The moment hierarchy in polynomial optimization: introduction and application to power systems

MADSTAT seminar, Toulouse School of Economics

Adrien Le Franc

15^{th} of December, 2022

LAAS CNRS, Toulouse, France

What is polynomial optimization ?

$$\rho = \min_{x \in \mathcal{X}} f(x) \qquad \text{(POP)}$$

- $\mathcal{X} = \{x \in \mathbb{R}^n \mid g_j(x) \geq 0 \,, \;\; \forall j \in [\![1, m]\!]\,\}$
- f and all g_j are polynomial functions
- we assume that (POP) has a solution
  (e.g. $\mathcal{X}$ is nonempty and compact)

Example (a nonconvex QCQP)

$$\begin{aligned} \rho = \min_{x \in \mathbb{R}^2} \quad & x_1 \\ s.t. \quad & 2x_1 - x_2 + 1 \geq 0 \\ & 2x_1 + x_2 + 1 \geq 0 \\ & x_1^2 + x_2^2 = 1 \end{aligned}$$

All the world is a POP! (...almost)

Example (a MILP — minimum vertex cover)

$$\begin{aligned} \rho = \min_{x \in \mathbb{R}^{|V|}} \quad & \sum_{v \in V} x_v \\ s.t. \quad & x_u + x_v \geq 1 \ , \ \ \forall (u,v) \in E \\ & x_v \in \{0,1\} \ , \ \ \forall v \in V \quad \rightarrow \quad x_v = x_v^2 \end{aligned}$$

Example (an exotic POP)

$$\begin{aligned} \rho = \min_{x \in \mathbb{R}^2} \quad & x_1 x_2 - x_1^3 \\ s.t. \quad & 5 - x_1^4 - 2x_2^2 \geq 0 \\ & 2x_1^2 + x_2 + 1 = 0 \end{aligned}$$

Outline of the presentation

1. The moment-SOS hierarchy for polynomial optimization
2. The AC-OPF problem: POP formulation and scalability issues
3. Conclusion and perspectives

Outline of the section

1. **The moment-SOS hierarchy for polynomial optimization**

   Background notions

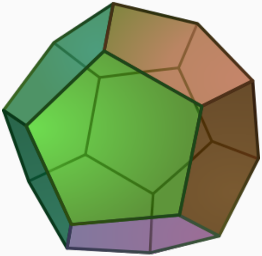
   The moment hierarchy

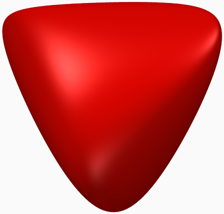
   The sum-of-squares hierarchy

   Some important results

Linear vs semidefinite programming

$$\min_{x \in \mathbb{R}^n} \langle x \, , c \rangle \quad \text{s.t.} \quad \begin{cases} Ax = 0 \\ x \geq 0 \end{cases} \qquad \text{vs} \qquad \min_{X \in \mathcal{S}^n} \langle X \, , C \rangle \quad \text{s.t.} \quad \begin{cases} \mathcal{A}(X) = 0 \\ X \succeq 0 \end{cases}$$





(Images from Wikipedia)

Polynomials

For a polynomial $f \in \mathbb{R}[x_1, \dots, x_n]$ we write

$$f = \sum_{\alpha \in \mathbb{N}^n} f_\alpha x^\alpha$$

where

- $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$ is the monomial labeled by $\alpha \in \mathbb{N}^n$
- $f_\alpha \in \mathbb{R}$ is the corresponding coefficient
- $\{x^\alpha\}_{\alpha \in \mathbb{N}^n}$ is the monomial basis of $\mathbb{R}[x_1, \dots, x_n]$
- for polynomials of degree at most k we truncate $\{x^\alpha\}_{\alpha \in \mathbb{N}^n}$ to labels in $\mathbb{N}_k^n = \left\{ \alpha \in \mathbb{N}^n \;\middle|\; \sum_{i=1}^n \alpha_i \leq k \right\}$

Quadratic module

- The polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ is SOS if $f = \sum_{i \in [\![1,p]\!]} f_i^2$ for some polynomials $f_i \in \mathbb{R}[x_1, \ldots, x_n]$
- The quadratic module of $\mathcal{X}$ is the set

$$\mathcal{Q}(\mathcal{X}) = \left\{ \sigma_0 + \sum_{j \in [\![1,m]\!]} \sigma_j g_j \;\middle|\; \{\sigma_j\}_{j \in [\![0,m]\!]} \text{ are SOS polynomials} \right\}$$

