Solving QCQPs (joint with G. Muñoz)

Daniel Bienstock

March 2015
Quadratically constrained, quadratic programming:

$$\min \ f_0(x)$$

s.t. \quad f_i(x) \leq 0, \quad 1 \leq i \leq m

\quad x \in \mathbb{R}^n

Here,

$$f_i(x) = x^T M_i x + c_i^T x + d_i$$

is a general quadratic

Each $M_i$ is $n \times n$, wlog symmetric
Application: power flows in electrical transmission systems

Setting:

• A “grid,” given by a set of “buses” (nodes) and “lines” (arcs).

• Some buses represent generators, some other buses represent “loads”.

• **What we can control:** the behavior of generators (voltage, output).

• **What we cannot control:** most. System obeys the laws of physics.

• **What we can want:** to operate the grid in a safe and economic manner.
Optimal power flow problem in rectangular coordinates, simplest form

Variables:
- Complex voltages $e_k + jf_k$, power flows $P_{km}, Q_{km}$, auxiliary variables

Notation: For a bus $k$, $\delta(k) =$ set of lines incident with $k$; $V =$ set of buses

**Basic problem**

$$\begin{align*}
\text{min} \quad & \sum_{k \in V} C_k \\
\text{s.t.} \quad & \forall km:\quad P_{km} = g_{km}(e_k^2 + f_k^2) - g_{km}(e_ke_m + f_kf_m) + b_{km}(e_kf_m - f_ke_m) \quad (1a) \\
& \forall km:\quad Q_{km} = -b_{km}(e_k^2 + f_k^2) + b_{km}(e_ke_m + f_kf_m) + g_{km}(e_kf_m - f_ke_m) \quad (1b) \\
& \forall km:\quad |P_{km}|^2 + |Q_{km}|^2 \leq U_{km} \quad (1c) \\
& \forall k:\quad P_{k}^{\text{min}} \leq \sum_{km \in \delta(k)} P_{km} \leq P_{k}^{\text{max}} \quad (1d) \\
& \forall k:\quad Q_{k}^{\text{min}} \leq \sum_{km \in \delta(k)} Q_{km} \leq Q_{k}^{\text{max}} \quad (1e) \\
& \forall k:\quad V_{k}^{\text{min}} \leq e_k^2 + f_k^2 \leq V_{k}^{\text{max}}, \quad (1f) \\
& \forall k:\quad C_k = F_k \left( \sum_{km \in \delta(k)} P_{km} \right). \quad (1g)
\end{align*}$$

Here, $F_k$ is a quadratic function for each $k$. 
Optimal power flow problem in rectangular coordinates, simplest form

Variables:
- Complex voltages \( e_k + jf_k \), power flows \( P_{km}, Q_{km} \), auxiliary variables

Notation: For a bus \( k \), \( \delta(k) \) = set of lines incident with \( k \); \( V \) = set of buses

Basic problem

\[
\min \sum_{k \in V} C_k \\
\text{s.t.} \quad \forall km : \quad P_{km} = g_{km}(e_k^2 + f_k^2) - g_{km}(e_ke_m + f_kf_m) + b_{km}(e_kf_m - f_ke_m) \tag{2a}
\]

\[
\forall km : \quad Q_{km} = -b_{km}(e_k^2 + f_k^2) + b_{km}(e_ke_m + f_kf_m) + g_{km}(e_kf_m - f_ke_m) \tag{2b}
\]

\[
\forall km : \quad |P_{km}|^2 + |Q_{km}|^2 \leq U_{km} \tag{2c}
\]

\[
\forall k : \quad P_k^{\min} \leq \sum_{km \in \delta(k)} P_{km} \leq P_k^{\max} \tag{2d}
\]

\[
\forall k : \quad Q_k^{\min} \leq \sum_{km \in \delta(k)} Q_{km} \leq Q_k^{\max} \tag{2e}
\]

\[
\forall k : \quad V_k^{\min} \leq e_k^2 + f_k^2 \leq V_k^{\max}, \tag{2f}
\]

\[
\forall k : \quad C_k = G_k \left( \sum_{km \in \delta(k)} Q_{km} \right). \tag{2g}
\]

Here, \( G_k \) is a quadratic function for each \( k \).
Optimal power flow problem in rectangular coordinates, simplest form

Variables:

- Complex voltages $e_k + j f_k$, power flows $P_{km}, Q_{km}$, auxiliary variables

