LP formulations for mixed-integer polynomial optimization problems

Daniel Bienstock and Gonzalo Muñoz, Columbia University
An **application**: the Optimal Power Flow problem (ACOPF)

**Input**: an undirected graph $G$.

- For every vertex $i$, **two** variables: $e_i$ and $f_i$
- For every edge $\{k, m\}$, **four** (specific) quadratics:

\[
H^P_{k,m}(e_k, f_k, e_m, f_m), \quad H^Q_{k,m}(e_k, f_k, e_m, f_m)
\]

\[
H^P_{m,k}(e_k, f_k, e_m, f_m), \quad H^Q_{m,k}(e_k, f_k, e_m, f_m)
\]

\[
\min \sum_k F_k \left( \sum_{\{k,m\} \in \delta(k)} H^P_{k,m}(e_k, f_k, e_m, f_m) \right)
\]

\[
\text{s.t.} \quad L^P_k \leq \sum_{\{k,m\} \in \delta(k)} H^P_{k,m}(e_k, f_k, e_m, f_m) \leq U^P_k \quad \forall k
\]

\[
L^Q_k \leq \sum_{\{k,m\} \in \delta(k)} H^Q_{k,m}(e_k, f_k, e_m, f_m) \leq U^Q_k \quad \forall k
\]

\[
V^L_k \leq \| (e_k, f_k) \| \leq V^U_k \quad \forall k.
\]

Function $F_k$ in the objective: **convex quadratic**
Complexity

**Theorem** (2011) Lavaei and Low: OPF is (weakly) NP-hard on trees.


**Complexity**

**Theorem** (2011) Lavaei and Low: OPF is (weakly) NP-hard on trees.


**Recent insight:** use the SDP relaxation (Lavaei and Low, 2009 + many others)

\[
\begin{align*}
& \min \sum_k F_k \left( \sum_{\{k,m\} \in \delta(k)} H_{k,m}^P(e_k, f_k, e_m, f_m) \right) \\
& \text{s.t. } L_k^P \leq \sum_{\{k,m\} \in \delta(k)} H_{k,m}^P(e_k, f_k, e_m, f_m) \leq U_k^P \quad \forall k \\
& L_k^Q \leq \sum_{\{k,m\} \in \delta(k)} H_{k,m}^Q(e_k, f_k, e_m, f_m) \leq U_k^Q \quad \forall k \\
& V_k^L \leq \|(e_k, f_k)\| \leq V_k^U \quad \forall k.
\end{align*}
\]
Complexity

Theorem (2011) Lavaei and Low: OPF is (weakly) NP-hard on trees.


Recent insight: use the SDP relaxation (Lavaei and Low, 2009 + many others)

Reformulation of ACOPF:

\[
\begin{align*}
\min \quad & F \bullet W \\
\text{s.t.} \quad & A_i \bullet W \leq b_i \quad i = 1, 2, \ldots \\
& W \succeq 0, \quad W \text{ of rank 1.}
\end{align*}
\]
Complexity

Theorem (2011) Lavaei and Low: OPF is (weakly) NP-hard on trees.


Recent insight: use the SDP relaxation (Lavaei and Low, 2009 + many others)

SDP Relaxation of OPF:

\[
\begin{align*}
\min & \quad F \cdot W \\
\text{s.t.} & \quad A_i \cdot W \leq b_i \quad i = 1, 2, \ldots \\
& \quad W \succeq 0.
\end{align*}
\]
Complexity

**Theorem** (2011) Lavaei and Low: OPF is (weakly) NP-hard on trees.


**Recent insight:** use the SDP relaxation (Lavaei and Low, 2009 + many others)

**SDP Relaxation** of OPF:

\[
\begin{align*}
\text{min} & \quad F \bullet W \\
\text{s.t.} & \quad A_i \bullet W \leq b_i \quad i = 1, 2, \ldots \\
& \quad W \succeq 0.
\end{align*}
\]

**Fact:** The SDP relaxation is often good! (“near” rank 1 solution).
**But:** the SDP relaxation is always slow on large graphs

- Real-life grids $\rightarrow > 10^4$ vertices
- SDP relaxation of OPF does not terminate

**But...**
**But:** the SDP relaxation is always slow on large graphs

- Real-life grids $\rightarrow > 10^4$ vertices
- SDP relaxation of OPF does not terminate

**But...**

**Fact?** Real-life grids have **small tree-width**

**Definition 1:** A graph has treewidth $\leq w$ if it has a chordal supergraph with clique number $\leq w + 1$
But: the SDP relaxation is always slow on large graphs

- Real-life grids \( \rightarrow > 10^4 \) vertices
- SDP relaxation of OPF does not terminate

But...

