Cutting Planes for Mixed-Integer Programming: Theory and Practice

Oktay Günlük

Math Sciences, IBM Research

April 2018 – ORF523, Princeton
Mathematical optimization

- A generic mathematical optimization problem:

\[
\begin{align*}
\text{min} : & \quad f(x) \\
\text{subject to:} & \quad g_i(x) \leq 0 \quad i = 1, \ldots, m \\
& \quad x \in X
\end{align*}
\]

- Computationally tractable cases:

  - If \( f(x) \) and all \( g_i(x) \) are linear, and \( X = \mathbb{R}^n_+ \) \( \Rightarrow \) LP
  
  - If \( f(x) \) and all \( g_i(x) \) are linear, and \( X = \mathbb{Z}^{n_1}_+ \times \mathbb{R}^{n_2}_+ \) \( \Rightarrow \) MILP
  
  - If \( f(x) \) is quadratic and all \( g_i(x) \) are linear, and \( X = \mathbb{R}^n_+ \) \( \Rightarrow \) QP
  
  - If \( f(x) \) and all \( g_i(x) \) are quadratic, and \( X = \mathbb{R}^n_+ \) \( \Rightarrow \) QCQP

- Only LP can be solved in polynomial time. Even Box QP is hard!

\[
\begin{align*}
\text{min} : & \quad x^T Q x \\
\text{subject to:} & \quad 1 \geq x \geq 0
\end{align*}
\]
A generic Mixed Integer Linear Program has the form:

$$\min \{ c^T x : Ax \geq b, x \geq 0, \ x_j \text{ integer}, \ j \in I \}$$

where matrix $A$ does not necessarily have a special structure.

A very large number of practical problems can be modeled in this form:

- Production planning,
- Airline scheduling (routing, staffing, etc.),
- Telecommunication network design,
- Classroom scheduling,
- Combinatorial auctions,
- ...

In theory, MIP is NP-hard: not much hope for efficient algorithms.

But in practice, even very large MIPs can be solved to optimality in reasonable time.
A generic Mixed Integer Linear Program has the form:

\[ \min \{ c^T x : Ax \geq b, x \geq 0, \ x_j \text{ integer}, \ j \in \mathcal{I} \} \]

where matrix \( A \) does not necessarily have a special structure.
Overview of the talk

• Introduction
  – Mixed-integer programming, branch-and-cut

• Commercial Software (Cplex)
  – Evolution, main components

• Cutting planes
  – Mixed-integer rounding

• A new approach to cutting planes
  – Lattice free cuts, multi-branch split cuts

• A finite cutting-plane algorithm
Solving Mixed Integer Linear Programs

- In practice MIPs are solved via enumeration:
  
  - The branch-and-bound algorithm, Land and Doig (1960)
  - The branch-and-cut scheme proposed by Padberg and Rinaldi (1987)

- Given an optimization problem $z^* = \min \{ f(x) : x \in P \}$,

  (i) Partitioning: Let $P = \bigcup_{i=1}^{p} P_i$ (division), then

  $$z^* = \min_i \{ z_i \} \text{ where } z_i = \min \{ f(x) : x \in P_i \} ,$$

  (ii) Lower bounding: For $i = 1, \ldots, p$, let $P_i \subseteq P_i^R$ (relaxation), then

  $$z_i \geq z_i^R = \min \{ f(x) : x \in P_i^R \}, \text{ and } z^* \geq \min_i \{ z_i^R \} .$$

  (iii) Upper bounding: If $\bar{x} \in P_i \subseteq P$ then $f(\bar{x}) \geq z^*$.

  [Same framework is used to solve non-convex QP’s, for example.]
Relaxation step

**Mixed Integer Program:**

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad Ax \geq b \\
& \quad x \geq 0 \\
& \quad x_j \in \mathbb{Z} \text{ for } j \in \mathcal{I}
\end{align*}
\]

**LP Relaxation:**

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad Ax \geq b \\
& \quad x \geq 0
\end{align*}
\]
Partitioning step

Initial problem

Subproblem 1: \( x_1 \leq \lfloor x_1^* \rfloor \)

Subproblem 2: \( x_1 \geq \lceil x_1^* \rceil \)
Next: Commercial Solvers
MIP Evolution, early days

- Early MIP solvers focused on developing fast and reliable LP solvers for branch-and-bound schemes. (eg. $10^6$-fold improvement in Cplex from 1990 to 2004!).

- Remarkable exceptions are:
  - 1983 Crowder, Johnson & Padberg: PIPX, pure 0/1 MIPs
  - 1987 Van Roy & Wolsey: MPSARX, mixed 0/1 MIPs

- When did the early days end?
  
  A crucial step has been the computational success of cutting planes for TSP
  - Padberg and Rinaldi (1987)
  - Applegate, Bixby, Chvátal, and Cook (1994)

- In addition for general MIPs:
  - 1994 Balas, Ceria & Cornuéjols: Lift-and-project
  - 1996 Balas, Ceria, Cornuéjols & Natraj: Gomory cuts revisited
Evolution of MIP Solvers by numbers

- Bixby & Achterberg compared all Cplex versions (with MIP capability)
- 1,734 MIP instances
- Computing times are geometric means normalized wrt Cplex 11.0

