Graph-Theoretic Algorithm for Arbitrary Polynomial Optimization Problems

Javad Lavaei

Department of Electrical Engineering
Columbia University

Joint work with:
Ramtin Madani, Ghazal Fazelnia and Abdulrahman Kalbat (Columbia University)
Somayeh Sojoudi (New York University)
Ross Baldick (University of Texas-Austin)
Decision making towards real-time operation:

- Slow time scale: Optimize the operating cost → Centralized or distributed optimization
- Fast time scale: Regulate the signals → Distributed control
- Robustness: Make the system robust and resilient → Constraint identification

Javad Lavaei, Columbia University
Contingency Identification: Contingency Analysis

- **Contingency Analysis:**
  - Assume a line is disconnected.
  - Many generators cannot change productions quickly.
  - The flows over other lines would increase.
  - This triggers a cascading failure.

- **N-1 contingency:** Design the operating point such that the flow limits are satisfied under every single fault.

- **Challenge:** Number of constraints on the order of $l^2$.

- **Example:** Number of contingency constraints for Polish system with ~3300 buses = ~17 millions
Contingency Analysis

- **Bad news**: There are \( l^2 \) contingency constraints.
- **Good news**: Just \( n \) of them would be binding.
- **Strategy**: Identify constraints that are implied by others and then remove them.

1- Constraint for line \((i,j)\) under normal conditions
2- Constraint for line \((i,j)\) under contingency 1
3- Constraint for line \((i,j)\) under contingency 2
Contingency Analysis

- **Parameters:** $P_1, P_2, ..., P_n$
- **Constraints:** $l^2$ linear constraints in terms of $P_1, P_2, ..., P_n$.
- **Feasible injection region:**
- Some constraints do not define a face.
- **Observation:** A majority of constraints are not important because:
  
  *Normal Rating* < *Long-term emergency* < *Short-term emergency*

**Main idea behind our geometric approach:**

- Mapping 1: injections $\rightarrow$ flows with no fault
- Mapping 2: injections $\rightarrow$ flows with one specific fault
- Transformation from mapping 1 to mapping 2 is not too far from identity.
Many constraints are always unimportant, independent of the on/off status of generators, production levels, and load profiles.

<table>
<thead>
<tr>
<th>System</th>
<th>Total constraints</th>
<th>Important constraints</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 14</td>
<td>400</td>
<td>95</td>
<td>0.2375</td>
</tr>
<tr>
<td>Case 30</td>
<td>1681</td>
<td>271</td>
<td>0.1612</td>
</tr>
<tr>
<td>Case 39</td>
<td>2116</td>
<td>253</td>
<td>0.1195</td>
</tr>
<tr>
<td>Case 57</td>
<td>6400</td>
<td>642</td>
<td>0.1003</td>
</tr>
<tr>
<td>Case 118</td>
<td>34596</td>
<td>934</td>
<td>0.0269</td>
</tr>
<tr>
<td>Case 300</td>
<td>168921</td>
<td>2492</td>
<td>0.0148</td>
</tr>
</tbody>
</table>
Optimization

- Optimization:
  - Optimal power flow (OPF)
  - Security-constrained OPF
  - State estimation
  - Network reconfiguration
  - Unit commitment
  - Dynamic energy management

- Issue of non-convexity:
  - Discrete parameters
  - Nonlinearity in continuous variables

- Challenge: ~90% of decisions are made in day ahead and ~10% are updated iteratively during the day so a local solution remains throughout the day.
Penalized Semidefinite Programming (SDP) Relaxation

- Quadratic optimization in $\mathbf{v}$
  - Linear optimization in $\mathbf{vv}^*$
  - Remove the rank constraint and possibly penalize its effect
  - Replace $\mathbf{vv}^*$ with a matrix $\mathbf{W}$ subject to $\mathbf{W} \succeq 0$ and $\text{rank}\left\{\mathbf{W}\right\} = 1$

- Exactness of SDP relaxation:
  - Existence of a rank-1 solution
  - Implies finding a global solution

\[
\begin{align*}
\min_{\mathbf{v} \in \mathbb{R}^n} & \quad \mathbf{v}^* \mathbf{M}_0 \mathbf{v} \\
\text{s.t.} & \quad \mathbf{v}^* \mathbf{M}_i \mathbf{v} \leq 0, \quad i = 1, 2, \ldots, t
\end{align*}
\]

\[
\begin{align*}
\min_{\mathbf{W} \in \mathbb{S}_n^+} & \quad \text{trace}\{\mathbf{M}_0 \mathbf{W}\} \\
\text{s.t.} & \quad \text{trace}\{\mathbf{M}_i \mathbf{W}\} \leq 0, \quad i = 1, \ldots, t \\
& \quad \mathbf{W} \succeq 0
\end{align*}
\]
Optimal Power Flow