Fact? Real-life grids have small \textit{tree-width}

\textbf{Definition 2:} A graph has treewidth \( \leq w \) if it is a subgraph of an intersection graph of subtrees of a tree, with \( \leq w + 1 \) subtrees overlapping at any vertex
But: the SDP relaxation is always slow on large graphs

- Real-life grids $\rightarrow > 10^4$ vertices
- SDP relaxation of OPF does not terminate

But...
Fact? Real-life grids have small tree-width

Definition 2: A graph has treewidth $\leq w$ if it is a subgraph of an intersection graph of subtrees of a tree, with $\leq w + 1$ subtrees overlapping at any vertex

(Seymour and Robertson, late 1980s)
Tree-width

Let $G$ be an undirected graph with vertices $V(G)$ and edges $E(G)$.

A tree-decomposition of $G$ is a pair $(T, Q)$ where:

- $T$ is a tree. **Not** a subtree of $G$, just a tree
- For each vertex $t$ of $T$, $Q_t$ is a subset of $V(G)$. These subsets satisfy the two properties:
  1. For each vertex $v$ of $G$, the set $\{t \in V(T) : v \in Q_t\}$ is a subtree of $T$, denoted $T_v$.
  2. For each edge $\{u, v\}$ of $G$, the two subtrees $T_u$ and $T_v$ intersect.
- The **width** of $(T, Q)$ is $\max_{t \in T} |Q_t| - 1$.

→ two subtrees $T_u, T_v$ may overlap even if $\{u, v\}$ is **not** an edge of $G$
Tree-width

Let $G$ be an undirected graph with vertices $V(G)$ and edges $E(G)$.

A tree-decomposition of $G$ is a pair $(T, Q)$ where:

- $T$ is a tree. Not a subtree of $G$, just a tree
- For each vertex $t$ of $T$, $Q_t$ is a subset of $V(G)$. These subsets satisfy the two properties:
  1. For each vertex $v$ of $G$, the set $\{t \in V(T) : v \in Q_t\}$ is a subtree of $T$, denoted $T_v$.
  2. For each edge $\{u, v\}$ of $G$, the two subtrees $T_u$ and $T_v$ intersect.
- The width of $(T, Q)$ is $\max_{t \in T} |Q_t| - 1$. 
**But**: the SDP relaxation is always slow on large graphs

- Real-life grids → > 10\(^4\) vertices
- SDP relaxation of OPF does not terminate

**But...**

**Fact?** Real-life grids have *small tree-width*

**Matrix-completion Theorem**

   gives fast SDP implementations:

Real-life grids with \(\approx 3 \times 10^3\) vertices: → 20 minutes runtime
But: the SDP relaxation is always slow on large graphs
  - Real-life grids $\rightarrow > 10^4$ vertices
  - SDP relaxation of OPF does not terminate

But...
  Fact? Real-life grids have small tree-width

Matrix-completion Theorem

gives fast SDP implementations:

Real-life grids with $\approx 3 \times 10^3$ vertices: $\rightarrow$ 20 minutes runtime

$\rightarrow$ Perhaps low tree-width yields direct algorithms for ACOPF itself?

That is to say, not for a relaxation?
Much previous work using structured sparsity

- Bienstock and Özbay
- Wainwright and Jordan
- Grimm, Netzer, Schweighofer
- Laurent
- Lasserre et al
- Waki, Kim, Kojima, Muramatsu

older work ...

- Lauritzen (1996): tree-junction theorem
- Bertele and Briosti (1972): nonserial dynamic programming
- Bounded tree-width in combinatorial optimization (too many authors)
- Fulkerson and Gross (1965): matrices with consecutive ones
ACOPF, again

Input: an undirected graph $G$.