<table>
<thead>
<tr>
<th>Cplex versions</th>
<th>year</th>
<th>better</th>
<th>worse</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.0</td>
<td>2007</td>
<td>0</td>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td>10.0</td>
<td>2005</td>
<td>201</td>
<td>650</td>
<td>1.91</td>
</tr>
<tr>
<td>9.0</td>
<td>2003</td>
<td>142</td>
<td>793</td>
<td>2.73</td>
</tr>
<tr>
<td>8.0</td>
<td>2002</td>
<td>117</td>
<td>856</td>
<td>3.56</td>
</tr>
<tr>
<td>7.1</td>
<td>2001</td>
<td>63</td>
<td>930</td>
<td>4.59</td>
</tr>
<tr>
<td>6.5</td>
<td>1999</td>
<td>71</td>
<td>997</td>
<td>7.47</td>
</tr>
<tr>
<td>6.0</td>
<td>1998</td>
<td>55</td>
<td>1060</td>
<td>21.30</td>
</tr>
<tr>
<td>5.0</td>
<td>1997</td>
<td>45</td>
<td>1069</td>
<td>22.57</td>
</tr>
<tr>
<td>4.0</td>
<td>1995</td>
<td>37</td>
<td>1089</td>
<td>26.29</td>
</tr>
<tr>
<td>3.0</td>
<td>1994</td>
<td>34</td>
<td>1107</td>
<td>34.63</td>
</tr>
<tr>
<td>2.1</td>
<td>1993</td>
<td>13</td>
<td>1137</td>
<td>56.16</td>
</tr>
<tr>
<td>1.2</td>
<td>1991</td>
<td>17</td>
<td>1132</td>
<td>67.90</td>
</tr>
</tbody>
</table>

- The key feature of Cplex v. 6.5 was extensive cutting plane generation.
• This problem has 2756 binary variables and 755 constraints

• Hardest instance in Crowder, Johnson and Padberg (1983)

• Solving with Cplex 11:
  – without cuts it takes 3,414,408 nodes
  – with cuts it takes 11 nodes!

• Cplex reduces the root optimality gap from 13.5% to 0.2% with
  – 22 Gomory mixed-integer cuts, and
  – 23 cover inequalities
    (both are "mixed-integer rounding" inequalities.)

• This and many other MIPLIB instances are available at http://miplib.zib.de
Strengthening the LP relaxation by cutting planes

- Given the optimal solution \( \bar{x} \) of the LP relaxation (not integral)
- Do not branch right away
- Find a valid inequality for the MIP \( a^T x \geq b \) such that \( a^T \bar{x} < b \).
Branch and cut
Main components of a MIP solver

- **Preprocessing**
  - *Clean up the model* (empty/implied rows, fixed variables, . . .)
  - *Coefficient reduction* (ex: p0033, all variables binary)

\[
-230x_{10} - 200x_{16} - 400x_{17} \leq -5 \implies x_{10} + x_{16} + x_{17} \geq 1
\]

- **Cutting plane generation:**
  - *Gomory Mixed Integer cuts, MIR inequalities, cover cuts, flow covers, . . .*
- **Branching strategies:**
  - *strong branching, pseudo-cost branching, (not most fractional!)*
- **Primal heuristics:**
  - *rounding heuristics, diving heuristics, local search, . . .*
- **Node selection strategies:**
  - *a combination of best-bound and diving.*
Some features of a good MIP solver

- **Solving a MIP to optimality is only one aspect for many applications (sometimes not the most important one)**
  - Detect infeasibility in the model early on and report its source to help with modeling.
  - Feasible (integral) solutions
    * Find good solutions quickly
    * Find many solutions and store them

- **Not all MIPs are the same**
  - Recognize problem structure and adjust parameters/strategies accordingly (there are too many parameters/options for hand-tuning.)
  - Deal with both small and very large scale problems
  - Handle numerically difficult instances with care (**very important**)

- **Not all users are the same**
  - Allow user to take over some of the control (callbacks)
Beyond MIP

- Not all non-convex optimization problems are MIPs :)
- But it is possible to extend the capability of the MIP framework. For example:
  1. **Bonmin** (Basic Open-source Nonlinear Mixed INteger programming, [Bonami et. al.])
     - For Convex MINLP within the framework of the MIP solver Cbc [Forrest].
  2. **GloMIQO** (Global mixed-integer quadratic optimizer, [Misener])
     - Spatial branch-and-bound algorithm for non-convex QP.
  3. **Couenne** (Convex Over and Under ENvelopes for Nonlinear Estimation, [Belotti])
     - Spatial and integer branch-and-bound algorithm for non-convex MINLP.
  4. **SCIP** (Solving Constraint Integer Programs, [Achterberg et. al.])
     - Tight integration of CP and SAT techniques within a MIP solver.
     - Significant recent progress for non-convex MINLP.
- All codes are open source and can be obtained free of charge.
Next: Cutting planes
Cutting planes for IP

Integer Program:

\[
\begin{align*}
\text{min} & \quad c^T x \\
Ax & \geq b, \\
x & \geq 0,
\end{align*}
\]

\(x\) integral

LP Relaxation:

\[
\begin{align*}
\text{min} & \quad c^T x \\
Ax & \geq b \\
x & \geq 0
\end{align*}
\]

Tighten:

\[
\begin{align*}
\text{min} & \quad c^T x \\
Ax & \geq b \\
x & \geq 0 \\
\alpha_1 x & \geq d_1 \\
\alpha_2 x & \geq d_2
\end{align*}
\]
Convex hull of mixed-integer sets

- Any MIP can be solved by linear programming (without branching) by finding the "right" cuts (i.e. by finding the convex hull).

- **LP Relaxation:**
  - **Strong LP Relaxation:**
  - **Convex hull of solutions:**

- Gomory proposed a finite cutting plane algorithm for pure IPs (1958).
- Dash, Dobbs, Gunluk, Nowicki, and Swirszcz, did the same for MIPs (2014).

