

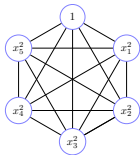
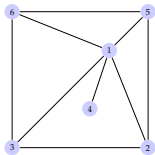
Sparse polynomial optimization

Victor Magron
LAAS CNRS

Optimization and Numerical Analysis Seminar

The University of Hong Kong

17 November 2021



What is sparse polynomial optimization?

Looks like a regular polynomial optimization problem (POP):

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

What is sparse polynomial optimization?

Looks like a regular polynomial optimization problem (POP):

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

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

What is sparse polynomial optimization?

Looks like a regular polynomial optimization problem (POP):

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

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

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

What is sparse polynomial optimization?

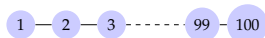
Looks like a regular polynomial optimization problem (POP):

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

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

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

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



What is sparse polynomial optimization?

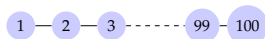
Looks like a regular polynomial optimization problem (POP):

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

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

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

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



Term sparsity: few terms in f, g_j

What is sparse polynomial optimization?

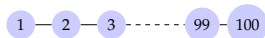
Looks like a regular polynomial optimization problem (POP):

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

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

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

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



Term sparsity: few terms in f, g_j

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

Where do we find sparse POPs?

Where do we find sparse POPs?

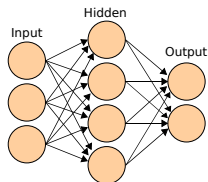
Everywhere (almost)!

Where do we find sparse POPs?

Everywhere (almost)!

Deep learning

~> robustness, computer vision

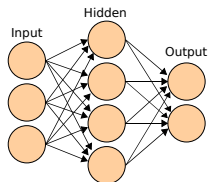


Where do we find sparse POPs?

Everywhere (almost)!

Deep learning

~> robustness, computer vision



Power systems

~> AC optimal power-flow, stability

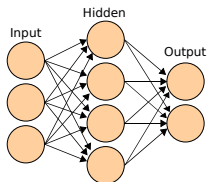


Where do we find sparse POPs?

Everywhere (almost)!

Deep learning

↪ robustness, computer vision



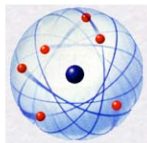
Power systems

↪ AC optimal power-flow, stability



Quantum Systems

↪ condensed matter, entanglement



The Moment-SOS Hierarchy for POP

Correlative sparsity

Term sparsity

The Moment-SOS Hierarchy for POP

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

Theory

(Primal)		(Dual)
$\inf \int f d\mu$		$\sup \lambda$
with μ proba \Rightarrow	INFINITE LP	\Leftarrow with $f - \lambda \geq 0$

The Moment-SOS Hierarchy for POP

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

Practice

(Primal **Relaxation**)

moments $\int x^\alpha d\mu$

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 $\implies \binom{n+2d}{n}$ **SDP** VARIABLES



The Moment-SOS Hierarchy for POP

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

Practice

(Primal **Relaxation**)

moments $\int x^\alpha d\mu$

finite number \Rightarrow



SDP

(Dual **Strengthening**)

$f - \lambda = \text{sum of squares}$

\Leftarrow fixed degree

LASSERRE'S HIERARCHY of **CONVEX PROBLEMS** $\uparrow f^*$

[Lasserre '01]

degree d & n vars $\implies \binom{n+2d}{n}$ **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. $\mathbf{M}_{r-r_j}(g_j \mathbf{y}) \succeq 0$		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. $\mathbf{M}_{r-r_j}(g_j \mathbf{y}) \succeq 0$		s.t. $\lambda \in \mathbb{R}$
$y_0 = 1$		$f - \lambda \in \mathcal{Q}(\mathbf{X})_r$

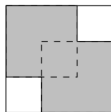
What is the primal-dual “SPARSE” variant?

