Near-Global Solutions of Nonlinear Power Optimization Problems: Theory, Numerical Algorithm, and Case Studies

Javad Lavaei

Department of Electrical Engineering
Columbia University

Joint work with:
Ramtin Madani, Abdulrahman Kalbat , Morteza Ashraphijuo, Salar Fattahi (Columbia University)
Somayeh Sojoudi (New York University)
Ross Baldick (University of Texas-Austin)

Polynomial Optimization

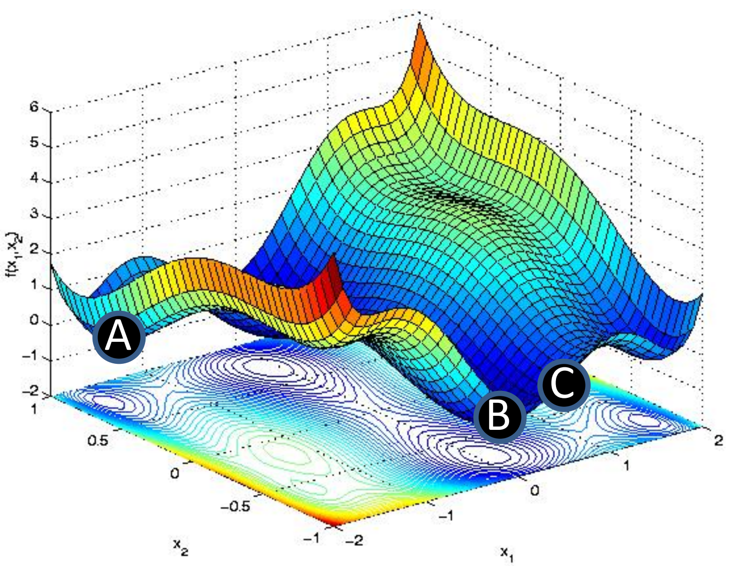
- **Polynomial Optimization:**

$$
\begin{aligned}
\min \quad & x^T M x \\
\text{s.t.} \quad & x_i^2 = 1, \quad i = 1, 2, ..., n
\end{aligned}
$$

Special case: Combinatorial optimization and integer programming problems

Very hard to solve

- **Different types of solutions:**



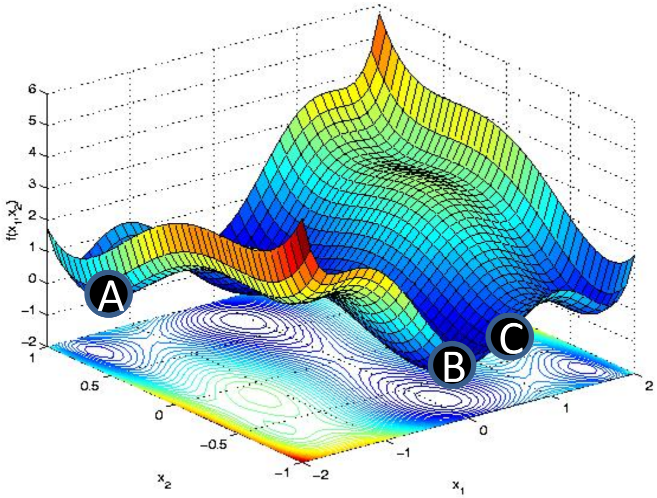
Point A: Local solution

Point B: Global solution

Point C: Near-global solution

Focus of this talk

Objective



$$\text{Optimality Guarantee} \geq \frac{\text{Global cost}}{\text{Near-global cost}} \times 100$$

A number between 0 % and 100 %

❖ **Focus of talk:** Find a near-global solution with a high optimality guarantee (close to 100%).

Problem 1: Convexification
Design a convex problem whose solution is near global for original problem.

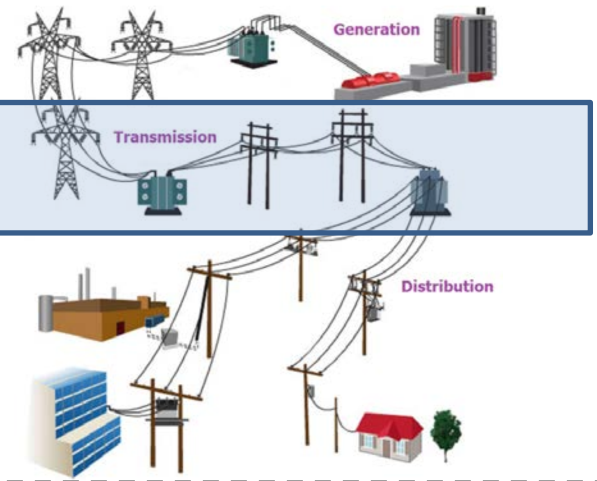
Problem 2: Numerical Algorithm
Design an algorithm to solve the (high-dim) convex program numerically.

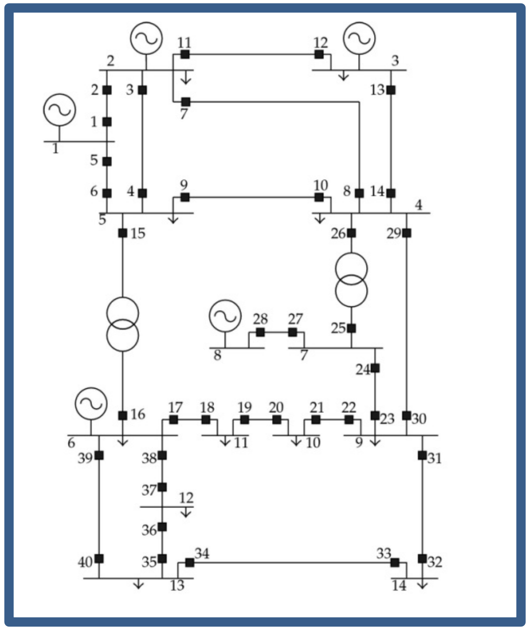
❖ **Approach:** Low-rank optimization, matrix completion, graph theory, convexification

Power Systems

❑ Power system:

- ❖ A large-scale system consisting of generators, loads, lines, etc.
- ❖ Used for generating, transporting and distributing electricity.





