Strengthening of Lower Bounds in the Global Optimization of Bilinear Generalized Disjunctive Programs

Ignacio E. Grossmann and Juan P. Ruiz
Center for Advanced Process Decision-making
Carnegie Mellon University
Pittsburgh, PA 15213, USA

European Workshop on Mixed-Integer Nonlinear Programming
Marseille, France
April 12-16, 2010
Objectives http://www.minlp.org:
- Create a library of optimization problems that are formulated as MINLP models.
- Provide high level descriptions of the problems with one or several model formulations with corresponding input files for one or several instances.
Formulation of models is emphasized which allows comparison and evaluation of numerical performance of different codes.

Models in diverse areas: engineering, physics, biology, finance.

- Supports discussion through forum

Future:
- Guidelines for modeling
- Contribute open problems
Outline

Introduction Generalized Disjunctive Programming

Linear Relaxations for Bilinear and Concave GDPs

Strengthening of lower bounds through basic steps

Rules to apply basic steps

Computational results

Bilinearities in process networks

Vector cuts

Computational results
Generalized Disjunctive Programs

Raman & Grossmann (1994)

- Combination of algebraic equations, disjunctions and logic propositions
- “Natural” representation of engineering problems

\[
\begin{align*}
\text{Min} & \quad Z = f(x) + \sum_{k \in K} c_k \\
\text{s.t.} & \quad g(x) \leq 0 \\
& \quad \bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ r_{ik}(x) \leq 0 \\ c_k = \gamma_{ik}(x) \end{bmatrix} \\
& \quad \Omega(Y) = \text{True}
\end{align*}
\]

<table>
<thead>
<tr>
<th>Non Convex</th>
<th>Bilinear</th>
<th>Concave</th>
</tr>
</thead>
</table>

Goal:
Develop efficient algorithms for bilinear and concave GDP with improved relaxations => stronger lower bounds global optimum
Nonconvex Generalized Disjunctive Programs

**Bilinear GDP**

\[
\begin{align*}
\text{Min} & \quad Z = d^T x + \sum_{k \in K} c_k \\
\text{s.t.} & \quad x^T Q^l x + a^l x \leq b^l \quad l \in L \\
& \quad \bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ x^T Q^j x + a^j x \leq b^j \quad j \in J_{ik} \\ c_k \equiv \gamma_{ik} \end{bmatrix} \quad k \in K \\
\Omega(Y) = \text{True} \\
x^{lo} \leq x \leq x^{up} \\
x \in \mathbb{R}^n, c_k \in \mathbb{R}^1, Y_{ik} \in \{\text{True, False}\}
\end{align*}
\]

**Remarks**

- If \(Q^l\) and \(Q_{i}^{jk}\) not SPD, then **indefinite matrices** with zero diagonal elements
- Defining \(X = xx^T\) and \(\bullet\) denoting the scalar product, \(x^T Q^l x = Q^l \bullet X\)

Carnegie Mellon
Nonconvex Generalized Disjunctive Programs

Linear GDP relaxation of Bilinear GDP

\[ \text{Min } Z = d^T x + \sum_{k \in K} c_k \]
\[ \text{s.t. } Q^i \cdot X + d^j x \leq b^i \quad i \in L \]
\[ \vee_{i \in D_k} \begin{bmatrix} Y_{sk} \\ Q_{ik}^j \cdot X + a^j_{ik} x \leq b^j_{ik} \\ c_k = \gamma_{sk} \end{bmatrix} \quad k \in K \]
\[ \Omega(\Gamma) = \text{True} \]

- Previous approaches (Lee & Grossmann, 2003) solve the hull relaxation of the LGDP to predict lower bounds. Can we obtain \textit{stronger lower bounds}?

