

moments $\int \mathbf{x}^\alpha \, d\mu$

(Dual **Strengthening**)

$f - \lambda =$ **sum of squares**

finite number $\Rightarrow$ **SDP** $\Leftarrow$ **fixed** degree

LASSERRE'S HIERARCHY of **CONVEX PROBLEMS** $\uparrow f^*$
[Lasserre '01]

degree d & n vars $\Longrightarrow \binom{n+2d}{n}$ **SDP** VARIABLES

HOW TO OVERCOME THIS NO-FREE LUNCH RULE?

The Moment-SOS Hierarchy for POP

NP-hard NON CONVEX Problem $f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$

- space $\mathcal{M}_+(\mathbf{X})$ of probability measures supported on $\mathbf{X}$
- quadratic module $\mathcal{Q}(\mathbf{X}) = \left\{ \sigma_0 + \sum_j \sigma_j g_j, \text{ with } \sigma_j \text{ SOS} \right\}$

Infinite-dimensional linear programs (LP)

(Primal)		(Dual)
$\inf \int_{\mathbf{X}} f \, d\mu$	$=$	$\sup \lambda$
s.t. $\mu \in \mathcal{M}_+(\mathbf{X})$		s.t. $\lambda \in \mathbb{R}$
		$f - \lambda \in \mathcal{Q}(\mathbf{X})$

The Moment-SOS Hierarchy for POP

NP-hard NON CONVEX Problem $f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$

- Pseudo-moment sequences $\mathbf{y}$ up to order r
- Truncated quadratic module $\mathcal{Q}(\mathbf{X})_r$

Finite-dimensional semidefinite programs (SDP)

(Moment)		(SOS)
$\inf \sum_\alpha f_\alpha y_\alpha$	$=$	$\sup \lambda$
s.t. $\mathbf{M}_{r-r_j}(g_j \, \mathbf{y}) \succcurlyeq 0$		s.t. $\lambda \in \mathbb{R}$
$y_0 = 1$		$f - \lambda \in \mathcal{Q}(\mathbf{X})_r$

The Moment-SOS Hierarchy for POP

NP-hard NON CONVEX Problem $f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$

- Pseudo-moment sequences $\mathbf{y}$ up to order r
- Truncated quadratic module $\mathcal{Q}(\mathbf{X})_r$

Finite-dimensional semidefinite programs (SDP)

(Moment)		(SOS)
$\inf \sum_\alpha f_\alpha y_\alpha$	$=$	$\sup \lambda$
s.t. $\mathbf{M}_{r-r_j}(g_j \, \mathbf{y}) \succcurlyeq 0$		s.t. $\lambda \in \mathbb{R}$
$y_0 = 1$		$f - \lambda \in \mathcal{Q}(\mathbf{X})_r$

What is the primal-dual “SPARSE/SYMMETRIC” variant?

The Moment-SOS Hierarchy for POP

Correlative sparsity

Term sparsity

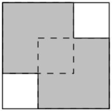
Symmetries

Sparse matrices

Symmetric matrices indexed by graph vertices

Sparse matrices

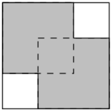
Symmetric matrices indexed by graph vertices



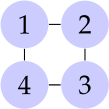
💡 no edge between 1 and 3 $\iff$ 0 entry in the $(1, 3)$ entry

Sparse matrices

Symmetric matrices indexed by graph vertices

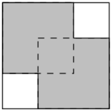


no edge between 1 and $3 \iff$ 0 entry in the $(1, 3)$ entry

cycle = 

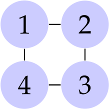
Sparse matrices

Symmetric matrices indexed by graph vertices



no edge between 1 and $3 \iff$ 0 entry in the $(1, 3)$ entry

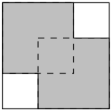
cycle =



chord = edge between two nonconsecutive vertices in a cycle

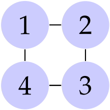
Sparse matrices

Symmetric matrices indexed by graph vertices



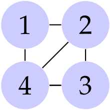
💡 no edge between 1 and $3 \iff$ 0 entry in the $(1, 3)$ entry

cycle =



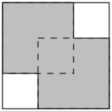
chord = edge between two nonconsecutive vertices in a cycle

chordal graph = all cycles of length $\geqslant$ 4 have at least one chord



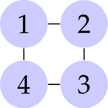
Sparse matrices

Symmetric matrices indexed by graph vertices



no edge between 1 and $3 \iff$ 0 entry in the $(1,3)$ entry

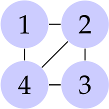
cycle =



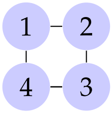
chord = edge between two nonconsecutive vertices in a cycle

chordal graph = all cycles of length $\geqslant$ 4 have at least one chord

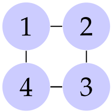
clique = a fully connected subset of vertices



Chordal extensions



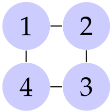
Chordal extensions



Fact

Any non-chordal graph can always be extended to a chordal graph, by adding appropriate edges

Chordal extensions

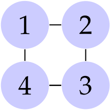


Fact

Any non-chordal graph can always be extended to a chordal graph, by adding appropriate edges

💡 Chordal extension is not unique!

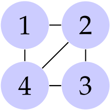
Chordal extensions

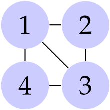


Fact

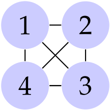
Any non-chordal graph can always be extended to a chordal graph, by adding appropriate edges

Chordal extension is not unique!





approximately minimal



maximal

Theorem [Gavril ’72, Vandenberghe & Andersen ’15]

The maximal cliques of a chordal graph can be enumerated in linear time in the number of nodes and edges.

Running intersection property (RIP)

RIP Theorem for chordal graphs [Blair & Peyton '93]

For a chordal graph with maximal cliques $I_1, \ldots, I_p$:

$$\text{(RIP)} \quad \boxed{\forall k < p \quad I_{k+1} \cap \bigcup_{j \leqslant k} I_j \subseteq I_i \quad \text{for some } i \leqslant k}$$

(possibly after reordering)

Running intersection property (RIP)

RIP Theorem for chordal graphs [Blair & Peyton '93]

For a chordal graph with maximal cliques $I_1, \ldots, I_p$:

$$\text{(RIP)} \quad \boxed{\forall k < p \quad I_{k+1} \cap \bigcup_{j \leqslant k} I_j \subseteq I_i \quad \text{for some } i \leqslant k}$$

(possibly after reordering)

RIP always holds for $p = 2$

Running intersection property (RIP)

RIP Theorem for chordal graphs [Blair & Peyton '93]

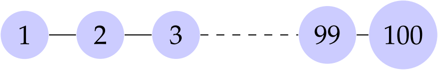
For a chordal graph with maximal cliques $I_1, \dots, I_p$:

$$\text{(RIP)} \quad \boxed{\forall k < p \quad I_{k+1} \cap \bigcup_{j \leqslant k} I_j \subseteq I_i \quad \text{for some } i \leqslant k}$$

(possibly after reordering)

- RIP always holds for $p = 2$

- RIP holds for chains



Running intersection property (RIP)

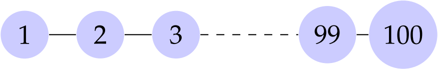
RIP Theorem for chordal graphs [Blair & Peyton '93]

For a chordal graph with maximal cliques $I_1, \ldots, I_p$:

$$\text{(RIP)} \quad \boxed{\forall k < p \quad I_{k+1} \cap \bigcup_{j \leqslant k} I_j \subseteq I_i \quad \text{for some } i \leqslant k}$$

(possibly after reordering)

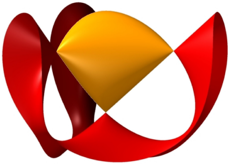
- RIP always holds for $p = 2$
- RIP holds for chains



- RIP holds for numerous applications!

Semidefinite Programming (SDP)

$$\min_{\mathbf{y}} \quad \mathbf{c}^\top \mathbf{y}$$
$$\text{s.t.} \quad \sum_i \mathbf{F}_i \, y_i \succcurlyeq \mathbf{F}_0$$



Spectrahedron

- Linear cost $\mathbf{c}$
- Symmetric matrices $\mathbf{F}_0, \mathbf{F}_i$
- Linear matrix inequalities "$\mathbf{F} \succcurlyeq 0$"
  ($\mathbf{F}$ has nonnegative eigenvalues)

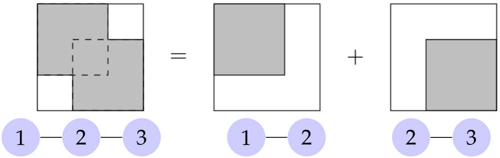
Sparse SDP matrices

Theorem [Griewank Toint '84, Agler et al. '88]

Chordal graph G with n vertices & maximal cliques I_1, I_2

$Q_G \succcurlyeq 0$ with nonzero entries corresponding to edges of G

$\Longrightarrow \; Q_G = P_1{}^T Q_1 P_1 + P_2{}^T Q_2 P_2$ with $Q_k \succcurlyeq 0$ indexed by I_k



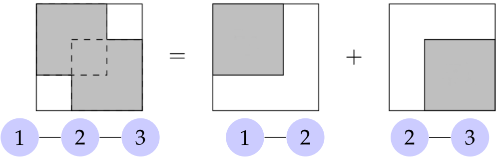
Sparse SDP matrices

Theorem [Griewank Toint '84, Agler et al. '88]

Chordal graph G with n vertices & maximal cliques I_1, I_2

$Q_G \succcurlyeq 0$ with nonzero entries corresponding to edges of G

$\Longrightarrow \; Q_G = P_1{}^T Q_1 P_1 + P_2{}^T Q_2 P_2$ with $Q_k \succcurlyeq 0$ indexed by I_k



What are P_1, P_2?

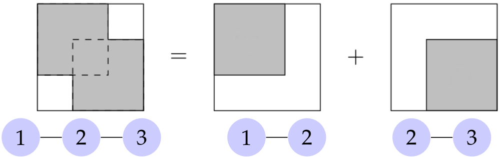
Sparse SDP matrices

Theorem [Griewank Toint '84, Agler et al. '88]

Chordal graph G with n vertices & maximal cliques I_1, I_2

$Q_G \succcurlyeq 0$ with nonzero entries corresponding to edges of G

$\Longrightarrow Q_G = {P_1}^T Q_1 P_1 + {P_2}^T Q_2 P_2$ with $Q_k \succcurlyeq 0$ indexed by I_k



What are P_1, P_2? $P_1 \in \mathbb{R}^{|I_1| \times n}$

$$P(i,j) = \begin{cases} 1 & \text{if } I(i) = j \\ 0 & \text{otherwise} \end{cases}$$

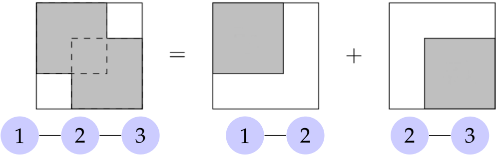
Sparse SDP matrices

Theorem [Griewank Toint '84, Agler et al. '88]

Chordal graph G with n vertices & maximal cliques I_1, I_2

$Q_G \succcurlyeq 0$ with nonzero entries corresponding to edges of G

$\implies Q_G = P_1{}^T Q_1 P_1 + P_2{}^T Q_2 P_2$ with $Q_k \succcurlyeq 0$ indexed by I_k



What are P_1, P_2? $P_1 \in \mathbb{R}^{|I_1| \times n}$

$$P(i,j) = \begin{cases} 1 & \text{if } I(i) = j \\ 0 & \text{otherwise} \end{cases}$$

$$I_1 = (1,2) \implies P_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

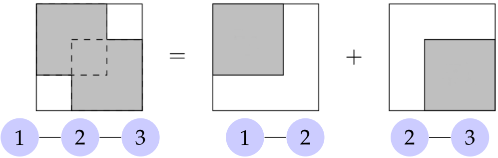
Sparse SDP matrices

Theorem [Griewank Toint '84, Agler et al. '88]

Chordal graph G with n vertices & maximal cliques I_1, I_2

$Q_G \succcurlyeq 0$ with nonzero entries corresponding to edges of G

$\implies Q_G = {P_1}^T Q_1 P_1 + {P_2}^T Q_2 P_2$ with $Q_k \succcurlyeq 0$ indexed by I_k



What are P_1, P_2? $P_1 \in \mathbb{R}^{|I_1| \times n}$

$$P(i,j) = \begin{cases} 1 & \text{if } I(i) = j \\ 0 & \text{otherwise} \end{cases}$$

$$I_1 = (1,2) \implies P_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

${P_1}^T Q_1 P_1$ inflates a $|I_1| \times |I_1|$ matrix Q_1 into a sparse $n \times n$ matrix

What is correlative sparsity?

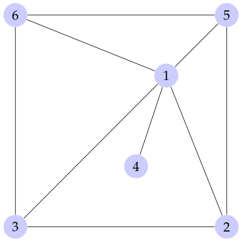
Exploit few links between **variables** [Lasserre, Waki et al. '06]

$$f(\mathbf{x}) = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

Correlative sparsity pattern (csp) graph G

Vertices = $\{1, \ldots, n\}$

$(i, j) \in$ Edges $\iff x_i x_j$ appears in f



What is correlative sparsity?

