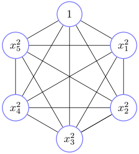
The quest of modeling, certification and efficiency in polynomial optimization

Victor Magron, Chargé de recherche
MAC Team, LAAS CNRS

Soutenance pour l'habilitation à diriger les recherches
LAAS/Visio 25 Mai 2021

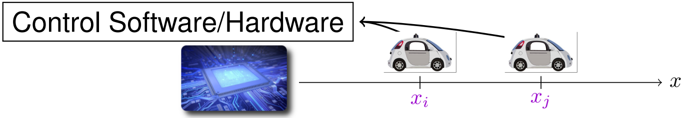


Why optimizing over polynomials?

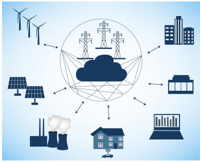
VERIFICATION/ANALYSIS OF COMPLEX NONLINEAR SYSTEMS . . .

SAFETY of critical parts for **computing** $\oplus$ **physical** devices

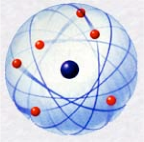
Cars



Smart Grids



Quantum Systems

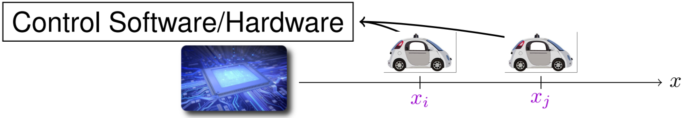


Why optimizing over polynomials?

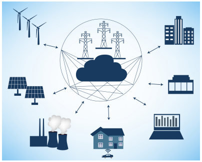
VERIFICATION/ANALYSIS OF COMPLEX NONLINEAR SYSTEMS ...

SAFETY of critical parts for **computing** $\oplus$ **physical** devices

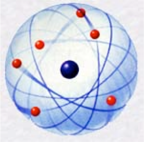
Cars



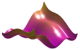
Smart Grids



Quantum Systems



... CAST AS OPTIMIZATION PROBLEM $\rightsquigarrow$ SOLVE **OFFLINE**

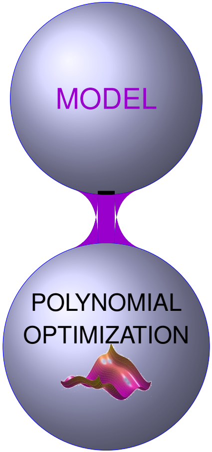
Input: linear  semidefinite  polynomial 

Output: value + numerical/symbolic/formal **certificate**

The quest of modeling: applications

STATIC **Optimization** $f, g \in \mathbb{R}[\mathbf{x}]$

$f =$ sum of squares $\sigma \implies \inf f \geqslant 0$



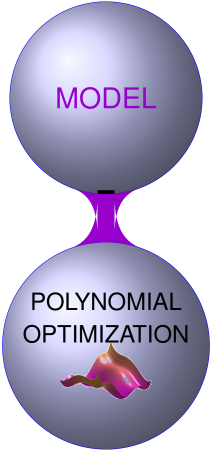
The quest of modeling: applications

STATIC **Optimization** $f, g \in \mathbb{R}[\mathbf{x}]$

f = sum of squares $\sigma \implies \inf f \geqslant 0$

Semialgebraic constraints $\mathbf{X} = \{\mathbf{x} : g(\mathbf{x}) \geqslant 0\}$

$f = \sigma_0 + \sigma_1 g \implies f \geqslant 0$ on $\mathbf{X}$



The quest of modeling: applications

STATIC **Optimization** $f, g \in \mathbb{R}[\mathbf{x}]$

f = sum of squares $\sigma \implies \inf f \geqslant 0$

Semialgebraic constraints $\mathbf{X} = \{\mathbf{x} : g(\mathbf{x}) \geqslant 0\}$

$f = \sigma_0 + \sigma_1 g \implies f \geqslant 0$ on $\mathbf{X}$

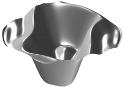
DYNAMICAL **Optimization**

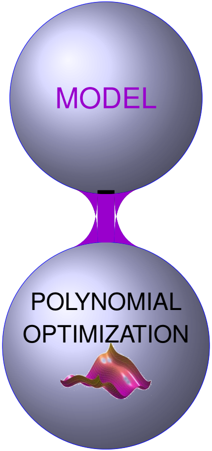
Optimal control

[Henrion Lasserre Prieur Trelat ’08]

Regions of attraction

[Henrion Korda ’14]





The quest of modeling: applications

STATIC **Optimization** $f, g \in \mathbb{R}[\mathbf{x}]$

f = sum of squares $\sigma \implies \inf f \geqslant 0$

Semialgebraic constraints $\mathbf{X} = \{\mathbf{x} : g(\mathbf{x}) \geqslant 0\}$

$f = \sigma_0 + \sigma_1 g \implies f \geqslant 0$ on $\mathbf{X}$

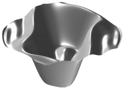
DYNAMICAL **Optimization**

Optimal control

[Henrion Lasserre Prieur Trelat '08]

Regions of attraction

[Henrion Korda '14]

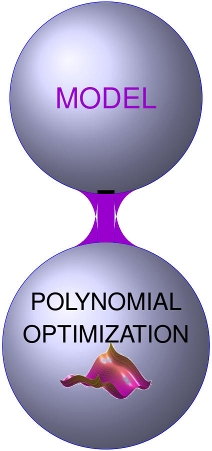


NONCOMMUTATIVE **Optimization** $(x_1 x_2 \neq x_2 x_1)$

Minimal eigenvalue/trace

Useful in quantum information (Bell inequalities)

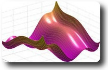
[Navascués Pironio Acín '08]



The quest of modeling: the hierarchy

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

Theory

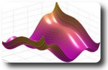
(Primal)		(Dual)
$\inf \int f \, d\mu$		$\sup \; b$
with μ proba $\Rightarrow$	**INFINITE LP**	$\Leftarrow$ with $f - b \geqslant 0$

The quest of modeling: the hierarchy

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

Practice

(Primal **Relaxation**)



(Dual **Strengthening**)

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

$f - b =$ **sum of squares**

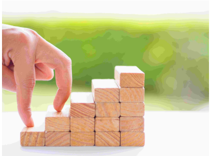
finite number $\Rightarrow$ **SDP** $\Leftarrow$ **fixed** degree

[Lasserre '01] HIERARCHY of **CONVEX PROBLEMS** $\uparrow f_{\min}$

Based on representation of positive polynomials [Putinar '93]

Numerical Solvers $\Longrightarrow$ **Approximate** Certificate

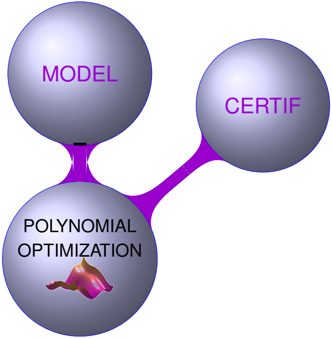
degree d
n vars $\Longrightarrow \binom{n+d}{n}$ **SDP** VARIABLES



The quest of certification (past)

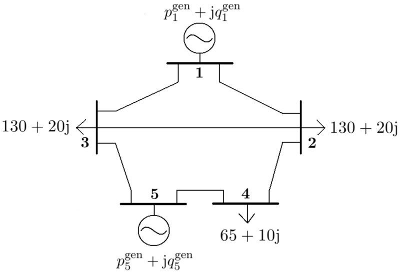
Kepler's conjecture (1611): the max density of sphere packings is $\pi/\sqrt{18}$

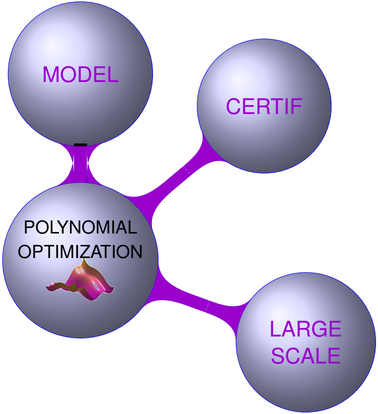
Flyspeck : **F**ormalizing the **p**roof of **K**epler [Hales et al. '94]
Certification of thousands of "tight" non-linear inequalities [Hales et al. '17]



The quest of efficiency (past)