- $\mathcal{Q}(\mathcal{X})$ is said to be Archimedean if for some $M > 0$ $M - \sum_{i \in [\![1,n]\!]} x_i^2 \in \mathcal{Q}(\mathcal{X})$

NB: $\mathcal{Q}(\mathcal{X} \cap \mathbb{B}_r)$ is Archimedean for any $r > 0$

Outline of the section

POP as a moment problem

$$\begin{aligned}\rho = \min_{x\in\mathcal{X}} f(x) &= \inf_{\mu\in\mathcal{M}(\mathcal{X})} \int_{\mathbb{R}^n} f(x)\mu(dx) \\ &= \inf_{\mu\in\mathcal{M}(\mathcal{X})} \sum_{\alpha\in\mathbb{N}^n} f_\alpha \int_{\mathbb{R}^n} x^\alpha \mu(dx) \\ &= \inf\left\{ \sum_{\alpha\in\mathbb{N}^n} f_\alpha y_\alpha \,\middle|\, \text{“}y \text{ has a representing measure on } \mathcal{X}\text{”} \right\}\end{aligned}$$

Proposition (necessary condition)

If $y \in \mathbb{R}^{\mathbb{N}^n_{2d}}$ is the sequence of moments (up to order $2d$) of a measure supported by the set $\mathcal{X} = \{x \in \mathbb{R}^n \mid g_j(x) \geq 0\ ,\ \ \forall j \in [\![1, m]\!]\}$, then

- $M_d(y) \succeq 0$ *(moment matrix)*
- $M_{d-d_j}(g_j y) \succeq 0\ ,\ \ \forall j \in [\![1, m]\!]$ *(localizing matrices)*

Moment matrices ?

$$M_d(y) = \left(y_{\alpha+\beta}\right)_{\alpha\in\mathbb{N}^n_d,\beta\in\mathbb{N}^n_d}$$

$$M_{d-d_j}(g_j y) = \left(\sum_{\gamma\in\mathrm{supp}(g_j)} g_{j,\gamma} y_{\alpha+\beta+\gamma}\right)_{\alpha\in\mathbb{N}^n_{d-d_j},\beta\in\mathbb{N}^n_{d-d_j}} \qquad \left(d_j = \left\lceil \frac{\deg(g_j)}{2} \right\rceil\right)$$

Example

For $n = 2$ and $d = 1$, $M_d(y) \succeq 0$ writes as

$$\begin{pmatrix} y_{00} & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{pmatrix} \succeq 0$$

NB: we need $d \geq \underline{d} = \max\left(\{d_j\}_{j\in[\![1,m]\!]}, \left\lceil \frac{\deg(f)}{2} \right\rceil\right)$ for (POP)

The truncated moment hierarchy

$$\begin{aligned} \rho_d^{\text{MOM}} = \inf_{y \in \mathbb{R}^{|\mathbb{N}^n_{2d}|}} \quad & \sum_{\alpha \in \mathbb{N}^n_{2d}} f_\alpha y_\alpha \\ \text{s.t.} \quad & M_d(y) \succeq 0 \\ & M_{d-d_j}(g_j y) \succeq 0 \,, \ \ \forall j \in [\![1, m]\!] \\ & y_{(0,\dots,0)} = 1 \end{aligned} \tag{MOM$_d$}$$

Theorem (Lasserre [2001])

If the set $\mathcal{X}$ is compact and the quadratic module $\mathcal{Q}(\mathcal{X})$ is Archimedean, then the monotonous non-decreasing sequence of values $\{\rho_d^{MOM}\}_{d \geq \underline{d}}$ of (MOM$_d$) converges to the value ρ of (POP).