- In practice,
  - These algorithms are hopeless except some very easy cases.
  - But, getting closer to the convex hull helps.
Gomory cuts

Let

\[ Q^0 = \{ y \in \mathbb{Z} : y \geq b_1, y \leq b_2 \} \]

then, the following inequalities:

\[ y \geq \lfloor b_1 \rfloor \quad \text{and} \quad y \leq \lceil b_2 \rceil \]

are valid for \( Q^0 \) and

\[ \text{conv}(Q^0) = \left\{ y \in \mathbb{R} : \lfloor b_2 \rfloor \geq y \geq \lceil b_1 \rceil \right\}. \]

- \( y \) can be replace with any integer expression to obtain a valid cut.
- These cuts are also called Chvatal-Gomory cuts
Basic mixed-integer rounding set (Wolsey ’98)

Let

\[ Q^1 = \left\{ v \in R, \ y \in Z : v + y \geq b, \ v \geq 0 \right\} \]

then, MIR Inequality:

\[ v \geq \hat{b}(\lceil b \rceil - y) \]

where \( \hat{b} = b - \lfloor b \rfloor \), is valid for \( Q^1 \) and

\[ \text{conv}(Q^1) = \left\{ v, y \in R : v + y \geq b, \ v + \hat{b}y \geq \hat{b} \lceil b \rceil, \ v \geq 0 \right\}. \]
Basic mixed-integer rounding set – example

Let

\[ Q^1 = \left\{ v \in R, \ y \in Z : v + y \geq 7.3, \ v \geq 0 \right\} \]

then, MIR Inequality:

\[ v \geq 0.3(8 - y) \]

where \( 0.3 = 7.3 - 7 \), is valid for \( Q^1 \) and

\[ \text{conv}(Q^1) = \left\{ v, y \in R : v + y \geq 7.3, \ v + 0.3y \geq 0.3 \times 8, \ v \geq 0 \right\}. \]
Next: MIR Inequalities
MIR inequalities for single constraint sets

Let

\[ P^1 = \left\{ v \in R^{|C|}, \ y \in Z^{|I|} : \sum_{j \in C} c_j v_j + \sum_{j \in I} a_j y_j \geq b, \ v, \ y \geq 0 \right\} \]

Re-write:

\[
\sum_{c_j < 0} c_j v_j + \sum_{c_j > 0} c_j v_j + \sum_{\hat{a}_j < \hat{b}} \hat{a}_j y_j + \sum_{\hat{a}_j \geq \hat{b}} \hat{a}_j y_j + \sum_{j \in I} \lfloor a_j \rfloor y_j \geq b = \hat{b} + \lfloor b \rfloor
\]

Relax:

\[
\underbrace{\sum_{c_j > 0} c_j v_j + \sum_{\hat{a}_j < \hat{b}} \hat{a}_j y_j}_{\geq 0} + \underbrace{\sum_{\hat{a}_j \geq \hat{b}} y_j + \sum_{j \in I} \lfloor a_j \rfloor y_j}_{\in Z} \geq b
\]

MIR cut:

\[
\sum_{c_j > 0} c_j v_j + \sum_{\hat{a}_j < \hat{b}} \hat{a}_j y_j + \hat{b} \left( \sum_{\hat{a}_j \geq \hat{b}} y_j + \sum_{j \in I} \lfloor a_j \rfloor y_j \right) \geq \hat{b} \lfloor b \rfloor
\]

(Applying MIR to the simplex tableau rows gives the Gomory mixed-integer cut.)
MIR inequalities for multiple constraint sets

Let
\[ P = \left\{ v \in \mathbb{R}^{|C|}, \ y \in \mathbb{Z}^{|I|} : Cv + Ay \geq d, \ v, y \geq 0 \right\} \]

where \( C \in \mathbb{R}^{m \times |C|}, \ A \in \mathbb{R}^{m \times |I|}, \ d \in \mathbb{R}^m \).

- Obtain a “base” inequality using \( \lambda \in \mathbb{R}_+^m : \lambda Cv + \lambda Ay \geq \lambda d \)

- Write the corresponding MIR inequality:
\[
\sum_{j \in C} (\lambda C_j)^+ v_j + \hat{b} \sum_{j \in I} \lfloor \lambda A_j \rfloor y + \sum_{j \in I} \min\{\lambda A_j - \lfloor \lambda A_j \rfloor, \hat{b}\} y \geq \hat{b} \lceil \lambda d \rceil
\]

where \( \hat{b} = \lambda d - \lfloor \lambda d \rfloor \).
Better MIR inequalities for multiple constraint sets

• Add (non-negative) slack variables to the defining inequalities:

\[ Cv + Ay - Is = d \]

• Obtain a “base” equation using \( \lambda \in \mathbb{R}^m \):

\[ \lambda Cv + \lambda Ay - \lambda Is = \lambda d \]

• Write the corresponding MIR inequality:

\[
\sum_{j \in C} (\lambda C_j)^+ v_j + \hat{b} \sum_{j \in I} \lfloor \lambda A_j \rfloor y + \sum_{j \in I} \min \{ \lambda A_j - \lfloor \lambda A_j \rfloor, \hat{b} \} y + \sum_{\lambda_i < 0} |\lambda_i| s_i \geq \hat{b} \lceil \lambda d \rceil
\]

• Substitute out slacks to obtain

\[
\sum_{j \in C} (\lambda C_j)^+ v_j + \hat{b} \sum_{j \in I} \lfloor \lambda A_j \rfloor y + \sum_{j \in I} \min \{ \ldots, \hat{b} \} y + \sum_{\lambda_i < 0} |\lambda_i| (Cv + Ax - d)_i \geq \hat{b} \lceil \lambda d \rceil
\]
Consider the set

\[ T = \{ v \in \mathbb{R}, x \in \mathbb{Z} : -v - 4x \geq -4, -v + 4x \geq 0, v, x \geq 0 \} \]

Any base inequality generated by \( \lambda_1, \lambda_2 \) has the form

\[ (-\lambda_1 - \lambda_2)v + (-4\lambda_1 + 4\lambda_2)x \geq -4\lambda_1 \]

If \( \lambda_1, \lambda_2 \geq 0 \), \( v \) has a negative coefficient and does not appear in the cut.