Remarks

- Previous approaches (Lee & Grossmann, 2003) solve the hull relaxation of the LGDP to predict lower bounds. Can we obtain \textit{stronger lower bounds}?
Nonconvex Generalized Disjunctive Programs

Concave GDP

\[
\begin{align*}
\text{Min} & \quad Z = d^T x + \sum_{k \in K} c_k \\
\text{s.t.} & \quad g_1^l(x) \leq 0 \quad l \in L \\
& \quad \bigwedge_{i \in D_k} \begin{bmatrix}
Y_{ik} \\
\gamma_k(x) \\
r_k j(x) \leq 0 \\
\end{bmatrix} \\
& \quad \Omega(Y) = \text{True} \\
& \quad x^{lo} \leq x \leq x^{up} \\
& \quad x \in \mathbb{R}^n, c_k \in \mathbb{R}^l, Y_{ik} \in \{\text{True}, \text{False}\}\]
\end{align*}
\]

where \( r_{ik}^l \), \( \gamma_k \) and/or \( g_1^l \) are concave

\textbf{Remarks}

- Economies of scale are often defined through the function \( \gamma_{ik}(x) \) which, in general, is concave univariate.

Carnegie Mellon
Nonconvex Generalized Disjunctive Programs

**Linear GDP relaxation of Concave GDP**

\[
\begin{align*}
\text{Min} \quad & Z = d^T x + \sum_{k \in \mathcal{K}} c_k \\
\text{s.t.} \quad & A_{\mathcal{E}}^i x \leq b_{\mathcal{E}}^i \quad & l \in L \\
& \bigvee_{i \in \mathcal{I}_k} \begin{bmatrix} Y_{ik} \\
A_{\mathcal{I}_k}^i x \leq b_{\mathcal{I}_k}^i \quad i \in J_{ik} \\
A_{\mathcal{J}_k}(x, c_k) \leq b_{\mathcal{J}_k}^i \end{bmatrix} \quad & k \in \mathcal{K} \\
& \Omega(Y) = \text{True} \\
& x^{lo} \leq x \leq x^{up} \\
& x \in \mathbb{R}^n, c_k \in \mathbb{R}, Y_{ik} \in \{\text{True}, \text{False}\}
\end{align*}
\]

**Remarks**
- If \( \gamma_{ik}(x) \) is concave univariate the **secant** can be used as underestimator, namely

\[
\gamma_{ik}^\Phi = \frac{\gamma_{ik}(x^{up}) - \gamma_{ik}(x^{lo})}{x^{up} - x^{lo}} (x - x^{lo}) + \gamma_{ik}(x^{lo})
\]

- **Previous approaches** (Lee & Grossmann, 2003) solve the **hull relaxation of the LGDP** to predict lower bounds. Can we obtain **stronger lower bounds**?
Proposed framework to obtain stronger relaxations for nonconvex GDP
(Bilinear and Concave GDP)  
Ruiz, Grossmann (2009)

Basic idea:
Based on relaxation of the nonconvex GDP as a linear GDP exploit
the theory behind DP to obtain stronger relaxations.

The framework consists of two main phases:

1- Generate a valid Linear Generalized Disjunctive Program relaxation for the nonconvex GDP problem (e.g. bilinear and concave).

2- Strengthen the continuous relaxation of the linear GDP obtained in phase 1 by applying a set of basic steps

The hierarchy of relaxations obtained by the application of basic steps is valid for the original nonconvex GDP problem
MILP Reformulation of Linear GDP

Lee and Grossmann (2000)

Linear GDP

\[ \begin{align*}
\text{Min } Z &= d^T x + \sum_k c_k \\
\text{s.t. } Bx &\leq b \\
\bigvee_{i \in \Omega_k} \begin{bmatrix} Y_{ik} \\ A_{ik} x \leq a_{ik} \\ c_k = \gamma_{ik} \end{bmatrix} &= \chi_{ik} \\
\Omega(Y) &= \text{True} \\
x_{\text{lo}} \leq x \leq x_{\text{up}} \\
x &\in \mathbb{R}^n, c_k \in \mathbb{R}^1, Y_{ik} \in \{\text{True}, \text{False}\}
\end{align*} \]