Moment relaxations are semidefinite programs

Example (nonconvex QCQP continued)

$$
\begin{aligned}
\rho_1^{\mathrm{MOM}} = \min_{y \in \mathbb{R}^6} \quad & y_{10} \\
\text{s.t.} \quad & \begin{pmatrix} y_{00} & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{pmatrix} \succeq 0 \\
& 2y_{10} - y_{01} + 1 \geq 0 \\
& 2y_{10} + y_{01} + 1 \geq 0 \\
& y_{20} + y_{02} - 1 = 0 \\
& y_{00} = 1
\end{aligned}
$$

Outline of the section

1. **The moment-SOS hierarchy for polynomial optimization**

   Background notions

   The moment hierarchy

   The sum-of-squares hierarchy

   Some important results

POPs and sum-of-squares polynomials

$$\rho = \min_{x \in \mathcal{X}} f(x) = \sup_{\gamma \in \mathbb{R}} \ \gamma\,, \quad \text{s.t. } f(x) - \gamma \geq 0\,, \ \forall x \in \mathcal{X}$$

$$\geq \sup_{\gamma \in \mathbb{R}} \ \gamma\,, \quad \text{s.t. } f - \gamma \in \mathcal{Q}(\mathcal{X})$$

Proposition (Putinar's Positivstellensatz)

If the quadratic module $\mathcal{Q}(\mathcal{X})$ is Archimedean, then $f - \gamma > 0$ on $\mathcal{X}$ implies that $f - \gamma \in \mathcal{Q}(\mathcal{X})$.

The truncated sum-of-squares hierarchy

$$\rho_d^{\mathrm{SOS}} = \sup_{\substack{\gamma \in \mathbb{R} \\ Z_j \in \mathcal{S}^{|\mathbb{N}^n_{d-d_j}|}}} \gamma \qquad (\mathrm{SOS}_d)$$

$$\text{s.t.} \quad f_0 - \gamma = \sum_{j \in [\![0,m]\!]} \langle A_{j,0} \, , Z_j \rangle$$

$$f_\alpha = \sum_{j \in [\![0,m]\!]} \langle A_{j,\alpha} \, , Z_j \rangle \ , \ \forall \alpha \in \mathbb{N}^n_{2d} \setminus \{0\}$$

$$Z_j \succeq 0 \ , \ \forall j \in [\![1, m]\!]$$

Proposition (Josz and Henrion [2016])

In general $\rho_d^{SOS} \leq \rho_d^{MOM}$ (weak duality). If Moreover, the set $\mathcal{X}$ is Archimedean, then $\rho_d^{SOS} = \rho_d^{MOM}$.

Sum-of-squares polynomials ?

σ_j of degree $2d$ is SOS

$$\text{iff } \sigma_j(x) = \left\langle Z_j \, , \left(x^{\alpha+\beta}\right)_{\alpha \in \mathbb{N}^n_d, \beta \in \mathbb{N}^n_d} \right\rangle \text{ for some } Z_j \succeq 0$$

Example

For $n = 2$ and $d = 1$, σ_0 is SOS writes as

$$\sigma_0(x) = \left\langle Z_0 \, , \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1 x_2 \\ x_2 & x_1 x_2 & x_2^2 \end{pmatrix} \right\rangle \text{ for some } Z_0 \succeq 0$$

SOS relaxations are semidefinite programs

Example (nonconvex QCQP continued)

$$
\begin{aligned}
\rho_1^{\mathrm{SOS}} = \max_{\substack{(\gamma, z_1, z_2, z_3) \in \mathbb{R}^4 \\ Z_0 \in \mathcal{S}^3}} \quad & \gamma \\
s.t. \quad & -\gamma = Z_0^{1,1} + z_1 + z_2 - z_3 \\
& 1 = (Z_0^{1,2} + Z_0^{2,1}) + 2z_1 + 2z_2 \\
& 0 = (Z_0^{1,3} + Z_0^{3,1}) - z_1 + z_2 \\
& 0 = (Z_0^{2,3} + Z_0^{3,2}) \\
& 0 = Z_0^{2,2} + z_3 \\
& 0 = Z_0^{3,3} + z_3 \\
& z_1 \geq 0 \,, \quad z_2 \geq 0 \,, \quad z_3 \geq 0 \\
& Z_0 \succeq 0
\end{aligned}
$$

Outline of the section

Finite convergence

Theorem (Nie [2014])

Suppose that the quadratic module $\mathcal{Q}(\mathcal{X})$ is Archimedean.
If the constraint qualification, strict complementarity and second order sufficiency conditions hold at every global minimizer of (POP), then the sequence $\{\rho_d^{MOM}\}_{d \geq \underline{d}}$ converges in finitely many steps.