Using multipliers \( \lambda = [-1/8, 1/8] \)

\[ -\frac{1}{8}(-v - 4x - s_1 = -4) + \frac{1}{8}(-v + 4x - s_2 = 0) \downarrow (Base \ inequality) \]

\[ x + s_1/8 - s_2/8 \geq 1/2 \downarrow (MIR) \]

\[ 1/2x + s_1/8 \geq 1/2 \Rightarrow -v/8 \geq 0 \Rightarrow v \leq 0 \]

This inequality defines the only non-trivial facet of \( T \).
## Computational performance of MIR inequalities

| instance       | $|I|$  | $|J|$  | # iter | # cuts | % gap | time | % gap | time | % gap | time |
|----------------|------|------|-------|-------|-------|------|------|------|------|------|------|
| 10teams        | 1,800| 225  | 338   | 3341  | 100.00| 3,600| 57.14| 1,200| 100.00| 90   |
| arki001        | 538  | 850  | 14    | 124   | 33.93 | 3,600| 28.04| 1,200| 83.05*| 193,536|
| bell3a         | 71   | 62   | 21    | 166   | 98.69 | 3,600| 48.10| 65   | 65.35 | 102 |
| bell5          | 58   | 46   | 105   | 608   | 93.13 | 3,600| 91.73| 4    | 91.03 | 2,233|
| blend2         | 264  | 89   | 723   | 3991  | 32.18 | 3,600| 36.40| 1,200| 46.52 | 552 |
| dano3mip       | 552  | 13,321| 1     | 124   | 0.10  | 3,600| 0.00  | 1,200| 0.22  | 73,835|
| danoint        | 56   | 465  | 501   | 2480  | 1.74  | 3,600| 0.01  | 1,200| 8.20  | 147,427|
| dcmulti        | 75   | 473  | 480   | 4527  | 98.53 | 3,600| 47.25| 1,200| 100.00| 2,154|
| egout          | 55   | 86   | 37    | 324   | 100.00| 31   | 81.77| 7    | 100.00| 18,179|
| fiber          | 1,254| 44   | 98    | 408   | 96.00 | 3,600| 4.83  | 1,200| 99.68 | 163,802|
| fixnet6        | 378  | 500  | 761   | 4927  | 94.47 | 3,600| 67.51| 43   | 99.75 | 19,577|
| flugpl         | 11   | 7    | 11    | 26    | 93.68 | 3,600| 19.19| 1,200| 100.00| 26   |
| gen            | 150  | 720  | 11    | 127   | 100.00| 16   | 86.60| 1,200| 100.00| 46   |
| gesa2          | 408  | 816  | 433   | 1594  | 99.81 | 3,600| 94.84| 1,200| 99.02 | 22,808|
| gesa2_o        | 720  | 504  | 131   | 916   | 97.74 | 3,600| 94.93| 1,200| 99.97 | 8,861|
| gesa3          | 384  | 768  | 464   | 1680  | 81.84 | 3,600| 58.96| 1,200| 95.81 | 30,591|
| gesa3_o        | 672  | 480  | 344   | 1278  | 69.74 | 3,600| 64.53| 1,200| 95.20 | 6,530|
| khb05250       | 24   | 1,326| 65    | 521   | 100.00| 113  | 4.70 | 3    | 100.00| 33   |

Table 1: MIPs of the MIPLIB 3.0.
Numerical and other practical issues

When implementing these ideas to solve mixed integer programs one has to be careful:

• How to obtain the base inequality?
  – Formulation rows
  – Simplex tableau rows
  – Aggregate formulation rows using different heuristics

• Numerical issues
  – LP-solvers are not numerically exact.

\[ b = 5.00001 \implies \lceil b \rceil = 6 \text{ and } \hat{b} = 0.00001 \]
\[ b = 4.99999 \implies \lceil b \rceil = 5 \text{ and } \hat{b} = 0.99999 \]

– Avoid large numbers: \( 1000000x_1 - 10000000x_2 \geq 0.3 \) is not a good cut.
– Avoid dense rows
Next:

_Beyond MIR Inequalities: Lattice free cuts, multi-branch split cuts_  

(joint work with Dash, Dobbs, Nowicki, and Świrszcz)
The region cut-off by the valid inequality is always strictly lattice-free.
Generating Cutting Planes Using Lattice Free Sets

- Relaxation minus a strictly lattice-free (convex) set gives a tighter relaxation.

Ex:

- We can also use non-convex lattice-free sets:

(but then need to convexify afterwards to obtain a nice relaxation)
Disjunctive cuts [Balas ’79]

• Let $D = \bigcup_{i \in K} D_i$ where

$$D_i = \{(x, y) \in \mathcal{R}^{n+l} : A^i x \leq b^i\}$$

• $D \subseteq \mathcal{R}^{n+l}$ is called a disjunction if $\mathcal{Z}^n \times \mathcal{R}^l \subseteq D$ (clearly $D = D^n \times \mathcal{R}^l$)

• Let $P = P^{LP} \cap (\mathcal{Z}^n \times \mathcal{R}^l)$ where

$$P^{LP} = \{(x, v) \in \mathcal{R}^n \times \mathcal{R}^l : Ax + Cv \geq d\}$$

• The disjunctive hull of $P$ with respect to $D$ is

$$P_D = \text{conv}\left(P^{LP} \cap D\right) = \text{conv}\left(\bigcup_{k \in K} (P^{LP} \cap D_k)\right)$$

• Notice that $P_D = \text{conv}\left(P^{LP} \setminus B\right)$ where $B = \mathcal{R}^{n+l} \setminus D$ is strictly lattice-free.
All valid inequalities are disjunctive cuts

Let $c^T x + d^T y \geq f$ be a valid inequality for $P$ and

$$V = \{(x, y) \in P^{LP} : c^T x + d^T y < f\}.$$

Clearly $V \cap (\mathbb{Z}^n \times \mathbb{R}^l) = \emptyset$, i.e. strictly lattice-free.