ISO, RTO, TSO

1. Optimal power flow (OPF)
2. Security-constrained OPF
3. State estimation
4. Network reconfiguration
5. Unit commitment
6. Dynamic energy management

NP-hard

(real-time operation and market)

Convexification

$$\operatorname{trace}\{M_0 x x^H\}$$

$$\begin{aligned} \min_{x \in \mathbf{C}^n} \quad & x^H M_0 x \\ \text{s.t.} \quad & x^H M_i x \le a_i, \quad i = 1, 2, \dots, m \end{aligned}$$

- **Transformation:** Replace xx^H with W.
- W is positive semidefinite and rank 1

SDP relaxation

$$\begin{aligned} \min_{W \in \mathbb{H}^n} \quad & \operatorname{trace}\{M_0 W\} \\ \text{s.t.} \quad & \operatorname{trace}\{M_i W\} \le a_i, \quad i = 1, 2, \dots, m \\ & W \succeq 0 \end{aligned}$$

- **Rank-1 SDP:** Recovery of a global solution x

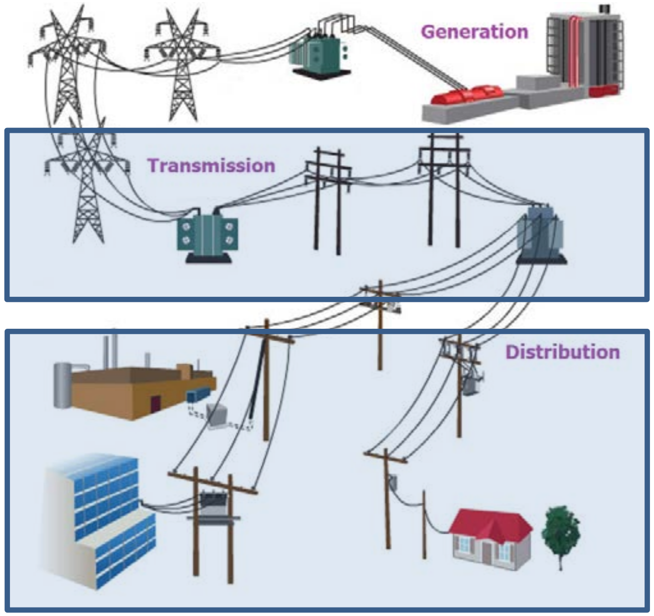
Penalized SDP

$$\begin{aligned} \min_{W} \quad & \operatorname{trace}\{M_0 W\} + \lambda\, g(W) \\ \text{s.t.} \quad & \operatorname{trace}\{M_i W\} \le a_i, \quad i = 1, 2, \dots, m \\ & W \succeq 0 \end{aligned}$$

- **Rank-1 penalized SDP:** Recovery of a near-global solution x

Exactness of Relaxation

- SDP is not exact in general.
- SDP is exact for IEEE benchmark examples and several real data sets.



cyclic → **Theorem:** Exact under positive LMPs with many transformers.

acyclic → **Theorem:** Exact under positive LMPs.

Physics of power networks (e.g., passivity) reduces computational complexity for power optimization problems.

1. S. Sojoudi and J. Lavaei, "Exactness of Semidefinite Relaxations for Nonlinear Optimization Problems with Underlying Graph Structure," SIOPT, 2014.
2. S. Sojoudi and J. Lavaei, "Physics of Power Networks Makes Hard Optimization Problems Easy to Solve," PES 2012.

Promises of SDP

- ❑ **Observation:** SDP may not be exact for ISOs' large-scale systems (some negative LMPs).
- ❑ **Remedy:** Design a penalized SDP to find a near-global solution.

Case	Cost	Guarantee	Time (sec)
Polish 2383wp	1874322.65	99.316%	529
Polish 2736sp	1308270.20	99.970%	701
Polish 2737sop	777664.02	99.995%	675
Polish 2746wop	1208453.93	99.985%	801
Polish 2746wp	1632384.87	99.962%	699
Polish 3012wp	2608918.45	99.188%	814
Polish 3120sp	2160800.42	99.073%	910

SDP looks very promising for energy applications

1. J. Lavaei and S. Low, "Zero Duality Gap in Optimal Power Flow Problem," IEEE Transactions on Power Systems, 2012.
2. J. Lavaei, D. Tse and B. Zhang, "Geometry of Power Flows and Optimization in Distribution Networks," IEEE Transactions on Power System, 2014.
3. R. Madani, S. Sojoudi and J. Lavaei, "Convex Relaxation for Optimal Power Flow Problem: Mesh Networks," IEEE Transactions on Power Systems, 2015.