Example (nonconvex QCQP continued)

bound	ρ_1^{MOM}	ρ_2^{MOM}	$\bar{\rho}$ (NLP)
value	-0.50	0.00	0.00

NB: finite convergence holds "generically"

Extracting a solution

Theorem (Curto and Fialkow [2000], Henrion and Lasserre [2005])

At step $d \geq \underline{d}$ of the moment hierarchy, if an optimal solution y^* of (MOM_d) satisfies $rank\big(M_{d-\underline{d}}(y^*)\big) = rank\big(M_d(y^*)\big)$ then $\rho_d^{MOM} = \rho$ and (POP) has at least $rank\big(M_d(y^*)\big)$ solutions that can be extracted by a linear algebra routine.

Example (nonconvex QCQP continued)

At step $d = 2$, we obtain a solution y^* satisfying

$$\mathrm{spec}\big(M_1(y^*)\big) \simeq \begin{pmatrix} 1.00 \\ 0.99 \\ 0.00 \end{pmatrix} \quad \text{and} \quad \mathrm{spec}\big(M_2(y^*)\big) \simeq \begin{pmatrix} 1.99 \\ 0.99 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \end{pmatrix}$$

and (POP) has exactly $d = 2$ solutions $(0, 1)$ and $(0, -1)$

Key takeaway

Convergence of the Moment-SOS hierarchy of semidefinite programs

$$\rho = \min_{x \in \mathcal{X}} f(x)$$

$$\begin{array}{ccc} \nearrow & & \nwarrow \\ \vdots & & \vdots \\ \rho_{d+1}^{\mathrm{MOM}} & \geq & \rho_{d+1}^{\mathrm{SOS}} \\ \geq & & \geq \\ \rho_{d}^{\mathrm{MOM}} & \geq & \rho_{d}^{\mathrm{SOS}} \end{array}$$

$$d \geq \max\left(\{d_j\}_{j \in [\![1,m]\!]}, \left\lceil \frac{\deg(f)}{2} \right\rceil\right)$$

Outline of the section

2. The AC-OPF problem: POP formulation and scalability issues

The alternative current - optimal power flow problem

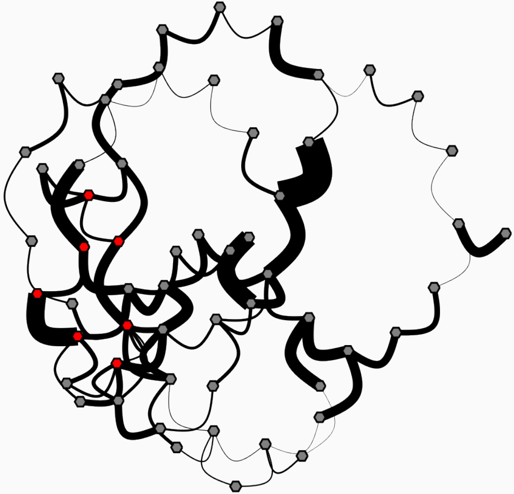
Scalability issues

Sparsity

First order methods

Power grids data

PGLib's case 57 IEEE



Notations for the AC-OPF problem

$$\min_{\substack{v\in\mathbb{C}^{|\mathcal{N}|}\\ s\in\mathbb{C}^{|\mathcal{G}|}\\ s^\ell\in\mathbb{C}^{2|\mathcal{E}|}}} \sum_{g\in\mathcal{G}} C_{2,g}\Re(s_g)^2 + C_{1,g}\Re(s_g) + C_{0,g} \tag{AC-OPF}$$

$$\begin{aligned}
\text{s.t.}\quad & \angle v_i = 0\ ,\ \ \forall i\in\mathcal{N}_r\\
& \underline{S}_g \le s_g \le \overline{S}_g\ ,\ \ \forall g\in\mathcal{G}\\
& \underline{V}_i \le |v_i| \le \overline{V}_i\ ,\ \ \forall i\in\mathcal{N}\\
& \sum_{g\in\mathcal{G}(i)} s_g - L_i - \left(Y_i^s\right)^*|v_i|^2 = \sum_{j\in\mathcal{N}(i)} s^\ell_{i,j}\ ,\ \ \forall i\in\mathcal{N}\\
& s^\ell_{i,j} = \left(Y_{i,j}+Y^c_{i,j}\right)^*\frac{|v_i|^2}{|T_{i,j}|^2} - Y^*_{i,j}\frac{v_i v_j^*}{T_{i,j}}\ ,\ \ \forall (i,j)\in\mathcal{E}\\
& s^\ell_{j,i} = \left(Y_{i,j}+Y^c_{j,i}\right)^*|v_j|^2 - Y^*_{i,j}\frac{v_i^* v_j}{T^*_{i,j}}\ ,\ \ \forall (i,j)\in\mathcal{E}\\
& |s^\ell_{i,j}| \le \overline{S}_{i,j}\ ,\ \ \forall j\in\mathcal{N}(i)\ ,\ \ \forall i\in\mathcal{N}\\
& \underline{\Theta}_{i,j} \le \angle v_i v_j^* \le \overline{\Theta}_{i,j}\ ,\ \ \forall (i,j)\in\mathcal{E}
\end{aligned}$$