Jörg (2007) observes that $V_x \subseteq \text{int}(B_x)$ where

- $V_x \subseteq \mathbb{R}^n$ is the orthogonal projection of $V$ in the space of the integer variables
- $B_x \subseteq \mathbb{R}^n$ is a polyhedral lattice-free set defined by rational (integral) data

$$B_x = \{x \in \mathbb{R}^n : \pi^T_i x \geq \gamma_i, \; i \in K\}$$

Therefore the cut is valid for

$$\text{conv}(P^{LP} \setminus (\text{int}(B_x) \times \mathbb{R}^l)) \subseteq \text{conv}(P^{LP} \setminus (V^x \times \mathbb{R}^l)) .$$

Based on this observation, Jörg then argues that $|K| \leq 2^n$ and

$$D = \bigcup_{i \in K} \{(x, y) \in \mathbb{R}^{n+l} : \pi^T_i x \leq \gamma_i\}$$

is a valid disjunction and $c^T x + d^T y \geq f$ can be derived from this disjunction.
Split cuts

- Let \( \pi \in \mathbb{Z}^n \) and \( \gamma \in \mathbb{Z} \) and consider the split set

\[
S(\pi, \gamma) = \{(x, y) \in \mathbb{R}^{n+1} : \gamma < \pi^T x < \gamma + 1\}
\]

(which is strictly lattice-free)

- A split cut is an inequality valid for \( P^{LP} \setminus S(\pi, \gamma) \):

- Split cuts are disjunctive cuts \( D_1 = \{\pi^T x \leq \gamma\} \) and \( D_2 = \{\pi^T x \geq \gamma + 1\} \)

- MIR cuts are split cuts with \( \pi = \lceil \lambda A \rceil \) and \( \gamma = \lfloor \lambda d \rfloor \).
A Generalization of Split Cuts: Cross cuts

$P^{LP}$  

the cross set

$P^{LP}\setminus$ cross set
## Computational experiments with cross cuts

<table>
<thead>
<tr>
<th>problem</th>
<th>GMI</th>
<th>DG</th>
<th>Split</th>
<th>Cross</th>
<th>Best Split</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>bell5</td>
<td>14.5</td>
<td>25.9</td>
<td>86.2</td>
<td>99.8</td>
<td>93.0</td>
<td>97.1</td>
</tr>
<tr>
<td>cap6000</td>
<td>41.7</td>
<td>64.6</td>
<td>65.2</td>
<td>67.2</td>
<td>65.2</td>
<td>5.7</td>
</tr>
<tr>
<td>gesa3</td>
<td>45.9</td>
<td>93.3</td>
<td>95.1</td>
<td>97.4</td>
<td>95.8</td>
<td>38.1</td>
</tr>
<tr>
<td>gesa3o</td>
<td>50.6</td>
<td>93.1</td>
<td>95.3</td>
<td>99.0</td>
<td>95.2</td>
<td>79.2</td>
</tr>
<tr>
<td>gt2</td>
<td>67.7</td>
<td>96.6</td>
<td>96.7</td>
<td>99.2</td>
<td>98.4</td>
<td>50.0</td>
</tr>
<tr>
<td>mas74</td>
<td>6.7</td>
<td>11.4</td>
<td>14.3</td>
<td>15.7</td>
<td>14.0</td>
<td>2.0</td>
</tr>
<tr>
<td>mas76</td>
<td>6.4</td>
<td>16.1</td>
<td>25.1</td>
<td>34.5</td>
<td>26.5</td>
<td>10.9</td>
</tr>
<tr>
<td>mkc</td>
<td>1.2</td>
<td>4.1</td>
<td>52.4</td>
<td>55.3</td>
<td>49.3</td>
<td>11.8</td>
</tr>
<tr>
<td>modglob</td>
<td>15.1</td>
<td>90.8</td>
<td>94.0</td>
<td>99.0</td>
<td>92.2</td>
<td>87.2</td>
</tr>
<tr>
<td>p0033</td>
<td>54.6</td>
<td>84.1</td>
<td>86.2</td>
<td>100.0</td>
<td>87.4</td>
<td>100.0</td>
</tr>
<tr>
<td>p0201</td>
<td>18.2</td>
<td>71.5</td>
<td>74.0</td>
<td>98.4</td>
<td>74.9</td>
<td>93.6</td>
</tr>
<tr>
<td>pp08a</td>
<td>52.9</td>
<td>96.0</td>
<td>96.6</td>
<td>98.4</td>
<td>97.0</td>
<td>46.7</td>
</tr>
<tr>
<td>pp08aCUTS</td>
<td>30.1</td>
<td>93.3</td>
<td>94.7</td>
<td>96.5</td>
<td>95.8</td>
<td>16.7</td>
</tr>
<tr>
<td>qiu</td>
<td>2.0</td>
<td>21.8</td>
<td>78.1</td>
<td>78.4</td>
<td>77.5</td>
<td>4.0</td>
</tr>
<tr>
<td>set1ch</td>
<td>38.1</td>
<td>88.0</td>
<td>88.7</td>
<td>98.6</td>
<td>89.7</td>
<td>86.4</td>
</tr>
<tr>
<td>vpm2</td>
<td>12.6</td>
<td>72.0</td>
<td>76.5</td>
<td>81.7</td>
<td>81.0</td>
<td>3.7</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>25.5</td>
<td>61.3</td>
<td>76.1</td>
<td>82.2</td>
<td>77.3</td>
<td>45.8</td>
</tr>
</tbody>
</table>

Table 2: Some MIPLIB Problems – 16 out of 32

(joint work with Dash and Vielma)


## Multi-branch split cuts

- Let

\[ P = \{(x, v) \in \mathbb{Z}^n \times \mathbb{R}^l : Ax + Cv = d, \ v \geq 0\} \]

be rational and let \( P^{LP} \) denote its continuous relaxation.

- Let \( \pi_i \in \mathbb{Z}^n \) and \( \gamma_i \in \mathbb{Z} \) for \( i = 1, \ldots, t \) and consider the split sets

\[ S(\pi_i, \gamma_i) = \{(x, y) \in \mathbb{R}^{n+k} : \gamma_i < \pi_i^T x < \gamma_i + 1\} \]

- A **multi-branch split cut** is an inequality valid for

\[ P^{LP} \setminus \bigcup_i S(\pi_i, \gamma_i) \]

(Li/Richard ('08) call these cuts **t-branch split cuts**)

- 2-branch split cuts are **cross** cuts.

- Multi-branch split cuts are disjunctive cuts [Balas '79].
Are all valid inequalities multi-branch split cuts?