MILP

(Hull Relaxation Reformulation)

\[ \begin{align*}
\text{Min } Z &= d^T x + \sum_k \sum_{i \in \Omega_k} \gamma_{ik} \lambda_{ik} \\
\text{s.t. } Bx &\leq b \\
x &= \sum_{i \in \Omega_k} \nu_{ik} \\
A_{ik} \nu_{ik} &\leq a_{ik} \lambda_{ik} \\
0 \leq \nu_{ik} \leq \lambda_{ik} U_v \\
\sum_{i \in \Omega_k} \lambda_{ik} &= 1 \\
A \lambda &\geq a \\
x_{\text{lo}} \leq x \leq x_{\text{up}} \\
x &\in \mathbb{R}^n, c_k \in \mathbb{R}^1, \lambda_{ik} \in \{0,1\}
\end{align*} \]

Does this formulation yield the strongest relaxation for \(0 \leq \lambda \leq 1?\)
**LGDP to Disjunctive Programming Reformulation**
*(Sawaya & Grossmann, 2008)*

**Linear GDP**

\[
\begin{align*}
\text{Min } & \quad Z = d^T x + \sum_k c_k \\
\text{s.t. } & \quad Bx \leq b \\
\end{align*}
\]

\[
\bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ A_{ik} x \leq a_{ik} \\ c_k = \gamma_{ik} \end{bmatrix} \quad k \in K
\]

\[
\bigvee_{i \in D_k} Y_{ik} \quad k \in K
\]

\[
\Omega(Y) = \text{True}
\]

\[
x_{lo} \leq x \leq x_{up}
\]

\[
x \in \mathbb{R}^n, c_k \in \mathbb{R}^1, Y_{ik} \in \{\text{True, False}\}
\]

**DP (Balas, 1979)**

\[
\begin{align*}
\text{Min } & \quad Z = d^T x + \sum_k c_k \\
\text{s.t. } & \quad Bx \leq b \\
\end{align*}
\]

\[
\bigvee_{i \in D_k} \begin{bmatrix} \lambda_{ik} = 1 \\ A_{ik} x \leq a_{ik} \\ c_k = \gamma_{ik} \end{bmatrix} \quad k \in K
\]

\[
\sum_{i \in D_k} \lambda_{ik} = 1 \quad k \in K
\]

\[
A \lambda \geq a
\]

\[
x_{lo} \leq x \leq x_{up}
\]

\[
x \in \mathbb{R}^n, c_k \in \mathbb{R}^1, 0 \leq \lambda_{ik} \leq 1
\]

**Proposition 1:** LGDP and LDP have equivalent solutions.

**Integration guaranteed for \( \lambda \)**
There exists many forms between CNF and DNF that are equivalent.

**Regular Form (RF):** form represented by intersection of unions of polyhedra

Thus the RF is:

\[ F = \bigcap_{t \in T} S_t \]

where for \( t \in T \), \( S_t = \bigcup_{i \in Q_t} P_i \), \( P_i \) a polyhedron, \( i \in Q_t \).

**Proposition 2 (Theorem 2.1 in Balas (1979)).** Let \( F \) be a disjunctive set in RF. Then \( F \) can be brought to DNF by \( |T| - 1 \) recursive applications of the following basic steps, which preserve regularity:

For some \( r, s \in T, r \neq s \), bring \( S_r \cap S_s \) to DNF, by replacing it with:

\[ S_{rs} = \bigcup_{i \in Q_r \cap Q_s} (P_i \cap P_i). \]
Illustrative Example: Basic Steps

\[ F = S_1 \cap S_2 \cap S_3 \]

\[ S_1 = (P_{11} \cup P_{21}) \]
\[ S_2 = (P_{12} \cup P_{22}) \]
\[ S_3 = (P_{13} \cup P_{23}) \]

Then \( F \) can be brought to DNF through 2 basic steps.

