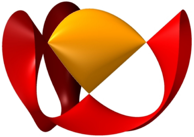
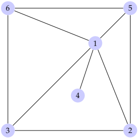
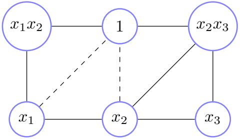
Exploiting Correlative and Term Sparsity in Noncommutative Polynomial Optimization

Victor Magron (LAAS CNRS)
Joint work with Igor Klep, Janez Povh and Jie Wang

SIAM Conference on Optimization
23 July 2021







What is a noncommutative polynomial?

An element of $\mathbb{R}\langle \underline{x} \rangle$

What is a noncommutative polynomial?

An element of $\mathbb{R}\langle \underline{x} \rangle$

Symmetric noncommutative variables $\underline{x} = (x_1, \ldots, x_n)$

$$f = x_1 x_2 x_1^2 - x_1 x_2 \in \mathbb{R}\langle \underline{x} \rangle$$

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

$$f^\star = x_1^2 x_2 x_1 - x_2 x_1$$

sums of hermitian squares (SOHS): $f^\star f$ hermitian square

What is a noncommutative polynomial?

An element of $\mathbb{R}\langle \underline{x} \rangle$

Symmetric noncommutative variables $\underline{x} = (x_1, \ldots, x_n)$

$$f = x_1 x_2 x_1^2 - x_1 x_2 \in \mathbb{R}\langle \underline{x} \rangle$$

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

$$f^\star = x_1^2 x_2 x_1 - x_2 x_1$$

sums of hermitian squares (SOHS): $f^\star f$ hermitian square

Minimal eigenvalue:

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

What is a noncommutative polynomial?

An element of $\mathbb{R}\langle \underline{x} \rangle$

Symmetric noncommutative variables $\underline{x} = (x_1, \ldots, x_n)$

$$f = x_1 x_2 x_1^2 - x_1 x_2 \in \mathbb{R}\langle \underline{x} \rangle$$

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

$$f^\star = x_1^2 x_2 x_1 - x_2 x_1$$

sums of hermitian squares (SOHS): $f^\star f$ hermitian square

Minimal eigenvalue:

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

Minimal (normalized) trace:

$$\operatorname{tr}(f) = \operatorname{tr}(x_1^3 x_2) - \operatorname{tr}(x_1 x_2) \in \mathbb{R}$$

What is a noncommutative polynomial?

An element of $\mathbb{R}\langle \underline{x} \rangle$

Symmetric noncommutative variables $\underline{x} = (x_1, \ldots, x_n)$

$$f = x_1 x_2 x_1^2 - x_1 x_2 \in \mathbb{R}\langle \underline{x} \rangle$$

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

$$f^\star = x_1^2 x_2 x_1 - x_2 x_1$$

sums of hermitian squares (SOHS): $f^\star f$ hermitian square

Minimal eigenvalue:

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

Minimal (normalized) trace:

$$\operatorname{tr}(f) = \operatorname{tr}(x_1^3 x_2) - \operatorname{tr}(x_1 x_2) \in \mathbb{R}$$

$S \subset \operatorname{Sym} \mathbb{R}\langle \underline{x} \rangle$ $\quad X_j \in$ finite von Neumann algebra

Constraints: $\mathcal{D}_S = \{ \underline{X} = (X_1, \ldots, X_n) : s(\underline{X}) \succcurlyeq 0, \quad \forall s \in S \}$

Motivation: quantum information

Entanglement in quantum mechanics
$\rightarrow$ **upper bounds** for violation levels of Bell inequalities

Conditional probabilities $p(i, j)$ generated by Alice & Bob after performing measurements on a shared entangled state

Motivation: quantum information

Entanglement in quantum mechanics
$\rightarrow$ **upper bounds** for violation levels of Bell inequalities

Conditional probabilities $p(i,j)$ generated by Alice & Bob after performing measurements on a shared entangled state
"In the classical world" $\rightarrow \sum c_{i,j} p(i,j) \leqslant C$

Motivation: quantum information

Entanglement in quantum mechanics
$\rightarrow$ **upper bounds** for violation levels of Bell inequalities

Conditional probabilities $p(i,j)$ generated by Alice & Bob after performing measurements on a shared entangled state
"In the classical world" $\rightarrow \sum c_{i,j} p(i,j) \leqslant C$

$p(i,j) = \mathbf{v} x_i y_j \mathbf{v}$ for unit $\mathbf{v}$ & self-adjoint (projection) x_i, y_j
Finding violation $\rightarrow$ solving a **maximal** eigenvalue problem!