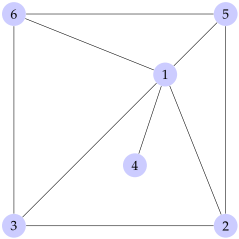
Exploit few links between **variables** [Lasserre, Waki et al. '06]

$$f(\mathbf{x}) = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

Correlative sparsity pattern (csp) graph G

Vertices = $\{1, \ldots, n\}$

$(i, j) \in$ Edges $\Longleftrightarrow x_i x_j$ appears in f

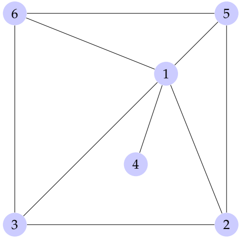


Similar construction with constraints $\mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\}$

What is correlative sparsity?

$$f(\mathbf{x}) = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

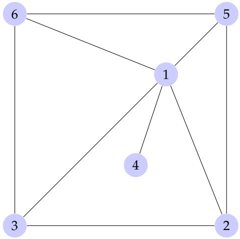
Chordal graph after adding edge $(3, 5)$



What is correlative sparsity?

$$f(\mathbf{x}) = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

Chordal graph after adding edge $(3, 5)$



maximal cliques $I_1 = \{1, 4\}$ $I_2 = \{1, 2, 3, 5\}$ $I_3 = \{1, 3, 5, 6\}$

$f = f_1 + f_2 + f_3$ where f_k involves **only** variables in I_k

💡 Let us index moment matrices and SOS with the cliques I_k

A sparse variant of Putinar's Positivstellensatz

Convergence of the Moment-SOS hierarchy is based on:

Theorem [Putinar '93] Positivstellensatz

If $\mathbf{X}$ contains a ball constraint $N - \sum_i x_i^2 \geqslant 0$ then

$$f > 0 \text{ on } \mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0\} \implies f = \sigma_0 + \sum_j \sigma_j g_j \text{ with } \sigma_j \text{ SOS}$$

A sparse variant of Putinar's Positivstellensatz

Convergence of the Moment-SOS hierarchy is based on:

Theorem [Putinar '93] Positivstellensatz

If $\mathbf{X}$ contains a ball constraint $N - \sum_i x_i^2 \geqslant 0$ then

$$f > 0 \text{ on } \mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0\} \implies f = \sigma_0 + \sum_j \sigma_j g_j \text{ with } \sigma_j \text{ SOS}$$

Theorem: Sparse Putinar's representation [Lasserre '06]

$f = \sum_k f_k$, f_k depends on $\mathbf{x}(I_k)$

$f > 0$ on $\mathbf{X}$

Each g_j depends on some I_k

RIP holds for (I_k)

ball constraints for each $\mathbf{x}(I_k)$

$\implies$

$$\boxed{f = \sum_k (\sigma_{0k} + \sum_{j \in J_k} \sigma_j g_j)}$$

SOS σ_{0k} "sees" vars in I_k

σ_j "sees" vars from g_j

A first key message

💡 **SUMS OF SQUARES PRESERVE SPARSITY** 💡

Sparse moment matrices

For each subset I_k, submatrix of $\mathbf{M}_r(\mathbf{y})$ corresponding of rows & columns indexed by monomials in $\mathbf{x}(I_k)$

Sparse moment matrices

For each subset I_k, submatrix of $\mathbf{M}_r(\mathbf{y})$ corresponding of rows & columns indexed by monomials in $\mathbf{x}(I_k)$

$f(\mathbf{x}) = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$

Sparse moment matrices

For each subset I_k, submatrix of $\mathbf{M}_r(\mathbf{y})$ corresponding of rows & columns indexed by monomials in $\mathbf{x}(I_k)$

$$f(\mathbf{x}) = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

$I_1 = \{1, 4\} \implies$ monomials in x_1, x_4

$$\mathbf{M}_1(\mathbf{y}, I_1) = \begin{pmatrix} 1 & \mid & y_{1,0,0,0,0,0} & y_{0,0,0,1,0,0} \\ & - & - & - \\ y_{1,0,0,0,0,0} & \mid & y_{2,0,0,0,0,0} & y_{1,0,0,1,0,0} \\ y_{0,0,0,1,0,0} & \mid & y_{1,0,0,1,0,0} & y_{0,0,0,2,0,0} \end{pmatrix}$$

Sparse moment matrices

For each subset I_k, submatrix of $\mathbf{M}_r(\mathbf{y})$ corresponding of rows & columns indexed by monomials in $\mathbf{x}(I_k)$

$$f(\mathbf{x}) = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

$I_1 = \{1,4\} \implies$ monomials in x_1, x_4

$$\mathbf{M}_1(\mathbf{y}, I_1) = \begin{pmatrix} 1 & | & y_{1,0,0,0,0,0} & y_{0,0,0,1,0,0} \\ & - & - & - \\ y_{1,0,0,0,0,0} & | & y_{2,0,0,0,0,0} & y_{1,0,0,1,0,0} \\ y_{0,0,0,1,0,0} & | & y_{1,0,0,1,0,0} & y_{0,0,0,2,0,0} \end{pmatrix}$$

💡 same for each localizing matrix $\mathbf{M}_r(g_j\mathbf{y})$

Sparse primal-dual Moment-SOS hierarchy

$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ with $\mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0\}$

Dense Moment-SOS hierarchy

(Moment)			(SOS)	
inf	$\sum_\alpha f_\alpha y_\alpha$	$=$	sup	λ
s.t.	$\mathbf{M}_r(\mathbf{y}) \succcurlyeq 0$		s.t.	$\lambda \in \mathbb{R}$
	$\mathbf{M}_{r-r_j}(g_j\, \mathbf{y}) \succcurlyeq 0$			$f - \lambda = \sigma_0 + \sum_j \sigma_j g_j$
	$y_0 = 1$			

Sparse primal-dual Moment-SOS hierarchy

$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ with $\mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0\}$

$f = \sum_k f_k$, with f_k depends on $\mathbf{x}(I_k)$

Each g_j depends on some I_k

Sparse Moment-SOS hierarchy

(Moment) (SOS)

$$\begin{aligned}
\inf \quad & \sum_\alpha f_\alpha y_\alpha & = \quad & \sup \quad \lambda \\
\text{s.t.} \quad & \mathbf{M}_r(\mathbf{y}, I_k) \succcurlyeq 0 & & \text{s.t.} \quad \lambda \in \mathbb{R} \\
& \mathbf{M}_{r-r_j}(g_j\, \mathbf{y}, I_k) \succcurlyeq 0 & & \qquad f - \lambda = \sum_k \Big(\sigma_{k0} + \sum_{j \in J_k} \sigma_j g_j\Big) \\
& y_0 = 1 & &
\end{aligned}$$

Sparse primal-dual Moment-SOS hierarchy

$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ with $\mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0\}$
$f = \sum_k f_k$, with f_k depends on $\mathbf{x}(I_k)$
Each g_j depends on some I_k

Sparse Moment-SOS hierarchy

(Moment) (SOS)

$$\begin{aligned} \inf \quad & \sum_\alpha f_\alpha \, y_\alpha & = \quad & \sup \quad \lambda \\ \text{s.t.} \quad & \mathbf{M}_r(\mathbf{y}, I_k) \succcurlyeq 0 & & \text{s.t.} \quad \lambda \in \mathbb{R} \\ & \mathbf{M}_{r-r_j}(g_j \, \mathbf{y}, I_k) \succcurlyeq 0 & & f - \lambda = \sum_k \Big(\sigma_{k0} + \sum_{j \in J_k} \sigma_j g_j\Big) \\ & y_0 = 1 & & \end{aligned}$$

RIP holds for (I_k) + ball constraints for each $\mathbf{x}(I_k) \implies$ Primal and dual optimal value converge to $f_{\min}$ by sparse Putinar

Computational cost

$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ with $\mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0, j \leqslant m\}$
$\tau = \max\{|I_1|, \ldots, |I_p|\}$

Sparse Moment-SOS hierarchy

(Moment) = (SOS)

$$\begin{aligned} \inf \quad & \sum_\alpha f_\alpha \, y_\alpha \\ \text{s.t.} \quad & \mathbf{M}_r(\mathbf{y}, I_k) \succcurlyeq 0 \\ & \mathbf{M}_{r-r_j}(g_j \, \mathbf{y}, I_k) \succcurlyeq 0 \\ & y_0 = 1 \end{aligned} \quad = \quad \begin{aligned} \sup \quad & \lambda \\ \text{s.t.} \quad & \lambda \in \mathbb{R} \\ & f - \lambda = \sum_k (\sigma_{k0} + \sum_{j \in J_k} \sigma_j g_j) \end{aligned}$$

Computational cost

$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ with $\mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0, j \leqslant m\}$

$\tau = \max\{|I_1|, \ldots, |I_p|\}$

Sparse Moment-SOS hierarchy

$$
\begin{array}{llcll}
 & \text{(Moment)} & & & \text{(SOS)} \\
\inf & \sum_\alpha f_\alpha \, y_\alpha & = & \sup & \lambda \\
\text{s.t.} & \mathbf{M}_r(\mathbf{y}, I_k) \succcurlyeq 0 & & \text{s.t.} & \lambda \in \mathbb{R} \\
 & \mathbf{M}_{r-r_j}(g_j \, \mathbf{y}, I_k) \succcurlyeq 0 & & & f - \lambda = \sum_k \Big(\sigma_{k0} + \sum_{j \in J_k} \sigma_j g_j\Big) \\
 & y_0 = 1 & & &
\end{array}
$$

$(m + p)$ SOS in at most τ vars of degree $\leqslant 2r$

Computational cost

$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ with $\mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0, j \leqslant m\}$
$\tau = \max\{|I_1|, \ldots, |I_p|\}$

Sparse Moment-SOS hierarchy

(Moment)		(SOS)
$\inf \quad \sum_\alpha f_\alpha \, y_\alpha$	$=$	$\sup \quad \lambda$
s.t. $\mathbf{M}_r(\mathbf{y}, I_k) \succcurlyeq 0$		s.t. $\lambda \in \mathbb{R}$
$\mathbf{M}_{r-r_j}(g_j \, \mathbf{y}, I_k) \succcurlyeq 0$		$f - \lambda = \sum_k (\sigma_{k0} + \sum_{j \in J_k} \sigma_j g_j)$
$y_0 = 1$		

$(m+p)$ SOS in at most τ vars of degree $\leqslant 2r$

💡 $(m+p) \, \mathcal{O}(r^{2\tau})$ SDP vars

Computational cost

$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ with $\mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0, j \leqslant m\}$
$\tau = \max\{|I_1|, \ldots, |I_p|\}$

Sparse Moment-SOS hierarchy

(Moment) = (SOS)

$$\begin{aligned} \inf \quad & \sum_\alpha f_\alpha y_\alpha \\ \text{s.t.} \quad & \mathbf{M}_r(\mathbf{y}, I_k) \succcurlyeq 0 \\ & \mathbf{M}_{r-r_j}(g_j\, \mathbf{y}, I_k) \succcurlyeq 0 \\ & y_0 = 1 \end{aligned} \quad = \quad \begin{aligned} \sup \quad & \lambda \\ \text{s.t.} \quad & \lambda \in \mathbb{R} \\ & f - \lambda = \sum_k \Big(\sigma_{k0} + \sum_{j \in J_k} \sigma_j g_j\Big) \end{aligned}$$

$(m + p)$ SOS in at most τ vars of degree $\leqslant 2r$

$(m + p)\, \mathcal{O}(r^{2\tau})$ SDP vars vs $(m + 1)\, \mathcal{O}(r^{2n})$ in the dense SDP

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1 x_2 + x_3 x_4$

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1 x_2 + x_3 x_4$
Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1 x_2 (1 + e_1) + x_3 x_4 (1 + e_2)](1 + e_3)$

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1 x_2 + x_3 x_4$
Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1 x_2 (1 + e_1) + x_3 x_4 (1 + e_2)](1 + e_3)$
$\mathbf{x} \in \mathbf{X}, \quad | e_i | \leqslant 2^{-\delta} \quad \delta = 24$ (single) or 53 (double)

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1 x_2 + x_3 x_4$
Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1 x_2 (1 + e_1) + x_3 x_4 (1 + e_2)](1 + e_3)$
$\mathbf{x} \in \mathbf{X}$, $\quad | e_i | \leqslant 2^{-\delta} \quad \delta = 24$ (single) or 53 (double)

1: Error $f(\mathbf{x}) - \hat{f}(\mathbf{x}, \mathbf{e}) = \ell(\mathbf{x}, \mathbf{e}) + h(\mathbf{x}, \mathbf{e})$, ℓ **linear** in $\mathbf{e}$

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1 x_2 + x_3 x_4$
Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1 x_2 (1 + e_1) + x_3 x_4 (1 + e_2)](1 + e_3)$
$\mathbf{x} \in \mathbf{X}$, $\quad | e_i | \leqslant 2^{-\delta} \quad \delta = 24$ (single) or 53 (double)

1: Error $f(\mathbf{x}) - \hat{f}(\mathbf{x}, \mathbf{e}) = \ell(\mathbf{x}, \mathbf{e}) + h(\mathbf{x}, \mathbf{e})$, ℓ **linear** in $\mathbf{e}$

2: Bound $h(\mathbf{x}, \mathbf{e})$ with interval arithmetic

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1 x_2 + x_3 x_4$
Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1 x_2 (1 + e_1) + x_3 x_4 (1 + e_2)](1 + e_3)$
$\mathbf{x} \in \mathbf{X}$, $\quad | e_i | \leqslant 2^{-\delta} \quad \delta = 24$ (single) or 53 (double)

1: Error $f(\mathbf{x}) - \hat{f}(\mathbf{x}, \mathbf{e}) = \ell(\mathbf{x}, \mathbf{e}) + h(\mathbf{x}, \mathbf{e})$, ℓ **linear** in $\mathbf{e}$

2: Bound $h(\mathbf{x}, \mathbf{e})$ with interval arithmetic

3: Bound $\ell(\mathbf{x}, \mathbf{e})$ with **SPARSE SUMS OF SQUARES**

$I_k \to \{\mathbf{x}, e_k\} \implies \boxed{m(n+1)^{2r} \text{ instead of } (n+m)^{2r}}$ SDP vars

Application to roundoff errors

$$f = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

$$\mathbf{x} \in [4.00, 6.36]^6 \,, \quad \mathbf{e} \in [-\epsilon, \epsilon]^{15} \,, \quad \epsilon = 2^{-53}$$

Dense SDP: $\binom{6+15+4}{6+15} = 12650$ variables $\rightsquigarrow$ **Out of memory**

Application to roundoff errors

$$f = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

$$\mathbf{x} \in [4.00, 6.36]^6 , \quad \mathbf{e} \in [-\epsilon, \epsilon]^{15} , \quad \epsilon = 2^{-53}$$

Dense SDP: $\binom{6+15+4}{6+15} = 12650$ variables $\rightsquigarrow$ **Out of memory**

Sparse SDP `Real2Float` tool: $15\binom{6+1+4}{6+1} = 4950 \rightsquigarrow 759\epsilon$

Application to roundoff errors

$$f = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

$$\mathbf{x} \in [4.00, 6.36]^6 , \quad \mathbf{e} \in [-\epsilon, \epsilon]^{15} , \quad \epsilon = 2^{-53}$$

Dense SDP: $\binom{6+15+4}{6+15} = 12650$ variables $\rightsquigarrow$ **Out of memory**

Sparse SDP `Real2Float` tool: $15\binom{6+1+4}{6+1} = 4950 \rightsquigarrow 759\epsilon$

Interval arithmetic: 922ϵ (10 $\times$ less CPU)