Exploiting sparsity
few terms [Reznick '78]
few correlations
[Lasserre, Waki et al. '06]





Optimal Powerflow $n \simeq \mathbf{10^3}$
[Josz et al. '18]

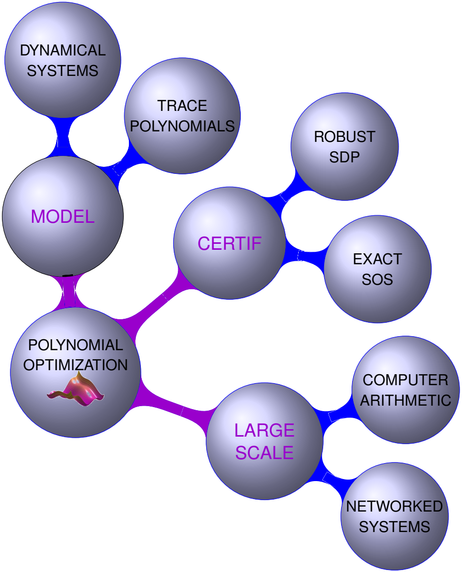
Contributions in polynomial optimization

51 papers =

23 journals

14 conf. proceedings

14 preprints



Introduction

The quest of modeling

The quest of certification

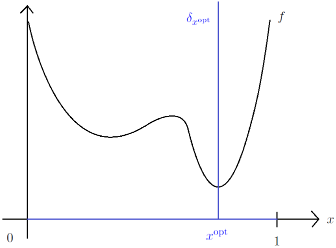
The quest of efficiency

Research projects in polynomial optimization

The quest of modeling: dynamical systems

Characterize a value

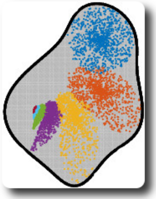
$$f_{\min} = \inf_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x}) = \inf_{\mu \in \mathscr{M}_+(\mathbf{X})} \int_{\mathbf{X}} f \, d\mu$$



Dirac measure at a minimizer

Characterize a set

?



Uniform measure on this set

The quest of modeling: dynamical systems

4 papers with Henrion

Polynomial map $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$

The quest of modeling: dynamical systems

4 papers with Henrion

Polynomial map $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$

Semialgebraic state set constraints $\mathbf{X}$ = either a box ■ or a ball ●

Discrete-time $\mathbf{x}_{t+1} = f(\mathbf{x}_t), \quad \mathbf{x}_t \in \mathbf{X}, \quad t \in \mathbb{N}$

The quest of modeling: dynamical systems

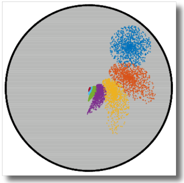
4 papers with Henrion

Polynomial map $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_n(\mathbf{x}))$

Semialgebraic state set constraints $\mathbf{X}$ = either a box ■ or a ball ●

Discrete-time $\mathbf{x}_{t+1} = f(\mathbf{x}_t), \quad \mathbf{x}_t \in \mathbf{X}, \quad t \in \mathbb{N}$

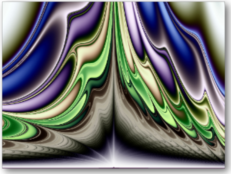
REACHABLE SET



Semialgebraic initial states $\mathbf{X}_0$

All admissible trajectories $\mathbf{X}^\infty$

ATTRACTORS



Support of invariant measure

$\mu(\mathbf{A}) = \mu(f^{-1}(\mathbf{A})) = f_{\#}\mu(\mathbf{A})$

$\forall$ Borel set $\mathbf{A} \in \mathcal{B}(\mathbf{X})$

The quest of modeling: dynamical systems

Reachable set $\mathbf{X}^\infty$

$\mu_0 \in \mathscr{M}_+(\mathbf{X}_0), \quad \mu_1 = f_\# \mu_0 \quad \ldots \quad \mu_t = f_\# \mu_{t-1}$

The quest of modeling: dynamical systems

Reachable set $\mathbf{X}^\infty$

$\mu_0 \in \mathscr{M}_+(\mathbf{X}_0), \quad \mu_1 = f_\# \mu_0 \quad \ldots \quad \mu_t = f_\# \mu_{t-1}$

Average $\nu_t = \sum_{i \leqslant t} \mu_i$

The quest of modeling: dynamical systems

REACHABLE SET $\mathbf{X}^\infty$

$\mu_0 \in \mathscr{M}_+(\mathbf{X}_0), \quad \mu_1 = f_\# \mu_0 \quad \ldots \quad \mu_t = f_\# \mu_{t-1}$

Average $\nu_t = \sum_{i \leqslant t} \mu_i$
The occupation measures μ_t, ν_t, μ_0 satisfy **Liouville's Equation**:

$$\mu_t + \nu_t = f_\# \nu_t + \mu_0 \leftarrow \text{linear in } \mu_t, \nu_t, \mu_0$$

The quest of modeling: dynamical systems

Reachable set $\mathbf{X}^\infty$

$$\mu_0 \in \mathscr{M}_+(\mathbf{X}_0), \quad \mu_1 = f_\# \mu_0 \quad \ldots \quad \mu_t = f_\# \mu_{t-1}$$

Average $\nu_t = \sum_{i \leqslant t} \mu_i$
The occupation measures μ_t, ν_t, μ_0 satisfy **Liouville's Equation**:

$$\mu_t + \nu_t = f_\# \nu_t + \mu_0 \leftarrow \text{linear in } \mu_t, \nu_t, \mu_0$$

💡 The uniform measure on $\mathbf{X}^\infty$ satisfies Liouville and is the (unique) solution of an LP over measures [Magron Henrion et al. '19]

The quest of modeling: dynamical systems

REACHABLE SET $\mathbf{X}^\infty$

$$\mu_0 \in \mathscr{M}_+(\mathbf{X}_0), \quad \mu_1 = f_\# \mu_0 \quad \ldots \quad \mu_t = f_\# \mu_{t-1}$$

Average $\nu_t = \sum_{i \leqslant t} \mu_i$
The occupation measures μ_t, ν_t, μ_0 satisfy **Liouville's Equation**:

$$\mu_t + \nu_t = f_\# \nu_t + \mu_0 \leftarrow \text{linear in } \mu_t, \nu_t, \mu_0$$

💡 The uniform measure on $\mathbf{X}^\infty$ satisfies Liouville and is the (unique) solution of an LP over measures [Magron Henrion et al. '19]

INVARIANT MEASURES → Similar story [Magron Forets Henrion '19]

The quest of modeling: dynamical systems

REACHABLE SET $\mathbf{X}^\infty$

$\mu_0 \in \mathscr{M}_+(\mathbf{X}_0), \quad \mu_1 = f_\# \mu_0 \quad \ldots \quad \mu_t = f_\# \mu_{t-1}$

Average $\nu_t = \sum_{i \leqslant t} \mu_i$
The occupation measures μ_t, ν_t, μ_0 satisfy **Liouville's Equation**:

$$\mu_t + \nu_t = f_\# \nu_t + \mu_0 \leftarrow \text{linear in } \mu_t, \nu_t, \mu_0$$

💡 The uniform measure on $\mathbf{X}^\infty$ satisfies Liouville and is the (unique) solution of an LP over measures [Magron Henrion et al. '19]

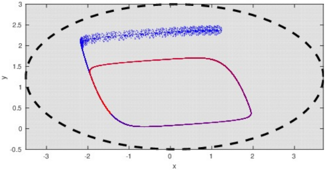
INVARIANT MEASURES $\rightarrow$ Similar story [Magron Forets Henrion '19]

💡 In both cases, the support of all measures is **bounded** to ensure that the LP has an **optimal solution**
💡 Zero duality gap follows from [Barvinok '02]

The quest of modeling: dynamical systems

💡 Use the moment-SOS hierarchy to relax the LP into a hierarchy of SDP

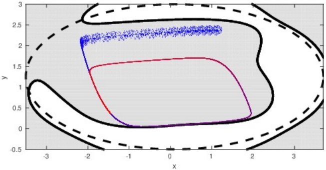
REACHABLE SET for FitzHugh-Nagumo Neuron model