Sparse matrices

Symmetric matrices indexed by graph vertices

Sparse matrices

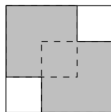
Symmetric matrices indexed by graph vertices



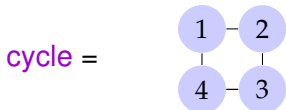
💡 no edge between 1 and 3 \iff 0 entry in the (1,3) entry

Sparse matrices

Symmetric matrices indexed by graph vertices

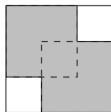


💡 no edge between 1 and 3 \iff 0 entry in the (1,3) entry

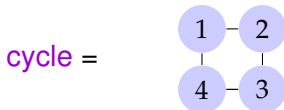


Sparse matrices

Symmetric matrices indexed by graph vertices



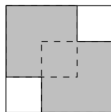
💡 no edge between 1 and 3 \iff 0 entry in the (1,3) entry



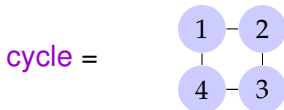
chord = edge between two nonconsecutive vertices in a cycle

Sparse matrices

Symmetric matrices indexed by graph vertices

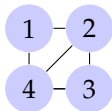


💡 no edge between 1 and 3 \iff 0 entry in the (1,3) entry



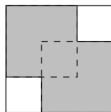
chord = edge between two nonconsecutive vertices in a cycle

chordal graph = all cycles of length ≥ 4 have at least one chord

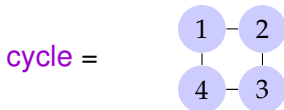


Sparse matrices

Symmetric matrices indexed by graph vertices



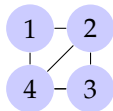
💡 no edge between 1 and 3 \iff 0 entry in the (1,3) entry



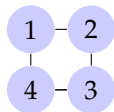
chord = edge between two nonconsecutive vertices in a cycle

chordal graph = all cycles of length ≥ 4 have at least one chord

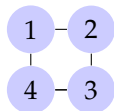
clique = a fully connected subset of vertices



Chordal extensions



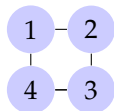
Chordal extensions



Fact

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

Chordal extensions

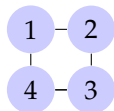


Fact

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

💡 Chordal extension is not unique!

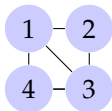
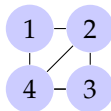
Chordal extensions



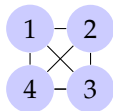
Fact

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, \dots, I_p :

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

(possibly after reordering)

Running intersection property (RIP)

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

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

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

(possibly after reordering)

💡 RIP always holds for $p = 2$

Running intersection property (RIP)

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

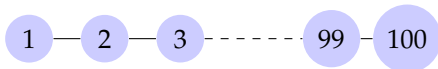
For a chordal graph with maximal cliques I_1, \dots, I_p :

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

(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]

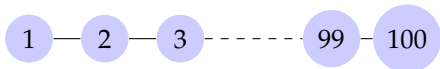
For a chordal graph with maximal cliques I_1, \dots, I_p :

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

(possibly after reordering)

💡 RIP always holds for $p = 2$

💡 RIP holds for chains



💡 RIP holds for numerous applications!

Semidefinite Programming (SDP)

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

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



Spectrahedron

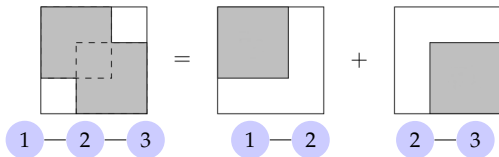
Sparse SDP matrices

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



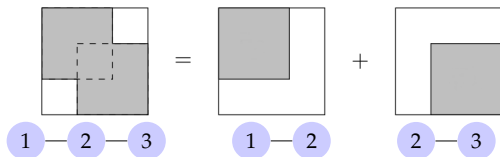
Sparse SDP matrices

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 ?

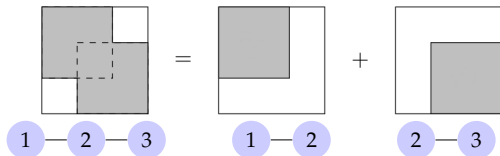
Sparse SDP matrices

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 ? $P_1 \in \mathbb{R}^{|I_1| \times n}$ $P(i, j) = \begin{cases} 1 & \text{if } I(i) = j \\ 0 & \text{otherwise} \end{cases}$

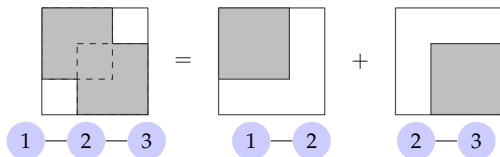
Sparse SDP matrices

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 ? $P_1 \in \mathbb{R}^{|I_1| \times n}$ $P(i, j) = \begin{cases} 1 & \text{if } I(i) = j \\ 0 & \text{otherwise} \end{cases}$

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

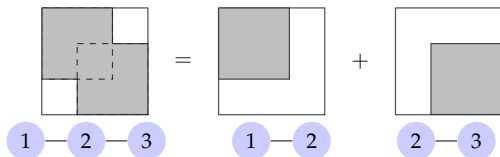
Sparse SDP matrices

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

The Moment-SOS Hierarchy for POP

Correlative sparsity

Term sparsity

What is correlative sparsity?