Outline

Arbitrary Real/Complex Polynomial Optimization

Conversion

$$\min_{x \in \mathbf{D}^n} \quad x^H M_0 x$$
$$\text{s.t.} \quad x^H M_i x \leq a_i, \quad i = 1, 2, ..., m$$

SDP/ Penalized SDP

$$\min_{W} \quad \text{trace}\{M_0 W\} + \lambda\, g(W)$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \leq a_i, \quad i = 1, 2, ..., m$$
$$W \succeq 0$$

Connection between sparsity and rank? → Proof of existence of low-rank solution using OS and treewidth

How to design penalized SDP? → Propose two methods to design penalty

Design scalable numerical algorithm? → Cheap iterations for large-scale problems

Outline

Arbitrary Real/Complex Polynomial Optimization

↓ Conversion

$$\min_{x \in \mathbf{D}^n} \quad x^H M_0 x$$
$$\text{s.t.} \quad x^H M_i x \leq a_i, \quad i = 1, 2, ..., m$$

↓ SDP/ Penalized SDP

$$\min_{W} \quad \text{trace}\{M_0 W\} + \lambda\, g(W)$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \leq a_i, \quad i = 1, 2, ..., m$$
$$W \succeq 0$$

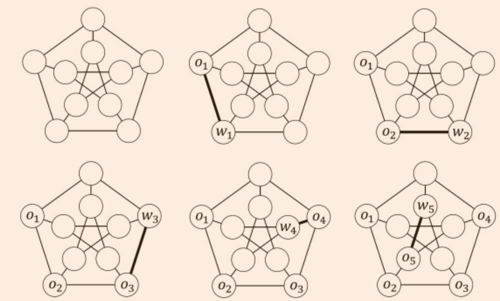
Connection between sparsity and rank? → Proof of existence of low-rank solution using OS and treewidth

How to design penalized SDP? → Propose two methods to design penalty

Design scalable numerical algorithm? → Cheap iterations for large-scale problems

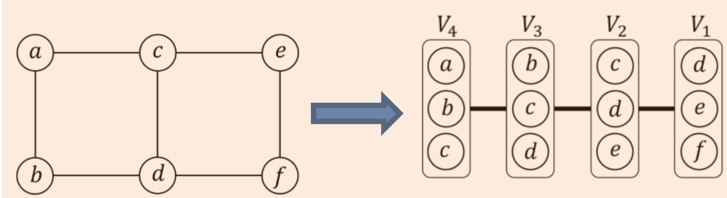
Graph Notions

- **OS-vertex sequence:** [Hackney et al, 2009]
  - Partial ordering of vertices
  - Assume $O_1, O_2, \ldots, O_m$ is a sequence.
  - O_i has a neighbor w_i not connected to the connected component of O_i in the subgraph induced by $O_1, \ldots, O_i$



OS: Maximum cardinality among all OS sequences

- **Tree decomposition:** Map the graph G into a tree T
  - Each node of T is a bag of vertices of G
  - Each edge of G appears in one node of T
  - If a vertex shows up in multiple nodes of T, those nodes should form a subtree
- **Width of T:** Max cardinality minus 1



Vertices

Bags of vertices

Treewidth of G: Minimum width

- Roughly speaking, <u>very</u> sparse graphs have high OS and low treewidth[1] (tree: OS=n-1, TW=1)

1. S. Sojoudi, R. Madani, G. Fazelnia and J. Lavaei, "Graph-Theoretic Algorithms for Solving Polynomial Optimization Problems," CDC 2014 (**Tutorial paper**).

Low-Rank Solution

$$\min_{x \in \mathbf{D}^n} \quad x^H M_0 x$$
$$\text{s.t.} \quad x^H M_i x \leq a_i, \quad i = 1, 2, \ldots, m$$

$\Big\Downarrow$ SDP

$$\min_{W} \quad \text{trace}\{M_0 W\}$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \leq a_i, \quad i = 1, 2, \ldots, m$$
$$W \succeq 0$$

$\Big\Downarrow$ Perturbed SDP

$$\min_{W} \quad \text{trace}\{M_0 W\} + \boxed{\sum_{(j,k) \in \mathcal{G}'} \varepsilon_{j,k} W_{jk}}$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \leq a_i, \quad i = 1, 2, \ldots, m$$
$$W \succeq 0$$

- **Sparsity Graph *G***: Generalized weighted graph with no weights.
- SDP may has infinitely many solutions.
- How to find a low-rank solution (if any)?
- Consider a supergraph *G'* of *G*.

Theorem: Every solution of perturbed SDP satisfies the following:

$$\text{Rank}\{W^{\text{opt}}\} \leq |\mathcal{G}'| - \min_{\mathcal{G}_s} \left\{ \text{OS}(\mathcal{G}_s) \;\middle|\; (\mathcal{G}' - \mathcal{G}) \subseteq \mathcal{G}_s \subseteq \mathcal{G}' \right\}$$

Equal bags: TW(*G*)+1 for a right choice of *G'*

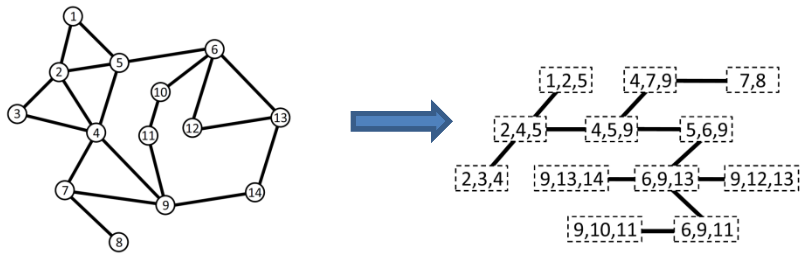
Unequal bags: Needs nonlinear penalty to attain TW(*G*)+1

- This result includes the recent work *Laurent and Varvitsiotis*, 2012.

1. R. Madani et al., "Low-Rank Solutions of Matrix Inequalities with Applications to Polynomial Optimization and Matrix Completion Problems," CDC 2014.
2. R. Madani et al., "Finding Low-rank Solutions of Sparse Linear Matrix Inequalities using Convex Optimization," Under review for SIOPT, 2014.