The quest of modeling: dynamical systems

💡 Use the moment-SOS hierarchy to relax the LP into a hierarchy of SDP

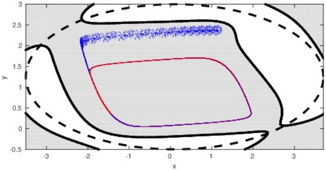
REACHABLE SET for FitzHugh-Nagumo Neuron model



The quest of modeling: dynamical systems

💡 Use the moment-SOS hierarchy to relax the LP into a hierarchy of SDP

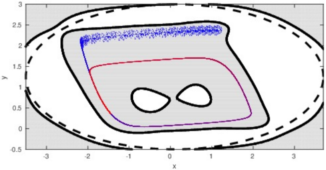
REACHABLE SET for FitzHugh-Nagumo Neuron model



The quest of modeling: dynamical systems

💡 Use the moment-SOS hierarchy to relax the LP into a hierarchy of SDP

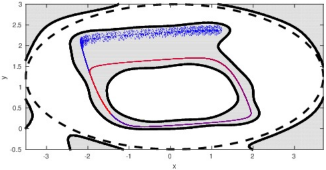
REACHABLE SET for FitzHugh-Nagumo Neuron model



The quest of modeling: dynamical systems

💡 Use the moment-SOS hierarchy to relax the LP into a hierarchy of SDP

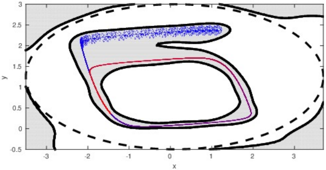
REACHABLE SET for FitzHugh-Nagumo Neuron model



The quest of modeling: dynamical systems

💡 Use the moment-SOS hierarchy to relax the LP into a hierarchy of SDP

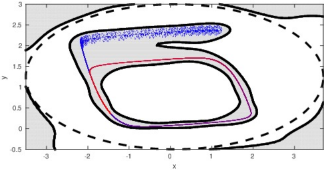
REACHABLE SET for FitzHugh-Nagumo Neuron model



The quest of modeling: dynamical systems

💡 Use the moment-SOS hierarchy to relax the LP into a hierarchy of SDP

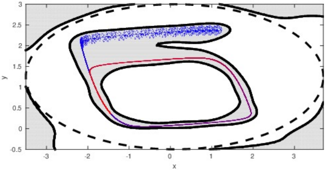
REACHABLE SET for FitzHugh-Nagumo Neuron model



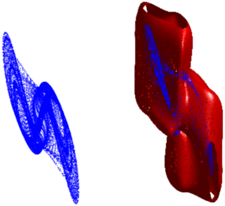
The quest of modeling: dynamical systems

💡 Use the moment-SOS hierarchy to relax the LP into a hierarchy of SDP

REACHABLE SET for FitzHugh-Nagumo Neuron model



ATTRACTOR of Arneodo-Coullet



The quest of modeling: trace polynomials

They arise from entanglement: Werner witnesses [Werner '89], polynomial Bell inequalities [Pozsgay et al. '17]

$$\mathbb{T} = \mathsf{T}\langle \underline{x} \rangle$$

The quest of modeling: trace polynomials

💡 They arise from entanglement: Werner witnesses [Werner '89], polynomial Bell inequalities [Pozsgay et al. '17]

$$\mathbb{T} = \mathsf{T}\langle \underline{x} \rangle$$

Symmetric noncommutative variables $\underline{x} = (x_1, \ldots, x_n)$
& sums of product traces T = pure trace polynomials

$$f = x_1 x_2 x_1^2 - \operatorname{tr}(x_2) \operatorname{tr}(x_1 x_2) \operatorname{tr}(x_1^2 x_2) x_2 x_1 \in \mathbb{T}$$

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

The quest of modeling: trace polynomials

💡 They arise from entanglement: Werner witnesses [Werner '89], polynomial Bell inequalities [Pozsgay et al. '17]

$$\mathbb{T} = \mathsf{T}\langle \underline{x} \rangle$$

Symmetric noncommutative variables $\underline{x} = (x_1, \ldots, x_n)$
& sums of product traces T = pure trace polynomials

$$f = x_1 x_2 x_1^2 - \operatorname{tr}(x_2) \operatorname{tr}(x_1 x_2) \operatorname{tr}(x_1^2 x_2) x_2 x_1 \in \mathbb{T}$$

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

$$\operatorname{tr}(f) = \operatorname{tr}(x_1^3 x_2) - \operatorname{tr}(x_2) \operatorname{tr}(x_1 x_2)^2 \operatorname{tr}(x_1^2 x_2) \in \mathsf{T}$$

The quest of modeling: trace polynomials

💡 They arise from entanglement: Werner witnesses [Werner '89], polynomial Bell inequalities [Pozsgay et al. '17]

$$\mathbb{T} = \mathsf{T}\langle \underline{x} \rangle$$

Symmetric noncommutative variables $\underline{x} = (x_1, \ldots, x_n)$
& sums of product traces T = pure trace polynomials

$$f = x_1 x_2 x_1^2 - \operatorname{tr}(x_2) \operatorname{tr}(x_1 x_2) \operatorname{tr}(x_1^2 x_2) x_2 x_1 \in \mathbb{T}$$

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

$$\operatorname{tr}(f) = \operatorname{tr}(x_1^3 x_2) - \operatorname{tr}(x_2) \operatorname{tr}(x_1 x_2)^2 \operatorname{tr}(x_1^2 x_2) \in \mathsf{T}$$

sums of hermitian squares ($f^\star f$)

The quest of modeling: trace polynomials

💡 They arise from entanglement: Werner witnesses [Werner '89], polynomial Bell inequalities [Pozsgay et al. '17]

$$\mathbb{T} = \mathsf{T}\langle \underline{x} \rangle$$

Symmetric noncommutative variables $\underline{x} = (x_1, \ldots, x_n)$
& sums of product traces T = pure trace polynomials

$$f = x_1 x_2 x_1^2 - \mathrm{tr}(x_2)\,\mathrm{tr}(x_1 x_2)\,\mathrm{tr}(x_1^2 x_2) x_2 x_1 \in \mathbb{T}$$

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

$$\mathrm{tr}(f) = \mathrm{tr}(x_1^3 x_2) - \mathrm{tr}(x_2)\,\mathrm{tr}(x_1 x_2)^2\,\mathrm{tr}(x_1^2 x_2) \in \mathsf{T}$$

sums of hermitian squares ($f^\star f$)
$S \subset \mathrm{Sym}\,\mathbb{T}$ $\quad X_j$ operators from finite von Neumann algebra
Constraints $\{\underline{X} = (X_1, \ldots, X_n) : g(\underline{X}) \succcurlyeq 0, \quad \forall g \in S\}$

The quest of modeling: trace polynomials

💡 Restrict the set of constraints to operators from "nice" von Neumann algebra of type II_1

The quest of modeling: trace polynomials

💡 Restrict the set of constraints to operators from "nice" von Neumann algebra of type II_1
$\implies$ One can minimize pure trace polynomials on such sets!

The quest of modeling: trace polynomials

💡 Restrict the set of constraints to operators from “nice” von Neumann algebra of type II_1
$\Longrightarrow$ One can minimize pure trace polynomials on such sets!

💡 $S[N] = S \cup \{N - x_j^2\}$: add “ball” constraints to ensure convergence

The quest of modeling: trace polynomials

💡 Restrict the set of constraints to operators from “nice” von Neumann algebra of type II_1
$\implies$ One can minimize pure trace polynomials on such sets!

💡 $S[N] = S \cup \{N - x_j^2\}$: add “ball” constraints to ensure convergence
$\implies$ Pure trace variant of Helton-McCullough representation

Theorem [Klep Magron Volcic ’21]

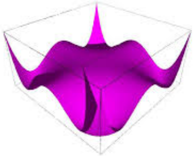
Let $S \subset \operatorname{Sym} \mathbb{T}$ and $f \in$ T. There is a hierarchy of SDP lower bounds converging to $f_{\min}$ on the II_1-von Neumann semialgebraic set associated to $S[N]$.