**Notation:** For a bus $k$, $\delta(k)$ = set of lines incident with $k$; $V$ = set of buses

**Basic problem**

$$\begin{align*}
\text{min} \quad & \sum_{k \in V} C_k \\
\text{s.t.} \quad & \forall km : \quad P_{km} = g_{km}(e_k^2 + f_k^2) - g_{km}(e_ke_m + f_kf_m) + b_{km}(e_kf_m - f_ke_m) \quad (3a) \\
& \forall km : \quad Q_{km} = -b_{km}(e_k^2 + f_k^2) + b_{km}(e_ke_m + f_kf_m) + g_{km}(e_kf_m - f_ke_m) \quad (3b) \\
& \forall km : \quad |P_{km}|^2 + |Q_{km}|^2 \leq U_{km} \quad (3c) \\
& \forall k : \quad P_{k}^{\min} \leq \sum_{km \in \delta(k)} P_{km} \leq P_{k}^{\max} \quad (3d) \\
& \forall k : \quad Q_{k}^{\min} \leq \sum_{km \in \delta(k)} Q_{km} \leq Q_{k}^{\max} \quad (3e) \\
& \forall k : \quad V_{k}^{\min} \leq e_k^2 + f_k^2 \leq V_{k}^{\max}, \quad (3f) \\
& \forall k : \quad C_k = F_k \left( \sum_{km \in \delta(k)} P_{km} \right) + G_k \left( \sum_{km \in \delta(k)} Q_{km} \right). \quad (3g)
\end{align*}$$

Here, $F_k, G_k$ are quadratic functions for each $k$. 
Optimal power flow problem in rectangular coordinates, simplest form

Variables:
- Complex voltages $e_k + jf_k$, power flows $P_{km}, Q_{km}$, auxiliary variables

Notation: For a bus $k$, $\delta(k) =$ set of lines incident with $k$; $V =$ set of buses

Basic problem

$$\begin{align*}
\text{min} & \quad \sum_{k \in V} C_k \\
\text{s.t.} & \quad \forall km: \quad P_{km} = g_{km}(e_k^2 + f_k^2) - g_{km}(e_ke_m + f_kf_m) + b_{km}(e_kf_m - f_ke_m) \\
& \quad \forall km: \quad Q_{km} = -b_{km}(e_k^2 + f_k^2) + b_{km}(e_ke_m + f_kf_m) + g_{km}(e_kf_m - f_ke_m) \\
& \quad \forall km: \quad |P_{km}|^2 + |Q_{km}|^2 \leq U_{km} \\
& \quad \forall k: \quad P_{k}^{\text{min}} \leq \sum_{km \in \delta(k)} P_{km} \leq P_{k}^{\text{max}} \\
& \quad \forall k: \quad Q_{k}^{\text{min}} \leq \sum_{km \in \delta(k)} Q_{km} \leq Q_{k}^{\text{max}} \\
& \quad \forall k: \quad V_{k}^{\text{min}} \leq e_k^2 + f_k^2 \leq V_{k}^{\text{max}} , \\
& \quad \forall k: \quad C_k = F_k \left( \sum_{km \in \delta(k)} P_{km} \right) + G_k \left( \sum_{km \in \delta(k)} Q_{km} \right). 
\end{align*}$$

Here, $F_k, G_k$ are quadratic functions for each $k$. Many possibilities, all structurally similar.
Optimal power flow problem in rectangular coordinates, simplest form

Variables:
- Complex voltages \( e_k + j f_k \), power flows \( P_{km}, Q_{km} \), auxiliary variables

Notation: For a bus \( k \), \( \delta(k) \) = set of lines incident with \( k \); \( V \) = set of buses

**Basic problem**

\[
\begin{align*}
\min & \quad \sum_{k \in V} C_k \\
\text{s.t.} & \quad \forall km: \ P_{km} = g_{km}(e_k^2 + f_k^2) - g_{km}(e_k e_m + f_k f_m) + b_{km}(e_k f_m - f_k e_m) \\
& \quad \forall km: \ Q_{km} = -b_{km}(e_k^2 + f_k^2) + b_{km}(e_k e_m + f_k f_m) + g_{km}(e_k f_m - f_k e_m) \\
& \quad \forall km: \ |P_{km}|^2 + |Q_{km}|^2 \leq U_{km} \\
& \quad \forall k: \ P_k^{\min} \leq \sum_{km \in \delta(k)} P_{km} \leq P_k^{\max} \\
& \quad \forall k: \ Q_k^{\min} \leq \sum_{km \in \delta(k)} Q_{km} \leq Q_k^{\max} \\
& \quad \forall k: \ V_k^{\min} \leq e_k^2 + f_k^2 \leq V_k^{\max} \\
& \quad C_k = F_k \left( \sum_{km \in \delta(k)} P_{km} \right) + G_k \left( \sum_{km \in \delta(k)} Q_{km} \right).
\end{align*}
\]

Here, \( F_k, G_k \) are quadratic functions for each \( k \). Many possibilities, all structurally similar.