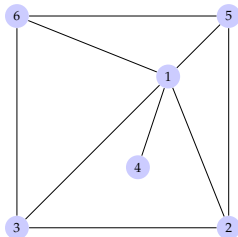
💡 Exploit few links between **variables** [Lasserre, Waki et al. '06]

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

Correlative sparsity pattern (csp) graph G

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

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



What is correlative sparsity?

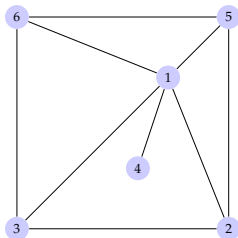
💡 Exploit few links between **variables** [Lasserre, Waki et al. '06]

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

Correlative sparsity pattern (csp) graph G

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

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

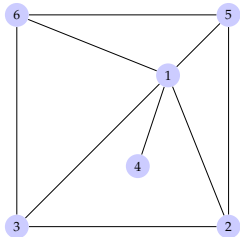


Similar construction with constraints $\mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0\}$

What is correlative sparsity?

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

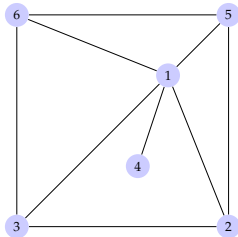
Chordal graph after adding edge (3,5)



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 $N - \sum_i x_i^2 \geq 0$ then

$f > 0$ on $\mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geq 0\} \implies f = \sigma_0 + \sum_j \sigma_j g_j$ with σ_j 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 $N - \sum_i x_i^2 \geq 0$ then

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

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 $\mathbf{x}(I_k)$

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

\implies 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 $\mathbf{x}(I_k)$

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

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

$$\mathbf{M}_1(\mathbf{y}, I_1) = \left(\begin{array}{c|cc} 1 & y_{1,0,0,0,0,0} & y_{0,0,0,1,0,0} \\ \hline y_{1,0,0,0,0,0} & y_{2,0,0,0,0,0} & y_{1,0,0,1,0,0} \\ y_{0,0,0,1,0,0} & y_{1,0,0,1,0,0} & y_{0,0,0,2,0,0} \end{array} \right)$$

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

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

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

Sparse primal-dual Moment-SOS hierarchy

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

Dense Moment-SOS hierarchy

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

Sparse primal-dual Moment-SOS hierarchy

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

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

Each g_j depends on some I_k

Sparse Moment-SOS hierarchy

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

Sparse primal-dual Moment-SOS hierarchy

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

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

Each g_j depends on some I_k

Sparse Moment-SOS hierarchy

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

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

$$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x}) \text{ with } \mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geq 0, j \leq m\}$$
$$\tau = \max\{|I_1|, \dots, |I_p|\}$$

Sparse Moment-SOS hierarchy

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

Computational cost

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

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

Sparse Moment-SOS hierarchy

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

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

Computational cost

$$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x}) \text{ with } \mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geq 0, j \leq m\}$$
$$\tau = \max\{|I_1|, \dots, |I_p|\}$$

Sparse Moment-SOS hierarchy

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

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

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

Computational cost

$$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x}) \text{ with } \mathbf{X} = \{\mathbf{x} : g_j(\mathbf{x}) \geq 0, j \leq m\}$$
$$\tau = \max\{|I_1|, \dots, |I_p|\}$$

Sparse Moment-SOS hierarchy

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

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

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

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1x_2 + x_3x_4$

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1x_2 + x_3x_4$

Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1x_2(1 + e_1) + x_3x_4(1 + e_2)](1 + e_3)$

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1x_2 + x_3x_4$

Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1x_2(1 + e_1) + x_3x_4(1 + e_2)](1 + e_3)$

$\mathbf{x} \in \mathbf{X}$, $|e_i| \leq 2^{-\delta}$ $\delta = 24$ (single) or 53 (double)

Application to roundoff errors

[Magron Constantinides Donaldson '17]

