A Strong Semidefinite Programming Relaxation of the Unit Commitment Problem

Javad Lavaei

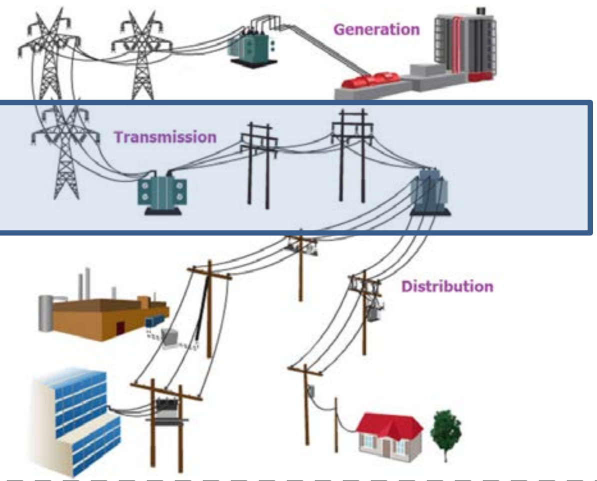
Industrial Engineering and Operations Research
University of California, Berkeley

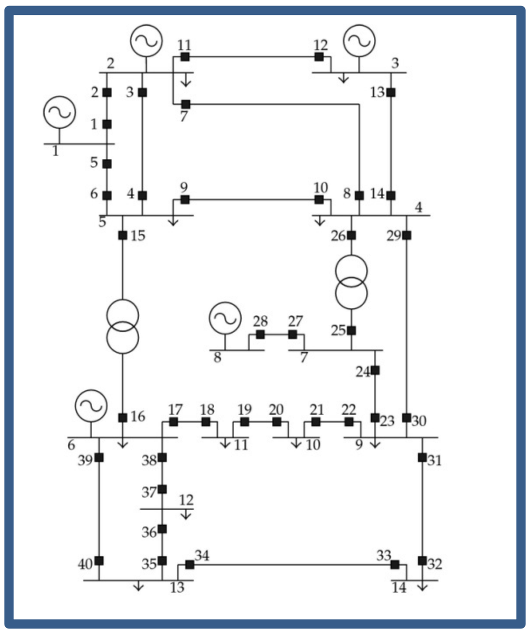
Joint work with:

Morteza Ashraphijuo, Salar Fattahi and Alper Atamturk (UC Berkeley)

Power Systems

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





ISO, RTO, TSO

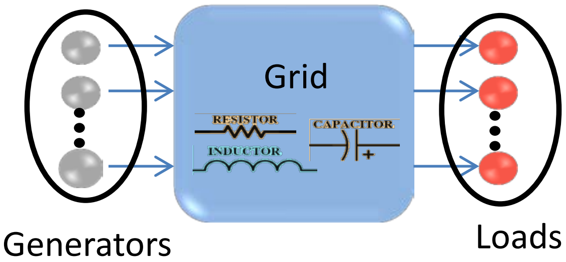
1. Optimal power flow (OPF)
2. Security-constrained OPF
3. State estimation
4. Network reconfiguration
5. Unit commitment
6. Dynamic energy management

NP-hard

(real-time operation and market)

Optimal Power Flow

Optimal Power Flow: Optimally match supply with demand



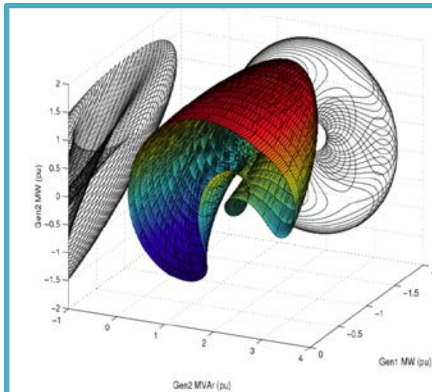
Vector of complex voltages

$$\min_{x\in\mathbf{C}^n} \quad x^H M_0 x$$

$$\text{s.t.} \quad x^H M_i x \le a_i, \quad i = 1, 2, \ldots, m$$

- **Real-time operation:** OPF is solved every 5-15 minutes.
- **Market:** Security-constrained unit-commitment OPF
- **Complexity:** Strongly NP-complete with long history since 1962.
- **Common practice:** Linearization

A multi-billion critical system depends on optimization.



OPF feasible set

(Ian Hisken et al. 2003)

Convexification

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

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

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

SDP relaxation

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

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

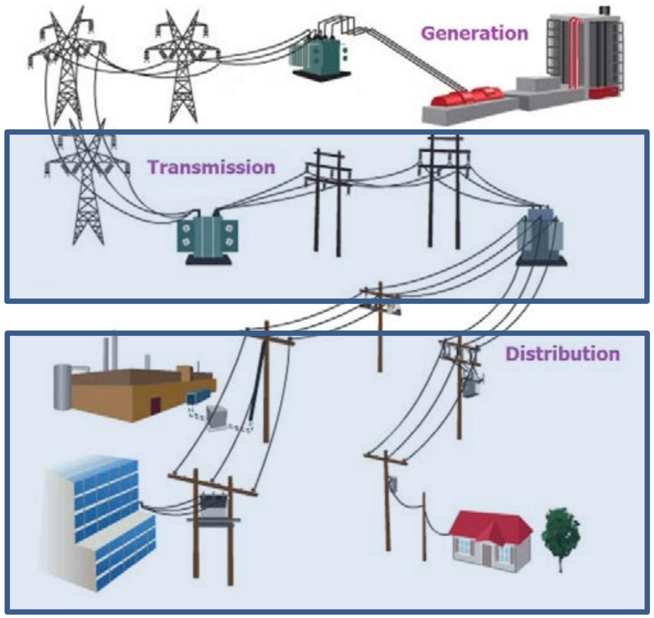
Penalized SDP

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

- **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

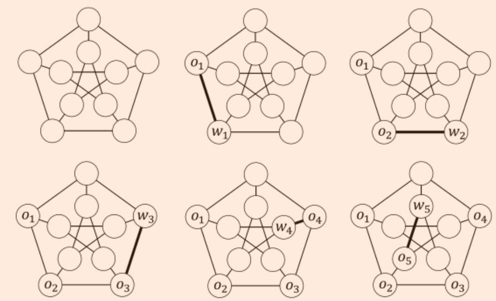
SDP looks very promising for energy applications

- **SDP revitalized the area:**
  - Follow-up work in academia (MIT, Stanford, Berkeley, Caltech, Gatech, UIUC, Cornell, JHU, Iowa State, CMU, UCLA, Wisconsin, UoM, Harvard, Michigan, ETH, EPFL, etc.)
  - Interest from industry

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.