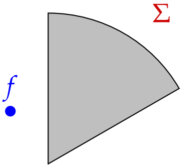
The quest of certification: two-player games

MOTZKIN POLYNOMIAL

$$f = \frac{1}{27} + x_1^2 x_2^4 + x_1^4 x_2^2 - x_1^2 x_2^2$$

$$f \geqslant 0 \text{ but } f \notin \Sigma = \text{ SOS}$$



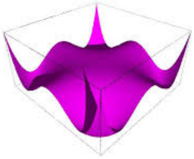


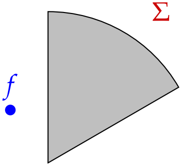
The quest of certification: two-player games

MOTZKIN POLYNOMIAL

$$f = \frac{1}{27} + x_1^2 x_2^4 + x_1^4 x_2^2 - x_1^2 x_2^2$$

$f \geqslant 0$ but $f \notin \Sigma =$ SOS





$$f_{\min} = \min_{x_i \in \mathbb{R}} f(x_1, x_2) = 0 \text{ for } |x_i| = \frac{\sqrt{3}}{3}$$

Moment-SOS hierarchy [Henrion-Lasserre '05]

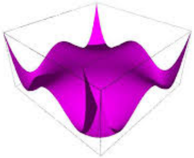
order 3 = "$-\infty$" unbounded SDP $\implies f \notin \Sigma$

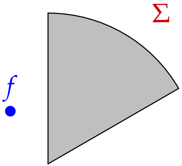
The quest of certification: two-player games

MOTZKIN POLYNOMIAL

$$f = \frac{1}{27} + x_1^2 x_2^4 + x_1^4 x_2^2 - x_1^2 x_2^2$$

$f \geqslant 0$ but $f \notin \Sigma =$ SOS





$$f_{\min} = \min_{x_i \in \mathbb{R}} f(x_1, x_2) = 0 \text{ for } |x_i| = \frac{\sqrt{3}}{3}$$

Moment-SOS hierarchy [Henrion-Lasserre '05]

order 3 = "$-\infty$" unbounded SDP $\Longrightarrow f \notin \Sigma$

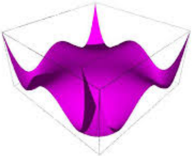
order 4 = "$-\infty$"

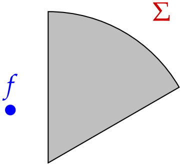
The quest of certification: two-player games

MOTZKIN POLYNOMIAL

$$f = \frac{1}{27} + x_1^2 x_2^4 + x_1^4 x_2^2 - x_1^2 x_2^2$$

$f \geqslant 0$ but $f \notin \Sigma =$ SOS





$$f_{\min} = \min_{x_i \in \mathbb{R}} f(x_1, x_2) = 0 \text{ for } |x_i| = \frac{\sqrt{3}}{3}$$

Moment-SOS hierarchy [Henrion-Lasserre '05]

order 3 = "$-\infty$" unbounded SDP $\implies f \notin \Sigma$

order 4 = "$-\infty$"

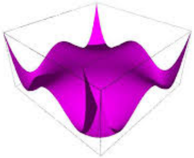
order 5 $\simeq -0.4$

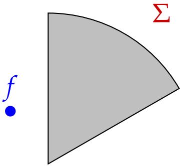
The quest of certification: two-player games

MOTZKIN POLYNOMIAL

$$f = \frac{1}{27} + x_1^2 x_2^4 + x_1^4 x_2^2 - x_1^2 x_2^2$$

$f \geqslant 0$ but $f \notin \Sigma =$ SOS





$$f_{\min} = \min_{x_i \in \mathbb{R}} f(x_1, x_2) = 0 \text{ for } |x_i| = \frac{\sqrt{3}}{3}$$

Moment-SOS hierarchy [Henrion-Lasserre '05]

order 3 = "$-\infty$" unbounded SDP $\Longrightarrow$ $f \notin \Sigma$

order 4 = "$-\infty$"

order 5 $\simeq -0.4$

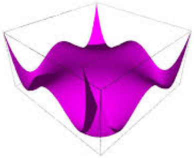
order 8 $\simeq -10^{-8} \oplus$ extraction of optimizers **Paradox**?!

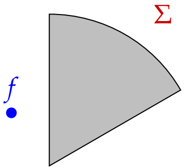
The quest of certification: two-player games

MOTZKIN POLYNOMIAL

$$f = \frac{1}{27} + x_1^2 x_2^4 + x_1^4 x_2^2 - x_1^2 x_2^2$$

$f \geqslant 0$ but $f \notin \Sigma =$ SOS





$$f_{\min} = \min_{x_i \in \mathbb{R}} f(x_1, x_2) = 0 \text{ for } |x_i| = \frac{\sqrt{3}}{3}$$

Moment-SOS hierarchy [Henrion-Lasserre '05]

order 3 = "$-\infty$" unbounded SDP $\implies f \notin \Sigma$

order 4 = "$-\infty$"

order 5 $\simeq -0.4$

order 8 $\simeq -10^{-8} \oplus$ extraction of optimizers **Paradox** ?!

Similar paradox in quantum information [Navascués et al. '13]

The quest of certification: two-player games

η perturbation of the SDP constraints $\implies \tilde{f} = f + \eta \sum_{\beta} \mathbf{x}^{2\beta}$

The quest of certification: two-player games

η perturbation of the SDP constraints $\implies \tilde{f} = f + \eta \sum_{\beta} \mathbf{x}^{2\beta}$

[Lasserre Magron '19] Inaccurate SDP Relaxations . . .

(Primal **Relaxation**)	(Dual **Strengthening**)
$\inf_{\mathbf{y}} \sum_{\alpha} \tilde{f}_\alpha y_\alpha$	$\sup b$
s.t. $\mathbf{M}_d(\mathbf{y}) \succcurlyeq 0$	$\tilde{f} - b = \sigma$
$y_0 = 1$	$\sigma \in \Sigma_d$

The quest of certification: two-player games

η perturbation of the SDP constraints $\implies \tilde{f} = f + \eta \sum_{\beta} \mathbf{x}^{2\beta}$

[Lasserre Magron '19] Inaccurate SDP Relaxations . . .

(Primal **Relaxation**)	(Dual **Strengthening**)
$\inf_{\mathbf{y}} \sum_{\alpha} \tilde{f}_{\alpha} y_{\alpha}$	$\sup \ b$
s.t. $\mathbf{M}_d(\mathbf{y}) \succcurlyeq 0$	$\tilde{f} - b = \sigma$
$y_0 = 1$	$\sigma \in \Sigma_d$

. . . of the **robust** problem $\max_{\tilde{f} \in \mathbf{B}_{\infty}(f,\eta)} \min_{\mathbf{x}} \ \tilde{f}(\mathbf{x})$

The quest of certification: two-player games

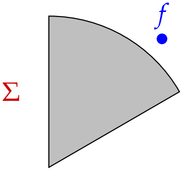
Theorem [Lasserre 06]

For fixed n, any $f \geqslant 0$ can be approximated arbitrarily closely by SOS polynomials.

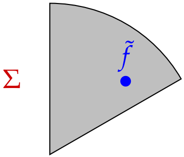
The quest of certification: two-player games

Theorem [Lasserre 06]

For fixed n, any $f \geqslant 0$ can be approximated arbitrarily closely by SOS polynomials.



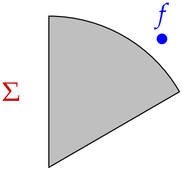
$$\tilde{f} = f + \eta \sum_{|\beta| \leqslant d} \mathbf{x}^{2\beta}$$



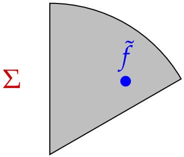
The quest of certification: two-player games

Theorem [Lasserre 06]

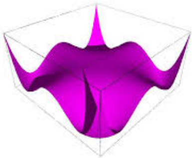
For fixed n, any $f \geqslant 0$ can be approximated arbitrarily closely by SOS polynomials.



$$\tilde{f} = f + \eta \sum_{|\beta| \leqslant d} \mathbf{x}^{2\beta}$$



At fixed η, when $d \nearrow$, $\tilde{f} \in \Sigma$!



$$f + 10^{-7} \sum_{|\beta| \leqslant 4} \mathbf{x}^{2\beta} \in \Sigma$$

Paradox Explanation

The quest of certification: two-player games

$\max - \min$ ROBUST OPTIMIZATION

Player 1 (solver) picks $\tilde{f} \in \mathbf{B}_\infty(f)$ ⇝ **SDP leads**

Player 2 (optimizer) picks an SOS ⇝ **User follows**

The quest of certification: two-player games

$\max - \min$ ROBUST OPTIMIZATION

Player 1 (solver) picks $\tilde{f} \in \mathbf{B}_\infty(f)$ $\rightsquigarrow$ **SDP leads**

Player 2 (optimizer) picks an SOS $\rightsquigarrow$ **User follows**

Convex SDP relaxations $\implies$ $\boxed{\max - \min = \min - \max}$

The quest of certification: two-player games

$\max - \min$ ROBUST OPTIMIZATION

Player 1 (solver) picks $\tilde{f} \in \mathbf{B}_\infty(f) \rightsquigarrow$ **SDP leads**

Player 2 (optimizer) picks an SOS $\rightsquigarrow$ **User follows**

Convex SDP relaxations $\Longrightarrow$ $\boxed{\max - \min = \min - \max}$

$\min - \max$ ROBUST OPTIMIZATION

Player 1 (robust optimizer) picks an SOS $\rightsquigarrow$ **User leads**

Player 2 (solver) picks $\tilde{f} \in \mathbf{B}_\infty(f) \rightsquigarrow$ **SDP follows**

The quest of certification: exact SOS

APPROXIMATE SOLUTIONS

sum of squares of $x_1^2 - 2x_1x_2 + x_2^2$?