These are QCQPs, quadratically constrained quadratic programs, with an underlying graph structure.
State-of-the-art

• Industry can solve routine problem instances easily.
State-of-the-art

• Industry can solve routine problem instances easily.

• These are instances where the grid is in an (intimately) familiar state.
State-of-the-art

• Industry can solve routine problem instances easily.

• These are instances where the grid is in an (intimately) familiar state.

• Unfamiliar and stressed states are difficult (impossible) to handle.
State-of-the-art

- Industry can solve routine problem instances easily.

- These are instances where the grid is in an (intimately) familiar state.

- Unfamiliar and stressed states are difficult (impossible) to handle.

As a result, what-if analyses become problematic. Agnostic what-if analysis are essentially impossible.
State-of-the-art

• Industry can solve routine problem instances easily.

• These are instances where the grid is in an (intimately) familiar state.

• Unfamiliar and stressed states are difficult (impossible) to handle. Agnostic what-if analysis are essentially impossible.

• Problem is strongly NP-hard. (A. Verma, 2009)
QCQPs

\[
\begin{align*}
\text{min} & \quad x^T M^0 x + 2c^T_0 x + d_0 \\
\text{s.t.} & \quad \forall km : \quad x^T M^i x + 2c^T_i x + d_i \geq 0, \quad 1 \leq i \leq m, \\
& \quad x \in \mathbb{R}^n.
\end{align*}
\]  

(6a) (6b) (6c)

Each matrix \( M^i \) symmetric.
This description includes linear inequalities, bounds on individual variables, quadratic/linear equations.
QCQPs

\[
\begin{align*}
\min & \quad x^T M^0 x + 2c_0^T x + d_0 \\
\text{s.t.} & \quad \forall km : x^T M^i x + 2c_i^T x + d_i \geq 0, \quad 1 \leq i \leq m, \\
& \quad x \in \mathbb{R}^n.
\end{align*}
\]

(7a)

(7b)

(7c)

Reformulation

observation: \( x^T M^i x + 2c_i^T x = (1 \ x^T) \begin{pmatrix} 0 & c_i^T \\ c_i & M^i \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} = (1 \ x^T) \tilde{M}^i \begin{pmatrix} 1 \\ x \end{pmatrix} \)

definition: for matrices \( A, B, \quad A \bullet B \doteq \sum_{i,j} a_{ij} b_{ij} \)

so for vector \( y \) and matrix \( A, \quad y^T A y = A \bullet yy^T \)

So QCQP can be rewritten as:

\[
\begin{align*}
Q^* & \doteq \min \quad \tilde{M}^0 \bullet X + d_0 \\
\text{s.t.} & \quad \forall km : M^i \bullet X + d_i \geq 0, \quad 1 \leq i \leq m, \\
& \quad X \in \mathbb{R}^{(n+1) \times (n+1)}, \quad X \succeq 0, \quad \text{of rank 1}.
\end{align*}
\]

(8a)

(8b)

(8c)

The semidefinite relaxation of this problem is:

\[
\begin{align*}
\tilde{Q} & \doteq \min \quad \tilde{M}^0 \bullet X + d_0 \\
\text{s.t.} & \quad \forall km : \tilde{M}^i \bullet X + d_i \geq 0, \quad 1 \leq i \leq m, \\
& \quad X \in \mathbb{R}^{(n+1) \times (n+1)}, \quad X \succeq 0.
\end{align*}
\]

(9a)

(9b)

(9c)

\[ \tilde{Q} \leq Q^* \]
The critical observation

• Lavaei and Low, 2011: the SDP relaxation of AC-OPF not infrequently is very tight
• This spurred much research
• Jabr, Hiskens and Molzahn, others
The critical observation

- Lavaei and Low, 2011: the SDP relaxation of AC-OPF not infrequently is very tight
- This spurred much research
- Jabr, Hiskens and Molzahn, others
- Under constrained conditions, the SDP relaxation can be weak
The critical observation