Motivation: quantum information

Entanglement in quantum mechanics
$\rightarrow$ **upper bounds** for violation levels of Bell inequalities

Conditional probabilities $p(i,j)$ generated by Alice & Bob after performing measurements on a shared entangled state
"In the classical world" $\rightarrow \sum c_{i,j} p(i,j) \leqslant C$

$p(i,j) = \mathbf{v} x_i y_j \mathbf{v}$ for unit $\mathbf{v}$ & self-adjoint (projection) x_i, y_j
Finding violation $\rightarrow$ solving a **maximal** eigenvalue problem!

[Werner '89] Werner witnesses to separate pure/entangled states

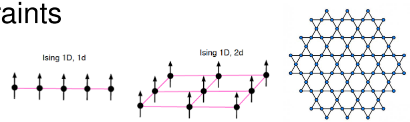
Motivation: condensed matter

Ground-state energy $\Leftrightarrow$ minimal eigenvalue of an Hamiltonian

$$f = \sum_{\langle i,j \rangle} (x_i \, x_j + y_i \, y_j + z_i \, z_j)$$

spin states (x_i, y_i, z_i), constraints

Lattices: 1D 2D Kagome



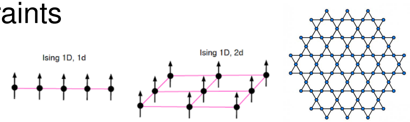
Motivation: condensed matter

Ground-state energy ⇔ minimal eigenvalue of an Hamiltonian

$$f = \sum_{\langle i,j \rangle} (x_i x_j + y_i y_j + z_i z_j)$$

spin states (x_i, y_i, z_i), constraints

Lattices: 1D 2D Kagome



First neighbors interactions : $f = \sum_{i=1}^{n} x_i x_{i+1} + y_i y_{i+1} + z_i z_{i+1}$

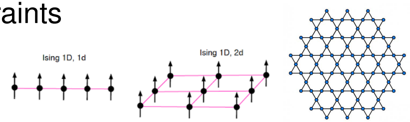
Motivation: condensed matter

Ground-state energy $\Leftrightarrow$ minimal eigenvalue of an Hamiltonian

$$f = \sum_{\langle i,j \rangle} (x_i\, x_j + y_i\, y_j + z_i\, z_j)$$

spin states (x_i, y_i, z_i), constraints

Lattices: 1D 2D Kagome



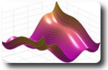
First neighbors interactions : $f = \sum_{i=1}^{n} x_i x_{i+1} + y_i y_{i+1} + z_i z_{i+1}$

Existing $\pm$ efficient techniques: quantum Monte Carlo & variational algorithms $\Rightarrow$ **upper bounds** on minimal energy

Fighting the scalability issue

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

Theory



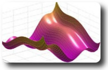
$$\sup \quad \lambda$$

INFINITE LP $\Leftarrow$ with $\quad f - \lambda \geqslant 0$

Fighting the scalability issue

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

Practice



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

FINITE SDP $\Leftarrow$ **fixed** degree

HIERARCHY of **CONVEX** (semidefinite) **PROBLEMS** $\uparrow f_{\min}$

[Lasserre '01]

[Navascués Pironio Acín '08] for NC problems

✓ Relaxations $\Longrightarrow$ **certified lower bounds**

✗ degree d & n vars $\Longrightarrow n^{2d}$ variables

Moment-sums of squares hierarchies

Correlative sparsity

Term sparsity

Moment-sums of squares hierarchies

Correlative sparsity

Term sparsity

Commutative Polynomial Optimization

NP hard General Problem: $f_{\min} := \min_{\mathbf{x} \in \mathcal{D}_S} f(\mathbf{x})$

Semialgebraic set $\mathcal{D}_S = \{\mathbf{x} \in \mathbb{R}^n : s_j(\mathbf{x}) \geqslant 0\}$