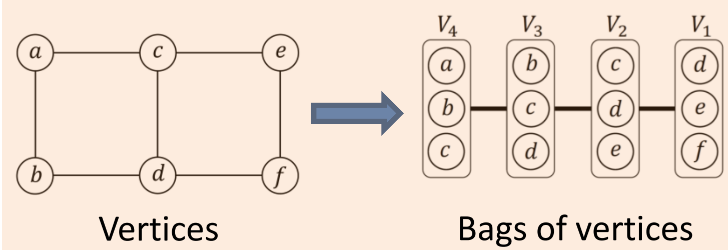
Graph Notions

- ❑ **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



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

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

↓ SDP

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

↓ Perturbed SDP

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

- **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'*

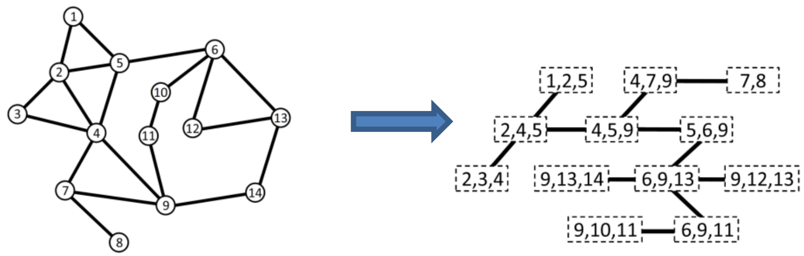
Unequal bags: Needs nonlinear penalty to attain TW(*G*)+1

- 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.

Simple UC Model Capturing Nonlinearity

- **Problem:** DC unit commitment (AC is similar)

- **Objective function:**

$$\begin{aligned}
&\text{Power generation cost}: && a_i^{(t)} \times (p_i^{(t)})^2 + b_i^{(t)} \times p_i^{(t)} + c_i^{(t)} \times x_i^{(t)}, \\
&\text{Startup cost}: && c_{i;\,\text{up}}^{(t)} \times (x_i^{(t)} - x_i^{(t-1)})^2 \times I(x_i^{(t)} \geq x_i^{(t-1)}), \\
&\text{Shutdown cost}: && c_{i;\,\text{down}}^{(t)} \times (x_i^{(t)} - x_i^{(t-1)})^2 \times I(x_i^{(t)} \leq x_i^{(t-1)})
\end{aligned}$$

- **Unit commitment constraints:**

$$\begin{aligned}
&\text{Generator statuses}: && x_i^{(t)} \in \{0, 1\}, \\
&\text{Generator limits}: && p_{\min}^{(t)} \times x_i^{(t)} \leq p_i^{(t)} \leq p_{\max}^{(t)} \times x_i^{(t)}, \\
&\text{Power balance}: && \sum_{i=1}^{n_g} p_i^{(t)} = p_d, \\
&\text{Line capacities}: && |p_{ij}^{(t)}| \leq p_{ij;\,\max}, \\
&\text{Ramp constraints}: && |p_i^{(t+1)} - p_i^{(t)}| \leq r^{(t)},
\end{aligned}$$

- **Decision variables:** commitment parameters and generator outputs

Strengthened SDP

- **SDP (sparse):** Relax quadratic equations using SDP constraints.
- **Bad news:** SDP = LP (and rounding is bad).
- **Current practice:** branch & bound + cutting plane
- **Question:** Find a convex model for the problem (good for pricing)
- **Strengthened SDP (dense):**
  - Flow constraints are linear
  - Multiply them pairwise to obtain valid quadratic constraints
  - Relax valid inequalities using SDP

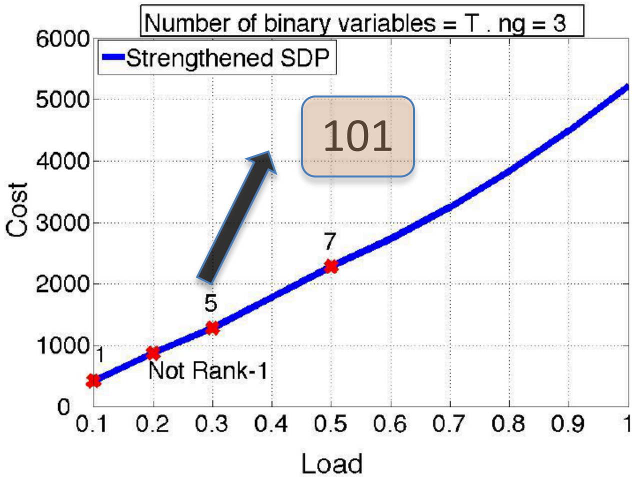
$$My - d \geq 0$$
$$(My - d)(My - d)^\top \geq 0$$
$$Myy^\top M^\top - dy^\top M^\top - Myd^\top + dd^\top \geq 0$$
$$MYM^\top - dy^\top M^\top - Myd^\top + dd^\top \geq 0$$

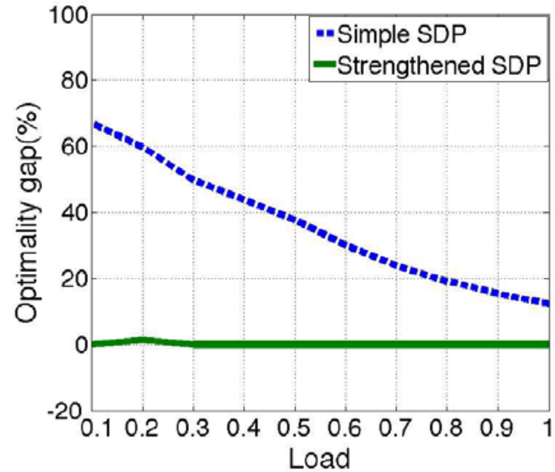
Strengthened SDP

- **Weakly strengthened SDP (sparse or dense):**
  - Solve SDP
  - Check how many valid inequalities are violated
  - First-order strengthened SDP: add those inequalities to SDP
  - Check how many valid inequalities are violated
  - Second-order strengthened SDP: add those inequalities to first-order SDP
  - …

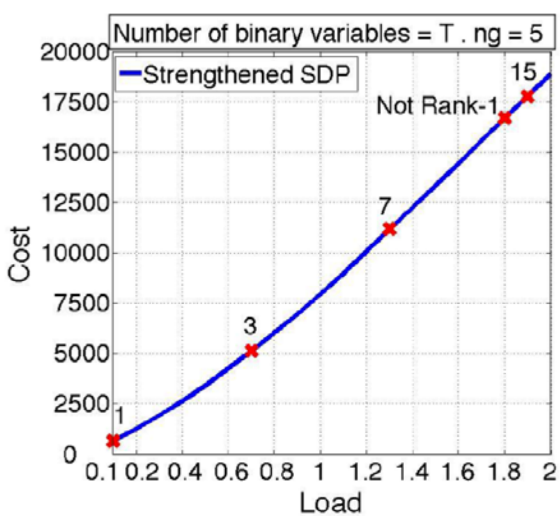
Single-Time UC Problem

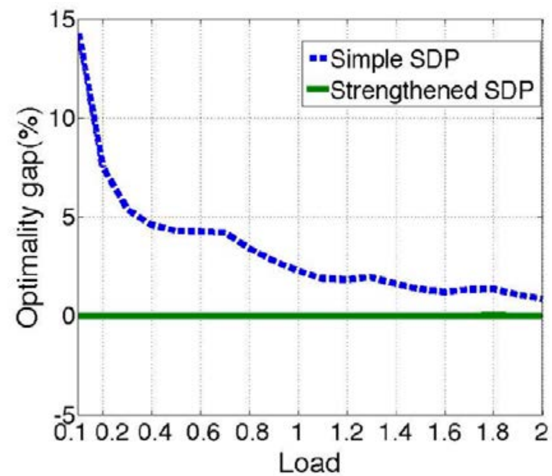
- Case 9 bus with 3 generators and 1 time slot:





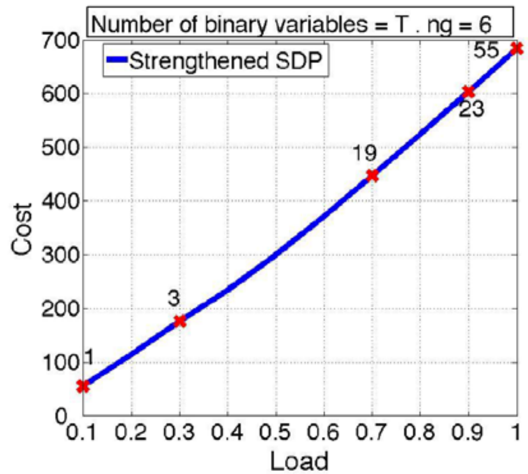
- Case 14 bus with 5 generators and 1 time slot:

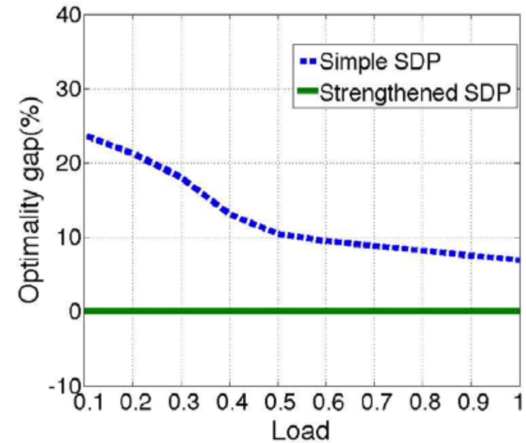




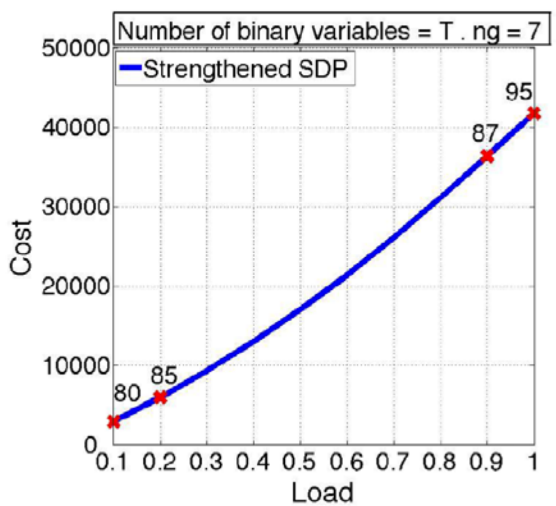
Single-Time UC Problem

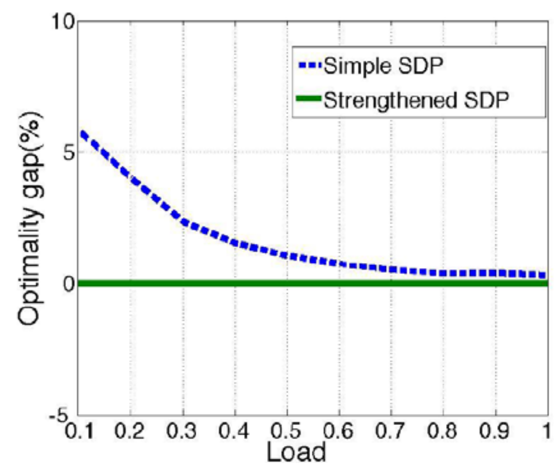
- Case 30 bus with 6 Generators and 1 time slot:





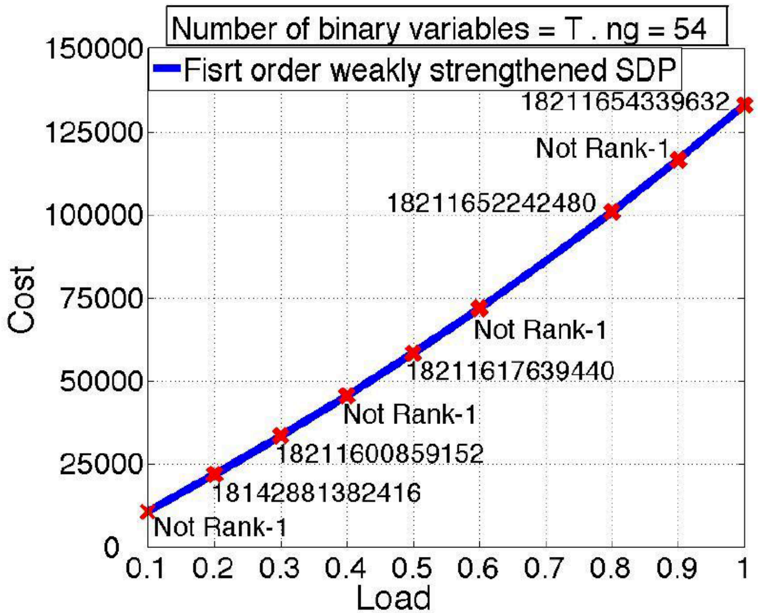
- Case 57 bus with 7 Generators and 1 time slot:

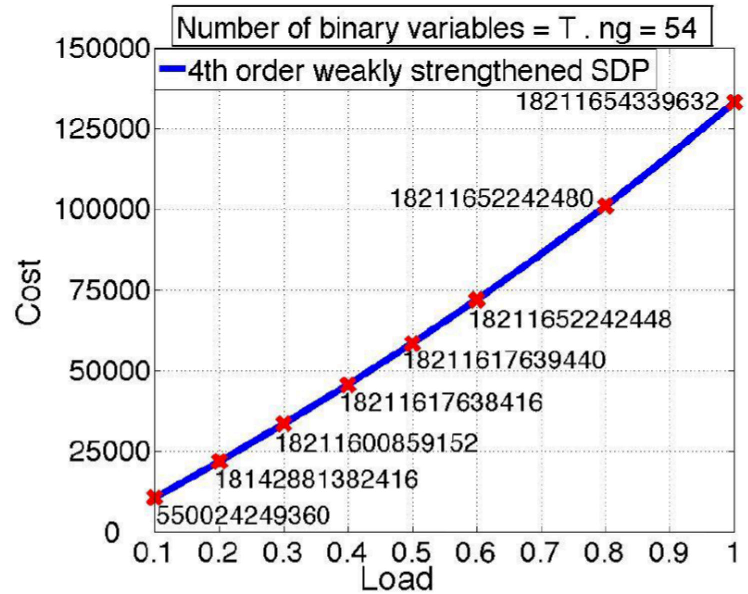




Single-Time UC Problem

❑ Case 118 bus with 54 Generators and 1 time slot:



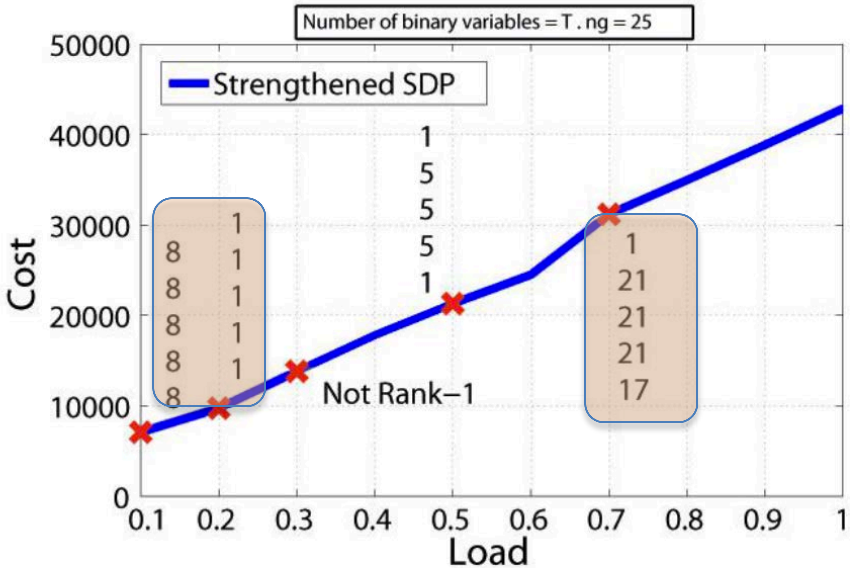


❑ **Observation:**

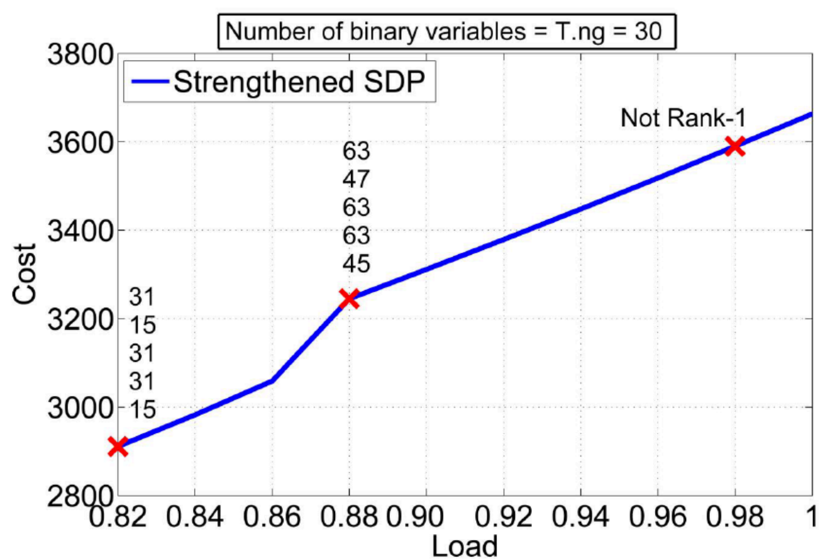
- SDP: really bad
- First-order strengthened SDP: 50% success
- Fourth-order strengthened SDP: 100% success

Multi-Period UC Problem

- Case 14 bus with 5 Generators and 5 time slots:

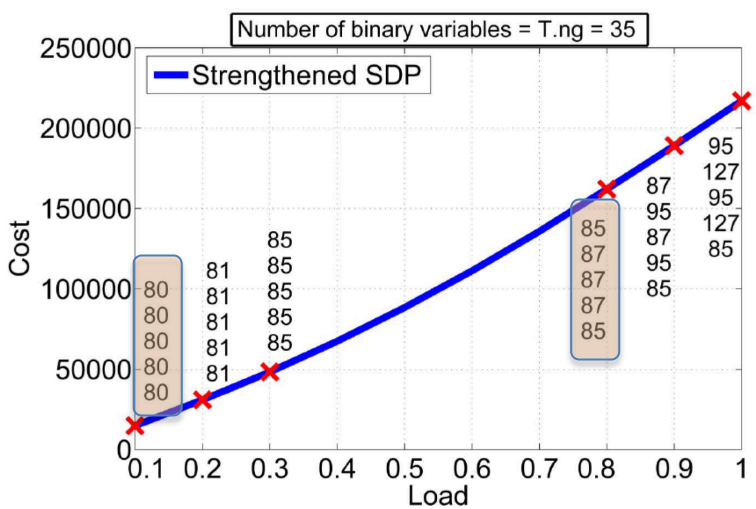


- Case 30 bus with 6 Generators and 5 time slots:

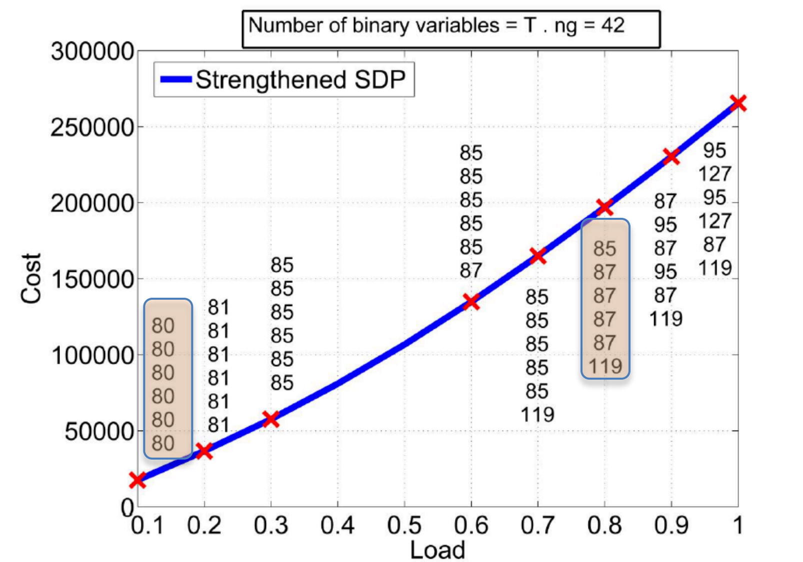


Multi-Period UC Problem

- ❑ Case 57 bus with 7 Generators and 5 time slots:



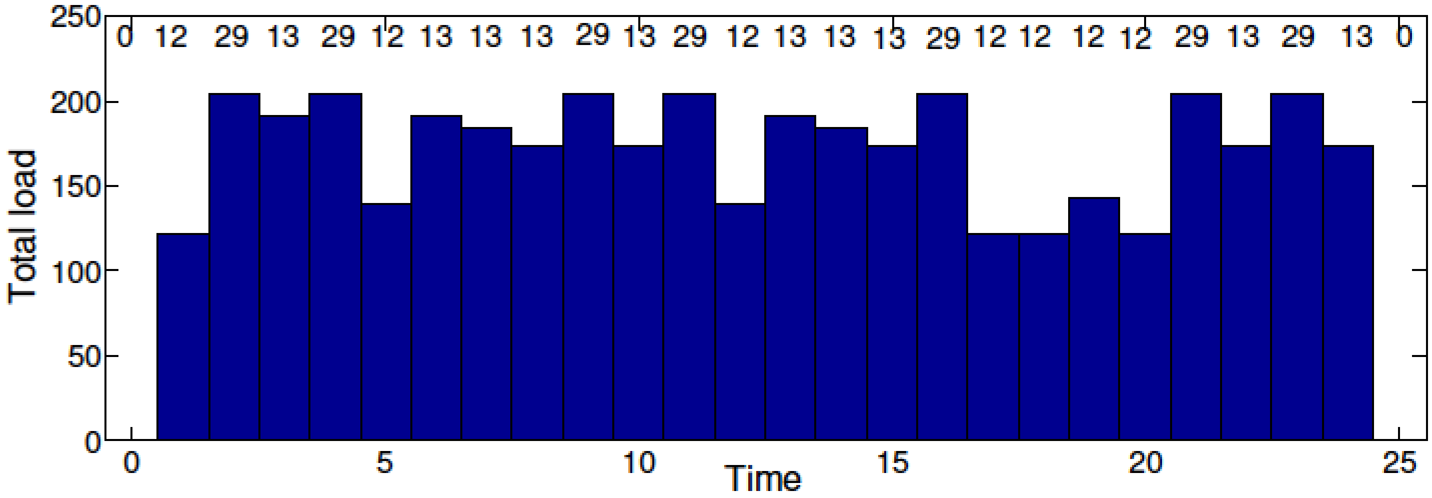
- ❑ Case 57 bus with 7 Generators and 6 time slots:



- ❑ It can be shown that as loads increase or flow constraints become tighter, the SDP becomes exact at some point.

Multi-Period UC Problem

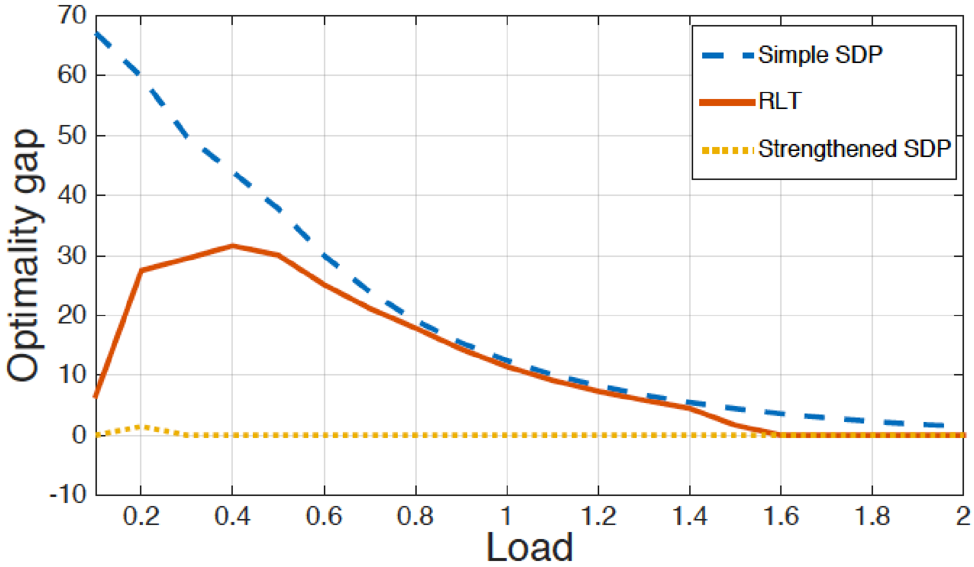
- ❑ IEEE 14-bus system over 24 hours:



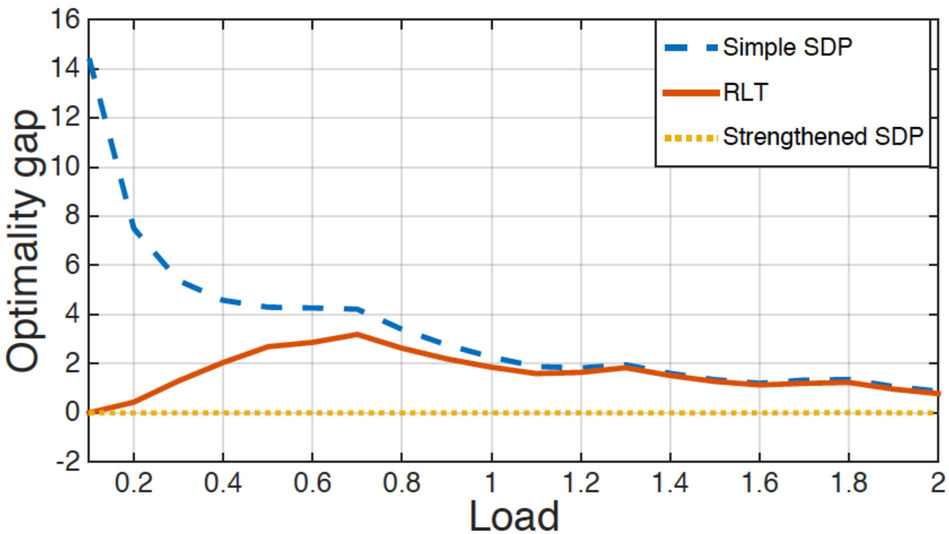
- ❑ SDP cost: 162600
- ❑ First-order strengthened SDP cost: 205838
- ❑ Strengthened SDP cost: 210159

SDP v.s. RLT v.s. Strengthened SDP

- ❑ IEEE 9-bus system:

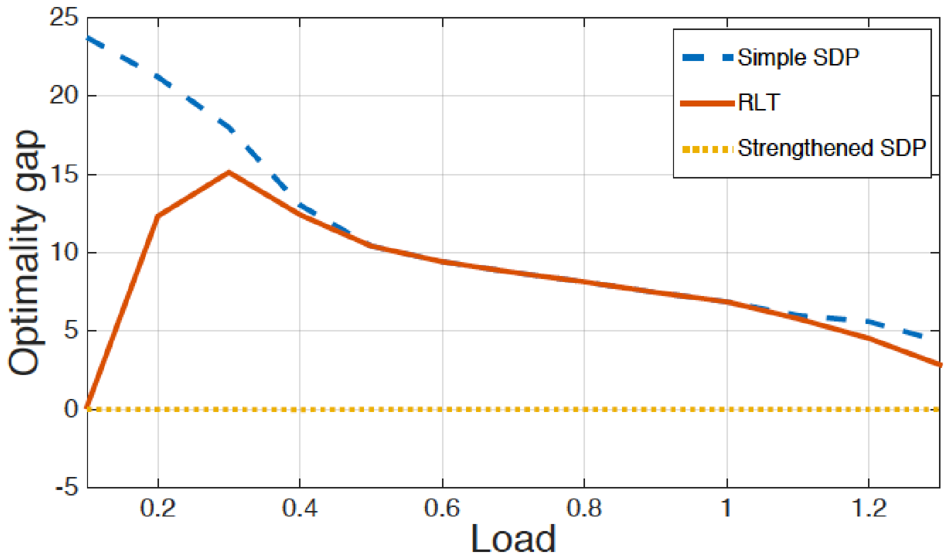


- ❑ IEEE 14-bus system:

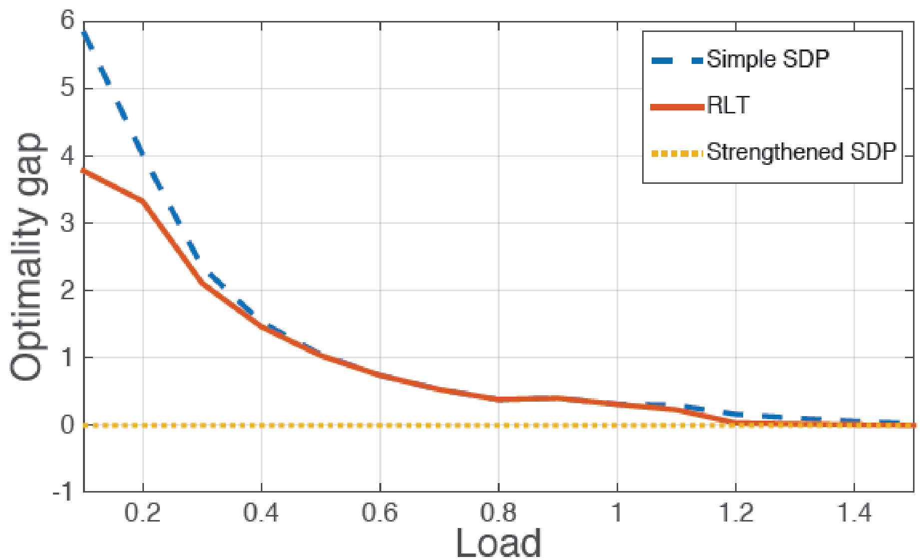


SDP v.s. RLT v.s. Strengthened SDP

❑ IEEE 30-bus system:

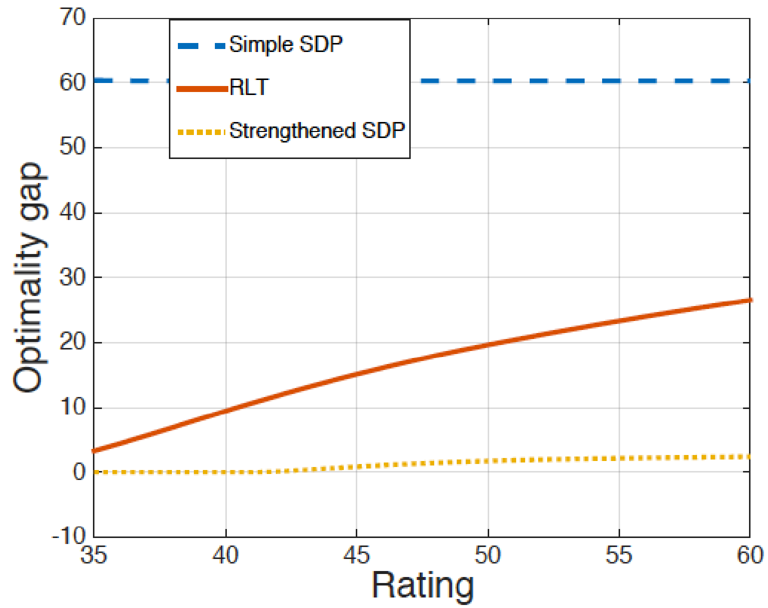


❑ IEEE 57-bus system:

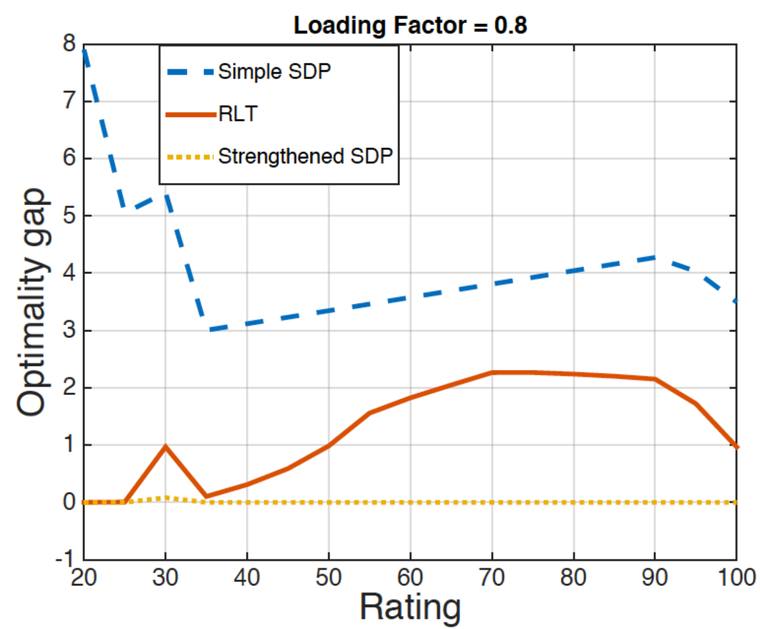


Role of Line Limits

❑ IEEE 14-bus system:



❑ IEEE 14-bus system:



State Estimation Problem Using Penalized Convex Program

- State estimation problem:

$$z_j = \mathbf{v}^* \mathbf{M}_j \mathbf{v} + \eta_j, \quad \forall j \in \mathcal{M} := \{1, 2, \dots, m\}$$

- Penalized convex problem:

$$\begin{aligned} \underset{\mathbf{X} \in \mathbb{H}^n, \boldsymbol{\nu} \in \mathbb{R}^m}{\text{minimize}} \quad & \rho f(\boldsymbol{\nu}) + \operatorname{Tr}(\mathbf{M}_0 \mathbf{X}) \\ \text{subject to} \quad & \operatorname{Tr}(\mathbf{M}_j \mathbf{X}) + \nu_j = z_j, \quad \forall j \in \mathcal{M} \\ & \mathbf{X} \succeq \mathbf{0}, \end{aligned}$$

- Assume measurements are voltage magnitudes and line flows over a spanning tree.

$$\|\mathbf{X}^{\text{opt}} - \beta \mathbf{v}\mathbf{v}^*\|_F \leq 2\sqrt{\frac{\rho \times f_{\text{WLAV}}(\boldsymbol{\eta}) \times \operatorname{Tr}(\mathbf{X}^{\text{opt}})}{\lambda}}$$

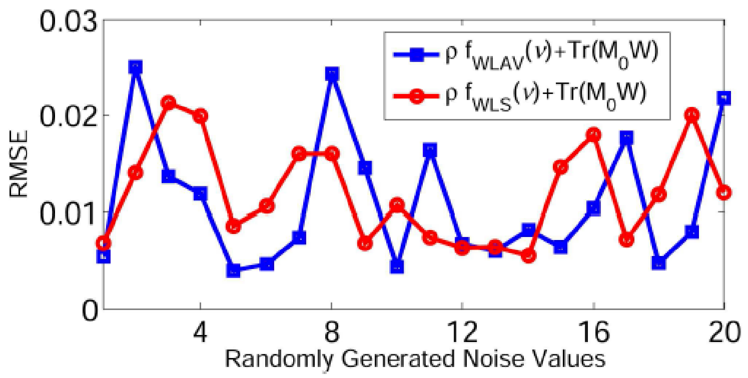
- The tail probability goes to zero exponentially fast (error: non-convexity and noise)

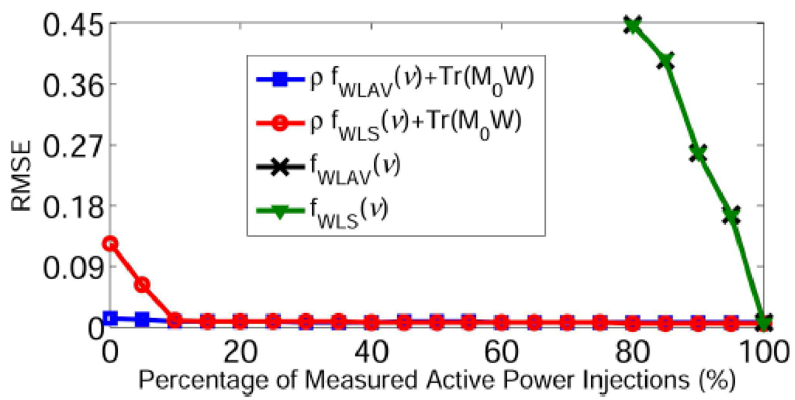
- The higher the number of measurements, the lower the estimation error.