Application to roundoff errors

$$f = x_2 x_5 + x_3 x_6 - x_2 x_3 - x_5 x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

$$\mathbf{x} \in [4.00, 6.36]^6\,, \quad \mathbf{e} \in [-\epsilon, \epsilon]^{15}\,, \quad \epsilon = 2^{-53}$$

Dense SDP: $\binom{6+15+4}{6+15} = 12650$ variables $\leadsto$ **Out of memory**

Sparse SDP `Real2Float` tool: $15\binom{6+1+4}{6+1} = 4950 \leadsto 759\epsilon$

Interval arithmetic: 922ϵ (10 $\times$ less CPU)

Symbolic Taylor `FPTaylor` tool: 721ϵ (21 $\times$ more CPU)

Application to roundoff errors

$$f = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$$

$$\mathbf{x} \in [4.00, 6.36]^6\,, \quad \mathbf{e} \in [-\epsilon, \epsilon]^{15}\,, \quad \epsilon = 2^{-53}$$

Dense SDP: $\binom{6+15+4}{6+15} = 12650$ variables $\rightsquigarrow$ **Out of memory**

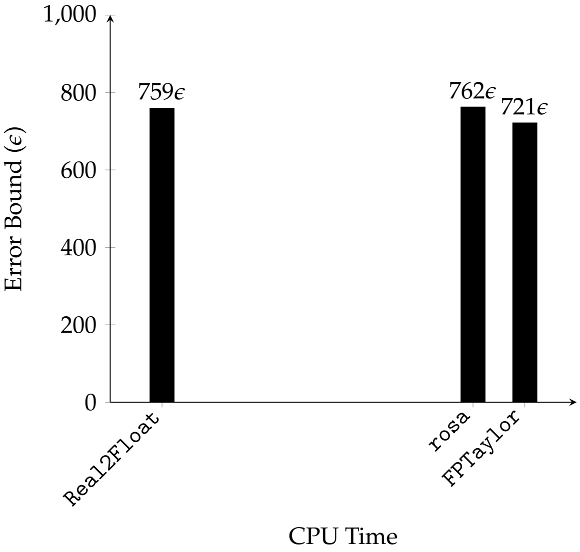
Sparse SDP `Real2Float` tool: $15\binom{6+1+4}{6+1} = 4950 \rightsquigarrow 759\epsilon$

Interval arithmetic: 922ϵ (10 $\times$ less CPU)

Symbolic Taylor `FPTaylor` tool: 721ϵ (21 $\times$ more CPU)

SMT-based `rosa` tool: 762ϵ (19 $\times$ more CPU)

Application to roundoff errors



Application to noncommutative optimization

Self-adjoint noncommutative variables x_i, y_j

$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - y_1 - 2y_1 - y_2$

with $x_1 x_2 \neq x_2 x_1$, **involution** $(x_1 y_3)^\star = y_3 x_1$

Application to noncommutative optimization

Self-adjoint noncommutative variables x_i, y_j

$$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - y_1 - 2y_1 - y_2$$

with $x_1 x_2 \neq x_2 x_1$, **involution** $(x_1 y_3)^\star = y_3 x_1$

Constraints $\mathbf{X} = \{(x, y) : x_i, y_j \succcurlyeq 0, x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

Application to noncommutative optimization

Self-adjoint noncommutative variables x_i, y_j

$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - y_1 - 2y_1 - y_2$

with $x_1 x_2 \neq x_2 x_1$, **involution** $(x_1 y_3)^\star = y_3 x_1$

Constraints $\mathbf{X} = \{(x, y) : x_i, y_j \succcurlyeq 0, x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

MINIMAL EIGENVALUE OPTIMIZATION

$$\lambda_{\min} = \inf \{ \langle f(x, y)\mathbf{v}, \mathbf{v} \rangle : (x, y) \in \mathbf{X}, \|\mathbf{v}\| = 1 \}$$

Application to noncommutative optimization

Self-adjoint noncommutative variables x_i, y_j

$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - y_1 - 2y_1 - y_2$

with $x_1 x_2 \neq x_2 x_1$, **involution** $(x_1 y_3)^\star = y_3 x_1$

Constraints $\mathbf{X} = \{(x, y) : x_i, y_j \succcurlyeq 0, x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

MINIMAL EIGENVALUE OPTIMIZATION

$$\begin{aligned} \lambda_{\min} &= \inf \{\langle f(x,y)\mathbf{v}, \mathbf{v}\rangle : (x,y) \in \mathbf{X}, \|\mathbf{v}\| = 1\} \\ &= \sup \quad \lambda \\ &\quad\ \text{s.t.} \quad f(x,y) - \lambda \mathbf{I} \succcurlyeq 0, \quad \forall (x,y) \in \mathbf{X} \end{aligned}$$

Application to noncommutative optimization

Ball constraint $N - \sum_i x_i^2 \succcurlyeq 0$ in $\mathbf{X}$

Theorem: NC Putinar's representation [Helton & McCullough '02]

$f \succ 0$ on $\mathbf{X} \implies \boxed{f = \sum_i s_i^\star s_i + \sum_j \sum_i t_{ji}^\star g_j t_{ji}}$ with s_i, $t_{ji} \in \mathbb{R}\langle \underline{x} \rangle$

Application to noncommutative optimization

Ball constraint $N - \sum_i x_i^2 \succcurlyeq 0$ in $\mathbf{X}$

Theorem: NC Putinar's representation [Helton & McCullough '02]

$f \succ 0$ on $\mathbf{X} \implies \boxed{f = \sum_i s_i^\star s_i + \sum_j \sum_i t_{ji}^\star g_j t_{ji}}$ with s_i, $t_{ji} \in \mathbb{R}\langle \underline{x} \rangle$

NC variant of Lasserre's Hierarchy for $\lambda_{\min}$:

replace "$f - \lambda \mathbf{I} \succcurlyeq 0$ on $\mathbf{X}$" by $f - \lambda \mathbf{I} = \sum_i s_i^\star s_i + \sum_j \sum_i t_{ji}^\star g_j t_{ji}$
with s_i, t_{ji} of **bounded** degrees = SDP

Application to noncommutative optimization

Self-adjoint noncommutative (NC) variables $\underline{x} = (x_1, \dots, x_n)$

Theorem [Helton & McCullough '02]

$f \succcurlyeq 0 \Leftrightarrow f$ SOS (all positive polynomials are sums of squares)

Application to noncommutative optimization

Self-adjoint noncommutative (NC) variables $\underline{x} = (x_1, \ldots, x_n)$

Theorem [Helton & McCullough '02]

$f \succcurlyeq 0 \Leftrightarrow f$ SOS (all positive polynomials are sums of squares)

BAD NEWS: there is **no** sparse analog! [Klep Magron Povh '21]

sparse f SOS $\not\Rightarrow f$ is a sparse SOS

Application to noncommutative optimization

Self-adjoint noncommutative (NC) variables $\underline{x} = (x_1, \ldots, x_n)$

Theorem [Helton & McCullough '02]

$f \succcurlyeq 0 \Leftrightarrow f$ SOS (all positive polynomials are sums of squares)

BAD NEWS: there is **no** sparse analog! [Klep Magron Povh '21]

sparse f SOS $\not\Rightarrow f$ is a sparse SOS Take $f = (x_1 + x_2 + x_3)^2$

Application to noncommutative optimization

Self-adjoint noncommutative (NC) variables $\underline{x} = (x_1, \dots, x_n)$

Theorem [Helton & McCullough '02]

$f \succcurlyeq 0 \Leftrightarrow f$ SOS (all positive polynomials are sums of squares)

BAD NEWS: there is **no** sparse analog! [Klep Magron Povh '21]

sparse f SOS $\not\Rightarrow f$ is a sparse SOS Take $f = (x_1 + x_2 + x_3)^2$

GOOD NEWS: there is an NC analog of the sparse Putinar's Positivstellensatz!

Application to noncommutative optimization

Self-adjoint noncommutative (NC) variables $\underline{x} = (x_1, \ldots, x_n)$

Theorem [Helton & McCullough '02]

$f \succcurlyeq 0 \Leftrightarrow f$ SOS (all positive polynomials are sums of squares)

BAD NEWS: there is **no** sparse analog! [Klep Magron Povh '21]

sparse f SOS $\not\Rightarrow f$ is a sparse SOS Take $f = (x_1 + x_2 + x_3)^2$

GOOD NEWS: there is an NC analog of the sparse Putinar's Positivstellensatz! Based on GNS construction & **amalgamation** [Blackadar '78, Voiculescu '85]

Application to noncommutative optimization

Self-adjoint noncommutative (NC) variables $\underline{x} = (x_1, \ldots, x_n)$

Theorem [Helton & McCullough '02]

$f \succcurlyeq 0 \Leftrightarrow f$ SOS (all positive polynomials are sums of squares)

BAD NEWS: there is **no** sparse analog! [Klep Magron Povh '21]

sparse f SOS $\not\Rightarrow f$ is a sparse SOS

Take $f = (x_1 + x_2 + x_3)^2$

GOOD NEWS: there is an NC analog of the sparse Putinar's Positivstellensatz! Based on GNS construction & **amalgamation** [Blackadar '78, Voiculescu '85]

Theorem: Sparse NC Positivstellensatz [Klep Magron Povh '21]

$f = \sum_k f_k$, f_k depends on $\mathbf{x}(I_k)$

$f > 0$ on $\mathbf{X}$

Each g_j depends on some I_k

RIP holds for (I_k)

ball constraints for each $\mathbf{x}(I_k)$

$\Longrightarrow$

$$f = \sum_{k,i} (s_{ki}^\star s_{ki} + \sum_{j \in J_k} t_{ji}{}^\star g_j t_{ji})$$

s_{ki} "sees" vars in I_k

t_{ji} "sees" vars from g_j

Application to noncommutative optimization

$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

$$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - x_1 - 2y_1 - y_2$$

Maximal violation levels $\to$ **upper bounds** on $\lambda_{\max}$ of f on $\{(x,y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

Application to noncommutative optimization

$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - x_1 - 2y_1 - y_2$

Maximal violation levels $\to$ **upper bounds** on $\lambda_{\max}$ of f on $\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

💡 $I_k \to \{x_1, x_2, x_3, y_k\}$

Application to noncommutative optimization

$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

$$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - x_1 - 2y_1 - y_2$$

Maximal violation levels $\to$ **upper bounds** on $\lambda_{\max}$ of f on $\{(x,y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

💡 $I_k \to \{x_1, x_2, x_3, y_k\}$

level	sparse	dense [Pál & Vértesi '18]
2	0.2550008	0.2509397

Application to noncommutative optimization

$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

$$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - x_1 - 2y_1 - y_2$$

Maximal violation levels $\rightarrow$ **upper bounds** on $\lambda_{\max}$ of f on $\{(x,y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

$I_k \rightarrow \{x_1, x_2, x_3, y_k\}$

level	sparse	dense [Pál & Vértesi ’18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756

Application to noncommutative optimization

$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

$$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - x_1 - 2y_1 - y_2$$

Maximal violation levels → **upper bounds** on $\lambda_{\max}$ of f on $\{(x,y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

💡 $I_k \to \{x_1, x_2, x_3, y_k\}$

level	sparse	dense [Pál & Vértesi '18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3'		0.2508754 (1 day)

Application to noncommutative optimization

$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

$$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - x_1 - 2y_1 - y_2$$

Maximal violation levels $\to$ **upper bounds** on $\lambda_{\max}$ of f on $\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

💡 $I_k \to \{x_1, x_2, x_3, y_k\}$

level	sparse	dense [Pál & Vértesi '18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3'		0.250875**4** (1 day)
4	0.2508917	

Application to noncommutative optimization

$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

$$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - x_1 - 2y_1 - y_2$$

Maximal violation levels → **upper bounds** on $\lambda_{\max}$ of f on $\{(x,y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

$I_k \rightarrow \{x_1, x_2, x_3, y_k\}$

level	sparse	dense [Pál & Vértesi '18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3'		0.2508754 (1 day)
4	0.2508917	
5	0.2508763	

Application to noncommutative optimization

$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

$$f = x_1(y_1 + y_2 + y_3) + x_2(y_1 + y_2 - y_3) + x_3(y_1 - y_2) - x_1 - 2y_1 - y_2$$

Maximal violation levels $\to$ **upper bounds** on $\lambda_{\max}$ of f on $\{(x,y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

$I_k \to \{x_1, x_2, x_3, y_k\}$

level	sparse	dense [Pál & Vértesi '18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3'		0.2508754 (1 day)
4	0.2508917	
5	0.2508763	
6	0.2508753977180	(1 hour)

PERFORMANCE  VS  ACCURACY

More and more applications!

Sparse positive definite forms [Mai, Lasserre & Magron '21]

More and more applications!

Sparse positive definite forms [Mai, Lasserre & Magron ’21]

Robust Geometric Perception [Yang & Carlone ’20]

More and more applications!

Sparse positive definite forms [Mai, Lasserre & Magron ’21]

Robust Geometric Perception [Yang & Carlone ’20]

Polynomial matrix inequalities [Zheng & Fantuzzi ’20]

More and more applications!

Sparse positive definite forms [Mai, Lasserre & Magron ’21]

Robust Geometric Perception [Yang & Carlone ’20]

Polynomial matrix inequalities [Zheng & Fantuzzi ’20]

Region of attraction [Tacchi et al., Schlosser et al. ’21]

More and more applications!

Sparse positive definite forms [Mai, Lasserre & Magron ’21]

Robust Geometric Perception [Yang & Carlone ’20]

Polynomial matrix inequalities [Zheng & Fantuzzi ’20]

Region of attraction [Tacchi et al., Schlosser et al. ’21]

Volume computation [Tacchi et al. ’21]

More and more applications!