Commutative Polynomial Optimization

NP hard General Problem: $f_{\min} := \min_{\mathbf{x} \in \mathcal{D}_S} f(\mathbf{x})$

Semialgebraic set $\mathcal{D}_S = \{\mathbf{x} \in \mathbb{R}^n : s_j(\mathbf{x}) \geqslant 0\}$
$\mathcal{D}_S = [0,1]^2 = \{\mathbf{x} \in \mathbb{R}^2 : x_1(1 - x_1) \geqslant 0, \quad x_2(1 - x_2) \geqslant 0\}$

Commutative Polynomial Optimization

NP hard General Problem: $f_{\min} := \min_{\mathbf{x} \in \mathcal{D}_S} f(\mathbf{x})$

Semialgebraic set $\mathcal{D}_S = \{\mathbf{x} \in \mathbb{R}^n : s_j(\mathbf{x}) \geqslant 0\}$
$\mathcal{D}_S = [0,1]^2 = \{\mathbf{x} \in \mathbb{R}^2 : x_1(1-x_1) \geqslant 0, \quad x_2(1-x_2) \geqslant 0\}$

$$\overbrace{x_1 x_2}^{f} =$$

$$-\frac{1}{8} + \overbrace{\frac{1}{2}\left(x_1 + x_2 - \frac{1}{2}\right)^2}^{\sigma_0} + \overbrace{\frac{1}{2}}^{\sigma_1} \overbrace{x_1(1-x_1)}^{s_1} + \overbrace{\frac{1}{2}}^{\sigma_2} \overbrace{x_2(1-x_2)}^{s_2}$$

Commutative Polynomial Optimization

NP hard General Problem: $f_{\min} := \min_{\mathbf{x} \in \mathcal{D}_S} f(\mathbf{x})$

Semialgebraic set $\mathcal{D}_S = \{\mathbf{x} \in \mathbb{R}^n : s_j(\mathbf{x}) \geqslant 0\}$
$\mathcal{D}_S = [0,1]^2 = \{\mathbf{x} \in \mathbb{R}^2 : x_1(1 - x_1) \geqslant 0, \quad x_2(1 - x_2) \geqslant 0\}$

$$\overbrace{x_1 x_2}^{f} =$$

$$-\frac{1}{8} + \overbrace{\frac{1}{2}\left(x_1 + x_2 - \frac{1}{2}\right)^2}^{\sigma_0} + \overbrace{\frac{1}{2}}^{\sigma_1} \overbrace{x_1(1 - x_1)}^{s_1} + \overbrace{\frac{1}{2}}^{\sigma_2} \overbrace{x_2(1 - x_2)}^{s_2}$$

Sums of squares (SOS) σ_j

Commutative Polynomial Optimization

NP hard General Problem: $f_{\min} := \min_{\mathbf{x} \in \mathcal{D}_S} f(\mathbf{x})$

Semialgebraic set $\mathcal{D}_S = \{\mathbf{x} \in \mathbb{R}^n : s_j(\mathbf{x}) \geqslant 0\}$
$\mathcal{D}_S = [0,1]^2 = \{\mathbf{x} \in \mathbb{R}^2 : x_1(1-x_1) \geqslant 0, \quad x_2(1-x_2) \geqslant 0\}$

$$\overbrace{x_1 x_2}^{f} =$$

$$-\frac{1}{8} + \overbrace{\frac{1}{2}\left(x_1 + x_2 - \frac{1}{2}\right)^2}^{\sigma_0} + \overbrace{\frac{1}{2}}^{\sigma_1} \overbrace{x_1(1-x_1)}^{s_1} + \overbrace{\frac{1}{2}}^{\sigma_2} \overbrace{x_2(1-x_2)}^{s_2}$$

Sums of squares (SOS) σ_j

Quadratic module: $\mathcal{M}(S)_d = \left\{ \sigma_0 + \sum_j \sigma_j s_j, \ \deg \sigma_j \, s_j \leqslant 2d \right\}$

Hierarchies for Polynomial Optimization

Hierarchy of SDP relaxations:

$$\lambda_d := \sup_{\lambda} \left\{ \lambda : f - \lambda \in \mathcal{M}(S)_d \right\}$$