- Lavaei and Low, 2011: the SDP relaxation of AC-OPF not infrequently is very tight
- This spurred much research
- Jabr, Hiskens and Molzahn, others
- Under constrained conditions, the SDP relaxation can be weak
- The SDP relaxation can prove unsolvable for larger grids
The critical observation

• Lavaei and Low, 2011: the SDP relaxation of AC-OPF not infrequently is very tight
• This spurred much research
• Jabr, Hiskens and Molzahn, others
• Under constrained conditions, the SDP relaxation can be weak
• The SDP relaxation can prove unsolvable for larger grids
• Factoid: there are polynomial-time algorithms for SDP, but require many assumptions
• There is no exact algorithm for SDP
The critical observation

- Lavaei and Low, 2011: the SDP relaxation of AC-OPF not infrequently is very tight
- This spurred much research
- Jabr, Hiskens and Molzahn, others
- Under constrained conditions, the SDP relaxation can be weak
- The SDP relaxation can prove unsolvable for larger grids
- Factoid: there are polynomial-time algorithms for SDP, but require many assumptions
- There is no exact algorithm for SDP
- Lavaei, Low, Hiskens-Molzahn:
  - when the underlying network has low tree-width, the SDP relaxation can be solved much faster
    - why: standard SDP solvers can leverage low tree-width
- What exactly is tree-width?
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.
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:
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 $T_v$.  


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**.
**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 $\max_{t \in T} |Q_t| - 1$. 

![Diagram of tree-decomposition](image)
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 $\max_{t \in T} |Q_t| - 1$.

![Diagram](image1.png)  
![Diagram](image2.png)

→ two subtrees $T_u, T_v$ may overlap even if $\{u, v\}$ is **not** an edge of $G$
Fulkerson and Gross (1965), binary packing integer programs

\[
\begin{align*}
\text{IP} & = \max \ c^T x \tag{10a} \\
\text{s.t.} \quad Ax & \leq b, \tag{10b} \\
x & \in \{0, 1\}^n \tag{10c}
\end{align*}
\]

Here, \( A \) is has 0, 1-valued entries. **Idea:** use the structure of \( A \).

The **intersection graph of \( A \), \( G_A \),** has:

- A vertex for each column of \( A \).
- An edge between two columns \( j, k \) if there is a row \( i \) with \( a_{ij} \neq 0, \ a_{ik} \neq 0 \).
History

Fulkerson and Gross (1965), binary packing integer programs

\[
\text{IP} = \max \quad c^T x \\
\text{s.t.} \quad Ax \leq b, \\
\quad x \in \{0, 1\}^n
\]  

(11a) (11b) (11c)

Here, \( A \) is has 0, 1-valued entries. \textbf{Idea:} use the structure of \( A \).

The \textbf{intersection graph of} \( A \), \( G_A \), has:

- A vertex for each column of \( A \).
- An edge between two columns \( j, k \) if there is a row \( i \) with \( a_{ij} \neq 0, \ a_{ik} \neq 0 \).

Each row of \( A \) induces a clique of \( G_A \).
History

Fulkerson and Gross (1965), binary packing integer programs

\[
\text{IP} = \max \quad c^T x \\
\text{s.t.} \quad Ax \leq b, \\
x \in \{0, 1\}^n
\]

Here, \( A \) is has 0, 1-valued entries. **Idea:** use the structure of \( A \).

The **intersection graph** of \( A \), \( G_A \), has:

- A vertex for each column of \( A \).
- An edge between two columns \( j, k \) if there is a row \( i \) with \( a_{ij} \neq 0, a_{ik} \neq 0 \).

**Theorem.** If \( G_A \) is an **interval graph**, then

\[
\text{IP} = \text{LP} = \max \quad c^T x \\
\text{s.t.} \quad Ax \leq b, \\
x \in [0, 1]^n
\]

(13a)

(13b)

(13c)

(so IP = value of its continuous relaxation).

A graph \( G = (V, E) \) is an interval graph, if there is a **path** \( P \), and a family of subpaths \( P_v \) (one for each \( v \in V \)), such that

- For each **pair of vertices** \( u \) and \( v \) of \( G \), we have \( \{u, v\} \in E \) whenever \( P_u \) and \( P_v \) intersect.
- The largest clique size of \( G \) is \( \max_{p \in P} |\{v \in V : p \in P_v\}|. \)

(The maximum number of subpaths that simultaneously overlap anywhere on \( P \))
The **intersection graph of** $A$, $G_A$, has:

- A vertex for each column of $A$, an edge between two columns $j,k$ if there is a row $i$ with $a_{ij} \neq 0$, $a_{ik} \neq 0$.