Illustration: Power Optimization

Tree decomposition for IEEE 14-bus system:



Case studies:

System $\mathcal{G}$	tw{$\mathcal{G}$}	System $\mathcal{G}$	Bound on tw{$\mathcal{G}$}
IEEE 14-bus	2	Polish 2383wp	23
IEEE 30-bus	3	Polish 2736sp	23
New England 39-bus	3	Polish 2746wop	23
IEEE 57-bus	5	Polish 3012wp	24
IEEE 118-bus	4	Polish 3120sp	24
IEEE 300-bus	6	Polish 3375wp	25

Treewidth of Poland < 30

Treewidth of NY < 40

SDP relaxation of every SC-UC-OPF problem solved over NY grid has rank less than 40 (size of *W* varies from 8500 to several millions).

1. R. Madani, S. Sojoudi and J. Lavaei, "Convex Relaxation for Optimal Power Flow Problem: Mesh Networks," IEEE Transactions on Power Systems, 2015.
2. R. Madani, M. Ashraphijuo and J. Lavaei, "Promises of Conic Relaxation for Contingency-Constrained Optimal Power Flow Problem," Allerton 2014.

Outline

Arbitrary Real/Complex Polynomial Optimization

↓ Conversion

$$\min_{x \in \mathbf{D}^n} \quad x^H M_0 x$$
$$\text{s.t.} \quad x^H M_i x \leq a_i, \quad i = 1, 2, ..., m$$

↓ SDP/ Penalized SDP

$$\min_{W} \quad \text{trace}\{M_0 W\} + \lambda\, g(W)$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \leq a_i, \quad i = 1, 2, ..., m$$
$$W \succeq 0$$

Connection between sparsity and rank? → Proof of existence of low-rank solution using OS and treewidth

How to design penalized SDP? → Propose two methods to design penalty

Design scalable numerical algorithm? Cheap iterations for large-scale problems

Non-convexity Localization

$$\min_{x \in \mathbf{D}^n} \quad x^H M_0 x$$
$$\text{s.t.} \quad x^H M_i x \le a_i, \quad i = 1, 2, \ldots, m$$

Sparse → SDP works if G has no edges:

$$x_i^2 \quad \Longrightarrow \quad y_i \qquad \text{(LP)}$$

$$\min_{W} \quad \text{trace}\{M_0 W\}$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \le a_i, \quad i = 1, 2, \ldots, m$$
$$W \succeq 0$$

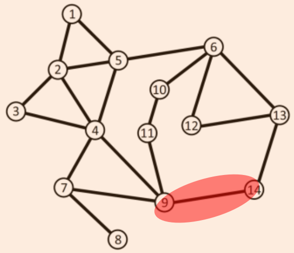
Low-rank →

- Assume SDP fails.
- **Can we identify what edges caused the failure?**
- Localized non-convexity v.s. uniform non-convexity?

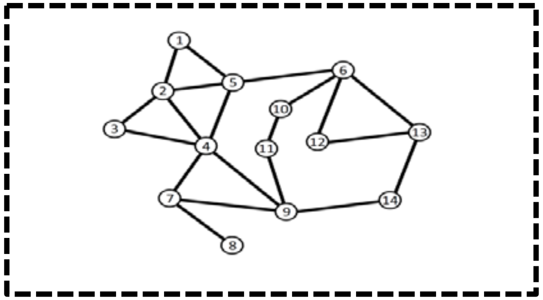
$$\min_{W} \quad \text{trace}\{M_0 W\} + \lambda\, g(W)$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \le a_i, \quad i = 1, 2, \ldots, m$$
$$W \succeq 0$$

Rank-1 →

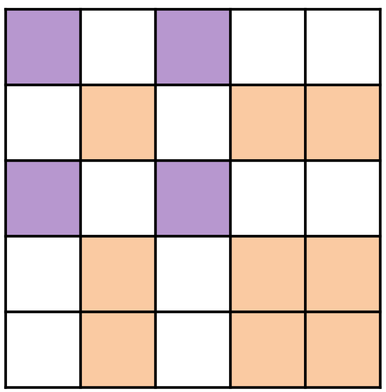
Approach for localized case:
Penalty over problematic edges



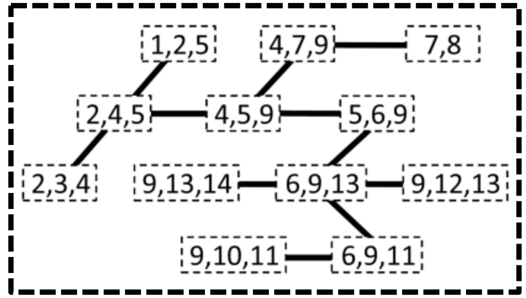
Problematic Edges



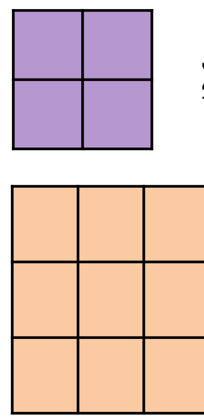
$\Downarrow$ *W*



$\succeq 0$



$\Downarrow$ Submatrices of *W*



$\succeq 0$

$\succeq 0$

Rank of *W* **=** **Max rank of submatrices**

Problematic edges: Identified based on high-rank submatrices

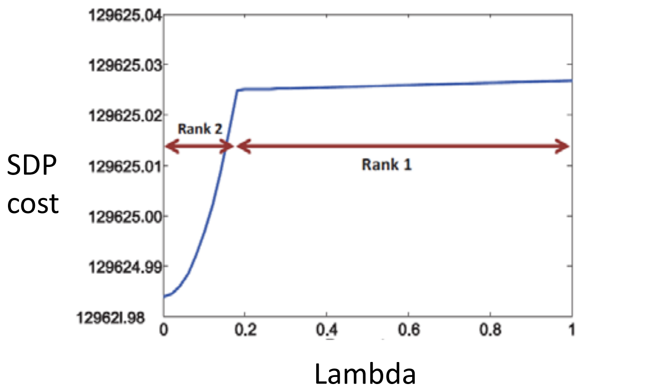
IEEE 300-bus: 2
Polish 2383-bus : 11

1. R. Madani et al., "Finding Low-rank Solutions of Sparse Linear Matrix Inequalities using Convex Optimization," Under review for SIOPT, 2014.
2. R. Madani, M. Ashraphijuo and J. Lavaei, "Promises of Conic Relaxation for Contingency-Constrained Optimal Power Flow Problem," Allerton 2014.