Hierarchies for Polynomial Optimization

Hierarchy of SDP relaxations:

$$\lambda_d := \sup_{\lambda} \Big\{ \lambda : f - \lambda \in \mathcal{M}(S)_d \Big\}$$

Convergence guarantees $\lambda_d \uparrow f_{\min}$ [Lasserre '01] when $\mathcal{M}(S)$ is Archimedean:

$N - \sum x_i^2 \in \mathcal{M}(S)$ for some $N > 0$

Hierarchies for Polynomial Optimization

Hierarchy of SDP relaxations:

$$\lambda_d := \sup_{\lambda} \Big\{ \lambda : f - \lambda \in \mathcal{M}(S)_d \Big\}$$

Convergence guarantees $\lambda_d \uparrow f_{\min}$ [Lasserre '01] when $\mathcal{M}(S)$ is Archimedean:

$N - \sum x_i^2 \in \mathcal{M}(S)$ for some $N > 0$

Can be computed with SDP solvers (`CSDP`, `SDPA`, `MOSEK`)

"No Free Lunch" Rule: $\binom{n+2r}{n}$ SDP variables

Eigenvalue optimization

$$\lambda_{\min} = \inf \{ \langle f(\underline{X}) \mathbf{v}, \mathbf{v} \rangle : \underline{X} \in \mathcal{D}_S^{\infty}, \|\mathbf{v}\| = 1 \}$$

Eigenvalue optimization

$$\lambda_{\min} = \inf \{ \langle f(\underline{X}) \mathbf{v}, \mathbf{v} \rangle : \underline{X} \in \mathcal{D}_S^\infty, \|\mathbf{v}\| = 1 \}$$

$$= \sup \quad \lambda$$
$$\text{s.t.} \quad f(\underline{X}) - \lambda \mathbf{I} \succcurlyeq 0, \quad \forall \underline{X} \in \mathcal{D}_S^\infty$$

Eigenvalue optimization

$$\lambda_{\min} = \inf \{ \langle f(\underline{X}) \mathbf{v}, \mathbf{v} \rangle : \underline{X} \in \mathcal{D}_S^\infty, \|\mathbf{v}\| = 1 \}$$

$$\begin{aligned} = \sup \quad & \lambda \\ \text{s.t.} \quad & f(\underline{X}) - \lambda \mathbf{I} \succcurlyeq 0, \quad \forall \underline{X} \in \mathcal{D}_S^\infty \end{aligned}$$

$\mathcal{M}(S)$ **Archimedean quadratic module**: $N - \sum_i x_i^2 \succcurlyeq 0$

Theorem: NC Putinar's Psatz [Helton-McCullough 02]

$f \succ 0$ on $\mathcal{D}_S^\infty \implies f \in \mathcal{M}(S)$

Eigenvalue optimization

$$\lambda_{\min} = \inf \{ \langle f(\underline{X}) \mathbf{v}, \mathbf{v} \rangle : \underline{X} \in \mathcal{D}_S^\infty, \|\mathbf{v}\| = 1 \}$$

$$= \sup \quad \lambda$$
$$\text{s.t.} \quad f(\underline{X}) - \lambda \mathbf{I} \succcurlyeq 0 \,, \quad \forall \underline{X} \in \mathcal{D}_S^\infty$$

$\mathcal{M}(S)$ **Archimedean quadratic module**: $N - \sum_i x_i^2 \succcurlyeq 0$

Theorem: NC Putinar's Psatz [Helton-McCullough 02]

$f \succ 0$ on $\mathcal{D}_S^\infty \implies f \in \mathcal{M}(S)$

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

replace "$f - \lambda \mathbf{I} \succcurlyeq 0$ on $\mathcal{D}_S^\infty$" by $f - \lambda \mathbf{1} \in \mathcal{M}(S)_d$

Eigenvalue optimization

$$\lambda_{\min} = \inf\{\langle f(\underline{X})\mathbf{v}, \mathbf{v}\rangle : \underline{X} \in \mathcal{D}_S^\infty, \|\mathbf{v}\| = 1\}$$

$$= \sup \quad \lambda$$
$$\text{s.t.} \quad f(\underline{X}) - \lambda \mathbf{I} \succcurlyeq 0, \quad \forall \underline{X} \in \mathcal{D}_S^\infty$$