**Definition:** (Gavril, 1974) A graph $G = (V,E)$ is **chordal**, if there exists

- A **tree** $T$, and a family of trees $P_v$ (one for each $v \in V$), such that
- For each **pair of vertices** $u$ and $v$ of $G$, we have $\{u,v\} \in E$ whenever $T_u$ and $T_v$ intersect.
- The largest clique size of $G$ is $\max_{t \in T} |\{v \in V : t \in T_v\}|$. (The maximum number of subtrees that simultaneously overlap anywhere on $T$)

(equivalent: a graph is chordal iff every cycle of length $> 3$ has a chord).
Contrast with tree-decompositions

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 $G$

So: A graph $G$ has a tree-decomposition of width $w$ iff there is a **chordal supergraph** of $G$ of clique number $w + 1$.
\[
\text{IP} = \max \quad c^T x \\
\text{s.t.} \quad Ax \leq b, \\
x \in \{0, 1\}^n
\]  
(15a)  
(15b)  
(15c)

The intersection graph of \( A \), \( G_A \), has:

- A vertex for each column of \( A \), an edge between two columns \( j, k \) if there is a row \( i \) with \( a_{ij} \neq 0, a_{ik} \neq 0 \).

**Definition:** (Gavril, 1974) A graph \( G = (V, E) \) is **chordal**, if there exists

- A **tree** \( T \), and a family of subtrees \( P_v \) (one for each \( v \in V \)), such that
- For each **pair of vertices** \( u \) and \( v \) of \( G \), we have \( \{u, v\} \in E \) iff \( \mathcal{T}_u \) and \( \mathcal{T}_v \) intersect.
- The **largest clique size of** \( G \) is \( \max_{t \in \mathcal{T}} |\{v \in V : t \in \mathcal{T}_v\}| \).

(The maximum number of subtrees that simultaneously overlap anywhere on \( T \))

(equivalent: a graph is chordal iff every cycle of length \( \geq 3 \) has a chord).

**Theorem.** If \( G_A \) is **chordal**, then

\[
\text{IP} = \text{LP} = \max \quad c^T x \\
\text{s.t.} \quad Ax \leq b, \\
x \in [0, 1]^n.
\]  
(16a)  
(16b)  
(16c)

(so IP = value of its continuous relaxation).

Chordal graphs are “nice.” In fact, they are **perfect**.
Why small tree-width helps

Cholesky factorization of:

\[ A = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * \\ * & * \\ * & * \end{pmatrix} \]
Cholesky factorization of:

\[
\begin{pmatrix}
1 & 2 & 3 & 4 \\
5 & 6 & 7 & 8
\end{pmatrix}
\]
Chordal supergraph:

Pivoting order: 1, 2, 5, 6, 7, 8, 3, 4
Graph Minors Project: Robertson and Seymour, 1983 - 2004

→ Tree-width as a measure of the complexity of a graph
CAUTION
CAUTION

\[ \text{sparsity} \neq \text{small tree-width} \]
CAUTION

sparsity $\neq$ small tree-width

$\exists$ graphs of max deg 3 and arbitrarily high tree-width
Graph Minors Project: Robertson and Seymour, 1983 - 2004

→ Tree-width as a measure of the complexity of a graph

• Algorithms community: small tree-width makes hard problems easy (late 1980s)

• Many NP-hard problems can be solved in polynomial time on graphs with small tree-width:
  TSP, max. clique, graph coloring, ...
Tree-width as a measure of the complexity of a graph

- Algorithms community: small tree-width makes hard problems easy (late 1980s)

- Many NP-hard problems can be solved in polynomial time on graphs with small tree-width: TSP, max. clique, graph coloring, ...

- Fellows & Langston; Bienstock & Langston; Arnborg, Corneil & Proskurowski; many other authors

- Common thread: exploit tree-decomposition to obtain good algorithms

- So-called “nonserial dynamic programming” (1972)
partial tree-decomposition
partial tree-decomposition

boundary
optimize in partial tree-decomposition
subject to ' boundary conditions '
partial tree-decomposition
enumerates several cases
optimize in partial tree-decomposition
subject to 'boundary conditions'
partial tree-decomposition enumerates several cases
also enumerate similar cases involving
the new set, and match
More recent history

- B. and Özbay. 2003. Tree-width and the Sherali-Adams reformulation operator. Implies that on graphs with tree-width $\leq \omega$, the Sherali-Adams reformulation for vertex packing, at level $\leq \omega$, is exact.