Polynomial optimization for AC-OPF

$$v_i = a_i + \mathbb{i} b_i \ , \ \ \forall i \in [\![1, n]\!]$$

Example (complex line power)

$$s^{\ell}_{i,j} = Z_{i,j}|v_i|^2 + Z'_{i,j} v_i v_j^*$$

$$\Longleftrightarrow$$

$$\begin{cases} \Re(s^{\ell}_{i,j}) = \Re(Z_{i,j})(a_i^2 + b_i^2) + \Re(Z'_{i,j})(a_i a_j + b_i b_j) - \Im(Z'_{i,j})(a_j b_i - a_i b_j) \\ \Im(s^{\ell}_{i,j}) = \Im(Z_{i,j})(a_i^2 + b_i^2) + \Im(Z'_{i,j})(a_i a_j + b_i b_j) + \Re(Z'_{i,j})(a_j b_i - a_i b_j) \end{cases}$$

The AC-OPF problem can be written in form (POP)!

Outline of the section

2. **The AC-OPF problem: POP formulation and scalability issues**

   The alternative current - optimal power flow problem

   Scalability issues

   Sparsity

   First order methods

Scalability issues

- AC-OPF IEEE case 57 (no line/angle limits) $\rightarrow$ POP

	(POP)
variables	128
eq. constraints	115
ineq. constraints	128

- POP $\rightarrow$ moment relaxation

	$d = 1$	$d = 2$
size(y)	8.385	12.082.785

ρ_2^{MOM} for PGLib's case 57 IEEE is intractable!
(with current computers and SDP solvers)

Further improvements needed for large scale problems

AC-OPF instances formulate as POPs but...
typical problems at RTE have over 6000 nodes!

what do we do ?

...some ongoing research ideas

- Exploit the sparsity of (POP)
- Circumvent interior point (second order) SDP solvers
- Combining both ? HPC implementations ?

Outline of the section

2. The AC-OPF problem: POP formulation and scalability issues

The alternative current - optimal power flow problem

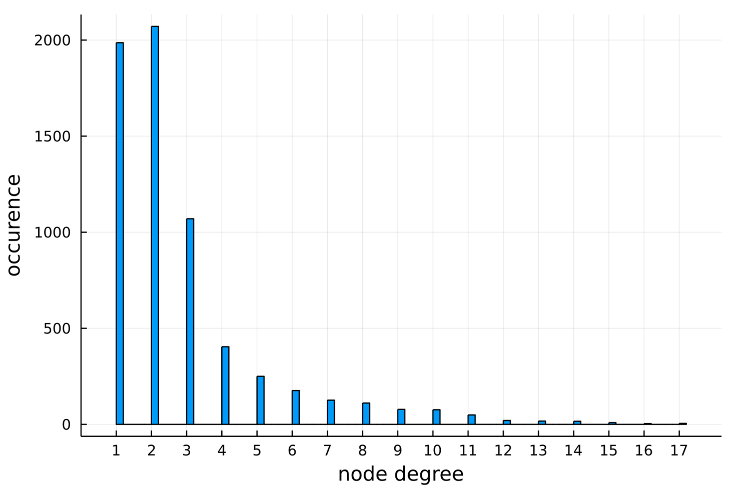
Scalability issues

Sparsity

First order methods

Sparsity structure in power grids

PGLib's case 6468 RTE



Exploiting correlative sparsity for POPs

$$\min_{x \in \mathbb{R}^3} x_1^2 + x_2^2 + x_3^2 + x_1 x_2 + x_2 x_3$$

Exploit absence of $x_1 x_3$ product ? Set $\mathcal{I}_1 = \{1, 2\}$, $\mathcal{I}_2 = \{2, 3\}$