$\mathcal{M}(S)$ **Archimedean quadratic module**: $N - \sum_i x_i^2 \succcurlyeq 0$

Theorem: NC Putinar's Psatz [Helton-McCullough 02]

$f \succ 0$ on $\mathcal{D}_S^\infty \implies f \in \mathcal{M}(S)$

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

replace "$f - \lambda \mathbf{I} \succcurlyeq 0$ on $\mathcal{D}_S^\infty$" by $f - \lambda \mathbf{1} \in \mathcal{M}(S)_d$

$f - \lambda \mathbf{1} = \sum_i h_i^\star h_i + \sum_j \sum_i t_{ji}^\star s_j t_{ji}$ with h_i, t_{ji} of **bounded** degrees

Trace optimization

$$\mathrm{tr}_{\min} = \inf\{\mathrm{tr}(f(\underline{X})) : \underline{X} \in \mathcal{D}_S\}$$

$$= \sup \quad m$$

$$\text{s.t.} \quad \mathrm{tr}(f(\underline{X}) - m) \geqslant 0\,, \quad \forall \underline{X} \in \mathcal{D}_S$$

Trace optimization

$$\mathrm{tr}_{\min} = \inf\{\mathrm{tr}(f(\underline{X})) : \underline{X} \in \mathcal{D}_S\}$$

$$\begin{aligned} = \sup \quad & m \\ \text{s.t.} \quad & \mathrm{tr}(f(\underline{X}) - m) \geqslant 0, \quad \forall \underline{X} \in \mathcal{D}_S \end{aligned}$$

$\mathrm{tr}_{\min}^{\mathrm{II}_1}$ = minimal trace over the union of type $-\,\mathrm{II}_1$ vN algebras

Disproving Connes' embedding conjecture: $\mathrm{tr}_{\min}^{\mathrm{II}_1} < \mathrm{tr}_{\min}$

Trace optimization

$$\mathrm{tr}_{\min} = \inf\{\mathrm{tr}(f(\underline{X})) : \underline{X} \in \mathcal{D}_S\}$$

$$\begin{aligned} = \sup \quad & m \\ \text{s.t.} \quad & \mathrm{tr}(f(\underline{X}) - m) \geqslant 0, \quad \forall \underline{X} \in \mathcal{D}_S \end{aligned}$$

$\mathrm{tr}_{\min}^{\mathrm{II}_1}$ = minimal trace over the union of type $-\mathrm{II}_1$ vN algebras

💡 Disproving Connes' embedding conjecture: $\mathrm{tr}_{\min}^{\mathrm{II}_1} < \mathrm{tr}_{\min}$

Converging hierarchy with cyclic quadratic modules:

💡 replace "$\mathrm{tr}(f - m) \geqslant 0$ on $\mathcal{D}_S^{\mathrm{II}_1}$" by $f - m\mathbf{1} \in \mathcal{M}^{\mathrm{cyc}}(S)_d$

$\mathcal{M}^{\mathrm{cyc}}(S)_d$ = polynomials with same trace as some from $\mathcal{M}(S)_d$

Moment-sums of squares hierarchies

Correlative sparsity

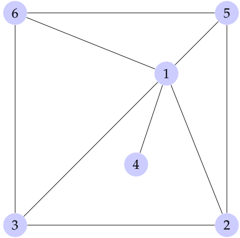
Term sparsity

Correlative sparsity

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

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

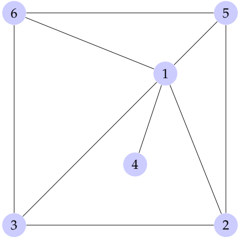
Chordal graph G



Correlative sparsity

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

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



Chordal graph G

maximal cliques I_k

Average size $\kappa \rightsquigarrow \kappa^{2d}$ vars

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

Dense SDP: 210 vars

Sparse SDP: 115 vars

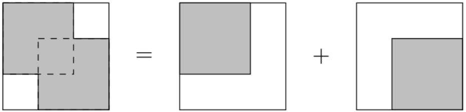
Correlative sparsity

Theorem [Griewank Toint '84]

Chordal graph G with maximal cliques I_1, I_2

$Q_G \succcurlyeq 0$ with nonzero entries at 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