Sparse positive definite forms [Mai, Lasserre & Magron '21]

Robust Geometric Perception [Yang & Carlone '20]

Polynomial matrix inequalities [Zheng & Fantuzzi '20]

Region of attraction [Tacchi et al., Schlosser et al. '21]

Volume computation [Tacchi et al. '21]

Robustness of implicit deep networks [Chen et al. '21]

The Moment-SOS Hierarchy for POP

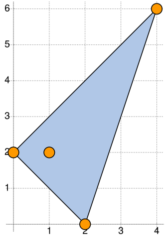
Correlative sparsity

Term sparsity

Symmetries

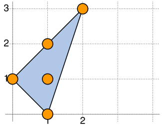
Term sparsity via Newton polytope

$f = 4x_1^4 x_2^6 + x_1^2 - x_1 x_2^2 + x_2^2$
spt(f) = $\{(4,6),(2,0),(1,2),(0,2)\}$



Newton polytope $\mathscr{B} = \operatorname{conv}(\text{spt}(f))$

Squares in SOS decomposition $\subseteq \frac{\mathscr{B}}{2} \cap \mathbb{N}^n$
[Reznick '78]



$$f = \begin{pmatrix} x_1 & x_2 & x_1x_2 & x_1x_2^2 & x_1^2x_2^3 \end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix} x_1 \\ x_2 \\ x_1x_2 \\ x_1x_2^2 \\ x_1^2x_2^3 \end{pmatrix}$$

Term sparsity: the unconstrained case

$$f = x_1^2 - 2x_1x_2 + 3x_2^2 - 2x_1^2x_2 + 2x_1^2x_2^2 - 2x_2x_3 + 6x_3^2 + 18x_2^2x_3 - 54x_2x_3^2 + 142x_2^2x_3^2$$

[Reznick '78] $\rightarrow f = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_2x_3 \end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$

$\rightsquigarrow \frac{6\times 7}{2} = 21$ "unknown" entries in Q

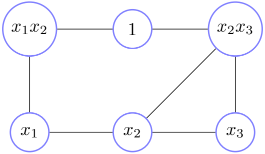
Term sparsity: the unconstrained case

$$f = x_1^2 - 2x_1x_2 + 3x_2^2 - 2x_1^2x_2 + 2x_1^2x_2^2 - 2x_2x_3 + 6x_3^2 + 18x_2^2x_3 - 54x_2x_3^2 + 142x_2^2x_3^2$$

[Reznick '78] $\to f = \begin{pmatrix}1 & x_1 & x_2 & x_3 & x_1x_2 & x_2x_3\end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix}1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3\end{pmatrix}$

$\rightsquigarrow \frac{6\times 7}{2} = 21$ "unknown" entries in Q

Term sparsity pattern graph G



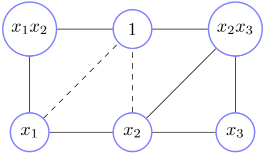
Term sparsity: the unconstrained case

$$f = x_1^2 - 2x_1x_2 + 3x_2^2 - 2x_1^2x_2 + 2x_1^2x_2^2 - 2x_2x_3 + 6x_3^2 + 18x_2^2x_3 - 54x_2x_3^2 + 142x_2^2x_3^2$$

[Reznick '78] $\to f = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_2x_3 \end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$

$\rightsquigarrow \frac{6\times 7}{2} = 21$ "unknown" entries in Q

💡 **Term** sparsity pattern graph G
+ chordal extension G'



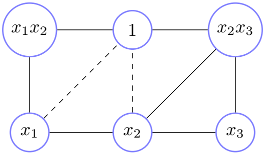
Term sparsity: the unconstrained case

$$f = x_1^2 - 2x_1x_2 + 3x_2^2 - 2x_1^2x_2 + 2x_1^2x_2^2 - 2x_2x_3 + 6x_3^2 + 18x_2^2x_3 - 54x_2x_3^2 + 142x_2^2x_3^2$$

[Reznick '78] $\rightarrow f = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_2x_3 \end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$

$\rightsquigarrow \frac{6\times 7}{2} = 21$ "unknown" entries in Q

Term sparsity pattern graph G
+ chordal extension G'



Replace Q by $Q_{G'}$ with nonzero entries at edges of G'

$\rightsquigarrow$ 6 + 9 = 15 "unknown" entries in $Q_{G'}$

Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes V = monomials of degree $\leqslant r$

Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes V = monomials of degree $\leqslant r$
Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \operatorname{supp} f \bigcup \operatorname{supp} g_j \bigcup_{|\alpha| \leqslant r} 2\alpha$$

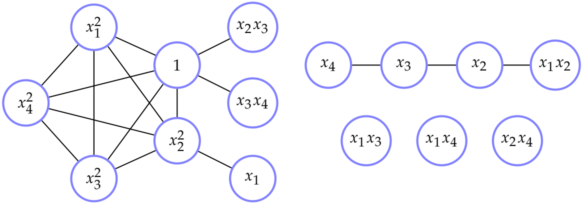
Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes V = monomials of degree $\leqslant r$
Edges E with

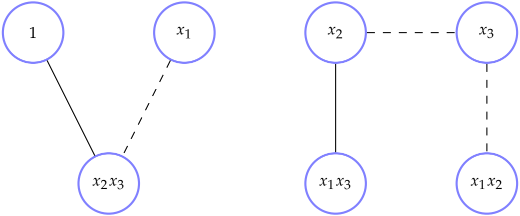
$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \operatorname{supp} f \bigcup \operatorname{supp} g_j \bigcup_{|\alpha| \leqslant r} 2\alpha$$

An example with $r = 2$
$f = x_1^4 + x_1 x_2^2 + x_2 x_3 + x_3^2 x_4^2$
$g_1 = 1 - x_1^2 - x_2^2 - x_3^2 \quad g_2 = 1 - x_3 x_4$



Term sparsity: support extension

$\alpha' + \beta' = \alpha + \beta$ and $(\alpha, \beta) \in E \Rightarrow (\alpha', \beta') \in E$



Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes V = monomials of degree $\leqslant r$

Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes V = monomials of degree $\leqslant r$
Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \operatorname{supp} f \bigcup \operatorname{supp} g_j \bigcup_{|\alpha| \leqslant r} 2\alpha$$

$\rightsquigarrow$ support extension

Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes V = monomials of degree $\leqslant r$
Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \operatorname{supp} f \bigcup \operatorname{supp} g_j \bigcup_{|\alpha| \leqslant r} 2\alpha$$

$\rightsquigarrow$ support extension $\rightsquigarrow$ chordal extension G'

Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes V = monomials of degree $\leqslant r$
Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \operatorname{supp} f \bigcup \operatorname{supp} g_j \bigcup_{|\alpha| \leqslant r} 2\alpha$$

$\rightsquigarrow$ support extension $\rightsquigarrow$ chordal extension G'

By iteratively performing support extension & chordal extension

$$G^{(1)} = G' \subseteq \cdots \subseteq G^{(s)} \subseteq G^{(s+1)} \subseteq \cdots$$

💡 Two-level hierarchy of lower bounds for $f_{\min}$, indexed by sparse order s and relaxation order r

Term sparsity: primal moment relaxations

Let G' be a chordal extension of G with maximal cliques (C_i)

$$C_i \longmapsto \mathbf{M}_{C_i}(\mathbf{y})$$

Term sparsity: primal moment relaxations

Let G' be a chordal extension of G with maximal cliques (C_i)

$$C_i \longmapsto \mathbf{M}_{C_i}(\mathbf{y})$$

In the moment relaxation,

$$\mathbf{M}_r(\mathbf{y}) \succcurlyeq 0 \longrightarrow \mathbf{M}_{C_i}(\mathbf{y}) \succcurlyeq 0$$

Term sparsity: primal moment relaxations

Let G' be a chordal extension of G with maximal cliques (C_i)

$$C_i \longmapsto \mathbf{M}_{C_i}(\mathbf{y})$$

In the moment relaxation,

$$\mathbf{M}_r(\mathbf{y}) \succcurlyeq 0 \longrightarrow \mathbf{M}_{C_i}(\mathbf{y}) \succcurlyeq 0$$

Similarly for the localizing matrices $\mathbf{M}_{r-r_j}(g_j\,\mathbf{y})$

Term sparsity: primal moment relaxations

Let G' be a chordal extension of G with maximal cliques (C_i)

$$C_i \longmapsto \mathbf{M}_{C_i}(\mathbf{y})$$

In the moment relaxation,

$$\mathbf{M}_r(\mathbf{y}) \succcurlyeq 0 \longrightarrow \mathbf{M}_{C_i}(\mathbf{y}) \succcurlyeq 0$$

Similarly for the localizing matrices $\mathbf{M}_{r-r_j}(g_j\, \mathbf{y})$

Each constraint $G_j \rightsquigarrow G_j^{(s)} \rightsquigarrow$ cliques $C_{j,i}^{(s)}$

Term sparsity: primal moment relaxations

Let $C_{j,i}^{(s)}$ be the maximal cliques of $G_j^{(s)}$. For each $s \geq 1$

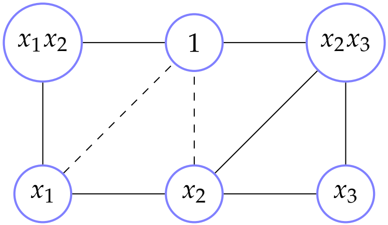
$$
\begin{aligned}
f_{\text{ts}}^{r,s} = \quad \inf \quad & \sum_\alpha f_\alpha y_\alpha \\
\text{s.t.} \quad & \mathbf{M}_{C_{0,i}^{(s)}}(\mathbf{y}) \succcurlyeq 0 \\
& \mathbf{M}_{C_{j,i}^{(s)}}(g_j\, \mathbf{y}) \succcurlyeq 0 \\
& y_0 = 1
\end{aligned}
$$

dual yields the TSSOS hierarchy

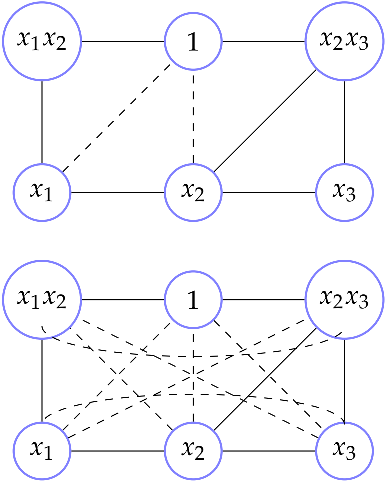
A two-level hierarchy of lower bounds

$$
\begin{array}{cccc}
f_{\text{ts}}^{r_{\min},1} \leq & f_{\text{ts}}^{r_{\min},2} \leq & \cdots \leq & f^{r_{\min}} \\
\wedge| & \wedge| & & \wedge| \\
f_{\text{ts}}^{r_{\min}+1,1} \leq & f_{\text{ts}}^{r_{\min}+1,2} \leq & \cdots \leq & f^{r_{\min}+1} \\
\wedge| & \wedge| & & \wedge| \\
\vdots & \vdots & \vdots & \vdots \\
\wedge| & \wedge| & & \wedge| \\
f_{\text{ts}}^{r,1} \leq & f_{\text{ts}}^{r,2} \leq & \cdots \leq & f^{r} \\
\wedge| & \wedge| & & \wedge| \\
\vdots & \vdots & \vdots & \vdots
\end{array}
$$

Different choices of chordal extensions



Different choices of chordal extensions



Term sparsity: convergence guarantees

Theorem [Lasserre Magron Wang '21]

Fixing a sparse order s, the sequence $(f_{\text{ts}}^{r,s})_{r \geq r_{\min}}$ is monotonically nondecreasing.

Term sparsity: convergence guarantees

Theorem [Lasserre Magron Wang '21]

Fixing a sparse order s, the sequence $(f_{\text{ts}}^{r,s})_{r \geq r_{\min}}$ is monotonically nondecreasing.

💡 Fixing a relaxation order r, the sequence $(f_{\text{ts}}^{r,s})_{s \geq 1} \uparrow f^r$ in finitely many steps if the maximal chordal extension is used.

Term sparsity: convergence guarantees

Theorem [Lasserre Magron Wang '21]

Fixing a sparse order s, the sequence $(f_{\mathrm{ts}}^{r,s})_{r \geq r_{\min}}$ is monotonically nondecreasing.

💡 Fixing a relaxation order r, the sequence $(f_{\mathrm{ts}}^{r,s})_{s \geq 1} \uparrow f^r$ in finitely many steps if the maximal chordal extension is used.

💡 The block structures converge to the one determined by the sign symmetries if the maximal chordal extension is used.

Term sparsity: convergence guarantees

Theorem [Lasserre Magron Wang '21]

Fixing a sparse order s, the sequence $(f_{\mathrm{ts}}^{r,s})_{r \geq r_{\min}}$ is monotonically nondecreasing.

💡 Fixing a relaxation order r, the sequence $(f_{\mathrm{ts}}^{r,s})_{s \geq 1} \uparrow f^r$ in finitely many steps if the maximal chordal extension is used.

💡 The block structures converge to the one determined by the sign symmetries if the maximal chordal extension is used.

$f = 1 + x_1^2 x_2^4 + x_1^4 x_2^2 + x_1^4 x_2^4 - x_1 x_2^2 - 3 x_1^2 x_2^2$

Newton polytope $\rightsquigarrow \mathscr{B} = (1 \quad x_1 x_2 \quad x_1 x_2^2 \quad x_1^2 x_2 \quad x_1^2 x_2^2)$

Term sparsity: convergence guarantees

Theorem [Lasserre Magron Wang '21]

Fixing a sparse order s, the sequence $(f_{\text{ts}}^{r,s})_{r \geq r_{\min}}$ is monotonically nondecreasing.

Fixing a relaxation order r, the sequence $(f_{\text{ts}}^{r,s})_{s \geq 1} \uparrow f^r$ in finitely many steps if the maximal chordal extension is used.

The block structures converge to the one determined by the sign symmetries if the maximal chordal extension is used.