$$M_1(y) = \begin{pmatrix} y_{000} & y_{100} & y_{010} & y_{001} \\ y_{100} & y_{200} & y_{110} & y_{101} \\ y_{010} & y_{110} & y_{020} & y_{011} \\ y_{001} & y_{101} & y_{011} & y_{002} \end{pmatrix} \succeq 0 \quad \text{vs} \quad \begin{cases} M_1(y|\mathcal{I}_1) \succeq 0 \\ M_1(y|\mathcal{I}_2) \succeq 0 \end{cases}$$

Reduce moment variables (MOM_d) / matrices size (SOS_d)

A sparse moment hierarchy

$$\begin{aligned} \rho_d^{\text{CS-MOM}} = \inf_{y \in \mathbb{R}^{n(\mathcal{I})}} \quad & \sum_{\alpha \in \text{supp}(f)} f_\alpha y_\alpha \\ \text{s.t.} \quad & M_d(y|\mathcal{I}_k) \succeq 0 \,, \ \forall k \in [\![1, p]\!] \\ & M_{d-d_j}(g_j y|\mathcal{I}_k) \succeq 0 \,, \ \forall j \in [\![1, m]\!] \,, \ \forall k \in [\![1, p]\!] \\ & y_{(0,\dots,0)} = 1 \end{aligned} \tag{CS-MOM$_d$}$$

Theorem (Lasserre [2006])

Suppose that the set $\mathcal{X}$ is compact and the quadratic module $\mathcal{Q}(\mathcal{X})$ is Archimedean. If the variable set $\mathcal{I} = \{\mathcal{I}_k\}_{k \in [\![1,p]\!]}$ satisfies the running intersection property (RIP), then the monotonous non-decreasing sequence of values $\{\rho_d^{\text{CS-MOM}}\}_{d \geq \underline{d}}$ of (CS-MOM$_d$) converges to the value ρ of (POP).

NB: the maximum cliques of a chordal graph satisfy the RIP

Application to AC-OPF

IEEE case 57 after chordal extension + cliques: $\begin{cases} |\mathcal{I}| = 52 \\ \max_{k \in [\![1,p]\!]} |\mathcal{I}_k| = 26 \end{cases}$

- POP $\rightarrow$ sparse moment relaxation

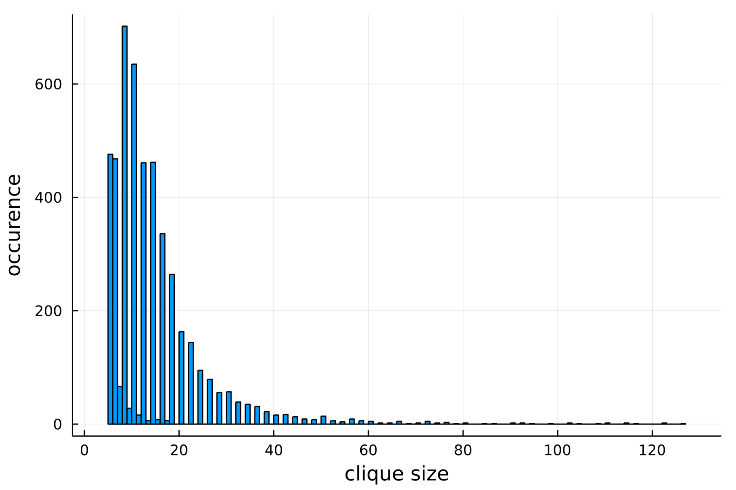
	$d = 1$	$d = 2$
size(y)	1.950	122.286

- numerical result (IEEE case 57 perturbed):

	value	gap to $\bar{\rho}$ (%)	time (s)
$\bar{\rho}$	2433.89	-	4.18
$\rho_2^{\text{CS-MOM}}$	2433.89	0.00	8356.59
$\rho_1^{\text{CS-MOM}}$	2359.58	3.05	0.75

Rip the RIP ?