- Wainwright and Jordan, 2004. (Constraint satisfaction community). On an all-binary polynomial optimization problem whose constraint graph has tree-width $\leq \omega$, the Sherali-Adams reformulation for vertex packing, at level $\leq \omega$, is exact.

- Lasserre, Waki, others (2006-). Polynomial-size relaxations for continuous polynomial optimization problem if the underlying constraint graph has bounded tree-width.
More recent history

- B. and Özbay. 2003. Tree-width and the Sherali-Adams reformulation operator. Implies that on graphs with tree-width $\leq \omega$, the Sherali-Adams reformulation for vertex packing, at level $\leq \omega$, is exact.

- Wainwright and Jordan, 2004. (Constraint satisfaction community). On an all-binary polynomial optimization problem whose constraint graph has tree-width $\leq \omega$, the Sherali-Adams reformulation for vertex packing, at level $\leq \omega$, is exact.

- Lasserre, Waki et al, others (2006-). Polynomial-size relaxations for continuous polynomial optimization problem if the underlying constraint graph has bounded tree-width.

**Question:** Can we use bounded tree-width to obtain good provably accurate, polynomial-size formulations for polynomial optimization?
Theorem: Given an instance of AC-OPF on a graph with a tree-decomposition of width \( \omega \), and \( n \) nodes, and \( 0 < \epsilon < 1 \),

there is a linear program \( \text{LP} \) such that:

(a) The number of variables and constraints is \( O(2^{2\omega} \omega n \epsilon^{-(\omega+1)} \log_2 \epsilon^{-1}) \).

(b) An optimal solution to \( \text{LP} \) solves AC-OPF, within tolerance \( \epsilon \).
**Theorem:** Given an instance of AC-OPF on a graph with a tree-decomposition of width $\omega$, and $n$ nodes, and $0 < \epsilon < 1$, there is a linear program $LP$ such that:

(a) The number of variables and constraints is $O(2^{2\omega} \omega n \epsilon^{-(\omega+1)} \log_2 \epsilon^{-1})$.

(b) An optimal solution to $LP$ solves AC-OPF, within tolerance $\epsilon$.

**Theorem:** Unless P=NP, above cannot be improved even for $\omega = 2$. 
More generic statement for AC-OPF

$$\min \sum_{k \in V} C_k$$

s.t. \( \forall km: \quad P_{km} = g_{km}(e_k^2 + f_k^2) - g_{km}(e_k e_m + f_k f_m) + b_{km}(e_k f_m - f_k e_m) \)

\( \forall km: \quad Q_{km} = -b_{km}(e_k^2 + f_k^2) + b_{km}(e_k e_m + f_k f_m) + g_{km}(e_k f_m - f_k e_m) \)

\( \forall km: \quad |P_{km}|^2 + |Q_{km}|^2 \leq U_{km} \)

\( \forall k: \quad P_k = \sum_{km \in \delta(k)} P_{km} \); \( P_k^{\min} \leq P_k \leq P_k^{\max} \)

\( \forall k: \quad Q_k = \sum_{km \in \delta(k)} Q_{km} \); \( Q_k^{\min} \leq Q_k \leq Q_k^{\max} \)

\( \forall k: \quad (V_k^{\min})^2 \leq e_k^2 + f_k^2 \leq (V_k^{\max})^2 \)

\( \forall k: \quad C_k = F_k(P_k, Q_k, e_k, f_k) + \sum_{km \in \delta(k)} H_{km}(P_{km}, Q_{km}, e_k, f_k, e_m, f_m) \)

Here, the \( F_k \) and \( H_{km} \) are quadratics.
A generalization: graphical QCQPs (abridged)

Inputs:

(1) An undirected graph $H$.

(2) For each vertex $v$ of $H$ a set $J(v)$, and for $j \in J(v)$ there is a real variable $x_j$. Write $V = \bigcup_{v \in V(H)} J(v)$.

(3) For each edge $\{v, u\}$ denote by $x_{v,u}$ the vector of all $x_j$ for $j \in J(v) \cup J(u)$.

(4) For each vertex $v$, and each edge $\{v, u\}$ a family of quadratics $p_{v,u}^k(x_{v,u})$ for $k = 1, \ldots, N(v)$.

(5) A vector $c \in \mathbb{R}^V$. 
A generalization: graphical QCQPs (abridged)

Inputs:

1. An undirected graph $H$.

2. For each vertex $v$ of $H$ a set $J(v)$, and for $j \in J(v)$ there is a real variable $x_j$. Write $V = \bigcup_{v \in V(H)} J(v)$.