Example: Near-Global Solutions

Strategy: Penalize reactive loss over problematic lines

- **Modified IEEE 118-bus:**
  - 3 local solutions
  - Costs: 129625, 177984, 195695



Case	TW	Cost	Guarantee	Time (sec)
Chow's 9 bus	2	5296.68	100%	≤ 5
IEEE 14 bus	2	8081.53	100%	≤ 5
IEEE 24 bus	4	63352.20	100%	≤ 5
IEEE 30 bus	3	576.89	100%	≤ 5
NE 39 bus	3	41864.40	99.994%	≤ 5
IEEE 57 bus	5	41737.78	100%	≤ 5
IEEE 118 bus	4	129660.81	99.995%	≤ 5
IEEE 300 bus	6	719725.10	99.998%	13.9
Polish 2383wp	23	1874322.65	99.316%	529
Polish 2736sp	23	1308270.20	99.970%	701
Polish 2737sop	23	777664.02	99.995%	675
Polish 2746wop	23	1208453.93	99.985%	801
Polish 2746wp	24	1632384.87	99.962 %	699
Polish 3012wp	24	2608918.45	99.188%	814
Polish 3120sp	24	2160800.42	99.073 %	910

Case	Minima	Cost	Guarantee
WB2	2	877.78	100%
WB3	2	417.25	100%
WB5	2	946.58	99.995%
WB5 Mod	3	1482.22	100%
LMBM3	5	5694.54	100%
LMBM3_50	2	5823.86	99.807%
case22loop	2	4538.80	100 %
case30loop	2	2863.06	100%
case30loop Mod	3	2861.88	100%
case39 Mod4	3	557.15	99.999%
case118 Mod1	3	129625.19	99.999%
case118 Mod2	2	85987.59	100 %
case300 Mod2	2	474643.46	99.996%

7000 simulations

1. R. Madani, S. Sojoudi and J. Lavaei, "Convex Relaxation for Optimal Power Flow Problem: Mesh Networks," IEEE Transactions on Power Systems, 2015.
2. R. Madani, M. Ashraphijuo and J. Lavaei, "Promises of Conic Relaxation for Contingency-Constrained Optimal Power Flow Problem," Allerton 2014.

Penalty Design

Why was penalty chosen as loss?

$$\min_{W} \quad \text{trace}\{M_0 W\} + \lambda\, g(W)$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \leq a_i, \quad i = 1, 2, \ldots, m$$
$$W \succeq 0$$

First try: $g(W) = \|W\|_*$

- Compressed sensing and phase retrieval
- Need *n log n* measurements for a much simpler problem [Candes and Recht, 2009].

Proposed penalty:

$$g(W) = \text{trace}\{MW\}$$

Algorithm design: Can we design an SDP to find the best M?

Good penalty: Minimization of penalty by itself ($\lambda = \infty$) leads to a rank-1 solution.

Study of a simpler case:

$$\min_{W} \quad \text{trace}\{MW\}$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} = a_i, \quad i = 1, 2, \ldots, n$$
$$W \succeq 0$$

Guess for solution of original QCQP: x_*

- $M \succeq 0$
- $M x_* = 0$
- Zero is a simple eig of M.

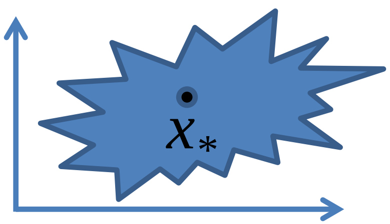
Penalty Design

Theorem: If Jacobian is nonsingular, then SDP is exact in a vicinity of x_*.

Local behavior: Linearization solves approximately but SDP solves exactly.

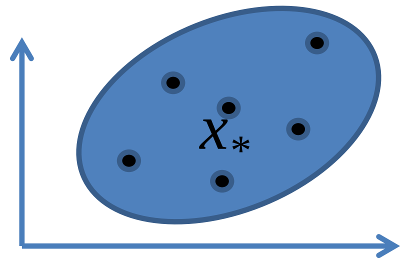
Global behavior: The region could be as big as the entire space.

Recoverable region for x :



$$R_M = \{x \mid g(x, M) \succeq 0\}$$

Design of M: Include x_* and a set of points



LMI

Power flow equations for power systems: *M* is a one-time design independent of loads.

1. M. Ashraphijuo and J. Lavaei, "SDP-Type Algorithm Design for Systems of Polynomials," Preprint, 2015.
2. R. Madani, R. Baldick and J. Lavaei, "Convexification of Power Flow Problem over Arbitrary Networks," Preprint, 2015.

Power Flow Problem Using SDP

Theorem: There is a region R_M with a non-empty interior and containing the vector 1 such that SDP solves PF if and only if PF has a solution in R_M .

- **Interpretations:**
  - SDP always solves PF precisely if PF has a solution with small angles.
  - It works for all networks (all topologies).
  - Unlike DC approximation, this accommodates equations for reactive power.