PGLib's case 6468 RTE



Replace the RIP with grid clustering

PGLib's case 57 IEEE perturbed

instances	value (k$/h)			time (s)	
	$\rho_1(\mathcal{I}^{\text{RTE}})$	$\rho_2(\mathcal{I}^{\text{RTE}})$	$\bar{\rho}$	$\rho_1(\mathcal{I}^{\text{RTE}})$	$\rho_2(\mathcal{I}^{\text{RTE}})$
84	2357.95	2433.88	2433.88	0.58	6834.19
260	3481.80	3542.64	3542.64	0.79	5122.87
267	5265.35	5333.00	5333.00	0.92	6291.01
299	6144.88	6267.34	6267.34	0.63	7302.58
628	2316.76	2483.23	2483.23	0.59	5746.87
683	3129.71	3205.65	3205.64	0.62	6532.20
829	3523.47	3597.40	3597.39	0.50	5414.41
868	3977.68	4067.72	4067.72	0.72	6807.93
974	4446.70	4535.65	4535.64	0.58	6557.98

Further thoughts

- What convergence guarantees when the RIP does not hold ?
- Correlative sparsity can be advantageously combined with term sparsity [Wang et al., 2020]
- Correlative & term sparsity successfully certifies 1%-optimality for AC-OPF instances with up to 6515 nodes [Wang et al., 2022]
- Still challenging for many instances with $+5000$ nodes

Outline of the section

2. The AC-OPF problem: POP formulation and scalability issues

The alternative current - optimal power flow problem

Scalability issues

Sparsity

First order methods

A dual decomposition method

New moment variables $y^k = \{y_\alpha \mid \mathrm{supp}(\alpha) \subseteq \mathcal{I}_k\}$

$$
\begin{aligned}
\rho_d^{\text{CS-MOM}} = \inf_{y \in \mathbb{Y}} \quad & \sum_{k \in [\![1,p]\!]} \left\langle y^k , f^k \right\rangle \\
\text{s.t.} \quad & M_{d-d_j}(g_j y^k) \succeq 0 \,, \;\; \forall j \in \mathcal{J}_k^0 \,, \;\; \forall k \in [\![1,p]\!] \\
& y^k_{(0,\ldots,0)} = 1 \,, \;\; \forall k \in [\![1,p]\!] \\
& B^k_{k,k'} y^k + B^{k'}_{k,k'} y^{k'} = 0 \,, \;\; \forall (k,k') \in \mathcal{P} \qquad (*)
\end{aligned}
$$

Next, we dualize the coupling constraints $(*)$

New dual problem

$$\theta_d^{\text{CS-MOM}} = \sup_{\lambda \in \mathbb{\Lambda}} \Phi(\lambda)$$

where $\Phi = \sum_{k \in [\![1,p]\!]} \phi^k$ and each ϕ^k is a (smaller) semidefinite program

$$\begin{aligned} \phi^k(\lambda) = \inf_{y^k \in \mathbb{Y}^k} \quad & \left\langle y^k \, , f^k + A^k \lambda \right\rangle \\ \text{s.t.} \quad & M_{d-d_j}(g_j y^k) \succeq 0 \, , \ \forall j \in \mathcal{J}_k^0 \\ & y^k_{(0,\dots,0)} = 1 \end{aligned}$$

When strong duality holds $\theta_d^{\text{CS-MOM}} = \rho_d^{\text{CS-MOM}}$

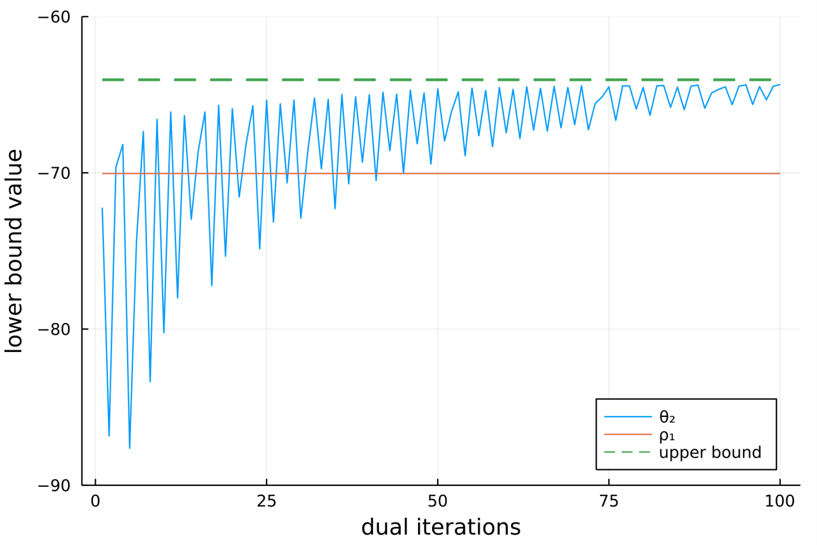
Numerical example (1/2)