\[
\begin{align*}
\min_{v, P_G, Q_G} & \quad \sum_{k \in \mathcal{G}} f_k(P_{G_k}) \\
\text{Subject to} & \quad P_{k}^{\text{min}} \leq P_{G_k} \leq P_{k}^{\text{max}} \quad (1b) \\
& \quad Q_{k}^{\text{min}} \leq Q_{G_k} \leq Q_{k}^{\text{max}} \quad (1c) \\
& \quad V_{k}^{\text{min}} \leq |V_{k}| \leq V_{k}^{\text{max}} \quad (1d) \\
& \quad \text{Re} \left\{ V_l (V_l - V_m)^* y_{lm}^* \right\} \leq P_{lm}^{\text{max}} \quad (1e) \\
& \quad \text{trace} \{ V V^* Y^* e_k e_k^* \} = P_{G_k} - P_{D_k} + (Q_{G_k} - Q_{D_k})i \quad (1f)
\end{align*}
\]

**Trick:** Replace \(VV^*\) with a matrix \(W \succeq 0\) subject to rank\(\{W\} = 1\).

- **Observation:** SDP relaxation works for almost all benchmark examples and several data sets for US and Europe.

- **Theory:** Lots of theories to support this observation.
Penalized SDP Relaxation

- What if we get a low-rank but not rank-1 solution?

Penalized SDP relaxation:

\[ \sum_{k \in G} f_k(P_{G_k}) \rightarrow \sum_{k \in G} f_k(P_{G_k}) + \varepsilon \sum_{k \in G} Q_{G_k} \]

- Penalized SDP relaxation aims to find a near-optimal solution.

- It worked for IEEE systems with over 7000 different cost functions.

- Near-optimal solution coincided with the IPM’s solution in 100%, 96.6% and 95.8% of cases for IEEE 14, 30 and 57-bus systems.
Penalized SDP Relaxation

- Let $\lambda_1$ and $\lambda_2$ denote the two largest eigenvalues of $W$.

<table>
<thead>
<tr>
<th>IEEE-14</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon$</td>
<td>0</td>
<td>0.012</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>15.1617</td>
<td>15.1340</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.0138</td>
<td>0</td>
</tr>
<tr>
<td>Cost</td>
<td>$316.08$</td>
<td>$316.13$</td>
</tr>
</tbody>
</table>

- Correction of active powers is negligible but reactive powers change noticeably.

- There is a wide range of values for $\epsilon$ giving rise to a nearly-global local solution.
Penalized SDP Relaxation

### IEEE-30

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>0</th>
<th>0.55</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>30.6789</td>
<td>30.8677</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.4986</td>
<td>0</td>
</tr>
<tr>
<td>Cost</td>
<td>$414.34$</td>
<td>$438.40$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$k$</th>
<th>$c_k$</th>
<th>$P_{G_k}$</th>
<th>$Q_{G_k}$</th>
<th>$P_{G_k}$</th>
<th>$Q_{G_k}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>80</td>
<td>11.11</td>
<td>80</td>
<td>-4.60</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0</td>
<td>39.16</td>
<td>0</td>
<td>-2.10</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>40</td>
<td>44.70</td>
<td>40</td>
<td>44.70</td>
</tr>
<tr>
<td>33</td>
<td>10</td>
<td>23.98</td>
<td>35.26</td>
<td>27.32</td>
<td>33.36</td>
</tr>
<tr>
<td>23</td>
<td>100</td>
<td>0</td>
<td>33.39</td>
<td>0</td>
<td>15.62</td>
</tr>
<tr>
<td>27</td>
<td>1</td>
<td>54.55</td>
<td>25.65</td>
<td>45.22</td>
<td>21.33</td>
</tr>
</tbody>
</table>

### IEEE-57

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>0</th>
<th>1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>57.1776</td>
<td>56.8887</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.0767</td>
<td>0</td>
</tr>
<tr>
<td>Cost</td>
<td>$259.70$</td>
<td>$272.73$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$k$</th>
<th>$c_k$</th>
<th>$P_{G_k}$</th>
<th>$Q_{G_k}$</th>
<th>$P_{G_k}$</th>
<th>$Q_{G_k}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>575.88</td>
<td>78.60</td>
<td>575.88</td>
<td>111.87</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>100</td>
<td>50</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>44.29</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>13.11</td>
<td>117.90</td>
<td>14.41</td>
<td>159.64</td>
</tr>
<tr>
<td>9</td>
<td>0.1</td>
<td>100</td>
<td>9</td>
<td>100</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>0.1</td>
<td>410</td>
<td>96.91</td>
<td>410</td>
<td>-6.29</td>
</tr>
</tbody>
</table>