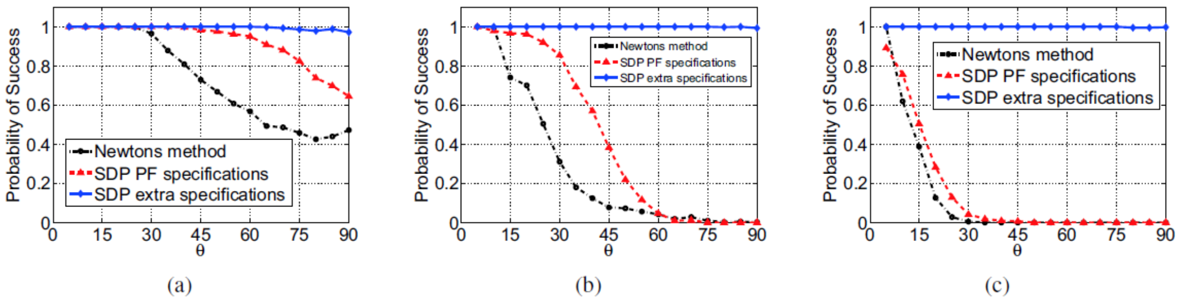


Fig. 4: These plots show the probability of success for Newton's method, SDP relaxation, and SDP relaxation with extra specifications for (a): IEEE 9-bus system, (b): New England 39-bus system, and (c): IEEE 57-bus system.

1. M. Ashraphijuo and J. Lavaei, "SDP-Type Algorithm Design for Systems of Polynomials," Preprint, 2015.
2. R. Madani, R. Baldick and J. Lavaei, "Convexification of Power Flow Problem over Arbitrary Networks," Preprint, 2015.

Outline

Arbitrary Real/Complex Polynomial Optimization

Conversion

$$\min_{x \in \mathbf{D}^n} \quad x^H M_0 x$$
$$\text{s.t.} \quad x^H M_i x \leq a_i, \quad i = 1, 2, ..., m$$

SDP/ Penalized SDP

$$\min_{W} \quad \text{trace}\{M_0 W\} + \lambda\, g(W)$$
$$\text{s.t.} \quad \text{trace}\{M_i W\} \leq a_i, \quad i = 1, 2, ..., m$$
$$W \succeq 0$$

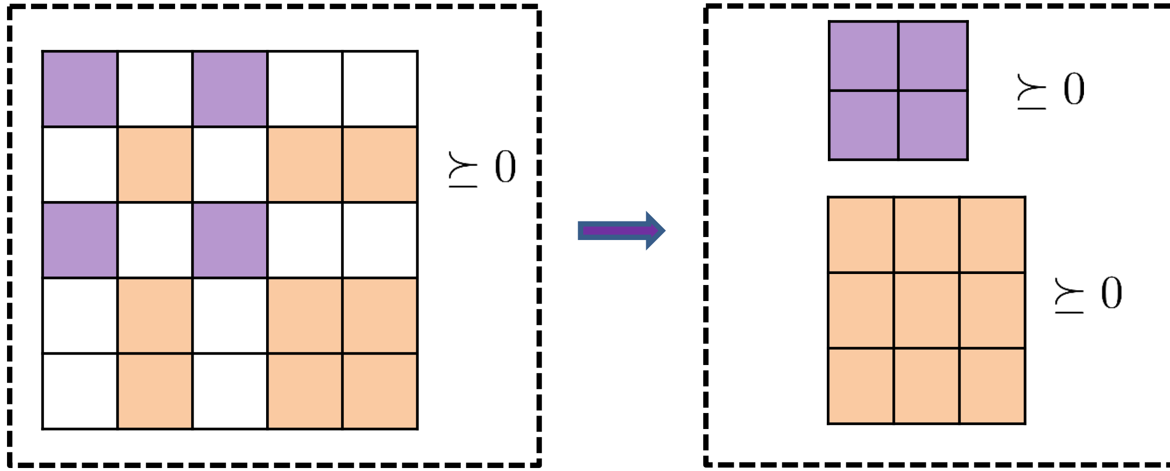
Connection between sparsity and rank? → Proof of existence of low-rank solution using OS and treewidth

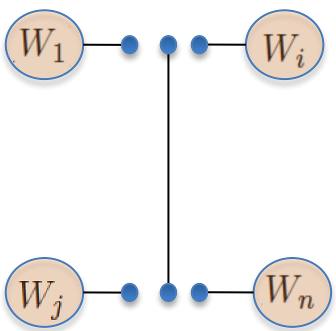
How to design penalized SDP? → Propose two methods to design penalty

Design scalable numerical algorithm? → Cheap iterations for large-scale problems

Low-Complex Algorithm

Goal: Design a low-complex algorithm for sparse LP/QP/QCQP/SOCP/SDP





$$
\begin{aligned}
\min \quad & \mathbf{tr}(A_1W_1) + \mathbf{tr}(A_2W_2) + \cdots + \mathbf{tr}(A_nW_n) \\
\text{subject to:} \quad & \\
& \mathbf{tr}(B_pW_i) = c_p, \quad p = 1, \ldots, k \\
& W_i \succeq 0 \\
& W_i(I_{ij}, I_{ij}) = W_j(I_{ji}, I_{ji}) \\
& \text{for every } i \in \mathcal{V} \text{ and } (i,j) \in \mathcal{E}.
\end{aligned}
$$