Exact $f(\mathbf{x}) = x_1x_2 + x_3x_4$

Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1x_2(1 + e_1) + x_3x_4(1 + e_2)](1 + e_3)$

$\mathbf{x} \in \mathbf{X}$, $|e_i| \leq 2^{-\delta}$ $\delta = 24$ (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 $f(\mathbf{x}) = x_1x_2 + x_3x_4$

Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1x_2(1 + e_1) + x_3x_4(1 + e_2)](1 + e_3)$

$\mathbf{x} \in \mathbf{X}$, $|e_i| \leq 2^{-\delta}$ $\delta = 24$ (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 $f(\mathbf{x}) = x_1x_2 + x_3x_4$

Floating-point $\hat{f}(\mathbf{x}, \mathbf{e}) = [x_1x_2(1 + e_1) + x_3x_4(1 + e_2)](1 + e_3)$

$\mathbf{x} \in \mathbf{X}$, $|e_i| \leq 2^{-\delta}$ $\delta = 24$ (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**

💡 $I_k \rightarrow \{\mathbf{x}, e_k\} \implies \boxed{m(n+1)^{2r} \text{ instead of } (n+m)^{2r}}$ 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 \leadsto **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 \leadsto **Out of memory**

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

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 \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)

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 \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**

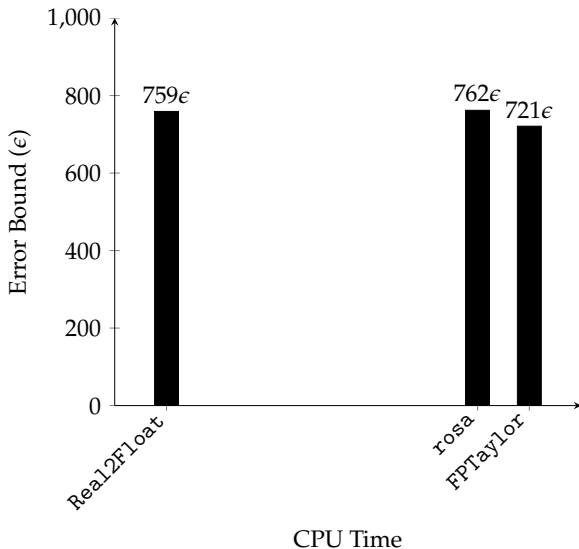
Sparse SDP Real2Float tool: $15 \binom{6+1+4}{6+1} = 4950 \rightsquigarrow 759\epsilon$

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

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

with $x_1x_2 \neq x_2x_1$, **involution** $(x_1y_3)^* = y_3x_1$

Application to noncommutative optimization

Self-adjoint noncommutative variables x_i, y_j

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

with $x_1x_2 \neq x_2x_1$, **involution** $(x_1y_3)^* = y_3x_1$

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

Application to noncommutative optimization

Self-adjoint noncommutative variables x_i, y_j

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

with $x_1x_2 \neq x_2x_1$, **involution** $(x_1y_3)^* = y_3x_1$

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

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

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

with $x_1x_2 \neq x_2x_1$, **involution** $(x_1y_3)^* = y_3x_1$

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

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]

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

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]

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

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

💡 replace “ $f - \lambda \mathbf{I} \succcurlyeq 0$ on \mathbf{X} ” by $f - \lambda \mathbf{I} = \sum_i s_i^* s_i + \sum_j \sum_i t_{ji}^* g_j t_{ji}$
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 $\underline{x} = (x_1, \dots, x_n)$

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

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

[Klep Magron Povh '21]
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!

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

[Klep Magron Povh '21]

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 $\underline{x} = (x_1, \dots, x_n)$

Theorem [Helton & McCullough '02]

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

BAD NEWS: there is **no** sparse analog!

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

[Klep Magron Povh '21]

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 $\underline{x} = (x_1, \dots, x_n)$

Theorem [Helton & McCullough '02]

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

BAD NEWS: there is **no** sparse analog!