$f = 1 + x_1^2 x_2^4 + x_1^4 x_2^2 + x_1^4 x_2^4 - x_1 x_2^2 - 3 x_1^2 x_2^2$

Newton polytope $\rightsquigarrow \mathscr{B} = (1 \quad x_1 x_2 \quad x_1 x_2^2 \quad x_1^2 x_2 \quad x_1^2 x_2^2)$

$x_2 \mapsto -x_2$

Sign-symmetries blocks $(1 \quad x_1 x_2^2 \quad x_1^2 x_2^2) \quad (x_1 x_2 \quad x_1^2 x_2)$

TSSOS blocks $(1 \quad x_1 x_2^2 \quad x_1^2 x_2^2) \quad (x_1 x_2) \quad (x_1^2 x_2)$

A second key message

TSSOS preserves the block structure related to sign-symmetries

Combining correlative & term sparsity

1. Partition the variables w.r.t. the maximal cliques of the csp graph

Combining correlative & term sparsity

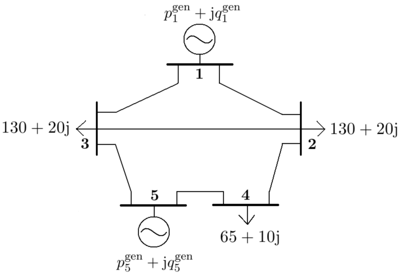
1. Partition the variables w.r.t. the maximal cliques of the csp graph
2. For each subsystem involving variables from one maximal clique, apply TSSOS

Combining correlative & term sparsity

1. Partition the variables w.r.t. the maximal cliques of the csp graph
2. For each subsystem involving variables from one maximal clique, apply TSSOS

💡 a two-level CS-TSSOS hierarchy of lower bounds for $f_{\min}$

Application to optimal power-flow



Optimal Powerflow $n \simeq 10^3$

[Josz et al. '18]

$$
\begin{cases}
\inf\limits_{V_i, S_s^g, S_{ij}} & \sum_{s\in G}\left(\mathbf{c}_{2s}(\Re(S_s^g))^2 + \mathbf{c}_{1s}\Re(S_s^g) + \mathbf{c}_{0s}\right) \\
\text{s.t.} & \angle V_{\text{ref}} = 0, \\
& \mathbf{S}_s^{gl} \leq S_s^g \leq \mathbf{S}_s^{gu} \ \forall s \in G, \quad \boldsymbol{v}_i^l \leq |V_i| \leq \boldsymbol{v}_i^u \ \forall i \in N \\
& \sum_{s\in G_i} S_s^g - \mathbf{S}_i^d - \mathbf{Y}_i^s |V_i|^2 = \sum_{(i,j)\in E_i \cup E_i^R} S_{ij}, \quad \forall i \in N \\
& S_{ij} = \left(\mathbf{Y}_{ij}^* - \mathbf{i}\frac{\mathbf{b}_{ij}^c}{2}\right)\frac{|V_i|^2}{|\mathbf{T}_{ij}|^2} - \mathbf{Y}_{ij}^* \frac{V_i V_j^*}{\mathbf{T}_{ij}}, \quad S_{ji} = \cdots \\
& |S_{ij}| \leq \mathbf{s}_{ij}^u, \ \boldsymbol{\theta}_{ij}^{\Delta l} \leq \angle(V_i V_j^*) \leq \boldsymbol{\theta}_{ij}^{\Delta u}, \quad \forall (i,j) \in E
\end{cases}
$$

Application to optimal power-flow

mb = the maximal size of blocks

m = number of constraints

n	m	CS ($r = 2$)			CS+TS ($r = 2, s = 1$)		
		mb	time (s)	gap	mb	time (s)	gap
114	315	66	5.59	0.39%	31	2.01	0.73%
348	1809	253	–	–	34	278	0.05%
766	3322	153	585	0.68%	44	33.9	0.77%
1112	4613	496	–	–	31	410	0.25%
4356	18257	378	–	–	27	934	0.51%
6698	29283	1326	–	–	76	1886	0.47%

Application to networked systems stability

Duffing oscillator Hamiltonian $V = \sum_{i=1}^{N} a_i (\frac{x_i^2}{2} - \frac{x_i^4}{4}) + \frac{1}{8} \sum_{i,k=1}^{N} b_{ik} (x_i - x_k)^4$

On which domain $V > 0$?

$$f = V - \sum_{i=1}^{N} \underbrace{\lambda_i}_{>0} x_i^2 (g - x_i^2) \geqslant 0$$

$\implies V > 0$ when $x_i^2 < g$

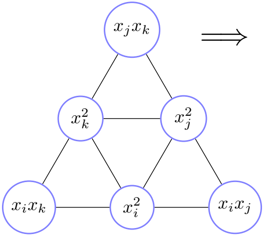
Application to networked systems stability

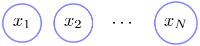
Duffing oscillator Hamiltonian $V = \sum_{i=1}^{N} a_i (\frac{x_i^2}{2} - \frac{x_i^4}{4}) + \frac{1}{8} \sum_{i,k=1}^{N} b_{ik} (x_i - x_k)^4$

On which domain $V > 0$?

$$f = V - \sum_{i=1}^{N} \underbrace{\lambda_i}_{>0} x_i^2 (g - x_i^2) \geqslant 0$$

$\implies$ $V > 0$ when $x_i^2 < g$





tsp graph G

Application to networked systems stability

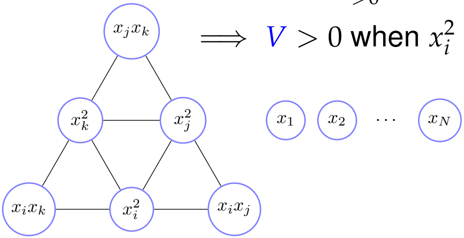
Duffing oscillator Hamiltonian $V = \sum_{i=1}^{N} a_i (\frac{x_i^2}{2} - \frac{x_i^4}{4}) + \frac{1}{8} \sum_{i,k=1}^{N} b_{ik} (x_i - x_k)^4$

On which domain $V > 0$?

$$f = V - \sum_{i=1}^{N} \underbrace{\lambda_i}_{>0} x_i^2 (g - x_i^2) \geqslant 0$$

$$\implies V > 0 \text{ when } x_i^2 < g$$

tsp graph G



$\rightsquigarrow \frac{N(N+1)}{2} + 6\binom{N}{2} + N$ "unknown" entries in $Q_G = 80$ for $N = 5$

Proof that $f \geqslant 0$ for $N = 50$ in ~ 1 second!

The Moment-SOS Hierarchy for POP

Correlative sparsity

Term sparsity

Symmetries

Primer on group representations

Let G be a finite group

1. A representation of G is a finite-dim vector space V with a homomorphism $\rho : G \to \mathrm{GL}(V)$, where $\mathrm{GL}(V)$ is the set of all invertible transformations of V

Primer on group representations

Let G be a finite group

1. A representation of G is a finite-dim vector space V with a homomorphism $\rho : G \to \mathrm{GL}(V)$, where $\mathrm{GL}(V)$ is the set of all invertible transformations of V
2. $\dim(V)$ is the degree of (V, ρ)

Primer on group representations

Let G be a finite group

1. A representation of G is a finite-dim vector space V with a homomorphism $\rho : G \to \mathrm{GL}(V)$, where $\mathrm{GL}(V)$ is the set of all invertible transformations of V
2. $\dim(V)$ is the degree of (V, ρ)
3. (V, ρ) isomorphic to (V', ρ') if there is an isomorphism $\theta : V \to V'$ s.t.

$$\rho'(g) = \theta \rho(g) \theta^{-1}, \forall g \in G$$

Primer on group representations

Let G be a finite group

1. A representation of G is a finite-dim vector space V with a homomorphism $\rho : G \to \mathrm{GL}(V)$, where $\mathrm{GL}(V)$ is the set of all invertible transformations of V
2. $\dim(V)$ is the degree of (V, ρ)
3. (V, ρ) isomorphic to (V', ρ') if there is an isomorphism $\theta : V \to V'$ s.t.

$$\rho'(g) = \theta \rho(g) \theta^{-1}, \forall g \in G$$

4. A basis of V gives a matrix representation of G, we identify G with a group $\mathbf{M}(G)$ of invertible matrices

Primer on group representations

V is a G-module if

1. $1 \cdot v = v$ and $g_1 \cdot (g_2 \cdot v) = (g_1 g_2) \cdot v$

Primer on group representations

V is a G-module if

1. $1 \cdot v = v$ and $g_1 \cdot (g_2 \cdot v) = (g_1 g_2) \cdot v$
2. $g \cdot (v_1 + v_2) = g \cdot v_1 + g \cdot v_2$ and $g \cdot (\lambda v) = \lambda g \cdot v$

Primer on group representations

V is a G-module if

1. $1 \cdot v = v$ and $g_1 \cdot (g_2 \cdot v) = (g_1 g_2) \cdot v$
2. $g \cdot (v_1 + v_2) = g \cdot v_1 + g \cdot v_2$ and $g \cdot (\lambda v) = \lambda g \cdot v$

$\rho : G \to \mathrm{GL}(V)$ with $\rho(g) = v \mapsto g \cdot v$ is a representation

Primer on group representations

V is a G-module if

1. $1 \cdot v = v$ and $g_1 \cdot (g_2 \cdot v) = (g_1 g_2) \cdot v$
2. $g \cdot (v_1 + v_2) = g \cdot v_1 + g \cdot v_2$ and $g \cdot (\lambda v) = \lambda g \cdot v$

$\rho : G \to \mathrm{GL}(V)$ with $\rho(g) = v \mapsto g \cdot v$ is a representation

1. $W \subseteq V$ is a G-submodule if $g \cdot w \in W$ for all $w \in W$ and $g \in G$
2. If V does not contain a non-trivial submodule then V is irreducible

Primer on group representations

V is a G-module if

1. $1 \cdot v = v$ and $g_1 \cdot (g_2 \cdot v) = (g_1 g_2) \cdot v$
2. $g \cdot (v_1 + v_2) = g \cdot v_1 + g \cdot v_2$ and $g \cdot (\lambda v) = \lambda g \cdot v$

$\rho : G \to \mathrm{GL}(V)$ with $\rho(g) = v \mapsto g \cdot v$ is a representation

1. $W \subseteq V$ is a G-submodule if $g \cdot w \in W$ for all $w \in W$ and $g \in G$
2. If V does not contain a non-trivial submodule then V is irreducible

Example: $G = S_2$ acting on the 2-dim vector space $V = \mathbb{R}e_1 \oplus \mathbb{R}e_2$ by permuting e_1 and e_2. Then V is reducible.

Primer on group representations

V is a G-module if

1. $1 \cdot v = v$ and $g_1 \cdot (g_2 \cdot v) = (g_1 g_2) \cdot v$
2. $g \cdot (v_1 + v_2) = g \cdot v_1 + g \cdot v_2$ and $g \cdot (\lambda v) = \lambda g \cdot v$

$\rho : G \to \mathrm{GL}(V)$ with $\rho(g) = v \mapsto g \cdot v$ is a representation

1. $W \subseteq V$ is a G-submodule if $g \cdot w \in W$ for all $w \in W$ and $g \in G$
2. If V does not contain a non-trivial submodule then V is irreducible

Example: $G = S_2$ acting on the 2-dim vector space $V = \mathbb{R}e_1 \oplus \mathbb{R}e_2$ by permuting e_1 and e_2. Then V is reducible. $W = \mathbb{R}(e_1 + e_2)$ is an S_2-submodule

Primer on group representations

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and G a finite group.

Theorem [Maschke]

If V is a finite-dim $\mathbb{K}$-vector space and a G-module then V is a direct sum of irreducible G-modules W_i

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

Primer on group representations

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and G a finite group.

Theorem [Maschke]

If V is a finite-dim $\mathbb{K}$-vector space and a G-module then V is a direct sum of irreducible G-modules W_i

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

ϕ is a G-homomorphism if $\phi(g \cdot v) = \phi(v)$ for all $g \in G$, $v \in V$.

Primer on group representations

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and G a finite group.

Theorem [Maschke]

If V is a finite-dim $\mathbb{K}$-vector space and a G-module then V is a direct sum of irreducible G-modules W_i

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

ϕ is a G-homomorphism if $\phi(g \cdot v) = \phi(v)$ for all $g \in G$, $v \in V$.

Theorem [Schur's lemma]

Let V, W be two irreducible G-modules. Then a G-homomorphism $\phi : V \to W$ is either zero or an isomorphism.

Primer on group representations

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and G a finite group.

Theorem [Maschke]

If V is a finite-dim $\mathbb{K}$-vector space and a G-module then V is a direct sum of irreducible G-modules W_i

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

ϕ is a G-homomorphism if $\phi(g \cdot v) = \phi(v)$ for all $g \in G$, $v \in V$.

Theorem [Schur's lemma]

Let V, W be two irreducible G-modules. Then a G-homomorphism $\phi : V \to W$ is either zero or an isomorphism.

💡 A G-homomorphism from V to V is equivalent to multiplication by a scalar.

Primer on group representations

Corollary

Let $V = m_1 W_1 \oplus \cdots \oplus m_k W_k$ be a complete decomposition of the representation V with $\dim W_i = d_i$. Then there is a basis of V such that the matrices of $\mathbf{M}(G)$ are of the form

$$\mathbf{M}(g) = \bigoplus_{l=1}^{k} \bigoplus_{j=1}^{m_i} \mathbf{M}^{(l)}(g)$$

where each $\mathbf{M}^{(l)}(G)$ represents W_l

Primer on group representations

Corollary

Let $V = m_1 W_1 \oplus \cdots \oplus m_k W_k$ be a complete decomposition of the representation V with $\dim W_i = d_i$. Then there is a basis of V such that the matrices of $\mathbf{M}(G)$ are of the form

$$\mathbf{M}(g) = \bigoplus_{l=1}^{k} \bigoplus_{j=1}^{m_i} \mathbf{M}^{(l)}(g)$$

where each $\mathbf{M}^{(l)}(G)$ represents W_l

Such a basis is called a **symmetry adapted basis**

Primer on group representations

Let $\rho : G \to \mathrm{GL}_n(\mathbb{K})$ and $\mathbf{Q} \in \mathbb{K}^{n \times n}$ with $\rho(g)\mathbf{Q} = \mathbf{Q}\rho(g)$ for all $g \in G$

Primer on group representations

Let $\rho : G \to \mathrm{GL}_n(\mathbb{K})$ and $\mathbf{Q} \in \mathbb{K}^{n \times n}$ with $\rho(g)\mathbf{Q} = \mathbf{Q}\rho(g)$ for all $g \in G$

Assume that $\rho = m_1\rho_1 \oplus \cdots \oplus m_k\rho_k$ with $d_i = \dim \rho_i$

Primer on group representations

Let $\rho : G \to \mathrm{GL}_n(\mathbb{K})$ and $\mathbf{Q} \in \mathbb{K}^{n\times n}$ with $\rho(g)\mathbf{Q} = \mathbf{Q}\rho(g)$ for all $g \in G$

Assume that $\rho = m_1\rho_1 \oplus \cdots \oplus m_k\rho_k$ with $d_i = \dim \rho_i$

💡 Use a symmetric adapted basis of $\mathbb{K}^n$ to block-diag $\mathbf{Q}$
$\implies \mathbf{N} = \mathbf{T}^{-1}\mathbf{Q}\mathbf{T}$ and

$$\begin{pmatrix} \mathbf{N}_1 & & 0 \\ & \ddots & \\ 0 & & \mathbf{N}_k \end{pmatrix} \quad \mathbf{N}_i = \begin{pmatrix} \mathbf{B}_i & & 0 \\ & \ddots & \\ 0 & & \mathbf{B}_i \end{pmatrix}$$