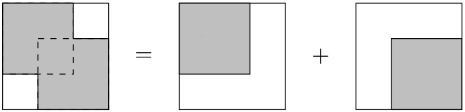
Correlative sparsity

Theorem [Griewank Toint '84]

Chordal graph G with maximal cliques I_1, I_2

$Q_G \succcurlyeq 0$ with nonzero entries at 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 $f = f_1 + f_2$ where f_k involves **only** variables in I_k

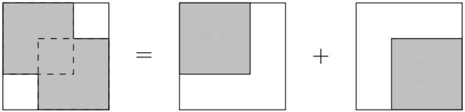
Correlative sparsity

Theorem [Griewank Toint '84]

Chordal graph G with maximal cliques I_1, I_2

$Q_G \succcurlyeq 0$ with nonzero entries at 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 $f = f_1 + f_2$ where f_k involves **only** variables in I_k

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

$f > 0$ on $\{\mathbf{x} : s_j(\mathbf{x}) \geqslant 0\}$
chordal graph G with cliques I_k $\Longrightarrow$
ball constraints for each $\mathbf{x}(I_k)$

$$\boxed{f = \sigma_{01} + \sigma_{02} + \sum_j \sigma_j s_j}$$

SOS σ_{0k} "sees" vars in I_k
σ_j "sees" vars from s_j

Correlative sparsity

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

Theorem [Helton-McCullough 02]

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

Correlative sparsity

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

Theorem [Helton-McCullough 02]

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

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

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

Correlative sparsity

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

Theorem [Helton-McCullough 02]

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

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

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

Correlative sparsity

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

Theorem [Helton-McCullough 02]

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

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

sparse $f \in \Sigma \not\Rightarrow f$ 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 representation!

Correlative sparsity

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

Theorem [Helton-McCullough 02]

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

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

sparse $f \in \Sigma \not\Rightarrow f$ 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 representation! Based on GNS construction & **amalgamation** [Blackadar '78, Voiculescu '85]

Correlative sparsity

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

Theorem [Helton-McCullough 02]

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

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

sparse $f \in \Sigma \not\Rightarrow f$ 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 representation! Based on GNS construction & **amalgamation** [Blackadar '78, Voiculescu '85]

Theorem [Klep Magron Povh '21]

$f \succ 0$ on $\{\underline{x} : s_j(\underline{x}) \succcurlyeq 0\}$

chordal graph G with cliques I_k $\Longrightarrow$

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

$$f = \sum_{k,i} s_{ki}^{\star} s_{ki} + \sum_{j,i} t_{ji}^{\star} s_j t_{ji}$$

s_{ki} "sees" vars in I_k

t_{ji} "sees" vars from s_j

Correlative sparsity

$\mathbf{I}_{3322}$ 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 $\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

Correlative sparsity

$\mathbf{I}_{3322}$ 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 $\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

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

Correlative sparsity

$\mathbf{I}_{3322}$ 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 $\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

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

d	sparse	dense [Pál-Vértesi 18]
2	0.2550008	0.2509397

Correlative sparsity

$\mathbf{I}_{3322}$ 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 $\{(x,y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

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

d	sparse	dense [Pál-Vértesi 18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756

Correlative sparsity

$\mathbf{I}_{3322}$ 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 $\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

💡 $I_k \to \{x_1, x_2, x_3, y_k\}$

d	sparse	dense [Pál-Vértesi 18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3’		0.250854 (1 day)

Correlative sparsity

$\mathbf{I}_{3322}$ 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 $\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

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

d	sparse	dense [Pál-Vértesi 18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3’		0.250875**4** (1 day)
4	0.2508917	

Correlative sparsity

$\mathbf{I}_{3322}$ 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 $\{(x, y) : x_i^2 = x_i, y_j^2 = y_j, x_i y_j = y_j x_i\}$

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

d	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	

Correlative sparsity

$\mathbf{I}_{3322}$ 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

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

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

d	sparse	dense [Pál-Vértesi 18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3'		0.250875**4** (1 day)
4	0.2508917	
5	0.2508763	
6	0.250875397718 0	(1 hour)