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

[Klep Magron Povh '21]

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 $\mathbf{x}(I_k)$

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

\Rightarrow

s_{ki} "sees" vars in I_k

t_{ji} "sees" vars from g_j

Application to noncommutative optimization

I₃₃₂₂ 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 λ_{\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

I₃₃₂₂ 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 λ_{\max} of f on

$$\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$$

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

Application to noncommutative optimization

I₃₃₂₂ 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 λ_{\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

Application to noncommutative optimization

I₃₃₂₂ 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 λ_{\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

I₃₃₂₂ 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 λ_{\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)

Application to noncommutative optimization

I₃₃₂₂ 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 λ_{\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	

Application to noncommutative optimization

I₃₃₂₂ 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 λ_{\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

I₃₃₂₂ 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 λ_{\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	
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

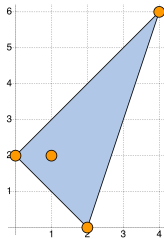
Term sparsity

Term sparsity via Newton polytope

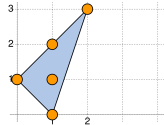
$$f = 4x_1^4x_2^6 + x_1^2 - x_1x_2^2 + x_2^2$$

$$\text{spt}(f) = \{(4,6), (2,0), (1,2), (0,2)\}$$

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



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



$$f = \left(x_1 \quad x_2 \quad x_1x_2 \quad x_1x_2^2 \quad x_1^2x_2^3 \right) \underbrace{Q}_{\succeq 0} \begin{pmatrix} x_1 \\ x_2 \\ x_1x_2 \\ x_1x_2^2 \\ x_1^2x_2^3 \end{pmatrix}$$

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] $\rightarrow f = (1 \quad x_1 \quad x_2 \quad x_3 \quad x_1x_2 \quad x_2x_3)$

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

$$\underbrace{\mathcal{Q}}_{\succeq 0} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$$

Term sparsity: the unconstrained case

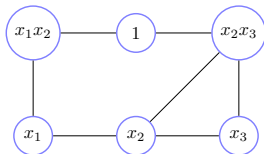
$$f = x_1^2 - 2x_1x_2 + 3x_2^2 - 2x_1^2x_2 + 2x_1^2x_2^2 - 2x_2x_3 + 6x_3^2 + 18x_2^2x_3 - 54x_2x_3^2 + 142x_2^2x_3^2$$

[Reznick '78] $\rightarrow f = (1 \ x_1 \ x_2 \ x_3 \ x_1x_2 \ x_2x_3) \underbrace{Q}_{\succeq 0}$

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

$$\begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$$

💡 **Term sparsity pattern graph G**



Term sparsity: the unconstrained case

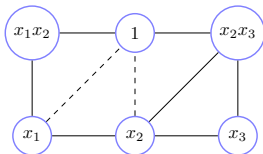
$$f = x_1^2 - 2x_1x_2 + 3x_2^2 - 2x_1^2x_2 + 2x_1^2x_2^2 - 2x_2x_3 + 6x_3^2 + 18x_2^2x_3 - 54x_2x_3^2 + 142x_2^2x_3^2$$

$$\text{[Reznick '78]} \rightarrow f = (1 \quad x_1 \quad x_2 \quad x_3 \quad x_1x_2 \quad x_2x_3) \underbrace{Q}_{\succeq 0}$$

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

$$\begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$$

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



Term sparsity: the unconstrained case

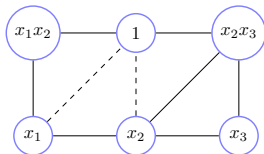
$$f = x_1^2 - 2x_1x_2 + 3x_2^2 - 2x_1^2x_2 + 2x_1^2x_2^2 - 2x_2x_3 + 6x_3^2 + 18x_2^2x_3 - 54x_2x_3^2 + 142x_2^2x_3^2$$

[Reznick '78] $\rightarrow f = (1 \ x_1 \ x_2 \ x_3 \ x_1x_2 \ x_2x_3) \underbrace{Q}_{\succeq 0}$

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

$$\begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$$

💡 **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 $\leq r$

Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes V = monomials of degree $\leq r$

Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \text{supp } f \cup \text{supp } g_j \cup \bigcup_{|\alpha| \leq r} 2\alpha$$

Term sparsity: the constrained case

At step r of the hierarchy, **tsp** graph G has

Nodes V = monomials of degree $\leq r$

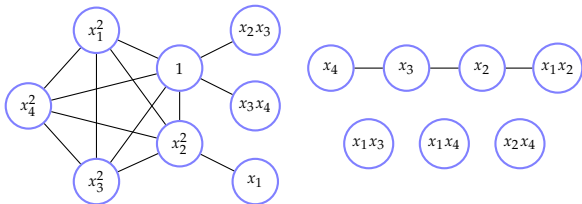
Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \text{supp } f \cup \bigcup_{|\alpha| \leq r} 2\alpha$$