Sum of agents' objectives

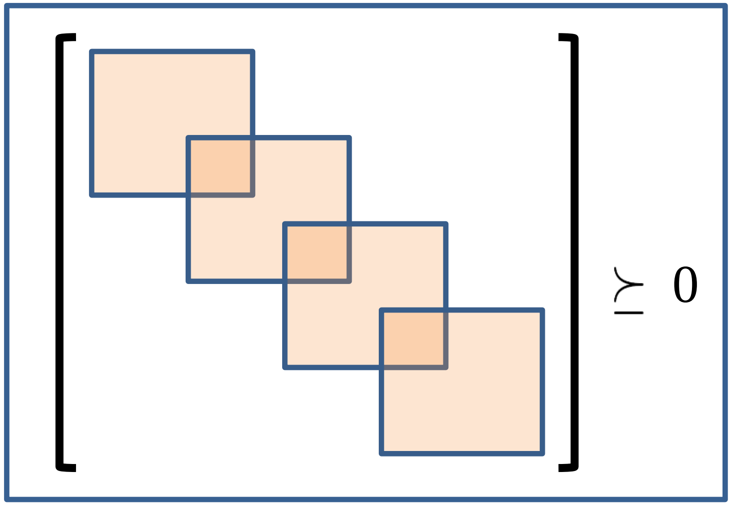
Local constraints

Overlapping constraints

- **Distributed Algorithm:** ADMM-based dual decomposed SDP (related work: [Parikh and Boyd, 2014], [Wen, Goldfarb and Yin, 2010], [Andersen, Vandenberghe and Dahl, 2010]).
- **Iterations:** Closed-form solution for every iteration (eigen-decomposition on submatrices)

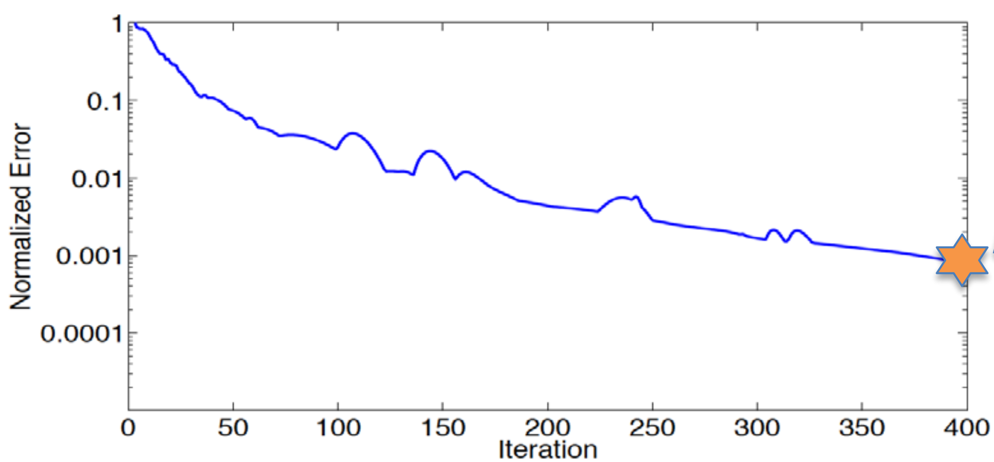
1. A. Kalbat and J. Lavaei, "Alternating Direction Method of Multipliers for Sparse Semidefinite Programs," Preprint, 2015.

Example: Large-Scale Random Problem



- **Number of blocks (agents):** 2000
- **Size of each block:** 40
- **Number of constraints per block:** 5
- **Overlapping degree:** 25%
- **Number of entries for full SDP:** 6.4B
- **Number of entries for decomposed SDP:** Over 3M
- **Number of constraints:** Several thousands

- **20 minutes in MATLAB with cold start (2.4 GHz and 8 GB):**



99.9% feasible and globally optimal

Example: IEEE Benchmark Systems

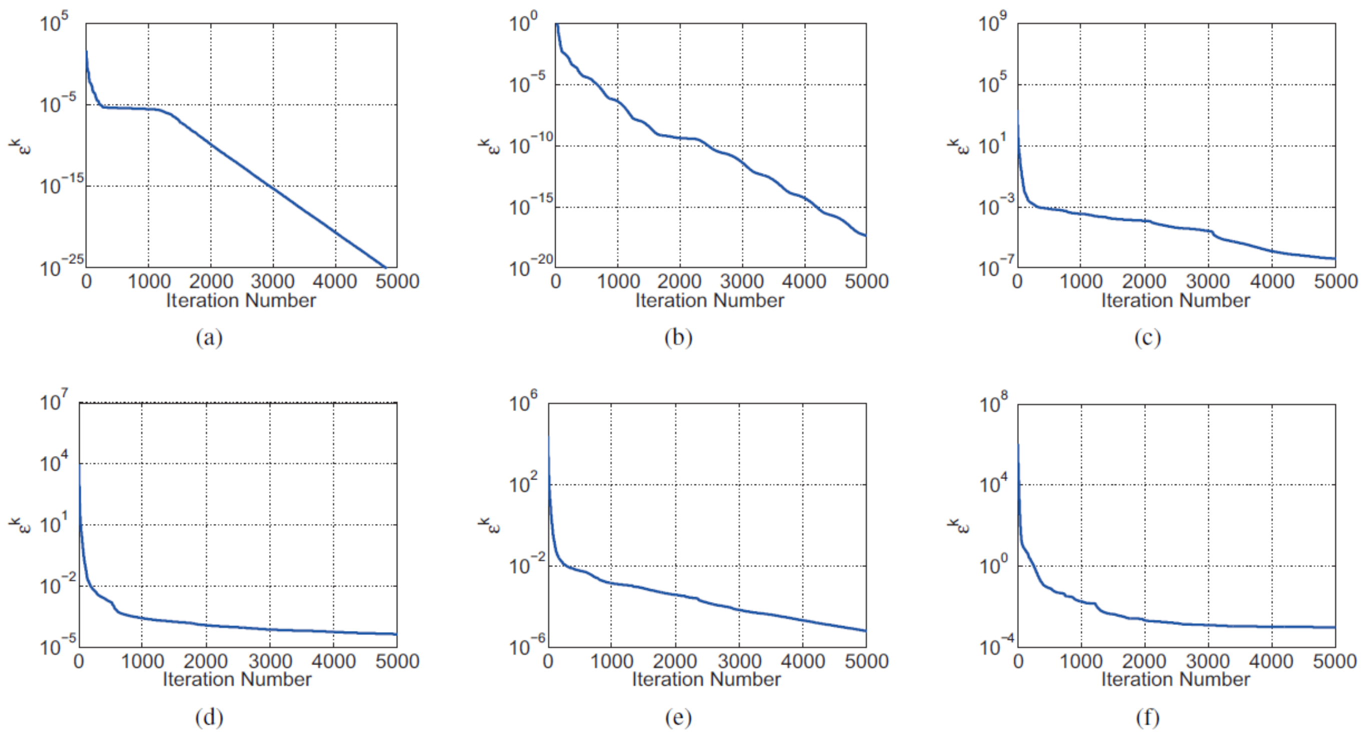
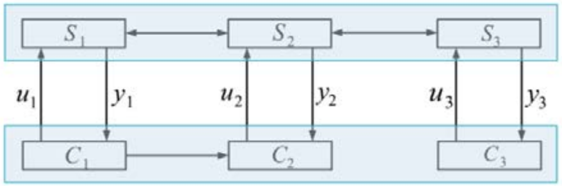