- For every vertex $i$, **two** variables: $e_i$ and $f_i$
- For every edge $\{k, m\}$, **four** (specific) quadratics:

\[
\begin{align*}
H_{k,m}^P(e_k, f_k, e_m, f_m), & \quad H_{k,m}^Q(e_k, f_k, e_m, f_m) \\
H_{m,k}^P(e_k, f_k, e_m, f_m), & \quad H_{m,k}^Q(e_k, f_k, e_m, f_m)
\end{align*}
\]

\[
\begin{align*}
\min \sum_k F_k \left( \sum_{\{k,m\} \in \delta(k)} H_{k,m}^P(e_k, f_k, e_m, f_m) \right) \\
\text{s.t.} \quad L_k^P \leq \sum_{\{k,m\} \in \delta(k)} H_{k,m}^P(e_k, f_k, e_m, f_m) \leq U_k^P \quad \forall k \\
L_k^Q \leq \sum_{\{k,m\} \in \delta(k)} H_{k,m}^Q(e_k, f_k, e_m, f_m) \leq U_k^Q \quad \forall k \\
V_k^L \leq \| (e_k, f_k) \| \leq V_k^U \quad \forall k.
\end{align*}
\]

Function $F_k$ in the objective: convex quadratic
ACOPF, again

Input: an undirected graph $G$.

- For every vertex $i$, two variables: $e_i$ and $f_i$
- For every edge $\{k, m\}$, four (specific) quadratics:

\[
\begin{align*}
H_{k,m}^P(e_k, f_k, e_m, f_m), & \quad H_{k,m}^Q(e_k, f_k, e_m, f_m) \\
H_{m,k}^P(e_k, f_k, e_m, f_m), & \quad H_{m,k}^Q(e_k, f_k, e_m, f_m)
\end{align*}
\]

\[
\begin{align*}
\min \sum_k w_k \\
\text{s.t.} \quad L_k^P \leq \sum_{\{k,m\} \in \delta(k)} H_{k,m}^P(e_k, f_k, e_m, f_m) \leq U_k^P \quad \forall k \\
L_k^Q \leq \sum_{\{k,m\} \in \delta(k)} H_{k,m}^Q(e_k, f_k, e_m, f_m) \leq U_k^Q \quad \forall k \\
V_k^L \leq \|(e_k, f_k)\| \leq V_k^U \quad \forall k \\\n\omega_k = F_k \left( \sum_{\{k,m\} \in \delta(k)} H_{k,m}^P(e_k, f_k, e_m, f_m) \right) \quad \forall k
\end{align*}
\]
Graphical QCQP

Input: an undirected graph \( G \).

- For every vertex \( k \), a set of variables: \( \{x_j : j \in I(k)\} \)
- For every edge \( e = \{k, m\} \), a quadratic
  \[
  H_e(x) = H_e(\{x_j : j \in I(k) \cup I(m)\}).
  \]
- For now, the sets \( I(k) \) are disjoint

\[
\begin{align*}
\min \quad & \sum_k \sum_{j \in I(k)} c_{k,j} x_j \\
\text{s.t.} \quad & \sum_{e \in \delta(k)} H_e(x) \leq b_k \quad \forall k \\
& 0 \leq x_j \leq 1, \quad \forall j
\end{align*}
\]

\( \rightarrow \) Easy to solve if graph has small tree-width?
Subset-sum problem

**Input:** positive integers $p_1, p_2, \ldots, p_n$.

**Problem:** find a solution to:

$$\sum_{j=1}^{n} p_j x_j = \frac{1}{2} \sum_{j=1}^{n} p_j$$

$$x_j \in \{0, 1\}, \quad \forall j$$

*(weakly) NP-hard (well...)*
Subset-sum problem

Input: positive integers $p_1, p_2, \ldots, p_n$.

Problem: find a solution to:

$$\sum_{j=1}^{n} p_j x_j = \frac{1}{2} \sum_{j=1}^{n} p_j$$

$$x_j (1 - x_j) = 0, \quad \forall j$$

(weakly) NP-hard (well...)