💡 Each column of $\mathbf{T}$ is an element of a symmetry adapted basis

💡 $\mathbf{B}_i$ has size m_i

A first key message

Whenever we have a linear group action on a vector space then

A NICE BASIS MAKES MATRICES SIMPLER

Symmetries in SDPs

$\mathrm{Sym}_n(\mathbb{K})$: Hermitian matrices

$$\begin{aligned} \inf_{\mathbf{Q}} \ & \langle \mathbf{C}, \mathbf{Q} \rangle \\ \text{s.t. } & \langle \mathbf{A}_i, \mathbf{Q} \rangle = f_i \\ & \mathbf{Q} \succcurlyeq 0 \, , \mathbf{Q} \in \mathrm{Sym}_n(\mathbb{K}) \end{aligned}$$

Symmetries in SDPs

$\mathrm{Sym}_n(\mathbb{K})$: Hermitian matrices

$$\inf_{\mathbf{Q}} \langle \mathbf{C}, \mathbf{Q} \rangle$$
$$\text{s.t. } \langle \mathbf{A}_i, \mathbf{Q} \rangle = f_i$$
$$\mathbf{Q} \succcurlyeq 0 \, , \mathbf{Q} \in \mathrm{Sym}_n(\mathbb{K})$$

Let us pick a representation $(\mathbb{K}^n, \rho)$ of G and an orthonormal basis for $\mathbb{K}^n$ w.r.t. a G-invariant inner product

The corresponding matrices are unitary: $\rho(g)\rho(g)^\star = \mathrm{Id}$

Symmetries in SDPs

$\mathrm{Sym}_n(\mathbb{K})$: Hermitian matrices

$$\begin{aligned} &\inf_{\mathbf{Q}} \langle \mathbf{C}, \mathbf{Q} \rangle \\ &\text{s.t. } \langle \mathbf{A}_i, \mathbf{Q} \rangle = f_i \\ &\quad \mathbf{Q} \succcurlyeq 0 \,, \mathbf{Q} \in \mathrm{Sym}_n(\mathbb{K}) \end{aligned}$$

Let us pick a representation $(\mathbb{K}^n, \rho)$ of G and an orthonormal basis for $\mathbb{K}^n$ w.r.t. a G-invariant inner product

💡 The corresponding matrices are unitary: $\rho(g)\rho(g)^\star = \mathrm{Id}$

💡 $\mathrm{Sym}_n(\mathbb{K})$ is a G-module via

$$\mathbf{Q}^g := \rho(g)\mathbf{Q}\rho(g)^\star$$

Symmetries in SDPs

$\mathrm{Sym}_n(\mathbb{K})$: Hermitian matrices

$$
\begin{aligned}
&\inf_{\mathbf{Q}} \ \langle \mathbf{C}, \mathbf{Q} \rangle \\
&\text{s.t. } \langle \mathbf{A}_i, \mathbf{Q} \rangle = f_i \\
&\qquad \mathbf{Q} \succcurlyeq 0 \,, \mathbf{Q} \in \mathrm{Sym}_n(\mathbb{K})
\end{aligned}
$$

Let us pick a representation $(\mathbb{K}^n, \rho)$ of G and an orthonormal basis for $\mathbb{K}^n$ w.r.t. a G-invariant inner product

💡 The corresponding matrices are unitary: $\rho(g)\rho(g)^\star = \mathrm{Id}$

💡 $\mathrm{Sym}_n(\mathbb{K})$ is a G-module via

$$\mathbf{Q}^g := \rho(g)\mathbf{Q}\rho(g)^\star$$

The above SDP is G-invariant if $\langle \mathbf{C}, \mathbf{Q} \rangle = \langle \mathbf{C}, \mathbf{Q}^g \rangle$ and $\langle \mathbf{A}_i, \mathbf{Q}^g \rangle = f_i$

Symmetries in SDPs

$$
\begin{aligned}
\inf_{\mathbf{Q}} \ & \langle \mathbf{C}, \mathbf{Q} \rangle \\
\text{s.t. } & \langle \mathbf{A}_i, \mathbf{Q} \rangle = f_i \\
& \mathbf{Q} = \mathbf{Q}^g \,, \forall g \in G \\
& \mathbf{Q} \succcurlyeq 0 \,, \mathbf{Q} \in \mathrm{Sym}_n(\mathbb{K})
\end{aligned}
$$

Symmetries in SDPs

$$\begin{aligned} &\inf_{\mathbf{Q}} \langle \mathbf{C}, \mathbf{Q} \rangle \\ &\text{s.t. } \langle \mathbf{A}_i, \mathbf{Q} \rangle = f_i \\ &\quad \mathbf{Q} = \mathbf{Q}^g \,, \forall g \in G \\ &\quad \mathbf{Q} \succcurlyeq 0 \,, \mathbf{Q} \in \mathrm{Sym}_n(\mathbb{K}) \end{aligned}$$

Theorem

The optimal value of the SDP is the same as the “dense” one if it is G-invariant.

Symmetries in SDPs

$$\inf_{\mathbf{Q}} \langle \mathbf{C}, \mathbf{Q} \rangle$$
$$\text{s.t. } \langle \mathbf{A}_i, \mathbf{Q} \rangle = f_i$$
$$\mathbf{Q} = \mathbf{Q}^g, \forall g \in G$$
$$\mathbf{Q} \succcurlyeq 0, \mathbf{Q} \in \mathrm{Sym}_n(\mathbb{K})$$

Theorem

The optimal value of the SDP is the same as the "dense" one if it is G-invariant.

Proof

Take a feasible $\mathbf{Q}$ and $g \in G$.
Since the feasible region is convex $\mathbf{Q}_G := \frac{1}{|G|} \sum_{g \in G} \mathbf{Q}^g$ is feasible for the "dense" SDP and $\langle \mathbf{C}, \mathbf{Q} \rangle = \langle \mathbf{C}, \mathbf{Q}_G \rangle$.

Symmetries in SDPs

💡 One can restrict to invariant matrices (i.e., the commutator)

Symmetries in SDPs

💡 One can restrict to invariant matrices (i.e., the commutator)

By Schur's lemma, we can find a basis that **block-diag** the matrices

Symmetries in SDPs

💡 One can restrict to invariant matrices (i.e., the commutator)

By Schur's lemma, we can find a basis that **block-diag** the matrices
Let $\mathbb{K}^n = W_1^1 \oplus \cdots \oplus W_{m_1}^1 \oplus \cdots \oplus W_{m_k}^k$ be an orthogonal decomposition into irreducibles, and choose an orthonormal basis $\{e_{l1}^j, \ldots, e_{ld_j}^j\}$ for each W_l^j
💡 orthonormal symmetry adapted basis $\mathbf{T}$

Symmetries in SDPs

💡 One can restrict to invariant matrices (i.e., the commutator)

By Schur's lemma, we can find a basis that **block-diag** the matrices

Let $\mathbb{K}^n = W_1^1 \oplus \cdots \oplus W_{m_1}^1 \oplus \cdots \oplus W_{m_k}^k$ be an orthogonal decomposition into irreducibles, and choose an orthonormal basis $\{e_{l1}^j, \ldots, e_{ld_j}^j\}$ for each W_l^j

💡 orthonormal symmetry adapted basis $\mathbf{T}$

$$\inf_{\mathbf{Q}_l} \sum_{l=1}^{k} d_l \langle \mathbf{C}_l, \mathbf{Q}_l \rangle$$

$$\text{s.t. } \langle \mathbf{A}_i, \mathbf{Q} \rangle = f_i \,, \quad \mathbf{T}^{-1}\mathbf{Q}\mathbf{T} = \operatorname{diag}\left(\mathbf{Q}_1, \ldots, \mathbf{Q}_k\right)$$

$$\mathbf{Q}_l \succcurlyeq 0 \,, \mathbf{Q}_l \in \mathrm{Sym}_{m_l}(\mathbb{K})$$

Symmetries in SDPs: an example

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} a & b & b \\ b & c_1 & d \\ b & d & c_2 \end{pmatrix}$$

Symmetries in SDPs: an example

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} a & b & b \\ b & c_1 & d \\ b & d & c_2 \end{pmatrix}$$

$\mathbf{C}, \mathbf{Q}$ invariant under S_2 permuting both the last 2 rows and columns

Symmetries in SDPs: an example

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} a & b & b \\ b & c_1 & d \\ b & d & c_2 \end{pmatrix}$$

$\mathbf{C}, \mathbf{Q}$ invariant under S_2 permuting both the last 2 rows and columns

💡 one can restrict to the fixed point subspace with $c_1 = c_2 = c$

Symmetries in SDPs: an example

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} a & b & b \\ b & c_1 & d \\ b & d & c_2 \end{pmatrix}$$

$\mathbf{C}, \mathbf{Q}$ invariant under S_2 permuting both the last 2 rows and columns

💡 one can restrict to the fixed point subspace with $c_1 = c_2 = c$

💡 S_2 has 2 irreducible representations of dimension $d_1 = d_2 = 1$, with multiplicities $m_1 = 2$, $m_2 = 1$

Symmetries in SDPs: an example

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} a & b & b \\ b & c_1 & d \\ b & d & c_2 \end{pmatrix}$$

$\mathbf{C}, \mathbf{Q}$ invariant under S_2 permuting both the last 2 rows and columns

💡 one can restrict to the fixed point subspace with $c_1 = c_2 = c$

💡 S_2 has 2 irreducible representations of dimension $d_1 = d_2 = 1$, with multiplicities $m_1 = 2$, $m_2 = 1$

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & \alpha \\ 0 & \alpha & -\alpha \end{pmatrix} \quad \alpha = \frac{1}{\sqrt{2}}$$

Symmetries in SDPs: an example

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} a & b & b \\ b & c_1 & d \\ b & d & c_2 \end{pmatrix}$$

$\mathbf{C}, \mathbf{Q}$ invariant under S_2 permuting both the last 2 rows and columns

💡 one can restrict to the fixed point subspace with $c_1 = c_2 = c$

💡 S_2 has 2 irreducible representations of dimension $d_1 = d_2 = 1$, with multiplicities $m_1 = 2$, $m_2 = 1$

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & \alpha \\ 0 & \alpha & -\alpha \end{pmatrix} \quad \alpha = \frac{1}{\sqrt{2}}$$

$$\mathbf{T}^{-1}\mathbf{C}\mathbf{T} = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix} \quad C_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad C_2 = 1$$

Symmetries in SDPs: an example

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} a & b & b \\ b & c_1 & d \\ b & d & c_2 \end{pmatrix}$$

$\mathbf{C}, \mathbf{Q}$ invariant under S_2 permuting both the last 2 rows and columns

- one can restrict to the fixed point subspace with $c_1 = c_2 = c$
- S_2 has 2 irreducible representations of dimension $d_1 = d_2 = 1$, with multiplicities $m_1 = 2$, $m_2 = 1$

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & \alpha \\ 0 & \alpha & -\alpha \end{pmatrix} \quad \alpha = \frac{1}{\sqrt{2}}$$

$$\mathbf{T}^{-1}\mathbf{C}\mathbf{T} = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix} \quad C_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad C_2 = 1$$

$$\mathbf{T}^{-1}\mathbf{Q}\mathbf{T} = \begin{pmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{Q}_2 \end{pmatrix} \quad \mathbf{Q}_1 = \begin{pmatrix} a & \sqrt{2}b \\ \sqrt{2}b & c + d \end{pmatrix} \quad \mathbf{Q}_2 = c - d$$

Symmetries in POPs

We come back to our initial POP:

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

Symmetries in POPs

We come back to our initial POP:

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

Finite group G and representation $\rho : G \to \mathrm{GL}_n(\mathbb{R})$
$f^g(\mathbf{x}) := f(\rho(g)^{-1}\mathbf{x})$

Symmetries in POPs

We come back to our initial POP:

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

Finite group G and representation $\rho : G \to \mathrm{GL}_n(\mathbb{R})$
$f^g(\mathbf{x}) := f(\rho(g)^{-1}\mathbf{x})$
💡 POP is G-invariant if $f^g = f$ and ${g_j}^g = g_j$

Symmetries in POPs

We come back to our initial POP:

$$\begin{aligned} \inf \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geqslant 0\} \end{aligned}$$

Finite group G and representation $\rho : G \to \mathrm{GL}_n(\mathbb{R})$
$f^g(\mathbf{x}) := f(\rho(g)^{-1}\mathbf{x})$
💡 POP is G-invariant if $f^g = f$ and $g_j{}^g = g_j$

The Reynolds Operator $\mathcal{R}_G : \mathbb{R}[\mathbf{x}] \to \mathbb{R}[\mathbf{x}]^G$ is

$$\mathcal{R}_G(f) := \frac{1}{|G|} \sum_{g \in G} f^g$$

Symmetries in POPs: a first hierarchy

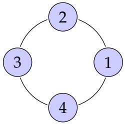
Dense vs Symmetric adapted hierarchy

(Dense)		(Symmetric)
$\inf \sum_\alpha f_\alpha \, y_\alpha$	$=$	$\inf \sum_\alpha f_\alpha \, y_\alpha^G$
s.t. $\mathbf{M}_r(\mathbf{y}) \succcurlyeq 0$		s.t. $\mathbf{M}_r(\mathbf{y}^G) \succcurlyeq 0$
$\mathbf{M}_{r-r_j}(g_j \, \mathbf{y}) \succcurlyeq 0$		$\mathbf{M}_{r-r_j}(g_j \, \mathbf{y}^G) \succcurlyeq 0$
$y_0 = 1$		$y_0^G = 1$

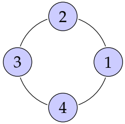
y_α^G is the pseudo-moment variable corresponding to the polynomial $\mathcal{R}_G(\mathbf{x}^\alpha)$

Symmetries in POPs: a first hierarchy

$G = C_4$ the cyclic group



Symmetries in POPs: a first hierarchy



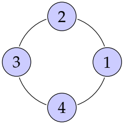
$G = C_4$ the cyclic group

Space of C_4-invariant polynomials of $\deg \leqslant 2$:

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1) \quad b_4 = \frac{1}{2}(x_1x_3 + x_2x_4)$$

Symmetries in POPs: a first hierarchy



$G = C_4$ the cyclic group

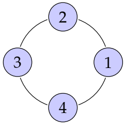
Space of C_4-invariant polynomials of $\deg \leqslant 2$:

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$
$$b_3 = \frac{1}{4}(x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_1) \quad b_4 = \frac{1}{2}(x_1 x_3 + x_2 x_4)$$