---

Javad Lavaei, Columbia University
Example with Multiple Solutions

- Example borrowed from Bukhsh et al.:
  - Modify IEEE 118-bus system has 3 local solutions with the optimal costs 129625.03, 177984.32 and 195695.54.
  - Our method finds the best one.
Tree decomposition:

We map a given graph $G$ into a tree $T$ such that:
- Each node of $T$ is a collection 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 a tree decomposition: The cardinality of largest node minus one

Treewidth of graph: The smallest width of all tree decompositions
Treewidth of a tree: 1

How about the treewidth of IEEE 14-bus system with multiple cycles?

<table>
<thead>
<tr>
<th>System $G$</th>
<th>$tw{G}$</th>
<th>System $G$</th>
<th>Bound on $tw{G}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE 14-bus</td>
<td>2</td>
<td>Polish 2383wp</td>
<td>23</td>
</tr>
<tr>
<td>IEEE 30-bus</td>
<td>3</td>
<td>Polish 2736sp</td>
<td>29</td>
</tr>
<tr>
<td>New England 39-bus</td>
<td>3</td>
<td>Polish 2746wop</td>
<td>29</td>
</tr>
<tr>
<td>IEEE 57-bus</td>
<td>5</td>
<td>Polish 3012wp</td>
<td>25</td>
</tr>
<tr>
<td>IEEE 118-bus</td>
<td>4</td>
<td>Polish 3120sp</td>
<td>26</td>
</tr>
<tr>
<td>IEEE 300-bus</td>
<td>6</td>
<td>Polish 3375wp</td>
<td>28</td>
</tr>
</tbody>
</table>
Real/complex optimization

Define $G$ as the sparsity graph

**Theorem:** There exists a solution with rank at most treewidth of $G + 1$

We proposed infinitely many optimizations to find that solution.

**Technique:**

- Choose a set of edges $E$ from the sparsity graph through our notion of enriched graph.
- For every $(i,j)$ in $E$, add a penalty term $t_{ij}X_{ij}$ to the objective, where $t_{ij}$ is arbitrary (nonzero).
- All solutions of the penalized problem have rank at most treewidth of $G + 1$. 

\[
\begin{align*}
\min & \quad \text{trace}\{F_0X\} \\
\text{s.t.} & \quad \text{trace}\{F_kX\} \leq 0 \quad \text{for} \quad k = 1, \ldots, p \\
& \quad X_{11} = 1 \\
& \quad X \succeq 0
\end{align*}
\]
Example: Consider the security-constrained unit-commitment OPF problem.

Use SDP relaxation for this mixed-integer nonlinear program.

What is the rank of $X^{opt}$?

1. IEEE 300-bus system: rank ≤ 6
2. Polish 3120-bus system: Rank ≤ 27

Use penalization to make them rank-1.

We have written a solver in Java and MATLAB (using CVX and MOSEK) to find a near-global solution.

Computation time for penalized SDP relaxation of ACOPF for Polish System with ~3300 buses (over 9,000,000 parameters): ~2.4 minutes.
A system with 41 lines and 21 contingencies (red lines)

Optimal cost of OPF: 576.89

Optimal cost of convex relaxation of SCOPF: 579.09

Near-optimal cost of SCOPF for the penalty factor 0.1: 579.10
Polynomial Optimization

\[
polynomial\ optimization \iff dense\ QCQP \iff sparse\ QCQP
\]

- **Vertex Duplication Procedure:**
  \[
  x_i \iff (x_{i1}, x_{i2}) \quad \text{s.t.} \quad x_{i1} = x_{i2}
  \]

- **Edge Elimination Procedure:**
  \[
  x_i x_j \iff z_1^2 - z_2^2 \quad \text{s.t.} \quad z_1 = \frac{x_i + x_j}{2}, \quad z_2 = \frac{x_i - x_j}{2}
  \]

- This gives rise to a sparse QCQP with a sparse graph.

- The treewidth can be reduced to 2.