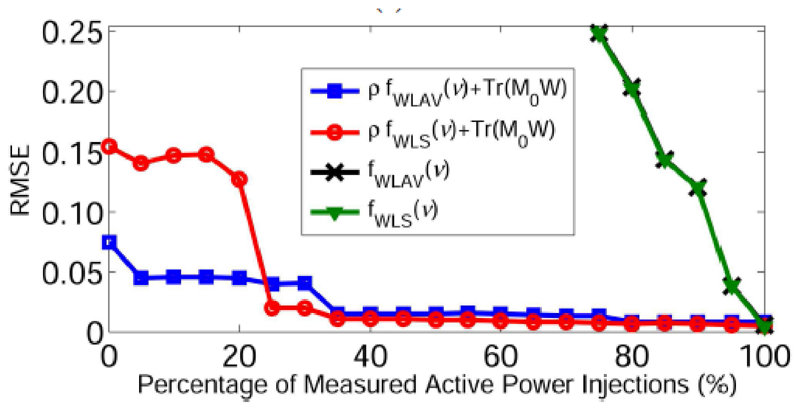
1. R. Madani, J. Lavaei and R. Baldick, "Convexification of Power Flow Problem over Arbitrary Networks," CDC, 2015.

Simulations

- **PEGASE 1354-bus system:**

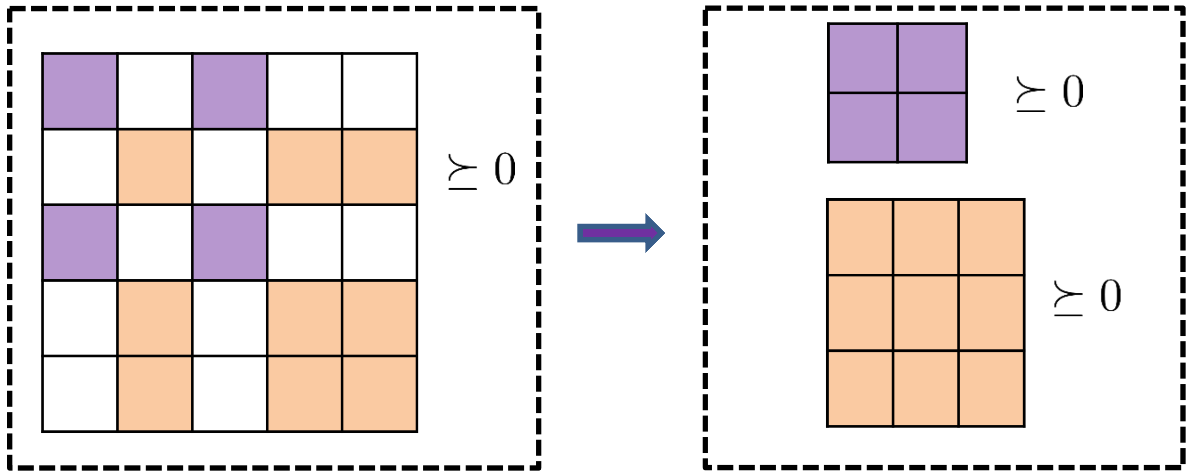


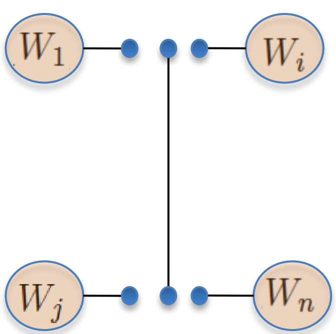




Low-Complex Algorithm

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





$$\min \quad \mathbf{tr}(A_1 W_1) + \mathbf{tr}(A_2 W_2) + \cdots + \mathbf{tr}(A_n W_n)$$

subject to:

$$\mathbf{tr}(B_p W_i) = c_p, \quad p = 1, \ldots, k$$
$$W_i \succeq 0$$
$$W_i(I_{ij}, I_{ij}) = W_j(I_{ji}, I_{ji})$$

for every $i \in \mathcal{V}$ and $(i,j) \in \mathcal{E}$.

→ Sum of agents' objectives

→ Local constraints

→ Overlapping constraints

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

Low-Complex Algorithm

- **Example of a three-agent SDP:**

- **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)

Iterations for Two Agents

Iterations of Agent 1:

$$R_1^{t+1} = \left(B_1^{\overset{t}{\mathrm{sum}}} + D_1^{\overset{t}{\mathrm{sum}}} + H_{1,2}^{\overset{t}{\mathrm{full}}} + A_1 - \frac{G_1^t}{\mu} \right)_+$$

$$u_1^{t+1} = \left(v_1^t + \frac{\lambda_1^t}{\mu} \right)_+$$

$$\overset{t+1}{H^{(1,2)}} = \frac{1}{2} \left(H_{1,2}^t + H_{2,1}^t + \frac{G_{1,2}^t}{\mu} + \frac{G_{2,1}^t}{\mu} \right)$$

$$(z_1, v_1, H_{1,2})^{t+1} = \mathrm{Lin}\left(u_1^{t+1}, R_1^{t+1}, \overset{t+1}{H^{(1,2)}}, G_1^t, G_{1,2}^t, \lambda_1^t \right)$$

$$G_1^{t+1} = G_1^t + \mu \left(-B_1^{\overset{t+1}{\mathrm{sum}}} - D_1^{\overset{t+1}{\mathrm{sum}}} + R_1^{t+1} - H_{1,2}^{\overset{t+1}{\mathrm{full}}} - A_1 \right)$$

$$G_{1,2}^{t+1} = G_{1,2}^t + \mu \left(H_{1,2}^{t+1} - \overset{t+1}{H^{(1,2)}} \right)$$

$$\lambda_1^{t+1} = \lambda_1^t + \mu \left(v_1^{t+1} - u_1^{t+1} \right)$$

Iterations of Agent 2:

$$R_2^{t+1} = \left(B_2^{\overset{t}{\mathrm{sum}}} + D_2^{\overset{t}{\mathrm{sum}}} + H_{2,1}^{\overset{t}{\mathrm{full}}} + A_2 - \frac{G_2^t}{\mu} \right)_+$$