Let \( \pi_i \) and \( \gamma_i \) be integral for \( i = 1, \ldots, t \) and consider the split sets

\[
S(\pi_i, \gamma_i) = \{(x, y) \in \mathbb{R}^{n+k} : \gamma_i < \pi_i^T x < \gamma_i + 1\}
\]

A **multi-branch split cut** is an inequality valid for \( P^{LP} \setminus \bigcup_i S(\pi_i, \gamma_i) \)

The corresponding disjunction is

\[
D = \bigcup_{S \subseteq \{1, \ldots, t\}} \{(x, y) \in \mathbb{R}^{n+k} : \pi_i^T x \leq \gamma_i \text{ if } i \in S, \pi_i^T x \geq \gamma_i + 1 \text{ if } i \not\in S\}
\]

**Question**: Are all facet defining inequalities \( t \)-branch split cuts for finite \( t \)?

Remember the points cut off by the valid inequality \( c^T x + d^T y \geq f \)

\[
V = \{(x, y) \in P^{LP} : c^T x + d^T y < f\}.
\]

**Fact**: Let \( S = \bigcup S_i \) be a collection of split sets in \( \mathbb{R}^{n+k} \). If \( V \subseteq S \), then \( c^T x + d^T y \geq f \) is a multi-branch split cut obtained from \( S \).
Lattice width

- Given a closed, bounded, convex set (or convex body) \( B \subseteq \mathbb{R}^n \) and a vector \( c \in \mathbb{Z}^n \),
  \[
  w(B, c) = \max\{ c^T x : x \in B \} - \min\{ c^T x : x \in B \},
  \]
  is the lattice width of \( B \) along the direction \( c \).

- The lattice width of \( B \) is
  \[
  w(B) = \min_{c \in \mathbb{Z}^n \setminus \{0\}} w(B, c)
  \]
  (If the set is not closed, we define its lattice width to be the lattice width of its closure)

- Khinchine’s flatness theorem: there exists a function \( f(\cdot) : \mathbb{Z}_+ \to \mathbb{R}_+ \) such that for any strictly lattice-free bounded convex set \( B \subseteq \mathbb{R}^n \),
  \[
  w(B) \leq f(n)
  \]
  where \( f(\cdot) \) depends on the dimension of \( B \) (not on the complexity of \( B \))

- Lenstra uses this result to construct a finite enumeration tree to solve the integer feasibility problem.
Bounding the lattice width

- Given a lattice free convex body $B \subseteq \mathbb{R}^n$ the lattice width is

$$w(B) = \min_{c \in \mathbb{Z}^n \setminus \{0\}} w(B, c) \leq f(n)$$

- Lenstra (1983) showed that $f(n) \leq 2^{n^2}$

- Kannan and Lovász (1988) showed that $f(n) \leq c_0(n + 1)n/2$ for some constant $c_0$ ($c_0 = \max\{1, 4/c_1\}$ where $c_1$ is another constant defined by Bourgain and Milman)

- Banaszczyk, Litvak, Pajor, and Szarek (1999) showed that $O(n^{3/2})$

- Rudelson (2000) showed that $O(n^{4/3} \log^c n)$ for some constant $c$. 
Lattice-free sets in $\mathbb{R}^2$

**Theorem**: [Hurkens (1990)] If $B \in \mathbb{R}^2$, then $w(B) \leq 1 + \frac{2}{\sqrt{3}} \approx 2.1547$. Furthermore, $w(B) = 1 + \frac{2}{\sqrt{3}}$ if and only if $B$ is a triangle with vertices $q_1, q_2, q_3$ such that:

$$\frac{1}{\sqrt{3}} q_i + (1 - \frac{1}{\sqrt{3}}) q_{i+1} = b_i, \text{ for } i = 1, 2, 3.$$

where $b_i \in \mathbb{Z}^2$ for $i = 1, 2, 3$. (and $q_4 := q_1$)

The lattice-free triangle $T$ when $b_1 = (0, 0)^T$, $b_2 = (0, 1)^T$, and $b_3 = (1, 0)^T$

(\textit{this is called a type 3 triangle})
Averkov, Wagner and Weismantel (2011) enumerated all maximal lattice-free bodies in $\mathbb{R}^3$ that are integral. These sets have the lattice width $\leq 3$.

There exists a tetrahedron $H$ with lattice width $2 + \frac{2}{\sqrt{3}} \approx 3.1547$:

where $s_4 = (0, 0, 2 + 2/\sqrt{3})$, and $q_1, \ldots, q_3 \in \mathbb{R}^2$ are the vertices of Hurken's triangle.

We can also show that $f(3) \leq 4.25$. 
Next: A finite cutting-plane algorithm for mixed-integer programming
Can MIP’s be solved only using cutting planes (without branching)?

History of finite cutting plane algorithms:

- Gomory (1958) developed the first finite cutting plane algorithm for pure IPs.
- Later, (1960) he extended this to MIPs with integer objective.
- Cook/Kannan/Schrijver (1990) gave an example in $\mathbb{Z}^2 \times \mathbb{R}$ which cannot be solved in finite time using split cuts.
- Later Dash and Gunluk (2013) generalized this to examples in $\mathbb{Z}^n \times \mathbb{R}$ that cannot be solved in finite time using $(n - 1)$-branch split cuts.
- For bounded polyhedra Jörg (2008) gave a finite cutting plane algorithm for MIPs.
- Using multi-branch split cuts, we recently gave a finite cutting plane algorithm for MIPs without assuming boundedness or integer objective. (“This algorithm is of purely theoretical interest, and is highly impractical.”)
Recap

- A multi-branch split cut is an inequality valid for $P^{LP} \setminus \bigcup_i S(\pi_i, \gamma_i)$ where

$$S(\pi_i, \gamma_i) = \{(x, y) \in \mathbb{R}^{n+k} : \gamma_i < \pi_i^T x < \gamma_i + 1\}$$

and $\pi_i$ and $\gamma_i$ are integral.