This is a graphical QCQP on a star – so treewidth 1.

(Perhaps) approximate solutions?

$\{0, 1\}$ solutions with error $\left(\frac{1}{2} \sum_{j=1}^{n} p_j\right) \epsilon$ in time polynomial in $\epsilon^{-1}$?
Graphical QCQP

Input: an undirected graph $G$.

- For every vertex $k$, a set of variables: $\{x_j : j \in I(k)\}$
- For every edge $e = \{k, m\}$, a quadratic
  \[
  H_e(x) = H_e(\{x_j : j \in I(k) \cup I(m)\}).
  \]
- For now, the sets $I(k)$ are disjoint

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad \sum_{e \in \delta(k)} H_e(x) \leq b_k \quad \forall k \\
& \quad 0 \leq x_j \leq 1, \quad \forall j
\end{align*}
\]
Graphical PCLP

Input: an undirected graph $G$.
- For every vertex $k$, a set of variables: $\{x_j : j \in I(k)\}$
- For every edge $e = \{k, m\}$, a polynomial
  $$P_e(x) = P_e(\{x_j : j \in I(k) \cup I(m)\}).$$

$$\min \quad c^T x$$
$$\text{s.t.} \quad \sum_{e \in \delta(k)} P_e(x) \leq b_k \quad \forall k$$
$$\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad 0 \leq x_j \leq 1, \quad \forall j$$

**Density** of a problem: size of largest set $I(k)$
Density of ACOPF problems: 3
Graphical, mixed-integer PCLP – or GMIPCLP

Input: an undirected graph $G$.

- For every vertex $k$, a set of variables: $\{x_j : j \in I(k)\}$
- For every edge $e = \{k, m\}$, a polynomial
  \[
P_e(x) = P_e(\{x_j : j \in I(k) \cup I(m)\}).
  \]

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad \sum_{e \in \delta(k)} P_e(x) \leq b_k \quad \forall k \\
& \quad x_j \in \{0, 1\} \quad \forall j \in I, \quad 0 \leq x_j \leq 1, \quad \text{otherwise}
\end{align*}
\]

- **Density** = size of largest $I(k)$
Graphical, mixed-integer PCLP – or GMIPCLP

Input: an undirected graph $G$.

- For every vertex $k$, a set of variables: $\{x_j : j \in I(k)\}$
- For every edge $e = \{k, m\}$, a polynomial

\[
P_e(x) = P_e(\{x_j : j \in I(k) \cup I(m)\}).
\]

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad \sum_{e \in \delta(k)} P_e(x) \leq b_k \quad \forall k \\
& \quad x_j \in \{0, 1\} \quad \forall j \in I, \quad 0 \leq x_j \leq 1, \quad \text{otherwise}
\end{align*}
\]

- **Density** = size of largest $I(k)$

**Theorem 3**

For any instance of GMIPCLP on a graph with treewidth $w$, density $d$, max. degree $\pi$, and any fixed $0 < \epsilon < 1$, there is a linear program of size (rows + columns) $O^*(\pi^{wd}\epsilon^{-w}n)$ whose feasibility and optimality error is $O(\epsilon)$ (abridged).
More general: MIPCLP (Basic polynomially-constrained mixed-integer LP)

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad P_i(x) \leq b_i \quad 1 \leq i \leq m \\
& \quad x_j \in \{0, 1\} \quad \forall j \in I, \quad 0 \leq x_j \leq 1, \quad \text{otherwise} \\
\end{align*}
\]

Each \( P_i(x) \) is a polynomial.
More general: MIPCLP (Basic polynomially-constrained mixed-integer LP)

\[
\begin{align*}
\text{min} & \quad c^T x \\
\text{s.t.} & \quad P_i(x) \leq b_i \quad 1 \leq i \leq m \\
& \quad x_j \in \{0, 1\} \quad \forall j \in I, \quad 0 \leq x_j \leq 1, \quad \text{otherwise}
\end{align*}
\]

Each \( P_i(x) \) is a polynomial.