PERFORMANCE  VS  ACCURACY

Term sparsity

[Postdoc Wang '19-21] ANR Tremplin-ERC

$$f = x_1^2 - x_1x_2 - x_2x_1 + 3x_2^2 - 2x_1x_2x_1 + 2x_1x_2^2x_1 - x_2x_3 - x_3x_2 + 6x_3^2 + 9x_2^2x_3 + 9x_3^2x_2 - 54x_3x_2x_3 + 142x_3x_2^2x_3$$

[Burgdorf Klep Povh '16] $\rightarrow$ Newton chip method

$$f = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & x_2x_1 & x_3x_2 \end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$$

$\rightsquigarrow \frac{6\times 7}{2} = 28$ "unknown" entries in Q

Term sparsity

[Postdoc Wang '19-21] ANR Tremplin-ERC

COPS

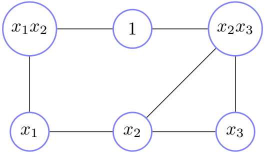
$$f = x_1^2 - x_1x_2 - x_2x_1 + 3x_2^2 - 2x_1x_2x_1 + 2x_1x_2^2x_1 - x_2x_3 - x_3x_2 + 6x_3^2 + 9x_2^2x_3 + 9x_3^2x_2 - 54x_3x_2x_3 + 142x_3x_2^2x_3$$

[Burgdorf Klep Povh '16] $\rightarrow$ Newton chip method

$$f = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & x_2x_1 & x_3x_2 \end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$$

$\rightsquigarrow \frac{6\times 7}{2} = 28$ "unknown" entries in Q

💡 **Term** sparsity pattern graph G



Term sparsity

[Postdoc Wang '19-21] ANR Tremplin-ERC

COPS

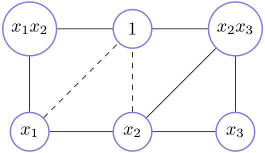
$$f = x_1^2 - x_1x_2 - x_2x_1 + 3x_2^2 - 2x_1x_2x_1 + 2x_1x_2^2x_1 - x_2x_3$$
$$- x_3x_2 + 6x_3^2 + 9x_2^2x_3 + 9x_3^2x_2 - 54x_3x_2x_3 + 142x_3x_2^2x_3$$

[Burgdorf Klep Povh '16] $\rightarrow$ Newton chip method

$$f = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & x_2x_1 & x_3x_2 \end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$$

$\rightsquigarrow \frac{6\times 7}{2} = 28$ "unknown" entries in Q

Term sparsity pattern graph G
+ chordal extension G'



Term sparsity

[Postdoc Wang '19-21] ANR Tremplin-ERC

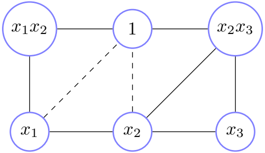
$$f = x_1^2 - x_1x_2 - x_2x_1 + 3x_2^2 - 2x_1x_2x_1 + 2x_1x_2^2x_1 - x_2x_3 - x_3x_2 + 6x_3^2 + 9x_2^2x_3 + 9x_3^2x_2 - 54x_3x_2x_3 + 142x_3x_2^2x_3$$

[Burgdorf Klep Povh '16] $\to$ Newton chip method

$$f = \begin{pmatrix} 1 & x_1 & x_2 & x_3 & x_2x_1 & x_3x_2 \end{pmatrix} \underbrace{Q}_{\succcurlyeq 0} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_2x_3 \end{pmatrix}$$

$\rightsquigarrow \frac{6\times 7}{2} = 28$ "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

At step d of NC hierarchy, one considers vector $\mathbf{W}_d$ of words with length at most d

The tsp graph G has edges E with

$$\{u, v\} \in E \Leftrightarrow u^\star v \in \operatorname{supp}(f) \bigcup \cup_{s \in S} \operatorname{supp}(s) \bigcup \{w^\star w \mid w \in \mathbf{W}_d\}$$

$\rightsquigarrow$ support extension

Term sparsity

At step d of NC hierarchy, one considers vector $\mathbf{W}_d$ of words with length at most d

The tsp graph G has edges E with

$$\{u, v\} \in E \Leftrightarrow u^\star v \in \operatorname{supp}(f) \bigcup \cup_{s \in S} \operatorname{supp}(s) \bigcup \{w^\star w \mid w \in \mathbf{W}_d\}$$