- Let $c^T x + d^T y \geq f$ be a valid inequality for $P$ and

$$V = \{(x, y) \in P^{LP} : c^T x + d^T y < f\}.$$

be the set points cut off by it. ($V \cap (\mathbb{Z}^n \times \mathbb{R}^l) = \emptyset$)

If $V \subseteq S$ where $S = \bigcup_{i=1}^t S_i$, then $c^T x + d^T y \geq f$ is a $t$-branch split cut.

Claim: All such $V$ can be covered by a bounded number of split sets.
Lemma: Let $B$ be a bounded, strictly lattice-free convex set in $\mathbb{R}^n$. Then $B$ is contained in the union of at most $h(n)$ split sets.

Proof: By Khinchine's flatness result.

- There is an integer vector $a \in \mathbb{Z}^n$ such that $f(n) \geq u - l$ where
  \[ u = \max\{a^T x : x \in B\} \quad \text{and} \quad l = \min\{a^T x : x \in B\} \]

- Therefore, $B \subseteq \{x \in \mathbb{R}^n : \lfloor l \rfloor \leq a^T x \leq \lceil u \rceil\}$.

- Let $U$ be the collection of the split sets $S(a,b)$ for $b \in W = \{\lfloor l \rfloor, \ldots, \lceil u \rceil - 1\}$
  \[ B \setminus \bigcup_{b \in W} S(a,b) = \bigcup_{b \in W} \{x \in B : a^T x = b\} \]
  where $\bar{W} = \{\lfloor l \rfloor, \ldots, \lceil u \rceil\}$.

- All $\{x \in B : a^T x = b\}$ are strictly lattice-free and have dimension at most $n - 1$

- Repeating the same argument proves the claim. ($h(n) \approx \Pi_{i=1}^n (2 + \lceil f(i) \rceil)$)
**Unbounded case**

**Lemma:** Let $B$ be a strictly lattice-free, convex, unbounded set in $\mathbb{R}^n$ which is contained in the interior of a maximal lattice-free convex set. Then $B$ can be covered by $h(n)$ split sets.

**Proof:**

- Let $B'$ be a maximal lattice free set containing $B$ in its interior.
- Lovász (1989) and Basu, Conforti, Cornuejols, Zambelli (2010) showed that
  \[ B' = Q + L \]
  where $Q$ is a polytope and $L$ a rational linear space.
- Let $\dim(Q) = d$ and $\dim(L) = n - d > 0$.
- After a unimodular transformation, $Q \subset \mathbb{R}^d$ and $L = \mathbb{R}^{n-d}$
- Use the result for the bounded case and the result follows.
Combining the two cases

**Theorem:** Every facet-defining inequality for $P$ is a $h(n)$-branch split cut.

- Let $c^T x + d^T y \geq f$ be valid for $\text{conv}(P)$ but not for $P^{LP}$.
- Let $V \subseteq \mathbb{R}^{n+l}$ be the set cut off by $c^T x + d^T y \geq f$ and let $V^x$ be its the projection on the space of the integer variables.
- $V^x$ is strictly lattice-free, and is non-empty.
- Jörg (2007) showed that $V^x$ is contained in the interior of a lattice-free rational polyhedron and therefore in the interior of a maximal lattice-free convex set.
- Depending on whether $V^x$ is bounded or unbounded, we can use either of the previous two lemmas to prove the claim.

**Note:**

- Jörg already observed that every facet-defining inequality is a disjunctive cut.
- We show that they can be derived as structured disjunctive cuts.
Solving mixed-integer programs

Theorem: The mixed-integer program

\[ \min \{ c^T x + d^T y : (x, y) \in \mathbb{Z}^n \times \mathbb{R}^l, \ Ax + Gy \geq b \} \]

where the data is rational, can be solved in finite time via a pure cutting-plane algorithm which generates only \( t \)-branch split cuts.

Proof: Let \( t = h(n) \approx \Pi_{i=1}^n (2 + \lceil f(i) \rceil) \).

- Represent any \( t \)-branch split disjunction \( D(\pi_1, \ldots, \pi_t, \gamma_1, \ldots, \gamma_t) \) by \( v \in \mathbb{Z}^{(n+1)t} \).
- Let \( \Omega = \mathbb{Z}^{(n+1)t} \) and arrange its members in a sequence \( \{\Omega_i\} \), (by increasing norm)
- Let \( D_i \) be the \( t \)-branch split disjunction defined by \( \Omega_i \).
- Any facet-defining inequality of \( \text{conv}(P) \), is a \( t \)-branch split cut defined by the disjunction \( D_k \) for some (finite) \( k \).
- Let \( k^* \) be the largest index of a disjunction associated with facet-defining inequalities.
- Solve the relaxation of the MIP for \( P_i = P_{i-1} \cap \text{conv}(P_0 \cap D_i) \) for \( i = 1, 2, \ldots \)

Note: Validity of a given inequality can also be checked by changing the termination criterion. Similarly, \( \text{conv}(P) \) can also be computed the same way.
How finite is this algorithm?

**Previous Theorem:** The mixed-integer program

\[
\min \{ c^T x + d^T y : (x, y) \in \mathbb{Z}^n \times \mathbb{R}^l, \ Ax + Gy \geq b \}
\]

*can be solved in finite time via a pure cutting-plane algorithm.*

**Proof:** The algorithm cannot run forever.

---

**Stronger result:** The runtime of this algorithm is bounded.

**Proof:** \( P^{LP} \) has bounded facet complexity (\#bits to represent facet defining inequalities)

\[ \Rightarrow \text{Therefore } \text{conv}(P) \text{ has bounded facet complexity.} \]

\[ \Rightarrow \text{Therefore } V \text{ (points cut-off by a facet) has bdd complexity.} \]

\[ \Rightarrow V \text{ has a "thin" direction along an integer vector of bdd complexity.} \]

\[ \text{(we prove this by formulating the lattice width problem as an IP with bdd complexity)} \]

\[ \Rightarrow \text{Therefore the multi-branch disjunction needed to generate a facet has bdd complexity.} \]

\[ \Rightarrow \text{It is possible to make a list of relevant split disjunctions in advance.} \]
Next: How finite is t?
• We showed that every facet-defining inequality for $P$ is a multi-branch split cut that uses at most $h(n) \approx \prod_{k=1}^{n} \left(2 + \lceil f(k) \rceil \right)$ split sets.