$\mathcal{R}_G(x_1) = b_1 \to y_1^G$

Symmetries in POPs: a first hierarchy

$G = C_4$ the cyclic group



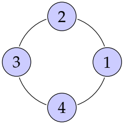
Space of C_4-invariant polynomials of $\deg \leqslant 2$:

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$
$$b_3 = \frac{1}{4}(x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1) \quad b_4 = \frac{1}{2}(x_1x_3 + x_2x_4)$$

$$\mathcal{R}_G(x_1) = b_1 \to y_1^G \quad \ldots \quad \mathcal{R}_G(x_4) = b_1 \to y_1^G$$

Symmetries in POPs: a first hierarchy

$G = C_4$ the cyclic group



Space of C_4-invariant polynomials of $\deg \leqslant 2$:

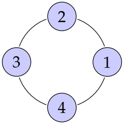
$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_1) \quad b_4 = \frac{1}{2}(x_1 x_3 + x_2 x_4)$$

$\mathcal{R}_G(x_1) = b_1 \to y_1^G \quad \dots \quad \mathcal{R}_G(x_4) = b_1 \to y_1^G$
$\mathcal{R}_G(x_i^2) = b_2 \to y_2^G$

Symmetries in POPs: a first hierarchy

$G = C_4$ the cyclic group



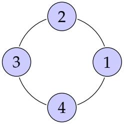
Space of C_4-invariant polynomials of $\deg \leqslant 2$:

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_1) \quad b_4 = \frac{1}{2}(x_1 x_3 + x_2 x_4)$$

$\mathcal{R}_G(x_1) = b_1 \to y_1^G \quad \dots \quad \mathcal{R}_G(x_4) = b_1 \to y_1^G$
$\mathcal{R}_G(x_i^2) = b_2 \to y_2^G \quad \mathcal{R}_G(x_1 x_2) = b_3 \to y_3^G$

Symmetries in POPs: a first hierarchy



$G = C_4$ the cyclic group

Space of C_4-invariant polynomials of $\deg \leqslant 2$:

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_1) \quad b_4 = \frac{1}{2}(x_1 x_3 + x_2 x_4)$$

$\mathcal{R}_G(x_1) = b_1 \to y_1^G \quad \ldots \quad \mathcal{R}_G(x_4) = b_1 \to y_1^G$
$\mathcal{R}_G(x_i^2) = b_2 \to y_2^G \quad \mathcal{R}_G(x_1 x_2) = b_3 \to y_3^G \quad \mathcal{R}_G(x_1 x_3) = b_4 \to y_4^G$

Symmetries in POPs: a first hierarchy

The symmetry-adapted moment matrix looks like this:

$$\mathbf{M}_1(\mathbf{y}) = \begin{pmatrix} 1 & y_1 & y_1 & y_1 & y_1 \\ y_1 & y_2 & y_3 & y_4 & y_3 \\ y_1 & y_3 & y_2 & y_3 & y_4 \\ y_1 & y_4 & y_3 & y_2 & y_3 \\ y_1 & y_3 & y_4 & y_3 & y_2 \end{pmatrix}$$

Symmetries in POPs: a first hierarchy

The symmetry-adapted moment matrix looks like this:

$$\mathbf{M}_1(\mathbf{y}) = \begin{pmatrix} 1 & y_1 & y_1 & y_1 & y_1 \\ y_1 & y_2 & y_3 & y_4 & y_3 \\ y_1 & y_3 & y_2 & y_3 & y_4 \\ y_1 & y_4 & y_3 & y_2 & y_3 \\ y_1 & y_3 & y_4 & y_3 & y_2 \end{pmatrix}$$

4 variables instead of 15

Symmetries in POPs: a first hierarchy

The symmetry-adapted moment matrix looks like this:

$$\mathbf{M}_1(\mathbf{y}) = \begin{pmatrix} 1 & y_1 & y_1 & y_1 & y_1 \\ y_1 & y_2 & y_3 & y_4 & y_3 \\ y_1 & y_3 & y_2 & y_3 & y_4 \\ y_1 & y_4 & y_3 & y_2 & y_3 \\ y_1 & y_3 & y_4 & y_3 & y_2 \end{pmatrix}$$

4 variables instead of 15

One can do even better!

Symmetries in POPs: a second hierarchy

The subset of $\mathbb{R}[\mathbf{x}]$ of degree at most r can be viewed as a real G-module

Symmetries in POPs: a second hierarchy

The subset of $\mathbb{R}[\mathbf{x}]$ of degree at most r can be viewed as a real G-module

$$\mathbb{R}[\mathbf{x}] \otimes \mathbb{C} = \bigoplus_{l=1}^{k} V_l = \bigoplus_{l=1}^{k} \bigoplus_{j \in J_l} W_{lj}$$

with complex irreducible components W_{lj}

Symmetries in POPs: a second hierarchy

The subset of $\mathbb{R}[\mathbf{x}]$ of degree at most r can be viewed as a real G-module

$$\mathbb{R}[\mathbf{x}] \otimes \mathbb{C} = \bigoplus_{l=1}^{k} V_l = \bigoplus_{l=1}^{k} \bigoplus_{j \in J_l} W_{lj}$$

with complex irreducible components W_{lj}

Pick a basis $\{s^l_{j,u}\}$ of W_{lj} and set $\mathcal{S}^l = \{s^l_{j,1} : j \in J_l\}$
💡 One selects the first basis elements of each W_{lj}

Symmetries in POPs: a second hierarchy

Truncation $\mathcal{S}_r^l = \{s_\alpha^l\}$ of $\mathcal{S}^l$ with basis elements of $\deg \leqslant r$

Symmetries in POPs: a second hierarchy

Truncation $\mathcal{S}_r^l = \{s_\alpha^l\}$ of $\mathcal{S}^l$ with basis elements of $\deg \leqslant r$

$\mathcal{B}_{2r}$ a basis of $\mathbb{R}[\mathbf{x}]_{2r}^G$ and y_b is the pseudo-moment variable corresponding to $b \in \mathcal{B}_{2r}$

Symmetries in POPs: a second hierarchy

Truncation $\mathcal{S}_r^l = \{s_\alpha^l\}$ of $\mathcal{S}^l$ with basis elements of $\deg \leqslant r$

$\mathcal{B}_{2r}$ a basis of $\mathbb{R}[\mathbf{x}]_{2r}^G$ and y_b is the pseudo-moment variable corresponding to $b \in \mathcal{B}_{2r}$

Theorem [Riener et al. '13]

(Dense) = (Symmetric)

$$\begin{aligned} \inf \quad & \sum_\alpha f_\alpha \, y_\alpha \\ \text{s.t.} \quad & \mathbf{M}_r(\mathbf{y}) \succcurlyeq 0 \\ & \mathbf{M}_{r-r_j}(g_j \, \mathbf{y}) \succcurlyeq 0 \\ & y_0 = 1 \end{aligned} \qquad = \qquad \begin{aligned} \inf \quad & \sum_\alpha f_\alpha \, y_\alpha \\ \text{s.t.} \quad & \mathbf{M}_r^G(\mathbf{y}) \succcurlyeq 0 \\ & \mathbf{M}_{r-r_j}^G(g_j \, \mathbf{y}) \succcurlyeq 0 \\ & y_0 = 1 \end{aligned}$$

$$\mathbf{M}_r^G(\mathbf{y}) = \bigoplus_{l=1}^{k} \mathbf{M}_{rl}^G(\mathbf{y})$$

Symmetries in POPs: a second hierarchy

Truncation $\mathcal{S}^l_r = \{s^l_\alpha\}$ of $\mathcal{S}^l$ with basis elements of $\deg \leqslant r$

$\mathcal{B}_{2r}$ a basis of $\mathbb{R}[\mathbf{x}]^G_{2r}$ and y_b is the pseudo-moment variable corresponding to $b \in \mathcal{B}_{2r}$

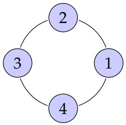
Theorem [Riener et al. '13]

(Dense)		(Symmetric)
$\inf \sum_\alpha f_\alpha y_\alpha$	$=$	$\inf \sum_\alpha f_\alpha y_\alpha$
s.t. $\mathbf{M}_r(\mathbf{y}) \succcurlyeq 0$		s.t. $\mathbf{M}^G_r(\mathbf{y}) \succcurlyeq 0$
$\mathbf{M}_{r-r_j}(g_j\, \mathbf{y}) \succcurlyeq 0$		$\mathbf{M}^G_{r-r_j}(g_j\, \mathbf{y}) \succcurlyeq 0$
$y_0 = 1$		$y_0 = 1$

$\mathbf{M}^G_r(\mathbf{y}) = \bigoplus_{l=1}^{k} \mathbf{M}^G_{rl}(\mathbf{y})$ $\quad (u, v)$ entry of $\mathbf{M}^G_{rl}(\mathbf{y})$ = $\mathcal{R}_G(s^l_u s^l_v)$

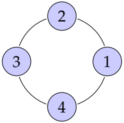
Symmetries in POPs: a second hierarchy

$G = C_4$ the cyclic group



Symmetries in POPs: a second hierarchy

$G = C_4$ the cyclic group

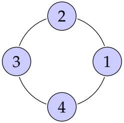


Space of C_4-invariant polynomials of $\deg \leqslant 2$:

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1) \quad b_4 = \frac{1}{2}(x_1x_3 + x_2x_4)$$

Symmetries in POPs: a second hierarchy



$G = C_4$ the cyclic group

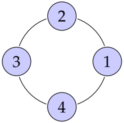
Space of C_4-invariant polynomials of $\deg \leqslant 2$:

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$
$$b_3 = \frac{1}{4}(x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1) \quad b_4 = \frac{1}{2}(x_1x_3 + x_2x_4)$$

All irreducible representations are 1-dim with symmetry adapted basis

$$\begin{pmatrix} 1 & i & -1 & -i \\ 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

Symmetries in POPs: a second hierarchy



$G = C_4$ the cyclic group

Space of C_4-invariant polynomials of $\deg \leqslant 2$:

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1) \quad b_4 = \frac{1}{2}(x_1x_3 + x_2x_4)$$

All irreducible representations are 1-dim with symmetry adapted basis

$$\begin{pmatrix} 1 & i & -1 & -i \\ 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \end{pmatrix} \rightsquigarrow \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 1 & 2 & 1 & 0 \\ 1 & 0 & -1 & -2 \\ 1 & -2 & 1 & 0 \end{pmatrix}$$

Symmetries in POPs: a second hierarchy

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_1) \quad b_4 = \frac{1}{2}(x_1 x_3 + x_2 x_4)$$

$\mathcal{S}_1^1 = \{\frac{1}{2}(x_1 + x_2 + x_3 + x_4)\} \quad \mathcal{S}_1^2 = \{x_2 - x_4\}$
$\mathcal{S}_1^3 = \{\frac{1}{2}(-x_1 + x_2 - x_3 + x_4)\} \quad \mathcal{S}_1^4 = \{x_1 - x_3\}$

Symmetries in POPs: a second hierarchy

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1) \quad b_4 = \frac{1}{2}(x_1x_3 + x_2x_4)$$

$$\mathcal{S}_1^1 = \{\tfrac{1}{2}(x_1 + x_2 + x_3 + x_4)\} \quad \mathcal{S}_1^2 = \{x_2 - x_4\}$$
$$\mathcal{S}_1^3 = \{\tfrac{1}{2}(-x_1 + x_2 - x_3 + x_4)\} \quad \mathcal{S}_1^4 = \{x_1 - x_3\}$$

$$\mathcal{R}_G(\tfrac{1}{2}(x_1 + x_2 + x_3 + x_4)) = 2b_1 \to 2y_1$$

Symmetries in POPs: a second hierarchy

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_1) \quad b_4 = \frac{1}{2}(x_1 x_3 + x_2 x_4)$$

$$\mathcal{S}_1^1 = \{\tfrac{1}{2}(x_1 + x_2 + x_3 + x_4)\} \quad \mathcal{S}_1^2 = \{x_2 - x_4\}$$
$$\mathcal{S}_1^3 = \{\tfrac{1}{2}(-x_1 + x_2 - x_3 + x_4)\} \quad \mathcal{S}_1^4 = \{x_1 - x_3\}$$

$$\mathcal{R}_G(\tfrac{1}{2}(x_1 + x_2 + x_3 + x_4)) = 2b_1 \to 2y_1$$

$$\mathcal{R}_G(x_2 - x_4) = \mathcal{R}_G(\tfrac{1}{2}(-x_1 + x_2 - x_3 + x_4)) = \mathcal{R}_G(x_1 - x_3) = 0$$

Symmetries in POPs: a second hierarchy

$$b_0 = 1 \quad b_1 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) \quad b_2 = \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

$$b_3 = \frac{1}{4}(x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1) \quad b_4 = \frac{1}{2}(x_1x_3 + x_2x_4)$$

$$\mathcal{S}_1^1 = \{\tfrac{1}{2}(x_1 + x_2 + x_3 + x_4)\} \quad \mathcal{S}_1^2 = \{x_2 - x_4\}$$
$$\mathcal{S}_1^3 = \{\tfrac{1}{2}(-x_1 + x_2 - x_3 + x_4)\} \quad \mathcal{S}_1^4 = \{x_1 - x_3\}$$

$$\mathcal{R}_G(\tfrac{1}{2}(x_1 + x_2 + x_3 + x_4)) = 2b_1 \to 2y_1$$

$$\mathcal{R}_G(x_2 - x_4) = \mathcal{R}_G(\tfrac{1}{2}(-x_1 + x_2 - x_3 + x_4)) = \mathcal{R}_G(x_1 - x_3) = 0$$

$$\mathcal{R}_G(\tfrac{1}{4}(x_1 + x_2 + x_3 + x_4)^2) = b_2 + 2b_3 + b_4 \to y_2 + 2y_3 + y_4$$

Symmetries in POPs: a second hierarchy

$$\mathbf{M}_1^G(\mathbf{y}) = \begin{pmatrix} 1 & 2y_1 & 0 & 0 & 0 \\ 2y_1 & y_2 + 2y_3 + y_4 & 0 & 0 & 0 \\ 0 & 0 & y_2 - y_4 & 0 & 0 \\ 0 & 0 & 0 & y_2 - 2y_3 + y_4 & 0 \\ 0 & 0 & 0 & 0 & y_2 - y_4 \end{pmatrix}$$