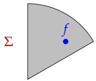
$(1.00001x_1 - 0.99998x_2)^2$!

$$x_1^2 - 2x_1x_2 + x_2^2 \simeq (1.00001x_1 - 0.99998x_2)^2$$
$$x_1^2 - 2x_1x_2 + x_2^2 \neq 1.0000200001x_1^2 - 1.9999799996x_1x_2 + 0.9999600004x_2^2$$

$\boxed{\simeq \quad \rightarrow \quad =}$?

The quest of certification: exact SOS

Win TWO-PLAYER GAME: given $f \in \mathbb{Q}[\mathbf{x}]$ compute $f_i \in \mathbb{Q}[\mathbf{x}], c_i \in \mathbb{Q}^{>0}$ s.t. $f = \sum_i c_i f_i^2$

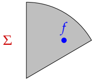


sum of squares of f?

$\simeq$ Output!

The quest of certification: exact SOS

Win TWO-PLAYER GAME: given $f \in \mathbb{Q}[\mathbf{x}]$ compute $f_i \in \mathbb{Q}[\mathbf{x}], c_i \in \mathbb{Q}^{>0}$
s.t. $f = \sum_i c_i f_i^2$



Hybrid Symbolic/Numeric Algorithms

sum of squares of $f - \varepsilon$?

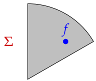
Error Compensation

$\simeq \quad \rightarrow \quad =$

$\simeq$ Output!

The quest of certification: exact SOS

Win TWO-PLAYER GAME: given $f \in \mathbb{Q}[\mathbf{x}]$ compute $f_i \in \mathbb{Q}[\mathbf{x}], c_i \in \mathbb{Q}^{>0}$
s.t. $f = \sum_i c_i f_i^2$



Hybrid Symbolic/Numeric Algorithms

sum of squares of $f - \varepsilon$?

$\simeq$ Output!

Error Compensation

$\simeq \quad \rightarrow \quad =$

PERTURBATION: approximate SOS $f(\mathbf{x})$ - $\varepsilon \sum_\alpha \mathbf{x}^{2\alpha} = \tilde{\sigma} + u$

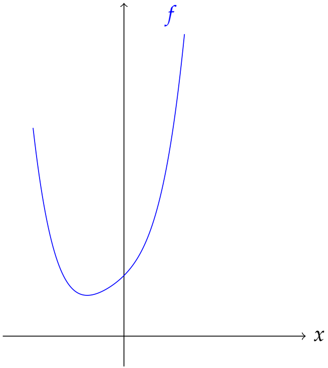
4 papers [Magron Safey El Din Schweighofer '17-21]

Software library: `RealCertify`

The quest of certification: exact SOS

[Chevillard et. al 11]

$f \in \mathbb{Q}[X]$, $\deg f = d = 2k$, $f > 0$



$f = 1 + x + x^2 + x^3 + x^4$

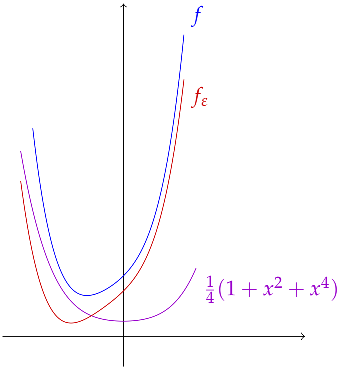
The quest of certification: exact SOS

[Chevillard et. al 11]

$f \in \mathbb{Q}[X]$, deg $f = d = 2k$, $f > 0$

PERTURB: find $\varepsilon \in \mathbb{Q}$ s.t.

$$f_\varepsilon = f - \varepsilon \sum_{i=0}^{k} x^{2i} > 0$$



$$f = 1 + x + x^2 + x^3 + x^4$$

$$\varepsilon = \frac{1}{4}$$

The quest of certification: exact SOS

[Chevillard et. al 11]

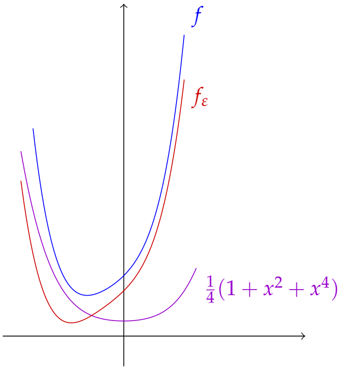
$f \in \mathbb{Q}[X]$, $\deg f = d = 2k$, $f > 0$

PERTURB: find $\varepsilon \in \mathbb{Q}$ s.t.

$$f_\varepsilon = f - \varepsilon \sum_{i=0}^{k} x^{2i} > 0$$

SDP Approximation

$$f - \varepsilon \sum_{i=0}^{k} x^{2i} = \tilde{\sigma} + u$$



$$f = 1 + x + x^2 + x^3 + x^4$$

$$\varepsilon = \frac{1}{4}$$

The quest of certification: exact SOS

[Chevillard et. al 11]

$f \in \mathbb{Q}[X]$, $\deg f = d = 2k$, $f > 0$

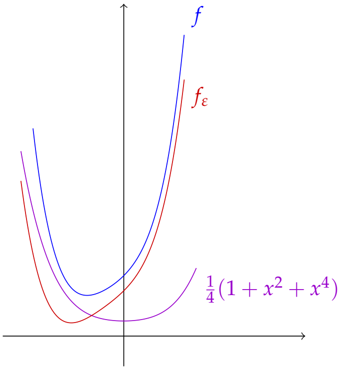
PERTURB: find $\varepsilon \in \mathbb{Q}$ s.t.

$$f_\varepsilon = f - \varepsilon \sum_{i=0}^{k} x^{2i} > 0$$

SDP Approximation

$$f - \varepsilon \sum_{i=0}^{k} x^{2i} = \tilde{\sigma} + u$$

ABSORB small enough u_i
$\implies \varepsilon \sum_{i=0}^{k} x^{2i} + u$ SOS



$f = 1 + x + x^2 + x^3 + x^4$

$\varepsilon = \frac{1}{4}$

The quest of certification: complexity

💡 Analysis weapons: quantifier elimination, root isolation [Cauchy 1832]

$\boxed{n = 1} \to$ polynomial in d, linear in τ = input bit size

The quest of certification: complexity

💡 Analysis weapons: quantifier elimination, root isolation [Cauchy 1832]

$\boxed{n = 1}$ $\rightarrow$ polynomial in d, linear in τ = input bit size

$\boxed{n > 1}$ $\rightarrow$ $\tau^2 d^{d^{\mathcal{O}(n)}}$, one stair higher than critical points [Grigoriev Vorobjov '88, Basu Pollack Roy '98]

The quest of certification: complexity

💡 Analysis weapons: quantifier elimination, root isolation [Cauchy 1832]

$\boxed{n = 1}$ $\rightarrow$ polynomial in d, linear in τ = input bit size

$\boxed{n > 1}$ $\rightarrow$ $\tau^2 d^{d^{\mathcal{O}(n)}}$, one stair higher than critical points [Grigoriev Vorobjov '88, Basu Pollack Roy '98]

💡 Similar algorithms for nonnegative circuits [Magron Wang '20], arithmetic-geometric-exponentials [Magron de Wolff Seidler '19]