$$u_2^{t+1} = \left(v_2^t + \frac{\lambda_2^t}{\mu} \right)_+$$

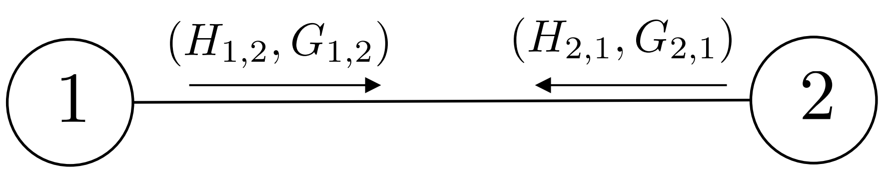
$$\overset{t+1}{H^{(1,2)}} = \frac{1}{2} \left(H_{1,2}^t + H_{2,1}^t + \frac{G_{1,2}^t}{\mu} + \frac{G_{2,1}^t}{\mu} \right)$$

$$(z_2, v_2, H_{2,1})^{t+1} = \mathrm{Lin}\left(u_2^{t+1}, R_2^{t+1}, \overset{t+1}{H^{(1,2)}}, G_2^t, G_{2,1}^t, \lambda_2^t \right)$$

$$G_2^{t+1} = G_2^t + \mu \left(-B_2^{\overset{t+1}{\mathrm{sum}}} - D_2^{\overset{t+1}{\mathrm{sum}}} + R_2^{t+1} - H_{2,1}^{\overset{t+1}{\mathrm{full}}} - A_2 \right)$$

$$G_{2,1}^{t+1} = G_{2,1}^t + \mu \left(H_{2,1}^{t+1} - \overset{t+1}{H^{(1,2)}} \right)$$

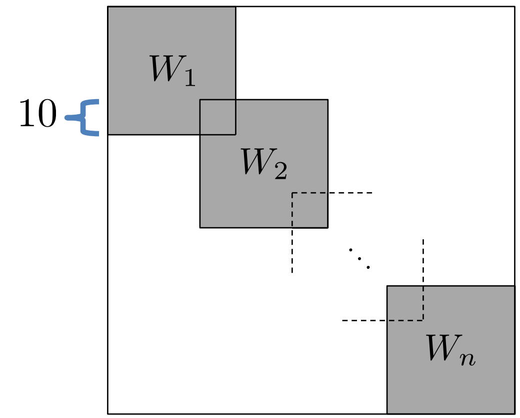
$$\lambda_2^{t+1} = \lambda_2^t + \mu \left(v_2^{t+1} - u_2^{t+1} \right)$$

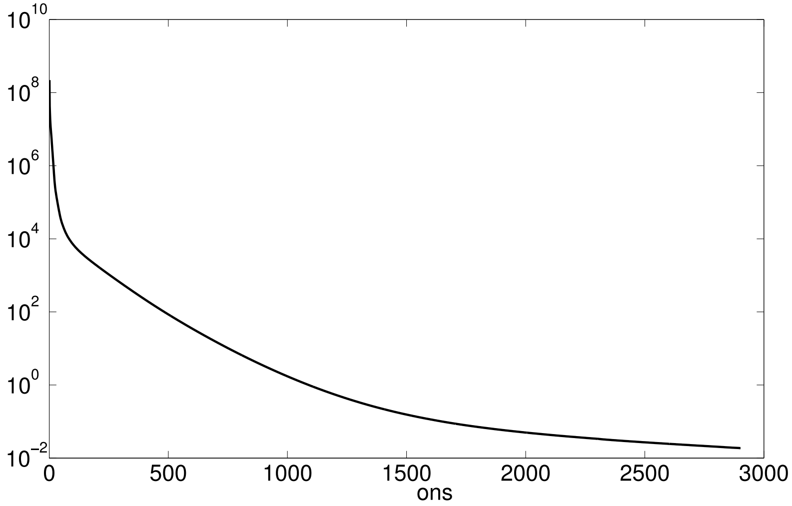


Simulations

❑ Large-scale SDP Simulations:

- 1000 agents: $N_{\text{Full}} = 0.9$ billion, $N_{\text{Decomp}} = 1.6$ million
- 2000 agents: $N_{\text{Full}} = 3.6$ billion, $N_{\text{Decomp}} = 3.2$ million
- 4000 agents: $N_{\text{Full}} = 14.4$ billion, $N_{\text{Decomp}} = 6.4$ million





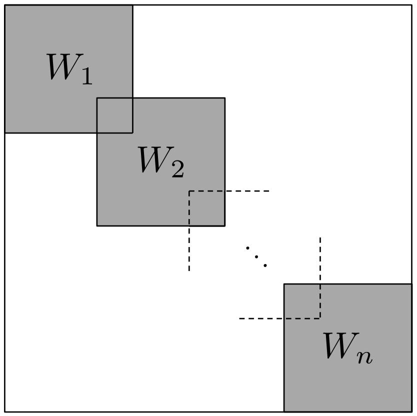
Aggregate residue for the case of 4000 agents with pi =qi =5

		1000	2000	4000
$p_i = 5$, $q_i = 5$	P_{obj}	1.192407e+06	2.373408e+06	4.741277e+06
	D_{obj}	1.192402e+06	2.373401e+06	4.741266e+06
	iter	2323	2754	2902
	t_{CPU} (min)	8.75	21.58	49
	t_{iter} (sec per iter)	0.23	0.47	1.01
	Optimality	99.9995%	99.9997%	99.9997%

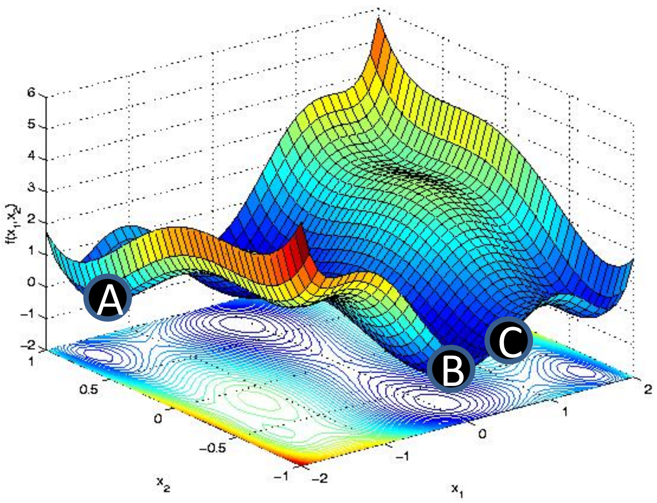
Simulation results for 1000, 2000 and 4000 agents

Amazon EC2 Simulations (36 cores, 60 GB RAM)

- Overlapping Cliques: n = 8000, p=5, q=5
  - 19.5 minutes
- Overlapping Cliques: n = 8000, p=5, q=0
  - 2.2 minutes
- Overlapping Cliques: n = 8000, p=0, q=5
  - 7.9 minutes
- Overlapping Cliques: n = 4000, p=5, q=5
  - 8.08 minutes (Amazon)
  - 49 minutes (Laptop)



Conclusions



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

Approach: Conic relaxation

- ❑ **Handling nonlinearity in continuous variables**
- ❑ **Handling discrete variables**
- ❑ **Handling noisy and imperfect information**
- ❑ **Numerical algorithms**