Fig. 1: These plots show the convergence behavior of the energy function ε^k for IEEE test cases. (a): Chow's 9 bus, (b): IEEE 14 bus, (c): IEEE 30 bus, (d): IEEE 57 bus, (e): IEEE 118 bus, (f): IEEE 300 bus.

Another Project: Distributed Control of Stochastic Systems



Distributed control

(NP-hard: Witsenhausen's example)

Stochastic Distributed Control: Design $u[\tau] = Ky[\tau]$ for

$$\begin{cases} x[\tau+1] = Ax[\tau] + Bu[\tau] + Ed[\tau] \\ y[\tau] = Cx[\tau] + Fv[\tau] \end{cases}$$

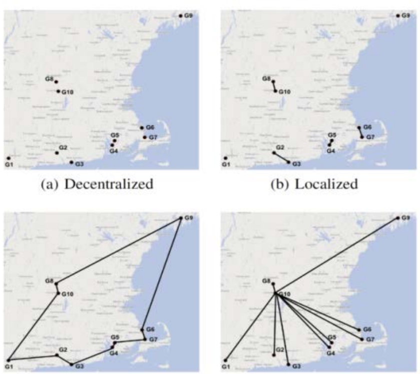
disturbance

noise

to minimize:

$$\lim_{\tau \to +\infty} \mathcal{E}\left(x[\tau]^T Q x[\tau] + u[\tau]^T R u[\tau]\right)$$

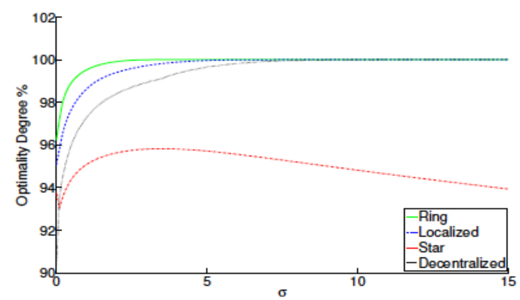
Theorem: Rank of SDP solution in the Lyapunov domain is 1, 2 or 3.



(a) Decentralized (b) Localized

(c) Ring (d) Star Topology (G_{10} in center)

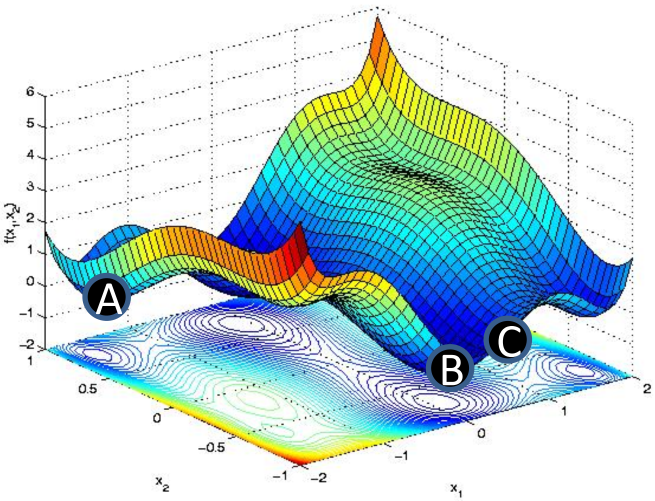
New England Test System



Optimality degree for stochastic ODC

1. G. Fazelnia et al., "Convex Relaxation for Optimal Distributed Control Problem — Part I: Time-Domain Formulation", Submitted to IEEE Transactions on Automatic Control, 2014 (conference version: CDC 2014).
2. G. Fazelnia et al., "Convex Relaxation for Optimal Distributed Control Problem — Part II: Lyapunov Formulation and Case Studies", Submitted to IEEE Transactions on Automatic Control, 2014 (conference version: Allerton 2014).
3. Salar Fatahi, Ghazal Fazelnia and Javad Lavaei, Transformation of Optimal Centralized Controllers Into Near-Global Static Distributed Controllers , Preprint, 2015.

Conclusions



Problem: Find a near-global solution together with a global optimality guarantee for energy problems

Approach: Graph-theoretic convexification

- ❑ **OS and treewidth:** Connection between rank and sparsity
- ❑ **Non-convexity diagnosis:** Graph-based localization
- ❑ **Penalized SDP:** Obtaining a near-global solution
- ❑ **Scalable algorithm:** High-dimensional sparse SDP

Funding Acknowledgements

- **ONR YIP:** Graph-theoretic and low-rank optimization

- **NSF CAREER:** Control and optimization for power systems

- **NSF EPCN:** Contingency analysis for power systems

- **Google:** Numerical algorithms for nonlinear optimization