The quest of certification: complexity

💡 Analysis weapons: quantifier elimination, root isolation [Cauchy 1832]

$\boxed{n = 1}$ $\rightarrow$ polynomial in d, linear in τ = input bit size

$\boxed{n > 1}$ $\rightarrow$ $\tau^2 d^{d^{\mathcal{O}(n)}}$, one stair higher than critical points [Grigoriev Vorobjov '88, Basu Pollack Roy '98]

💡 Similar algorithms for nonnegative circuits [Magron Wang '20], arithmetic-geometric-exponentials [Magron de Wolff Seidler '19]

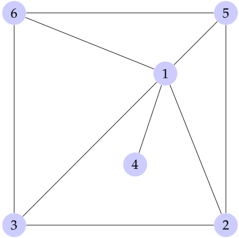
💡 Extension to non SOS polynomials, $\mathbb{C}[\mathbf{x}]$ [PhD Hieu '19-22]

The quest of efficiency: correlative sparsity

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

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

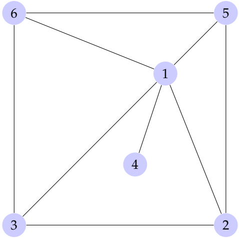
Chordal graph G



The quest of efficiency: 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

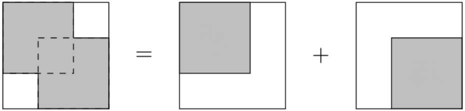
The quest of efficiency: correlative sparsity

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

$\Longrightarrow \; Q_G = P_{I_1}{}^T Q_1 P_{I_1} + P_{I_2}{}^T Q_2 P_{I_2}$ with $Q_k \succcurlyeq 0$ indexed by I_k



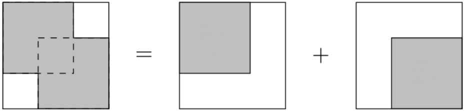
The quest of efficiency: correlative sparsity

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

$\Longrightarrow \; Q_G = {P_{I_1}}^T Q_1 P_{I_1} + {P_{I_2}}^T Q_2 P_{I_2}$ with $Q_k \succcurlyeq 0$ indexed by I_k



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

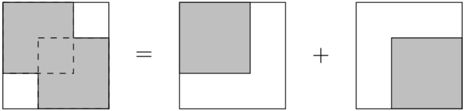
The quest of efficiency: correlative sparsity

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

$\Longrightarrow \; Q_G = {P_{I_1}}^T Q_1 P_{I_1} + {P_{I_2}}^T Q_2 P_{I_2}$ 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 $\{\mathbf{x} : g_j(\mathbf{x}) \geqslant 0\}$
chordal graph G with cliques I_k $\Longrightarrow$
ball constraints for each $\mathbf{x}(I_k)$

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

SOS σ_{0k} "sees" vars in I_k

σ_j "sees" vars from g_j

The quest of efficiency: roundoff errors

[Magron Constantinides Donaldson '17]

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

The quest of efficiency: roundoff errors

[Magron Constantinides Donaldson '17]

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

The quest of efficiency: roundoff errors

[Magron Constantinides Donaldson '17]

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

The quest of efficiency: roundoff errors

[Magron Constantinides Donaldson '17]

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

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

The quest of efficiency: roundoff errors

[Magron Constantinides Donaldson '17]

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

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

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

The quest of efficiency: roundoff errors

[Magron Constantinides Donaldson '17]

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

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

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

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

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

The quest of efficiency: roundoff errors

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

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

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

The quest of efficiency: 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: $15\binom{6+1+4}{6+1} = 4950 \rightsquigarrow 759\epsilon$

The quest of efficiency: 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)

The quest of efficiency: 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)

The quest of efficiency: 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**

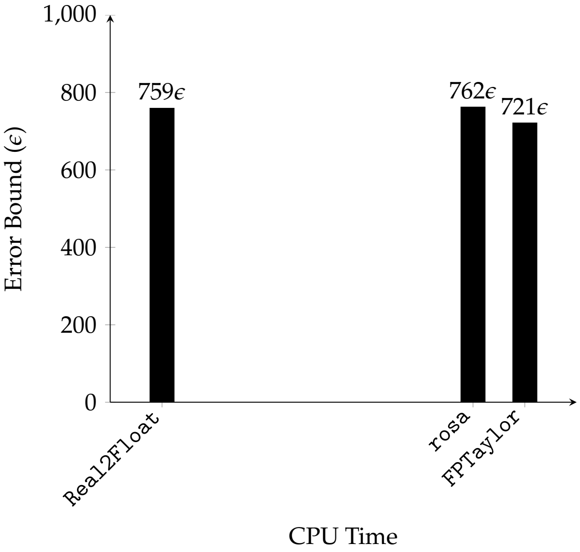
Sparse SDP `Real2Float` tool: $15\binom{6+1+4}{6+1} = 4950 \rightsquigarrow 759\epsilon$

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)

The quest of efficiency: roundoff errors



The quest of efficiency: back to the NC world

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)

The quest of efficiency: back to the NC world

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

The quest of efficiency: back to the NC world

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

The quest of efficiency: back to the NC world

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!

The quest of efficiency: back to the NC world

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]

The quest of efficiency: back to the NC world

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 $\{\underline{x} : g_j(\underline{x}) \succcurlyeq 0\}$

chordal graph G with cliques I_k $\Longrightarrow$

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

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

s_{ki} "sees" vars in I_k

t_{ji} "sees" vars from g_j

The quest of efficiency: back to the NC world

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

The quest of efficiency: back to the NC world

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

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

The quest of efficiency: back to the NC world

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

💡 $C_k \to \{x_1, x_2, x_3, y_k\}$

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

The quest of efficiency: back to the NC world

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

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

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

The quest of efficiency: back to the NC world

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

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

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

The quest of efficiency: back to the NC world

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

💡 $C_k \to \{x_1, x_2, x_3, y_k\}$

level	sparse	dense [Pál-Vértesi 18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3’		0.2508754 (1 day)
4	0.2508917	

The quest of efficiency: back to the NC world

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

💡 $C_k \to \{x_1, x_2, x_3, y_k\}$

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

The quest of efficiency: back to the NC world

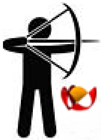
$\mathbf{I}_{3322}$ Bell inequality (entanglement in quantum information)

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

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

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

level	sparse	dense [Pál-Vértesi 18]
2	0.2550008	0.2509397
3	0.2511592	0.2508756
3'		0.2508754 (1 day)
4	0.2508917	
5	0.2508763	
6	0.25087539771800	(1 hour)

PERFORMANCE  VS  ACCURACY

The quest of efficiency: term sparsity

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

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

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

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

The quest of efficiency: term sparsity

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

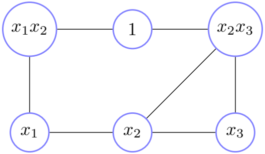
COPS

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

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

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

Term sparsity pattern graph G



The quest of efficiency: term sparsity

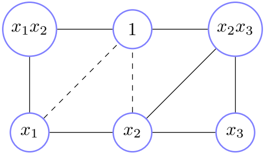
[Postdoc Wang '19-21] ANR Tremplin-ERC

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

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

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

Term sparsity pattern graph G
+ chordal extension



The quest of efficiency: term sparsity

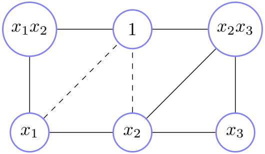
[Postdoc Wang '19-21] ANR Tremplin-ERC

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

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

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

Term sparsity pattern graph G
+ chordal extension



Replace Q by Q_G with nonzero entries at edges of G
$\rightsquigarrow$ 6 + 9 = 15 "unknown" entries in Q_G

The quest of efficiency: term sparsity

Lyapunov functions from NETWORKED SYSTEMS

$$f = \sum_{i=1}^{N} a_i (x_i^2 + x_i^4) - \sum_{i,k=1}^{N} b_{ik} x_i^2 x_k^2 \quad a_i \in [1,2] \quad b_{ik} \in [\frac{0.5}{N}, \frac{1.5}{N}]$$

$\rightsquigarrow \binom{N+2}{2}(\binom{N+2}{2}+1)/2 = 231$ "unknown" entries in Q for $N = 5$