A randomly generated QCQP with $n = 20$ variables

	value	gap to $\bar{\rho}$ (%)	time (s)
$\bar{\rho}$	-64.0435	–	3.129
$\theta_2^{\text{CS-MOM}}$	-64.3478	0.47	1258.811
$\rho_2^{\text{CS-MOM}}$	$*$	$*$	$*$
$\rho_1^{\text{CS-MOM}}$	-70.0464	9.37	0.02

- Computing $\rho_2^{\text{CS-MOM}}$ with Mosek and 8GB of RAM gives a memory error
- Computing $\theta_2^{\text{CS-MOM}}$ is less memory demanding (...but quite slow!)

Numerical example (2/2)



Further thoughts

- Smoothing the dual problem ?
- Avoid solving SDP subproblems ?
- Approximate subproblems ? [Dathathri et al., 2020]
- Other first order SDP approaches [Yurtsever et al., 2021]

Outline of the section

3. Conclusion and perspectives

Time to conclude !

Conclusion and perspectives

- POPs cover a large class of optimization problems
- The Moment-SOS hierarchy solves POPs via SDP relaxations
- Not covered in this talk: extensions to Control and PDE [Henrion et al., 2020], complex optimization [Josz and Molzahn, 2018], noncommutative optimization [Klep et al., 2022] etc...
- Large scale applications e.g. in AC-OPF are still challenging
- Combined use of sparsity and first order methods offer new perspectives

References I

Raúl Curto and Lawrence Fialkow. The truncated complex -moment problem. *Transactions of the American mathematical society*, 352(6): 2825–2855, 2000.

Sumanth Dathathri, Krishnamurthy Dvijotham, Alexey Kurakin, Aditi Raghunathan, Jonathan Uesato, Rudy R Bunel, Shreya Shankar, Jacob Steinhardt, Ian Goodfellow, Percy S Liang, et al. Enabling certification of verification-agnostic networks via memory-efficient semidefinite programming. *Advances in Neural Information Processing Systems*, 33: 5318–5331, 2020.

Didier Henrion and Jean-Bernard Lasserre. Detecting global optimality and extracting solutions in gloptipoly. In *Positive polynomials in control*, pages 293–310. Springer, 2005.

Didier Henrion, Milan Korda, and Jean Bernard Lasserre. *Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*, volume 4. World Scientific, 2020.

Cédric Josz and Didier Henrion. Strong duality in lasserres hierarchy for polynomial optimization. *Optimization Letters*, 10(1):3–10, 2016.

Cédric Josz and Daniel K Molzahn. Lasserre hierarchy for large scale polynomial optimization in real and complex variables. *SIAM Journal on Optimization*, 28(2):1017–1048, 2018.

Igor Klep, Victor Magron, and Janez Povh. Sparse noncommutative polynomial optimization. *Mathematical Programming*, 193(2): 789–829, 2022.

Jean B Lasserre. Convergent sdp-relaxations in polynomial optimization with sparsity. *SIAM Journal on Optimization*, 17(3):822–843, 2006.

References III

Jean-Bernard Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on optimization*, 2001.

Jiawang Nie. Optimality conditions and finite convergence of lasserres hierarchy. *Mathematical programming*, 146(1):97–121, 2014.

Jie Wang, Victor Magron, Jean B Lasserre, and Ngoc Hoang Anh Mai. Cs-tssos: Correlative and term sparsity for large-scale polynomial optimization. *arXiv preprint arXiv:2005.02828*, 2020.

Jie Wang, Victor Magron, and Jean B Lasserre. Certifying global optimality of ac-opf solutions via sparse polynomial optimization. *Electric Power Systems Research*, 213:108683, 2022.

Alp Yurtsever, Joel A Tropp, Olivier Fercoq, Madeleine Udell, and Volkan Cevher. Scalable semidefinite programming. *SIAM Journal on Mathematics of Data Science*, 3(1):171–200, 2021.