An example with $r = 2$

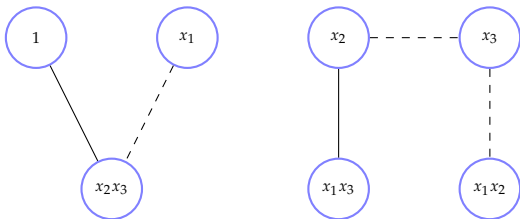
$$f = x_1^4 + x_1x_2^2 + x_2x_3 + x_3^2x_4^2$$

$$g_1 = 1 - x_1^2 - x_2^2 - x_3^2 \quad g_2 = 1 - x_3x_4$$



Term sparsity: support extension

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



Term sparsity: the constrained case

At step r of the hierarchy, tsp graph G has

Nodes $V =$ monomials of degree $\leq r$

Term sparsity: the constrained case

At step r of the hierarchy, **tsp** graph G has

Nodes V = monomials of degree $\leq r$

Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \text{supp } f \cup \bigcup_{|\alpha| \leq r} \text{supp } g_j \cup 2\alpha$$

\rightsquigarrow **support extension**

Term sparsity: the constrained case

At step r of the hierarchy, **tsp** graph G has

Nodes V = monomials of degree $\leq r$

Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \text{supp } f \cup \bigcup_{|\alpha| \leq r} \text{supp } g_j \cup 2\alpha$$

\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 $\leq r$

Edges E with

$$\{\alpha, \beta\} \in E \Leftrightarrow \alpha + \beta \in \text{supp } f \cup \bigcup_{|\alpha| \leq r} 2\alpha$$

\rightsquigarrow **support extension** \rightsquigarrow **chordal extension** G'

By iteratively performing **support extension** & **chordal extension**

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

💡 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,

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

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,

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

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

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,

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

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

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

Term sparsity: primal moment relaxations

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

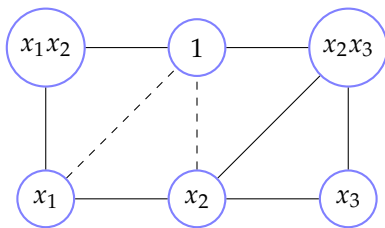
$$\begin{aligned} f_{\text{ts}}^{r,s} &= \inf \sum_{\alpha} f_{\alpha} y_{\alpha} \\ \text{s.t. } & \mathbf{M}_{C_{0,i}^{(s)}}(\mathbf{y}) \succeq 0 \\ & \mathbf{M}_{C_{j,i}^{(s)}}(\mathbf{g}_j \mathbf{y}) \succeq 0 \\ & y_0 = 1 \end{aligned}$$

💡 dual yields the **TSSOS** hierarchy

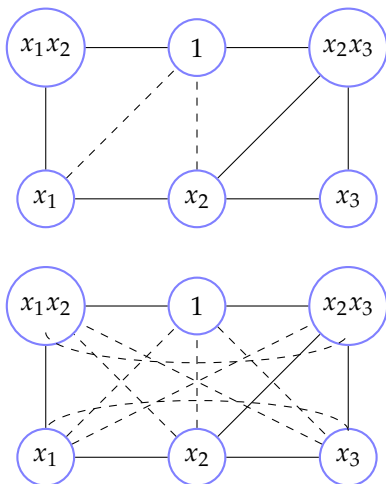
A two-level hierarchy of lower bounds

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

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 $(f_{ts}^{r,s})_{r \geq r_{\min}}$ is monotonically nondecreasing.

Term sparsity: convergence guarantees

Theorem [Lasserre Magron Wang '21]

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

💡 Fixing a relaxation order r , the sequence $(f_{ts}^{r,s})_{s \geq 1} \uparrow f^r$ 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 $(f_{ts}^{r,s})_{r \geq r_{\min}}$ is monotonically nondecreasing.

💡 Fixing a relaxation order r , the sequence $(f_{ts}^{r,s})_{s \geq 1} \uparrow f^r$ 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 $(f_{ts}^{r,s})_{r \geq r_{\min}}$ is monotonically nondecreasing.

💡 Fixing a relaxation order r , the sequence $(f_{ts}^{r,s})_{s \geq 1} \uparrow f^r$ 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.