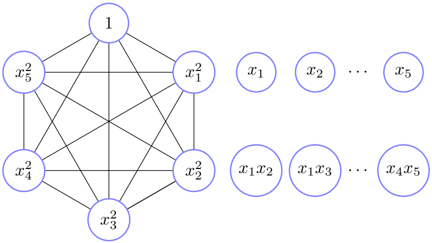
The quest of efficiency: term sparsity

Lyapunov functions from NETWORKED SYSTEMS

$$f = \sum_{i=1}^{N} a_i (x_i^2 + x_i^4) - \sum_{i,k=1}^{N} b_{ik} x_i^2 x_k^2 \quad a_i \in [1, 2] \quad b_{ik} \in [\frac{0.5}{N}, \frac{1.5}{N}]$$

$\rightsquigarrow \binom{N+2}{2}(\binom{N+2}{2}+1)/2 = 231$ "unknown" entries in Q for $N = 5$

term sparsity graph G



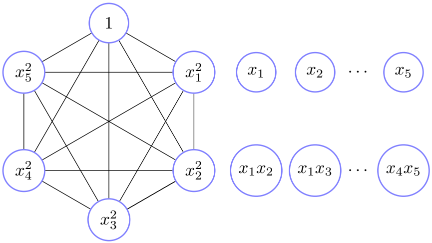
The quest of efficiency: term sparsity

Lyapunov functions from NETWORKED SYSTEMS

$$f = \sum_{i=1}^{N} a_i(x_i^2 + x_i^4) - \sum_{i,k=1}^{N} b_{ik} x_i^2 x_k^2 \quad a_i \in [1,2] \quad b_{ik} \in [\frac{0.5}{N}, \frac{1.5}{N}]$$

$\leadsto \binom{N+2}{2}(\binom{N+2}{2}+1)/2 = 231$ "unknown" entries in Q for $N = 5$

term sparsity graph G



$\leadsto (N+1)^2 = 36$ "unknown" entries in Q_G for $N = 5$

Proof that $f \geqslant 0$ for $N = 80$ in ~ 10 seconds!

The quest of efficiency: term sparsity

- CONVERGENCE GUARANTEES
- handles NC polynomials, $\mathbb{C}[\mathbf{x}]$, joint spectral radii, combo with correlative sparsity $\to$ 7 papers
- Julia libraries `TSSOS` & `NCTSSOS` $\to$ solve problems with $n = 10^3$!
- choice of the CHORDAL EXTENSION: min / max

Introduction

The quest of modeling

The quest of certification

The quest of efficiency

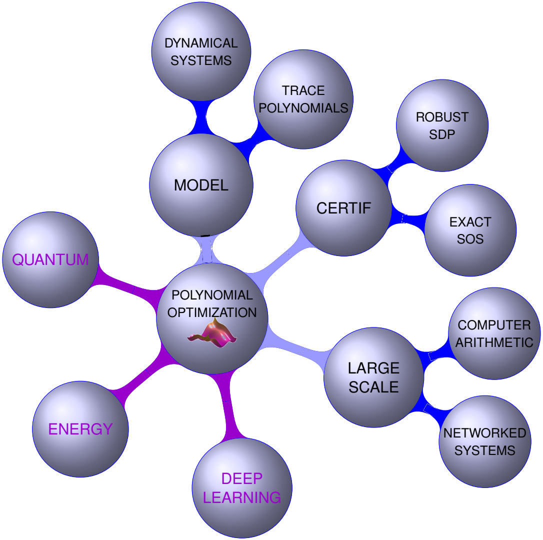
Research projects in polynomial optimization

Research projects in polynomial optimization

Embed polynomial optimization in academic & industrial frameworks

For each project: I will present

1. context + ideas
2. zoom on a specific application target



Quantum information & free probabilities

QUANTUM APPLICATIONS

Ground state energy, trace polynomials for Werner witnesses
💡 **symmetric** & **sparse**

Quantum information & free probabilities

QUANTUM APPLICATIONS

Ground state energy, trace polynomials for Werner witnesses
💡 **symmetric** & **sparse**

RESEARCH DIRECTIONS RELYING ON FREE PROBABILITIES

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

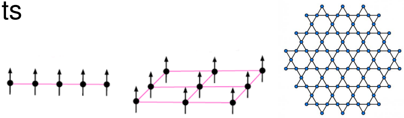
Zoom: condensed matter

Ground-state energy ⇔ minimal eigenvalue of an Hamiltonian

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

spin states (x_i, y_i, z_i), constraints

Lattices: 1D 2D Kagome



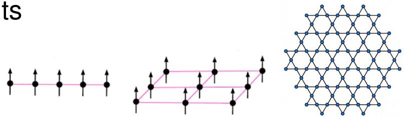
Zoom: condensed matter

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

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

spin states (x_i, y_i, z_i), constraints

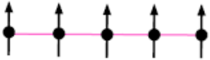
Lattices: 1D 2D Kagome

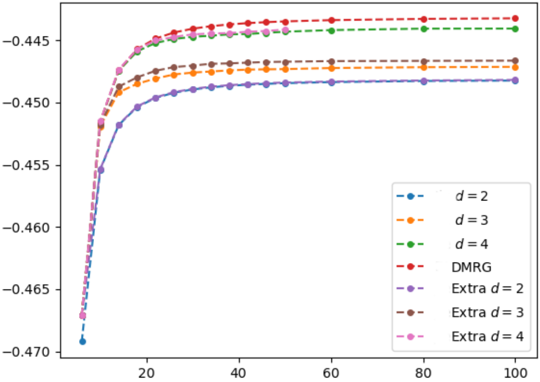


Existing $\pm$ efficient techniques: quantum Monte Carlo & variational algorithms $\Rightarrow$ **upper bounds** on minimal energy

Zoom: condensed matter

Lower bounds of the energy 1D lattice

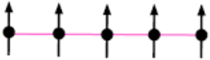


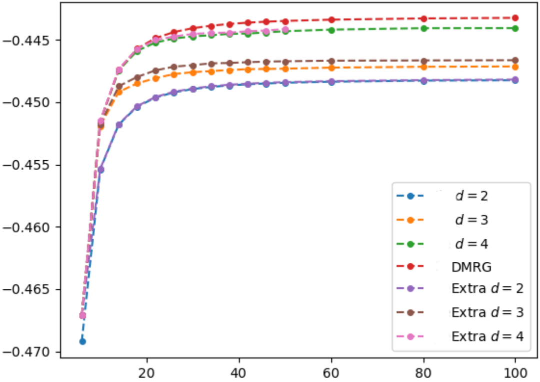


Dense $d = 4$, $n = 10^2 \Rightarrow 10^{19}$ variables (solvers handle $\simeq 10^4$)

Zoom: condensed matter

Lower bounds of the energy 1D lattice





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

Sparse solved within 1 hour on PFCALCUL at LAAS

Quantum & free probabilities: interaction

MODELING AND EFFICIENCY

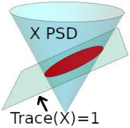
MAC Bhardwaj Korda Lasserre Mai Wang (ANR/CIMI funding)
IRSAMC, LPT Nechita
IMT Belinschi

ICFO Acín's group (Inst. Quantique Occitan funding)
U. Ljubljana Klep Povh (PHC funding)
Texas A&M Volčič
U. Ben Gourion Vinnikov
U. Krakow Huber

Mini-symposium with I. Klep *Computational aspects of commutative and noncommutative positive polynomials* at EUROPEAN CONGRESS OF MATHEMATICIANS

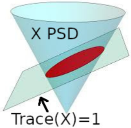
Energy networks

OPTIMAL POWER FLOW → large-scale problems with structured relaxations 💡 **sparse** & **constant trace**
[PhD Mai '19-22]



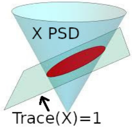
Energy networks

OPTIMAL POWER FLOW → large-scale problems with structured relaxations 💡 **sparse** & **constant trace** [PhD Mai '19-22]



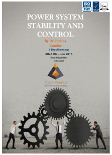
FINITE IMPULSE RESPONSE FILTERS → noise reduction for smart grids 💡 **Certification** [PhD Hieu '19-22]

Energy networks



OPTIMAL POWER FLOW → large-scale problems with structured relaxations **sparse** & **constant trace** [PhD Mai '19-22]