3. For each edge $\{v, u\}$ denote by $x^{v,u}$ the vector of all $x_j$ for $j \in J(v) \cup J(u)$.

4. For each vertex $v$, and each edge $\{v, u\}$ a family of quadratics $p_{v,u,k}^k(x^{v,u})$ for $k = 1, \ldots, N(v)$.

5. A vector $c \in \mathbb{R}^V$.

Problem:

(GQCQP): $\min c^T x$

subject to: $\sum_{u \in \delta(v)} p_{v,u,k}^k(x^{v,u}) \geq 0, \quad v \in V(H), \quad k = 1, \ldots, N(v)$

$0 \leq x_j \leq 1, \quad \forall j \in V.$
A generalization: mixed-integer graphical QCQPs (abridged)

Inputs:

(1) An undirected graph $H$.

(2) For each vertex $v$ of $H$ a set $J(v)$, and for $j \in J(v)$ there is a real variable $x_j$. Write $V = \bigcup_{v \in V(H)} J(v)$.

(3) For each edge $\{v, u\}$ denote by $x^{v,u}$ the vector of all $x_j$ for $j \in J(v) \cup J(u)$.

(4) For each vertex $v$, and each edge $\{v, u\}$ a family of quadratics $p_{v,u}^k(x^{v,u})$ for $k = 1, \ldots, N(v)$.

(5) A vector $c \in \mathbb{R}^V$.

(6) A partition $V = V_Z \cup V_R$. 
Problem:

(MGP): $\min \ c^T x$

subject to: $\sum_{u \in \delta(v)} p_{v,u,k}(x_{v,u}) \geq 0, \ v \in V(H), \ k = 1, \ldots, N(v)$

$0 \leq x_j \leq 1 \ \forall \ j \in V_R; \ x_j = 0 \ \textbf{or} \ 1 \ \forall \ j \in V_Z.$
(1) An undirected graph $H$.

(2) For each vertex $v$ of $H$ a set $J(v)$, and for $j \in J(v)$ there is a real variable $x_j$. Write $\mathcal{V} = \bigcup_{v \in V(H)} J(v)$.

(3) For each edge $\{v, u\}$ denote by $x^{v,u}$ the vector of all $x_j$ for $j \in J(v) \cup J(u)$.

(4) For each vertex $v$, and each edge $\{v, u\}$ a family of polynomials $p_{v,u,k}^k(x^{v,u})$ for $k = 1, \ldots, N(v)$.

(5) A vector $c \in \mathbb{R}^\mathcal{V}$.

(6) A partition $\mathcal{V} = V_Z \cup V_R$.

\[ \text{(MGP):} \quad \min c^T x \] 
\[ \text{subject to:} \quad \sum_{u \in \delta(v)} p_{v,u,k}(x^{v,u}) \geq 0, \quad v \in V(H), \quad k = 1, \ldots, N(v) \] 
\[ 0 \leq x_j \leq 1 \quad \forall j \in V_R; \quad x_j = 0 \text{ or } 1 \quad \forall j \in V_Z. \]  

\textbf{Theorem:} Given an instance of MGP on a graph with a tree-decomposition of width $\omega$, there is an equivalent instance of MGP on a graph

- With tree-width $\leq 2\omega + 1$
- Of maximum degree $3$.

\textbf{Remark.} If we start with an instance of AC-OPF, the equivalent problem is no longer an AC-OPF problem.
Let $x$ be a variable, with bounds $0 \leq x \leq 1$. Let $0 < \gamma < 1$. Then we can approximate

$$x \approx \sum_{i=1}^{L} 2^{-i} y_i$$

where each $y_i$ is a binary variable. In fact, choosing $L = \lceil \log_2 \gamma^{-1} \rceil$, we have

$$x \leq \sum_{i=1}^{L} 2^{-i} y_i \leq x + \gamma.$$

So: given an instance of $MGP$, approximate each continuous variable $x_j$ in this manner.
**Theorem:** Consider an instance $\mathcal{I}$ of problem MGP, with $n$ variables. Then there is another instance, $\mathcal{B}$ of MGP, with

1. $\mathcal{B}$ is defined on the same graph as $\mathcal{I}$.
2. All variables in $\mathcal{B}$ are binary.
3. For each continuous variable $x_j$ of $\mathcal{I}$, we now have $\log_2 J^* \log \epsilon^{-1}$ binary variables used to approximate $x_j$.
4. Solving $\mathcal{B}$ to exact optimality yields a solution to $\mathcal{I}$ within tolerance $\epsilon$.