Theorem 2

For any instance of MIPCLP whose intersection graph has treewidth \( w \), max. degree \( \pi \), and any fixed \( 0 < \epsilon < 1 \), there is a linear program of size (rows + columns) \( O^*(\pi^w \epsilon^{-w-1} n) \) whose feasibility and optimality error is \( O(\epsilon) \) (abridged).
More general: MIPCLP (Basic polynomially-constrained mixed-integer LP)

\[ \text{min } c^T x \]
\[ \text{s.t. } P_i(x) \leq b_i \quad 1 \leq i \leq m \]
\[ x_j \in \{0, 1\} \quad \forall j \in I, \quad 0 \leq x_j \leq 1, \quad \text{otherwise} \]

Each \( P_i(x) \) is a polynomial.

**Theorem 2**

For any instance of MIPCLP whose intersection graph has treewidth \( w \), max. degree \( \pi \), and any fixed \( 0 < \epsilon < 1 \), there is a linear program of size (rows + columns) \( O^*(\pi^w \epsilon^{-w-1} n) \) whose feasibility and optimality error is \( O(\epsilon) \) (abridged).

**Intersection graph of a constraint system:** (Fulkerson? (1962?))

- Has a vertex for every variably \( x_j \)
- Has an edge \( \{x_i, x_j\} \) whenever \( x_i \) and \( x_j \) appear in the same constraint
Theorem 2

For any instance of MIPCLP whose intersection graph has treewidth $w$, max. degree $\pi$, and any fixed $0 \leq \epsilon < 1$, there is a linear program of size (rows + columns) $O^*(\pi^w \epsilon^{-w-1} n)$ whose feasibility and optimality error is $O(\epsilon)$ (abridged).

Theorem 3

For any instance of GMIPCLP on a graph $G$ with treewidth $w$, density $d$, max. degree $\pi$, and any fixed $0 \leq \epsilon < 1$, there is a linear program of size (rows + columns) $O^*(\pi^{wd} \epsilon^{-wd} n)$ whose feasibility and optimality error is $O(\epsilon)$ (abridged).
Theorem 2

For any instance of MIPCLP whose intersection graph has treewidth \( w \), max. degree \( \pi \), and any fixed \( 0 < \epsilon < 1 \), there is a linear program of size (rows + columns) \( O^*(\pi^w \epsilon^{w-1} n) \) whose feasibility and optimality error is \( O(\epsilon) \) (abridged).

Theorem 3

For any instance of GMIPCLP on a graph \( G \) with treewidth \( w \), density \( d \), max. degree \( \pi \), and any fixed \( 0 < \epsilon < 1 \), there is a linear program of size (rows + columns) \( O^*(\pi^{wd} \epsilon^{-w} n) \) whose feasibility and optimality error is \( O(\epsilon) \) (abridged).

Two graphs:
• \( G \), the graph of the instance
• \( H \), the intersection graph of the constraints

→ Even if \( G \) has small treewidth, \( H \) might not

Example: subset sum problem. \( G \) is a star, \( H \) is a clique.
Theorem 0

Given an instance of graphical mixed-integer PCLP

- On a graph $G$ of treewidth $w$,
- with density $d$,
- For every vertex $k$, a set of variables: $\{x_j : j \in I(k)\}$, for every edge $e = \{k, m\}$, a polynomial

$$P_e(x) = P_e(\{x_j : j \in I(k) \cup I(m)\}).$$

$$\min c^T x$$
$$\text{s.t. } \sum_{e \in \delta(k)} P_e(x) \leq b_k \quad \forall k$$
$$x_j \in \{0, 1\}, \quad \forall j \in I, \quad 0 \leq x_j \leq 1, \quad \text{otherwise.}$$

**Density** of a problem: size of largest set $I(k)$

There is an **equivalent**

**mixed-integer polynomial optimization problem**

whose **intersection graph** has tree-width $O(wd)$. 
Theorem 0

Given an instance of graphical mixed-integer PCLP

- On a graph $G$ of treewidth $w$,
- with density $d$,

There is an equivalent mixed-integer polynomial optimization problem whose intersection graph has tree-width $O(wd)$.