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

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

Term sparsity: convergence guarantees

Theorem [Lasserre Magron Wang '21]

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

💡 Fixing a relaxation order r , the sequence $(f_{ts}^{r,s})_{s \geq 1} \uparrow f^r$ 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.

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



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

$$x_2 \mapsto -x_2$$

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

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

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

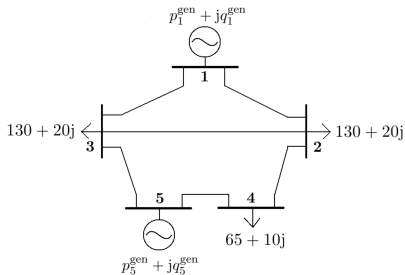
Combining correlative & term sparsity

- 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]

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

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

On which domain $V > 0$? $f = V - \sum_{i=1}^N \underbrace{\lambda_i}_{>0} x_i^2 (g - x_i^2) \geq 0$

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

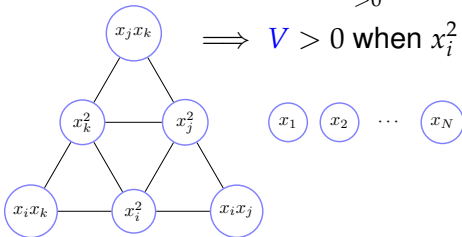
Application to networked systems stability

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

On which domain $V > 0$? $f = V - \sum_{i=1}^N \underbrace{\lambda_i}_{>0} x_i^2 (g - x_i^2) \geq 0$

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

💡 tsp graph G



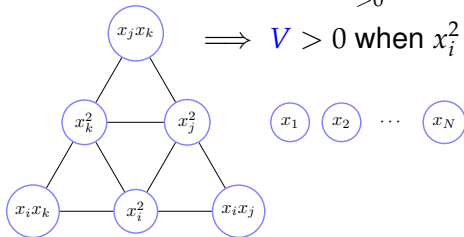
Application to networked systems stability

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

On which domain $V > 0$? $f = V - \sum_{i=1}^N \underbrace{\lambda_i}_{>0} x_i^2 (g - x_i^2) \geq 0$

$\implies V > 0$ 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 $f \geq 0$ for $N = 50$ in ~ 1 second!

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 sparsity

Further topics

Convergence rate of **SPARSE** hierarchies?

💡 (smart) solution extraction for term sparsity

Numerical conditioning of sparse SDP?



Further topics

Convergence rate of **SPARSE** hierarchies?

💡 (smart) solution extraction for term sparsity

Numerical conditioning of sparse SDP?

💡 Tons of applications . . .



Further topics: deep learning

[SIAM News March '21]

“Yet DL has an Achilles’ heel. Current implementations can be highly unstable, meaning that a certain small perturbation to the input of a trained neural network can cause substantial change in its output. This phenomenon is both a nuisance and a major concern for the safety and robustness of DL-based systems in critical applications—like healthcare—where reliable computations are essential”

Further topics: deep learning

💡 “Direct” certification of a classifier with 1 hidden layer

$$\begin{aligned} \max_{\mathbf{x}, \mathbf{z}} \quad & (\mathbf{C}^{i,:} - \mathbf{C}^{k,:})\mathbf{z} \\ \text{s.t.} \quad & \begin{cases} \mathbf{z} = \text{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \\ \|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon \end{cases} \end{aligned}$$

Further topics: deep learning

💡 “Direct” certification of a classifier with 1 hidden layer

$$\begin{aligned} \max_{\mathbf{x}, \mathbf{z}} \quad & (\mathbf{C}^{i,:} - \mathbf{C}^{k,:})\mathbf{z} \\ \text{s.t.} \quad & \begin{cases} \mathbf{z} = \text{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \\ \|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon \end{cases} \end{aligned}$$

💡 Monotone equilibrium networks [Winston Kolter '20]

$$\mathbf{z} = \text{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \rightarrow \mathbf{z} = \text{ReLU}(\mathbf{W}\mathbf{z} + \mathbf{A}\mathbf{x} + \mathbf{b})$$

Further topics: deep learning

💡 “Direct” certification of a classifier with 1 hidden layer

$$\begin{aligned} \max_{\mathbf{x}, \mathbf{z}} \quad & (\mathbf{C}^{i,:} - \mathbf{C}^{k,:})\mathbf{z} \\ \text{s.t.} \quad & \begin{cases} \mathbf{z} = \text{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \\ \|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon \end{cases} \end{aligned}$$