**Theorem:** Every polynomial optimization has a QCQP formulation whose SDP relaxation has a solution with rank 1 or 2.
Consider the time-varying system:

\[
\begin{align*}
    y[\tau] &= C[\tau]x[\tau]
\end{align*}
\]

The goal is to design a structured controller \( u[\tau] = K y[\tau] \) to minimize

\[
\sum_{\tau=0}^{p} (x[\tau]^T Q[\tau] x[\tau] + u[\tau]^T R[\tau] u[\tau]) + \mu \text{ trace}\{KK^T\}
\]

- **Optimal centralized control**: Easy (LQR, LQG, etc.)
- **Optimal distributed control (ODC)**: NP-hard (Witsenhausen’s example)
Two Quadratic Formulations in Static Case

- **Formulation in time domain:**
  - Stack the free parameters of $K$ in a vector $h$.
  - Define $v$ as:
    \[ v = \begin{bmatrix} 1 & h^* & x[0]^* & x[1]^* & \cdots & x[p]^* & y[0]^* & \cdots & y[p]^* & u[0]^* & \cdots & u[p]^* \end{bmatrix}^* \]

- **Formulation in Lypunov domain:**
  - Consider the BMI constraint:
    \[ \begin{bmatrix} P & P(A + BKC)^* \\ (A + BKC)P & P \end{bmatrix} \succ 0 \]
  - Define $v$ as:
    \[ v = \begin{bmatrix} 1 & h^* & P_{11} & P_{12} & \cdots & P_{1n} & \cdots & P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix}^* \]
The proposed SDP relaxation was obtained from a big vector.

**Computationally cheap SDP relaxation:** Define a relaxation in terms of a small super-vector.

**Strategy in time domain:**
- Solve a computationally cheap SDP relaxation.
- Recover the state parameters from the SDP matrix.
- Solve a second optimization to recover the controller.

**Strategy in Lyapunov domain:**
- Solve a computationally cheap “penalized” SDP relaxation.
- Recover the Lyapunov matrix from the SDP matrix.
- Solve a second optimization to recover the controller.
Problem: Adjust the mechanical power of each generator based on the angle and frequency of neighboring generators in an optimal way (discretization time: 0.05 sec).
Using Kron reduction, we look at the interaction among 10 generators.
Decentralized Communication Topology

Penalty: 61000
Performance loss: 6.86%
Localized Communication Topology

Penalty: 15000  Performance loss: 6.08%
Star Communication Topology

Penalty: 2300  Performance loss: 0.22%
Ring Communication Topology

Penalty: 670  Performance loss: 0.19%
Another Communication Topology

Penalty: 30  Performance loss: 0.009%
### Four Communication Topologies

- **Finite horizon problem over 80 samples:**

<table>
<thead>
<tr>
<th>Structure</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Normalized Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Fully Decentralized)</td>
<td>316.28</td>
<td>320.27</td>
<td>0.0125</td>
</tr>
<tr>
<td>2 (Localized)</td>
<td>316.28</td>
<td>319.62</td>
<td>0.0105</td>
</tr>
<tr>
<td>3 (Star Topology)</td>
<td>316.28</td>
<td>317.31</td>
<td>0.0033</td>
</tr>
<tr>
<td>4 (Complete Subgraphs)</td>
<td>316.28</td>
<td>316.55</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

- **Solution for the star communication topology:**

<table>
<thead>
<tr>
<th>Bus #</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>-0.21</td>
<td>1.04</td>
<td>-1.80</td>
<td>-0.70</td>
<td>3.88</td>
<td>2.17</td>
<td>-0.70</td>
<td>-2.08</td>
<td>-0.29</td>
<td>-1.41</td>
</tr>
<tr>
<td>31</td>
<td>-1.08</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>-0.83</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>-1.52</td>
<td></td>
<td>1.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>-1.48</td>
<td></td>
<td>1.33</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>-0.64</td>
<td></td>
<td></td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>-0.78</td>
<td></td>
<td></td>
<td></td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>-0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>-0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.57</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>2.20</td>
<td>1.90</td>
<td>-2.64</td>
<td>3.32</td>
<td>-1.97</td>
<td>-1.59</td>
<td>-0.59</td>
<td>0.58</td>
<td>-0.02</td>
<td>-1.75</td>
</tr>
<tr>
<td>31</td>
<td>-0.24</td>
<td>-0.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>-0.28</td>
<td>-0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>-0.49</td>
<td>-0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>-0.49</td>
<td>-0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>-0.63</td>
<td></td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>-0.42</td>
<td></td>
<td></td>
<td>-0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>-0.28</td>
<td></td>
<td></td>
<td></td>
<td>-0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>-0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>-0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Javad Lavaei, Columbia University
Conclusions

- **Focus**: Optimization and control
- **Goal**: Design efficient algorithms

- Three thrusts:
  - Global optimization
  - Distributed control
  - Contingency analysis

- We have written two solvers to be posted online soon.

Javad Lavaei, Columbia University