[best know bound $f(k) \leq O(k^{4/3} \log^c k)$ by Rudelson, 2000].

• Are there examples where one has to use a large number of split sets?
  – It is easy to show that $t \geq \Omega(n)$
  – With some work, we can also show that $t \geq \Omega(2^n)$
Theorem: For any $n \geq 3$ there exists a nonempty rational mixed-integer polyhedral set in $\mathbb{Z}^n \times \mathbb{R}$ with a facet-defining inequality that cannot be expressed as a $3 \times 2^{n-2}$-branch split cut.

Proof: (outline)

- Construct a full-dimensional rational, lattice-free polytope $B \subset \mathbb{R}^n$ such that
  - Its interior cannot be covered by $3 \times 2^{n-2}$ split sets
  - The integer hull of $B \subset \mathbb{R}^n$ has dimension $n$

- Define a mixed-integer polyhedral set $P_B$ as follows:
  \[ P_B = \{(x, y) \in \mathbb{Z}^n \times \mathbb{R} : (x, y) \in B'\} \]

  where
  \[ B' = \text{conv}((B \times \{-1\}) \cup (B \times \{0\}) \cup (\bar{x} \times \{1/2\})) \]

  and $\bar{x}$ is a point in the interior of $B$.

- $y \leq 0$ is a facet-defining inequality for $\text{conv}(P_B)$

- To cover $V = \{(x, y) \in P_B^{LP} : y > 0\}$, one needs at least $(3 \times 2^{n-2}) + 1$ split sets.
How to construct the lattice-free polytope $B \subset \mathcal{R}^n$

For $\Delta \in \{0, \ldots, 2^{n-2} - 1\}$, let $T_\Delta \in \mathcal{R}^2$ be a (rational) lattice-free triangle

Let $\Delta = \sum_{l=1}^{n-2} \delta_l 2^{l-1}$ with $\delta_l \in \{0, 1\}$

$T_\Delta = \{(\delta_1, \ldots, \delta_{n-2}, x, y) \in \mathcal{R}^n | (x, y) \in T_\Delta\}$

Define

$$B_\varepsilon = \text{conv}\left(\bigcup_{\Delta=0}^{2^{n-2}-1} (T_\Delta \cup \{p_{\varepsilon, \Delta}\})\right)$$

where

$$p_{\varepsilon, \Delta} = (\delta_1, \ldots, \delta_{n-2}, \text{cent}(T_\Delta)) + ((2\delta_1 - 1)\varepsilon, \ldots, (2\delta_{n-2} - 1)\varepsilon, 0, 0)$$

(For example, $p_{\varepsilon, 0} = (-\varepsilon, \ldots, -\varepsilon, \bar{x}, \bar{y})$ where $(\bar{x}, \bar{y}) = \text{cent}(T_0)$.)

**Fact:** $B_\varepsilon$ is full dimensional, rational, lattice-free and $\text{rel.int}(T_\Delta) \subset \text{int}(B_\varepsilon)$. 
How to construct the triangles $T_\Delta$

- $T_0 \in \mathcal{R}^2$ is a rational Hurken's triangle with $w(T_0) \geq 2.15$ that needs at least 3 split sets to cover.

- For $\Delta \in \{1, \ldots, 2^{n-2} - 1\}$,

$$T_\Delta = M_\Delta T_0$$

where $M_\Delta$ is a $2 \times 2$ unimodular matrix with the property that:

* If a split set is useful in covering some $T_\Delta$, it is not useful for $T'_\Delta$ unless $\Delta = \Delta'$

when $n = 3$
How to construct the unimodular matrices $M_{\Delta}$

1. **Useful split sets are finite.**

   For any compact set $K \subset \mathbb{R}^n$ and any number $\varepsilon > 0$, the collection of split sets $S(a, b)$ such that $\text{vol}(K \cap S(a, b)) \geq \varepsilon$ is finite.

2. **Useful split sets are really necessary for $T_0$.**

   For any fixed $l \geq 0$, there exists a finite collection of split sets $\Sigma_l$ such that whenever some $l$ split sets cover $T_0$, then at least 3 of them are contained in $\Sigma_l$.

3. **Bending the triangles.**

   Given any two finite sets of vectors $V, W \subseteq \mathbb{Z}^2 \setminus \{0\}$, there exists an unimodular matrix $M$ such that $MW \subseteq \mathbb{Z}^2 \setminus \{0\}$ and $MW \cap V = \emptyset$.

   **Proof**: Let $q = \max_{v \in W} ||v||_{\infty}$ then

   $$M = \begin{pmatrix} 1 \\ \mu \\ \mu^2 + 1 \end{pmatrix} \text{ where } \mu = 3q$$
Putting it together

- Start with $2^{n-2}$ copies of the rational Hurken’s triangle in $\mathcal{R}^2$.
- Bend the $k$th copy so that split sets useful for $T_0, \ldots, T_{k-1}$ are not useful for $T_k$.
- Extend the corners of a hypercube in $\mathcal{R}^{n-2}$ with the triangles to $\mathcal{R}^n$.
- Add apexes to make the triangles in the interior of $B$.
- To cover the interior of $B$, one needs to cover the triangles
- Last two coordinates of a split set in $\mathcal{R}^n$ gives a split set in $\mathcal{R}^2$.
- At least $3 \cdot 2^{n-2}$ split sets are necessary to cover $B$.
- To show that $y \leq 0$ is valid for $\text{conv}(P_B)$

$$P_B = \{(x, y) \in \mathbb{Z}^n \times \mathcal{R} : (x, y) \in B'\}.$$ 

where

$$B' = \text{conv}((B \times \{-1\}) \cup (B \times \{0\}) \cup (\bar{x} \times \{1/2\}))$$

one needs at least $(3 \times 2^{n-2})$ split sets.
thank you...