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

Term sparsity

At step d of NC hierarchy, one considers vector $\mathbf{W}_d$ of words with length at most d

The tsp graph G has edges E with

$$\{u, v\} \in E \Leftrightarrow u^\star v \in \operatorname{supp}(f) \bigcup \cup_{s \in S} \operatorname{supp}(s) \bigcup \{w^\star w \mid w \in \mathbf{W}_d\}$$

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

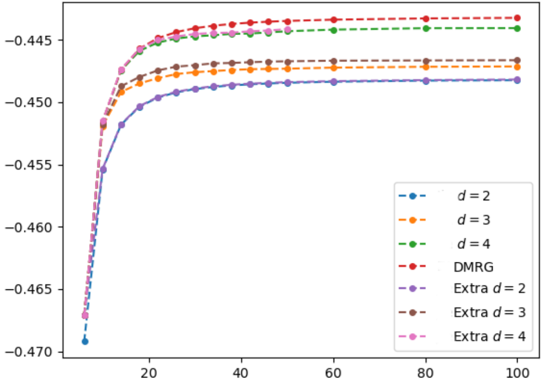
By iteratively performing support extension & chordal extension

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

Two-level hierarchy of lower bounds for $\lambda_{\min}(f)$ on $\mathcal{D}_S^\infty$, indexed by sparse order ℓ and relaxation order d

An example from condensed matter

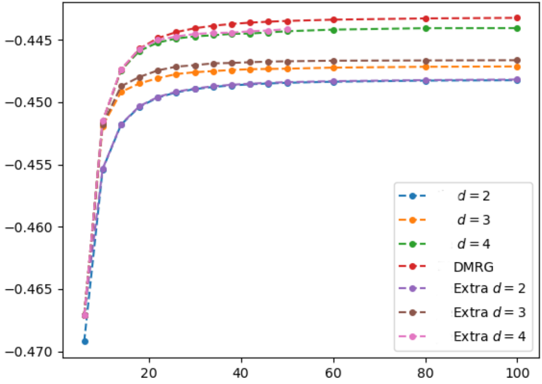
Hamiltonian ground-state energy   1D lattice 



Dense $d = 4$, $n = 10^2 \Rightarrow 10^{19}$ variables (solvers handle $\simeq 10^4$)

An example from condensed matter

Hamiltonian ground-state energy   1D lattice 



Dense $d = 4$, $n = 10^2 \Rightarrow 10^{19}$ variables (solvers handle $\simeq 10^4$)

Sparse solved in $\simeq$ 1 hour on the lab cluster

Conclusion and perspectives

SPARSITY EXPLOITING CONVERGING HIERARCHIES to minimize eigenvalue/trace

FAST IMPLEMENTATION IN JULIA: NCTSSOS

Combine correlative & term sparsity $\rightsquigarrow$ solves problems with 100-1000 variables

Conclusion and perspectives

SPARSITY EXPLOITING CONVERGING HIERARCHIES to minimize eigenvalue/trace

FAST IMPLEMENTATION IN JULIA: NCTSSOS

💡 Combine correlative & term sparsity $\rightsquigarrow$ solves problems with 100-1000 variables

💡 Extension to **trace polynomials** with APPLICATIONS IN QUANTUM INFORMATION: Werner states

Conclusion and perspectives

SPARSITY EXPLOITING CONVERGING HIERARCHIES to minimize eigenvalue/trace

FAST IMPLEMENTATION IN JULIA: NCTSSOS

💡 Combine correlative & term sparsity $\rightsquigarrow$ solves problems with 100-1000 variables

💡 Extension to **trace polynomials** with APPLICATIONS IN QUANTUM INFORMATION: Werner states

💡 SYMMETRIC noncommutative problems?
Ground state energy of hamiltonians 💡 **symmetric** & **sparse**

Thank you for your attention!

`https://homepages.laas.fr/vmagron`

- Wang & Magron. Exploiting term sparsity in Noncommutative Polynomial Optimization. *Computational Optimization & Applications*, arxiv:2010.06956 NCTSSOS
- Klep, Magron & Povh. Sparse Noncommutative Polynomial Optimization. *Math Prog. A*, arxiv:1909.00569 NCSOStools