$J^* = \text{size of largest set } J(v)$. (AC-OPF $\Rightarrow$ $J^* = 2$)
Review

(1) A mixed-integer, graphical polynomial optimization problem on a graph with a tree-decomposition of width $\omega$. 
(1) A mixed-integer, graphical polynomial optimization problem on a graph with a tree-decomposition of width $\omega$.

(2) An equivalent mixed-integer, graphical polynomial optimization problem on a graph with a tree-decomposition of width $O(\omega)$ and degree $\leq 3$. 
(1) A mixed-integer, graphical polynomial optimization problem on a graph with a tree-decomposition of width $\omega$.

(2) An equivalent mixed-integer, graphical polynomial optimization problem on a graph with a tree-decomposition of width $O(\omega)$ and degree $\leq 3$.

(3) An all-binary, graphical polynomial optimization problem on the same graph which is equivalent to the problem in (2) within tolerance $\epsilon$. The sets $J(v)$ have grown by a factor of $\log_2 J^* \log_2 \epsilon^{-1}$. 
Fulkerson and Gross (1965), binary packing integer programs

\[
\text{IP} = \max \ c^T x \\
\text{s.t.} \ Ax \leq b, \\
x \in \{0, 1\}^n
\]  

Here, \( A \) is has 0,1-valued entries. **Idea:** use the structure of \( A \).

The **intersection graph of \( A \), \( G_A \), has:**

- A vertex for each column of \( A \).
- An edge between two columns \( j, k \) if there is a row \( i \) with \( a_{ij} \neq 0, a_{ik} \neq 0 \).

Each row of \( A \) induces a clique of \( G_A \).
(1) A mixed-integer, graphical polynomial optimization problem on a graph with a tree-decomposition of width $\omega$.

(2) An equivalent mixed-integer, graphical polynomial optimization problem on a graph with a tree-decomposition of width $O(\omega)$ and degree $\leq 3$.

(3) An all-binary, graphical polynomial optimization problem on the same graph which is equivalent to the problem in (2) within tolerance $\epsilon$. The sets $J(v)$ have grown by a factor of $\log_2 J^* \log_2 \epsilon^{-1}$.

(4) Corollary. The intersection graph of the problem in (3) has a tree-decomposition of width at most

$$O(\omega J^* \log_2 J^* \log_2 \epsilon^{-1})$$

Note: There are two graphs. The initial graph used to define the problem, and the intersection graph for the constraints in (3).
Theorem. Given an all-binary problem on $n$ variables and whose intersection graph has a tree-decomposition of width $k$, then there is an exact linear programming representation using

$O(2^k n)$

variables and constraints.

Construction similar to Lovász-Schrijver, Sherali-Adams, Lasserre, Bienstock-Zuckerberg
Theorem. Given an all-binary problem on $n$ variables and whose intersection graph has a tree-decomposition of width $k$, then there is an exact linear programming representation using $O(2^k n)$ variables and constraints.

Construction similar to Lovász-Schrijver, Sherali-Adams, Lasserre, Bienstock-Zuckerberg

$(A)$ A mixed-integer, graphical polynomial optimization problem, with $N$ variables, on a graph with a tree-decomposition of width $\omega$. $J^* = \text{size of largest set } J(v)$. (AC-OPF $J^* = 2$)
Theorem. Given an all-binary problem on $n$ variables and whose intersection graph has a tree-decomposition of width $k$, then there is an exact linear programming representation using $O(2^k n)$ variables and constraints.

Construction similar to Lovász-Schrijver, Sherali-Adams, Lasserre, Bienstock-Zuckerberg

(A) A mixed-integer, graphical polynomial optimization problem, with $N$ variables, on a graph with a tree-decomposition of width $\omega$. $J^* = \text{size of largest set $J(v)$}$. (AC-OPF $J^* = 2$)

(B) A linear program that solves the problem in (A) within tolerance $\epsilon$, of size $O(2^{O(\omega J^*) \omega J^* \epsilon^{-1} N})$
Should we able to do better?

Probably.

But.

• There are trivial AC-OPF problems where there is a unique feasible solution and it is irrational. Under the bit model of computing we cannot produce an “exact” answer.

• AC-OPF is weakly NP-hard on trees. Lavaei and Low (2011), a more recent proof by Coffrin and van Hentenryck.

• AC-OPF is strongly NP-hard on general graphs. A. Verma (2009). So no strong approximation algorithms exist unless $P = NP$. 