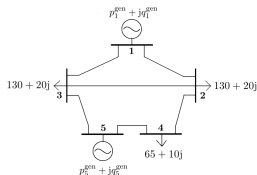
💡 Monotone equilibrium networks [Winston Kolter '20]

$$\mathbf{z} = \text{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \rightarrow \mathbf{z} = \text{ReLU}(\mathbf{W}\mathbf{z} + \mathbf{A}\mathbf{x} + \mathbf{b})$$

💡 “Indirect” with Lipschitz constant/ellipsoid approximation

Further topics: power systems

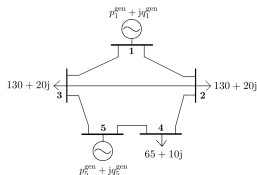
Solving Alternative Current OPF to **global optimality**
→ benchmarks [PGLIB '18] with up to **25 000 buses!**



Further topics: power systems

Solving Alternative Current OPF to **global optimality**
→ benchmarks [PGLIB '18] with up to **25 000 buses!**

COMPLEX vs REAL hierarchy of relaxations?
[D'Angelo Putinar '09, Jozs et al. '18, Magron Wang '21]



Further topics: power systems

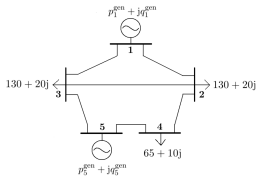
Solving Alternative Current OPF to **global optimality**
→ benchmarks [PGLIB '18] with up to **25 000 buses!**

COMPLEX vs REAL hierarchy of relaxations?

[D'Angelo Putinar '09, Jozs et al. '18, Magron Wang '21]

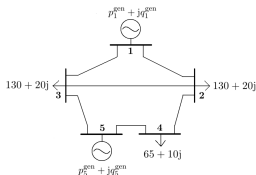
6515_RTE → $n = 7000$ complex variables (14000 real variables)

solved at 0.6% gap within 3 hours on PFCALCUL at LAAS



Further topics: power systems

Solving Alternative Current OPF to **global optimality**
→ benchmarks [PGLIB '18] with up to **25 000 buses!**



COMPLEX vs **REAL** hierarchy of relaxations?

[D'Angelo Putinar '09, Jozs et al. '18, Magron Wang '21]

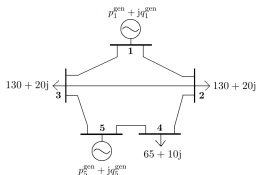
6515_RTE → $n = 7000$ complex variables (14000 real variables)
solved at 0.6% gap within 3 hours on PFCALCUL at LAAS

STABILITY OF LARGE-SCALE POWER SYSTEMS → reachability
analysis of continuous-time systems 💡 **Sparse** [Kundur '07]



Further topics: power systems

Solving Alternative Current OPF to **global optimality**
→ benchmarks [PGLIB '18] with up to **25 000 buses!**



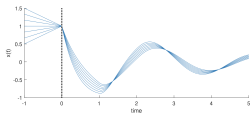
COMPLEX vs **REAL** hierarchy of relaxations?

[D'Angelo Putinar '09, Jozs et al. '18, Magron Wang '21]
6515_RTE → $n = 7000$ complex variables (14000 real variables)
solved at 0.6% gap within 3 hours on PFCALCUL at LAAS


STABILITY OF LARGE-SCALE POWER SYSTEMS → reachability
analysis of continuous-time systems 💡 **Sparse** [Kundur '07]




TIME DELAY SYSTEMS → deteriorate controllers of networked power
systems 💡 **occupation measures**



Further topics: quantum information & more

Ground state energy, trace polynomials for [Werner '89] witnesses
 **symmetric & sparse**

Further topics: quantum information & more

Ground state energy, trace polynomials for [Werner '89] witnesses
 **symmetric & sparse**

RESEARCH DIRECTIONS RELYING ON FREE PROBABILITIES






Minimizer approximation: noncommutative Christoffel-Darboux
kernels and the Siciak function [Beckermann et al. '20]

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. arxiv:2103.12444
-  Wang & Magron. Exploiting term sparsity in Noncommutative Polynomial Optimization. *Computational Optimization & Applications*, arxiv:2010.06956 [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