ACOPF problem on small treewidth graph $\rightarrow$ (generalize)

Graphical QCQP on small treewidth graph and small density $\rightarrow$ (generalize)

GMIPCLP on small treewidth graph and small density $\rightarrow$ (generalize, reduce)

Mixed-integer PCLP with small treewidth intersection graph

**Basic theorem:**

There is a polynomial-time $\epsilon$-approximate algorithm for such problems
Main technique: approximation through pure-binary problems

Glover, 1975 (abridged)

Let \(x\) be a variable, with bounds \(0 \leq x \leq 1\). Let \(0 < \gamma < 1\). Then we can approximate

\[
x \approx \sum_{h=1}^{L} 2^{-h} y_h
\]

where each \(y_h\) is a binary variable. In fact, choosing \(L = \lceil \log_2 \epsilon^{-1} \rceil\), we have

\[
x \leq \sum_{h=1}^{L} 2^{-h} y_h \leq x + \epsilon.
\]

→ Given a mixed-integer polynomially constrained LP (MIPCLP), apply this technique to each continuous variable \(x_j\).
Mixed-integer polynomially-constrained LP:

\[(P) \quad \min \ c^T x \]

s.t. \( P_i(x) \leq b_i \quad 1 \leq i \leq m \)

\( x_j \in \{0, 1\} \quad \forall j \in I, \quad 0 \leq x_j \leq 1, \quad \text{otherwise} \)

substitute: \( \forall j \notin I, \quad x_j \rightarrow \sum_{h=1}^{L} 2^{-h} y_{h,j} \), where each \( y_{h,j} \in \{0, 1\} \)

\( L \approx \log_2 \epsilon^{-1} \)
Mixed-integer polynomially-constrained LP:

\[(P) \quad \min \ c^T x \]
\[
\text{s.t.} \quad P_i(x) \leq b_i \quad 1 \leq i \leq m
\]
\[
x_j \in \{0, 1\} \quad \forall j \in I, \quad 0 \leq x_j \leq 1, \quad \text{otherwise}
\]

substitute: \( \forall j \notin I, \ x_j \rightarrow \sum_{h=1}^{L} 2^{-h} y_{h,j} \), where each \( y_{h,j} \in \{0, 1\} \)

\[L \approx \log_2 \epsilon^{-1}\]

obtain pure binary problem:

\[(Q) \quad \min \ \hat{c}^T z \]
\[
\text{s.t.} \quad \hat{P}_i(z) \leq \hat{b}_i \quad 1 \leq i \leq m
\]
\[
z_k \in \{0, 1\} \quad \forall k
\]

If \( (P) \) has intersection graph of treewidth \( w \),
then \( (Q) \) has intersection graph of treewidth \( Lw \).
Mixed-integer polynomially-constrained LP:

(P) \quad \min c^T x \\
\text{s.t.} \quad P_i(x) \leq b_i \quad 1 \leq i \leq m \\
\text{otherwise} \quad x_j \in \{0, 1\} \quad \forall j \in I, \quad 0 \leq x_j \leq 1 \\

substitute: \forall j \notin I, \quad x_j \rightarrow \sum_{h=1}^{L} 2^{-h} y_{h,j}, \text{ where each } y_{h,j} \in \{0, 1\} \\

L \approx \log_2 \epsilon^{-1}

obtain pure binary problem:

(Q) \quad \min \hat{c}^T z \\
\text{s.t.} \quad \hat{P}_i(z) \leq \hat{b}_i \quad 1 \leq i \leq m \\
\text{otherwise} \quad z_k \in \{0, 1\} \quad \forall k \\

If (P) has intersection graph of treewidth \( w \), 
then (Q) has intersection graph of treewidth \( Lw \).

Theorem

Consider a pure-binary PCLP with \( n \) variables.
If the intersection graph has treewidth \( \leq W \) then there is an exact linear programming formulation with

\( O(2^W n) \) variables and constraints.
Conclusion

Given an ACOPF problem on a graph of treewidth $\leq w$ and $n$ edges, and $0 \leq \epsilon \leq 1$ there is an LP formulation with the following properties:

- It has $O(\text{poly}(\epsilon^{-1})2^{O(w)n})$ variables and constraints
- It produces $\epsilon$-optimal and -feasible solutions.

Talk on Friday by Gonzalo on the pure-binary problems.