Symmetries in POPs: a second hierarchy

$$\mathbf{M}_1^G(\mathbf{y}) = \begin{pmatrix} 1 & 2y_1 & 0 & 0 & 0 \\ 2y_1 & y_2 + 2y_3 + y_4 & 0 & 0 & 0 \\ 0 & 0 & y_2 - y_4 & 0 & 0 \\ 0 & 0 & 0 & y_2 - 2y_3 + y_4 & 0 \\ 0 & 0 & 0 & 0 & y_2 - y_4 \end{pmatrix}$$

💡 4 variables instead of 15, 2×2 block + 3 elementary blocks instead of 5×5 block

Symmetries in POPs: special case of S_n

- Irreducible repr. of S_n isomorphic to the partitions of n

Symmetries in POPs: special case of S_n

Irreducible repr. of S_n isomorphic to the partitions of n

Young tableau of $\lambda = (4, 3, 1, 1, 1) \vdash 10$
with columns $\mathcal{C}_j$
Classe of equivalent Young tableaux = $\{t\}$

$$t = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 6 \\ \hline 5 & 7 & 8 & \\ \cline{1-3} 9 & & & \\ \cline{1-1} 2 & & & \\ \cline{1-1} 10 & & & \\ \cline{1-1} \end{array}$$

Symmetries in POPs: special case of S_n

Irreducible repr. of S_n isomorphic to the partitions of n

Young tableau of $\lambda = (4, 3, 1, 1, 1) \vdash 10$
with columns $\mathcal{C}_j$
Classe of equivalent Young tableaux = $\{t\}$

$$t = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 6 \\ \hline 5 & 7 & 8 & \multicolumn{1}{c}{} \\ \cline{1-3} 9 & \multicolumn{3}{c}{} \\ \cline{1-1} 2 & \multicolumn{3}{c}{} \\ \cline{1-1} 10 & \multicolumn{3}{c}{} \\ \cline{1-1} \end{array}$$

$\beta = (\beta_1, \ldots, \beta_n)$ with distinct components $b_1, \ldots, b_\ell$ ordered $\searrow$

Symmetries in POPs: special case of S_n

💡 Irreducible repr. of S_n isomorphic to the partitions of n

Young tableau of $\lambda = (4, 3, 1, 1, 1) \vdash 10$
with columns $\mathcal{C}_j$
Classe of equivalent Young tableaux = $\{t\}$

$t =$

1	3	4	6
5	7	8	
9			
2			
10			

$\beta = (\beta_1, \ldots, \beta_n)$ with distinct components $b_1, \ldots, b_\ell$ ordered $\searrow$
$\mu_j = |i : \beta_i = b_j| \Rightarrow \mu = (\mu_1, \ldots, \mu_\ell) \vdash n$ is the **shape** of β

Symmetries in POPs: special case of S_n

Irreducible repr. of S_n isomorphic to the partitions of n

Young tableau of $\lambda = (4, 3, 1, 1, 1) \vdash 10$
with columns $\mathcal{C}_j$
Classe of equivalent Young tableaux = $\{t\}$

$t =$

1	3	4	6
5	7	8	
9			
2			
10			

$\beta = (\beta_1, \dots, \beta_n)$ with distinct components $b_1, \dots, b_\ell$ ordered $\searrow$
$\mu_j = |i : \beta_i = b_j| \Rightarrow \mu = (\mu_1, \dots, \mu_\ell) \vdash n$ is the **shape** of β
$(0, 0, 0), (1, 0, 0), (2, 0, 0)$ have shape $(3), (2, 1), (2, 1)$
For each β, take pairs (t, T) where t is λ-tableau and T has shape λ and content μ to build:

$$\mathbf{x}^{t,T} = \prod_{i,j} x_{\mathcal{C}_j}^{b_{T(i,j)}}$$

Symmetries in POPs: special case of S_n

Column stabilizer $\mathrm{CStab}_t = S_{\mathcal{C}_1} \times \cdots \times S_{\mathcal{C}_\nu}$

Symmetries in POPs: special case of S_n

Column stabilizer $\mathrm{CStab}_t = S_{\mathcal{C}_1} \times \cdots \times S_{\mathcal{C}_\nu}$

Specht polynomial $\sum_{g \in \mathrm{CStab}_t} \mathrm{sgn}(g)(\mathbf{x}^{t,T})^g$

Symmetries in POPs: special case of S_n

Column stabilizer $\mathrm{CStab}_t = S_{\mathcal{C}_1} \times \cdots \times S_{\mathcal{C}_\nu}$

Specht polynomial $\sum_{g \in \mathrm{CStab}_t} \mathrm{sgn}(g)(\mathbf{x}^{t,T})^g \rightsquigarrow$ generalized Specht polynomial after summing over tableaux equivalent to T

Symmetries in POPs: special case of S_n

Column stabilizer $\mathrm{CStab}_t = S_{\mathcal{C}_1} \times \cdots \times S_{\mathcal{C}_\nu}$

Specht polynomial $\sum_{g \in \mathrm{CStab}_t} \mathrm{sgn}(g)(\mathbf{x}^{t,T})^g \rightsquigarrow$ generalized Specht polynomial after summing over tableaux equivalent to T

Theorem

β with shape $\mu \implies$

$$\mathbb{R}\{\mathbf{x}^\beta\} = \bigoplus_{\lambda \trianglerighteq \mu} \bigoplus_{T} \mathbb{R}\{S_{(t,T)}\}$$

t a λ-tableau with $\nearrow$ rows & columns

T with shape λ and content μ

Symmetries in POPs: special case of S_n

Column stabilizer $\mathrm{CStab}_t = S_{\mathcal{C}_1} \times \cdots \times S_{\mathcal{C}_v}$

Specht polynomial $\sum_{g \in \mathrm{CStab}_t} \mathrm{sgn}(g)(\mathbf{x}^{t,T})^g \rightsquigarrow$ generalized Specht polynomial after summing over tableaux equivalent to T

Theorem

β with shape $\mu \implies$

$$\mathbb{R}\{\mathbf{x}^\beta\} = \bigoplus_{\lambda \trianglerighteq \mu} \bigoplus_{T} \mathbb{R}\{S_{(t,T)}\}$$

t a λ-tableau with $\nearrow$ rows & columns

T with shape λ and content μ

💡 Gives a special block-structure for the moment matrix!

Symmetries in POPs: special case of S_3

$r = 2 \implies$ moment variables indexed by partitions of $\{1, 2, 3, 4\}$ with at most $n = 3$ parts:

$$y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_{11} \quad y_{22} \quad y_{21} \quad y_{111} \quad y_{211}$$

Symmetries in POPs: special case of S_3

$r = 2 \implies$ moment variables indexed by partitions of $\{1, 2, 3, 4\}$ with at most $n = 3$ parts:

$$y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_{11} \quad y_{22} \quad y_{21} \quad y_{111} \quad y_{211}$$

β should be $(0, 0, 0) \quad (1, 0, 0) \quad (2, 0, 0) \quad (1, 1, 0)$

Symmetries in POPs: special case of S_3

$r = 2 \implies$ moment variables indexed by partitions of $\{1, 2, 3, 4\}$ with at most $n = 3$ parts:

$$y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_{11} \quad y_{22} \quad y_{21} \quad y_{111} \quad y_{211}$$

β should be $(0, 0, 0) \quad (1, 0, 0) \quad (2, 0, 0) \quad (1, 1, 0)$
Possible shapes (3) and $(2, 1)$ with generalized Specht polynomials

Symmetries in POPs: special case of S_3

$r = 2 \implies$ moment variables indexed by partitions of $\{1, 2, 3, 4\}$ with at most $n = 3$ parts:

$$y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_{11} \quad y_{22} \quad y_{21} \quad y_{111} \quad y_{211}$$

💡 β should be $(0, 0, 0) \quad (1, 0, 0) \quad (2, 0, 0) \quad (1, 1, 0)$
Possible shapes (3) and $(2, 1)$ with generalized Specht polynomials

$$\{1 \quad x_1 + x_2 + x_3 \quad x_1^2 + x_2^2 + x_3^2 \quad x_1x_2 + x_2x_3 + x_3x_1\}$$

$$\{x_3 - x_2 - x_1 \quad x_3^2 - x_2^2 - x_1^2 \quad -x_1x_2 + x_2x_3 + x_3x_1\}$$

💡 Leads to 4×4 + 3×3-block moment matrices instead of 10×10!

Conclusion

SPARSITY EXPLOITING CONVERGING HIERARCHIES to minimize polynomials, eigenvalue/trace, joint spectral radius

Conclusion

SPARSITY EXPLOITING CONVERGING HIERARCHIES to minimize polynomials, eigenvalue/trace, joint spectral radius

FAST IMPLEMENTATION IN JULIA: TSSOS, NCTSSOS, SparseJSR

Conclusion

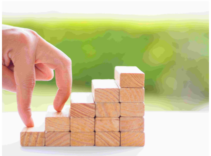
SPARSITY EXPLOITING CONVERGING HIERARCHIES to minimize polynomials, eigenvalue/trace, joint spectral radius

FAST IMPLEMENTATION IN JULIA: TSSOS, NCTSSOS, SparseJSR

Combine correlative & term sparsity for problems with $n = 10^3$

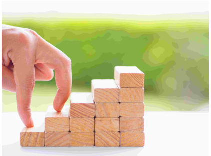
Further topics

Convergence rate of SPARSE hierarchies?



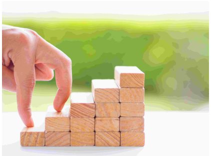
Further topics

Convergence rate of SPARSE hierarchies?



💡 (smart) solution extraction for term sparse/symmetric POPs

Further topics



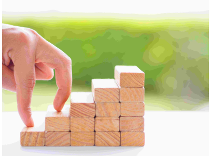
Convergence rate of SPARSE hierarchies?

💡 (smart) solution extraction for term sparse/symmetric POPs

Numerical conditioning of sparse/symmetric SDP relaxations?

Further topics

Convergence rate of SPARSE hierarchies?



💡 (smart) solution extraction for term sparse/symmetric POPs

Numerical conditioning of sparse/symmetric SDP relaxations?

💡 Tons of applications . . .

Thank you for your attention!

https://homepages.laas.fr/vmagron

GITHUB:TSSOS

References

- Gavril. Algorithms for minimum coloring, maximum clique, minimum covering by cliques, and maximum independent set of a chordal graph. SIAM Comp., 1972
- Griewank & Toint. Numerical experiments with partially separable optimization problems. Numerical analysis, 1984
- Agler, Helton, McCullough & Rodman. Positive semidefinite matrices with a given sparsity pattern. Linear algebra & its applications, 1988
- Blair & Peyton. An introduction to chordal graphs and clique trees. Graph theory & sparse matrix computation, 1993
- Vandenberghe & Andersen. Chordal graphs and semidefinite optimization. Foundations & Trends in Optim., 2015

References

- Lasserre. Convergent SDP-relaxations in polynomial optimization with sparsity. SIAM Optim., 2006
- Waki, Kim, Kojima & Muramatsu. Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. SIAM Optim., 2006
- Magron, Constantinides, & Donaldson. Certified Roundoff Error Bounds Using Semidefinite Programming. Trans. Math. Softw., 2017
- Magron. Interval Enclosures of Upper Bounds of Roundoff Errors Using Semidefinite Programming. Trans. Math. Softw., 2018
- Josz & Molzahn. Lasserre hierarchy for large scale polynomial optimization in real and complex variables. SIAM Optim., 2018
- Weisser, Lasserre & Toh. Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity. Math. Program., 2018
- Chen, Lasserre, Magron & Pauwels. A sublevel moment-sos hierarchy for polynomial optimization, arxiv:2101.05167

References

- Chen, Lasserre, Magron & Pauwels. Semialgebraic Optimization for Bounding Lipschitz Constants of ReLU Networks. NIPS, 2020
- Chen, Lasserre, Magron & Pauwels. Semialgebraic Representation of Monotone Deep Equilibrium Models and Applications to Certification. arxiv:2106.01453
- Mai, Lasserre & Magron. A sparse version of Reznick's Positivstellensatz. arxiv:2002.05101
- Tacchi, Weisser, Lasserre & Henrion. Exploiting sparsity for semi-algebraic set volume computation. Foundations of Comp. Math., 2021
- Tacchi, Cardozo, Henrion & Lasserre. Approximating regions of attraction of a sparse polynomial differential system. IFAC, 2020
- Schlosser & Korda. Sparse moment-sum-of-squares relaxations for nonlinear dynamical systems with guaranteed convergence. arxiv:2012.05572
- Zheng & Fantuzzi. Sum-of-squares chordal decomposition of polynomial matrix inequalities. arxiv:2007.11410

References

- Klep, Magron & Povh. Sparse Noncommutative Polynomial Optimization. *Math Prog. A*, arxiv:1909.00569 NCSOStools
- Reznick. Extremal PSD forms with few terms. Duke mathematical journal, 1978
- Wang, Magron & Lasserre. TSSOS: A Moment-SOS hierarchy that exploits term sparsity. SIAM Optim., 2021
- Wang, Magron & Lasserre. Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension. SIAM Optim., 2021
- Wang, Magron, Lasserre & Mai. CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization. arxiv:2005.02828
- Magron & Wang. TSSOS: a Julia library to exploit sparsity for large-scale polynomial optimization, MEGA, 2021
- Parrilo & Jadbabaie. Approximation of the joint spectral radius using sum of squares. Linear Algebra & its Applications, 2008
- Wang, Maggio & Magron. SparseJSR: A fast algorithm to compute joint spectral radius via sparse sos decompositions. ACC 2021

References

- Vreman, Pazzaglia, Wang, Magron & Maggio. Stability of control systems under extended weakly-hard constraints. arxiv:2101.11312
- Wang & Magron. Exploiting Sparsity in Complex Polynomial Optimization. JOTA, 2021
- Wang & Magron. Exploiting term sparsity in Noncommutative Polynomial Optimization. *Computational Optimization & Applications*, 2022 NCTSSOS
- Navascués, Pironio & Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. New Journal of Physics, 2008
- Klep, Magron & Volčič. Optimization over trace polynomials. Annales Henri Poincaré, 2021
- Serre. Linear representations of finite groups. Springer, 1977
- Gatterman & Parrilo. Symmetry groups, semidefinite programs, and sums of squares, Journal of Pure and Applied Algebra, 2004
- Riener, Theobald, Andrén & Lasserre. Exploiting symmetries in SDP-relaxations for polynomial optimization. MathOR, 2013
- Blekherman & Riener. Symmetric non-negative forms and sums of squares. Discrete & Computational Geometry, 2021