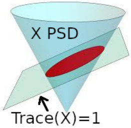
FINITE IMPULSE RESPONSE FILTERS → noise reduction for smart grids **Certification** [PhD Hieu '19-22]



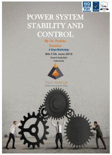
STABILITY OF LARGE-SCALE POWER SYSTEMS → reachability analysis of continuous-time systems **Sparse** [Kundur '07]

Energy networks

OPTIMAL POWER FLOW → large-scale problems with structured relaxations **sparse** & **constant trace** [PhD Mai '19-22]

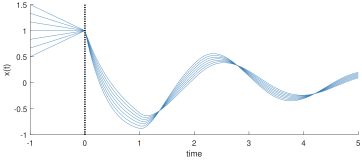


FINITE IMPULSE RESPONSE FILTERS → noise reduction for smart grids **Certification** [PhD Hieu '19-22]



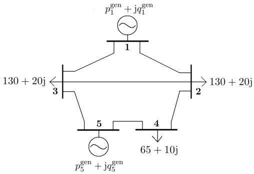
STABILITY OF LARGE-SCALE POWER SYSTEMS → reachability analysis of continuous-time systems **Sparse** [Kundur '07]

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



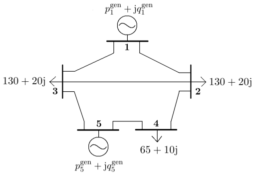
Zoom: optimal power flow

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



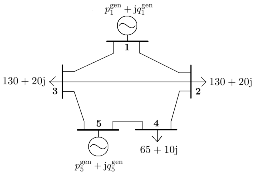
Zoom: optimal power flow

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, Josz et al. '18, Magron Wang '21]
6515_RTE $\rightarrow n = 7000$ complex variables (14000 real variables)
solved at 0.6% gap within 3 hours on PFCALCUL at LAAS

Zoom: optimal power flow

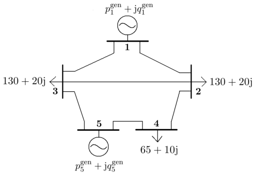


Solving Alternative Current OPF to **global optimality**
$\rightarrow$ benchmarks [PGLIB '18] with up to **25 000 buses**!

COMPLEX vs REAL hierarchy of relaxations?
[D'Angelo Putinar '09, Josz et al. '18, Magron Wang '21]
6515_RTE $\rightarrow n = 7000$ complex variables (14000 real variables) solved at 0.6% gap within 3 hours on PFCALCUL at LAAS

SDP have CONSTANT TRACE PROPERTY [PhD Mai '19-22]

Zoom: optimal power flow

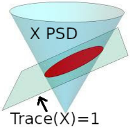


Solving Alternative Current OPF to **global optimality**
$\rightarrow$ benchmarks [PGLIB '18] with up to **25 000 buses**!

COMPLEX vs REAL hierarchy of relaxations?
[D'Angelo Putinar '09, Josz et al. '18, Magron Wang '21]
6515_RTE $\rightarrow n = 7000$ complex variables (14000 real variables)
solved at 0.6% gap within 3 hours on PFCALCUL at LAAS

SDP have CONSTANT TRACE PROPERTY [PhD Mai '19-22]

Replace interior-point solvers by 1st-order methods
$\Rightarrow$ handle matrices of size up to 2000 with more than 1.5 million constraints. . . in 1 hour!



Energy networks: interaction

MODELING, CERTIFICATION & EFFICIENCY

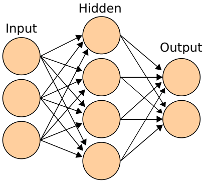
MAC Henrion Korda Lasserre Mai Schlosser Wang
RTE Maeght Panciatici Ruiz (AMIES/RTE funding)
Sorbonne Safey El Din Hieu (POEMA funding)

U. Konstanz Schweighofer (POEMA funding)
Northeastern U. Miller Sznaier

Deep learning

ROBUSTNESS CERTIFICATION OF NEURAL NETWORKS

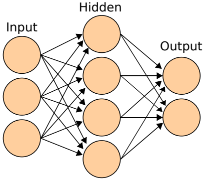
sparse [Chen Lasserre Magron Pauwels '20]



Deep learning

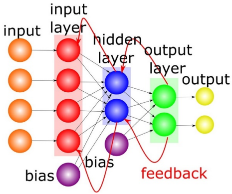
ROBUSTNESS CERTIFICATION OF NEURAL NETWORKS

💡 **sparse** [Chen Lasserre Magron Pauwels '20]



STABILITY ANALYSIS OF RECURRENT NETWORKS

💡 copositive program + integral quadratic constraints
[Megretski Rantzer '97]
[Ebihara Waki Mai Magron Peaucelle Tarbouriech '20]



💡 Formal proofs [Devadze Streif Magron '21]

Zoom: robustness of NN [PhD Chen '19-22]

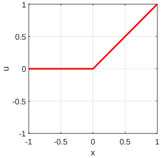
[SIAM News March '21]

"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"

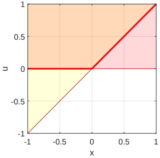
Zoom: robustness of NN [PhD Chen '19-22]

$$\mathbf{x}_0 \in \mathbb{R}^p \longrightarrow \mathbf{z}_0 \in \mathbb{R}^p \longrightarrow \mathbf{z}_1 \in \mathbb{R}^{p_1} \longrightarrow \cdots \longrightarrow \mathbf{z}_m \in \mathbb{R}^{p_m} \quad \mathbf{z}_i = \mathbf{A}_i \mathrm{ReLU}(\mathbf{z}_{i-1}) + \mathbf{b}_i$$

ReLU (left) & its "semialgebraicity" (right)



$$u = \max\{x, 0\}$$



$$u(u - x) = 0, u \geq x, u \geq 0$$

Zoom: robustness of NN [PhD Chen '19-22]

"Direct" certification of a classifier with 1 hidden layer

$$\max_{\mathbf{x},\mathbf{z}} \quad (\mathbf{C}^{i,:} - \mathbf{C}^{k,:})\mathbf{z}$$

$$\text{s.t.} \quad \begin{cases} \mathbf{z} = \mathrm{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \\ ||\mathbf{x} - \mathbf{x}_0|| \leq \epsilon \end{cases}$$

Zoom: robustness of NN [PhD Chen '19-22]

"Direct" certification of a classifier with 1 hidden layer

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

Monotone equilibrium networks [Winston Kolter '20]

$$\mathbf{z} = \mathrm{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \to \mathbf{z} = \mathrm{ReLU}(\mathbf{W}\mathbf{z} + \mathbf{A}\mathbf{x} + \mathbf{b})$$

Zoom: robustness of NN [PhD Chen '19-22]

- "Direct" certification of a classifier with 1 hidden layer

$$\max_{\mathbf{x},\mathbf{z}} \quad (\mathbf{C}^{i,:} - \mathbf{C}^{k,:})\mathbf{z}$$

$$\text{s.t.} \quad \begin{cases} \mathbf{z} = \mathrm{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \\ ||\mathbf{x} - \mathbf{x}_0|| \leq \epsilon \end{cases}$$

- Monotone equilibrium networks [Winston Kolter '20]

$$\mathbf{z} = \mathrm{ReLU}(\mathbf{A}\mathbf{x} + \mathbf{b}) \to \mathbf{z} = \mathrm{ReLU}(\mathbf{W}\mathbf{z} + \mathbf{A}\mathbf{x} + \mathbf{b})$$

- "Indirect" with Lipschitz constant/ellipsoid approximation

Deep learning: interaction

MODELING, CERTIFICATION & EFFICIENCY

MAC Chen Korda Lasserre Mai Peaucelle Tarbouriech (ANITI funding)
UT3 Pauwels

TU Chemnitz Devadze Streif
Kyushu U. Ebihara Waki

Take-away

Why should you do polynomial optimization?

Take-away

Why should you do polynomial optimization?

💡 powerful & accurate MODELING tool

Take-away

Why should you do polynomial optimization?

- powerful & accurate MODELING tool

- CERTIFICATION cost $\simeq$ optimization cost

Take-away

Why should you do polynomial optimization?

- powerful & accurate MODELING tool
- CERTIFICATION cost $\simeq$ optimization cost
- EFFICIENCY guaranteed on structured applications

Thanks to all of my collaborators

And thanks for your presence and attention today!