Apply Basic Step to:

\[ S_1 \cap S_2 = (P_{11} \cup P_{21}) \cap (P_{12} \cup P_{22}) \]
\[ S_{12} = (P_{11} \cap P_{12}) \cup (P_{11} \cap P_{22}) \cup (P_{21} \cap P_{12}) \cup (P_{21} \cap P_{22}) \]

We can then rewrite

\[ F = S_1 \cap S_2 \cap S_3 \quad \text{as} \quad F = S_{12} \cap S_3 \]

Apply Basic Step to:

\[ S_{12} \cap S_3 = ((P_{11} \cap P_{12}) \cup (P_{11} \cap P_{22}) \cup (P_{21} \cap P_{12}) \cup (P_{21} \cap P_{22})) \cap (P_{13} \cup P_{23}) \]
\[ S_{123} = \left( (P_{11} \cap P_{12} \cap P_{13}) \cup (P_{11} \cap P_{22} \cap P_{13}) \cup (P_{21} \cap P_{12} \cap P_{13}) \cup (P_{21} \cap P_{22} \cap P_{13}) \right) \]
\[ \cup (P_{11} \cap P_{12} \cap P_{23}) \cup (P_{11} \cap P_{22} \cap P_{23}) \cup (P_{21} \cap P_{12} \cap P_{23}) \cup (P_{21} \cap P_{22} \cap P_{23}) \]

We can then rewrite

\[ F = S_{12} \cap S_3 \quad \text{as} \quad F = S_{123} \quad \text{which is its equivalent DNF} \]
A Hierarchy of Relaxations for DP

Hull Relaxation (Balas, 1985):

Let us take the following disjunctive set:

\[ F = \bigcap_{j \in T} S_j \]

Then the hull-relaxation corresponds to:

\[ h - \text{rel } F := \bigcap_{j \in T} \text{clconv } S_j. \]

Proposition 3 (Theorem 4.3 in Balas (1979)): For \( i = 0, 1, \ldots, t \), let \( F_i = \bigcap_{j \in T_i} S_j \) be a sequence of regular forms of a disjunctive set, such that

i) \( F_0 \) is in CNF, with \( P_0 = \bigcap_{j \in T_0} S_j \);

ii) \( F_i \) is in DNF;

iii) for \( i = 1, \ldots, t \), \( F_i \) is obtained from \( F_{i-1} \) by a basic step.

Then,

\[ P_0 = h - \text{rel } F_0 \supseteq h - \text{rel } F_1 \supseteq \cdots \supseteq h - \text{rel } F_t = \text{clconv } F_t. \quad (\text{true convex hull}) \]
Proposed Framework (Phase 1)

Nonconvex GDP

\[
\begin{align*}
\text{Min } & \quad Z = d^T x + \sum_{k \in K} c_k \\
\text{s.t.} & \quad g^l(x) \leq 0 \quad l \in L \\
& \quad \sum_{i \in D_k} \left[ y_{ik} \right] r^j(x) \leq 0 \quad j \in J_k \\
& \quad c_k = \gamma_k(x) \\
& \quad \Omega(Y) = \text{True} \\
\end{align*}
\]

\[x \in \mathbb{R}^n, c_k \in \mathbb{R}^1, Y_k \in \{\text{True}, \text{False}\}\]

\[\text{GDP}_\text{NC}\]

Linear GDP Relaxation

\[
\begin{align*}
\text{Min } & \quad Z = d^T x + \sum_{k \in K} c_k \\
\text{s.t.} & \quad A_g^l x \leq b_g^l \quad l \in L \\
& \quad \sum_{i \in D_k} \left[ y_{ik} \right] A^j x \leq b_j^k \quad j \in J_k \\
& \quad A_k(x, c_k) \leq b_k^k \\
& \quad \Omega(Y) = \text{True} \\
\end{align*}
\]

\[x \in \mathbb{R}^n, c_k \in \mathbb{R}^1, Y_k \in \{\text{True}, \text{False}\}\]

\[\text{GDP}_\text{RLP}\]

\[\text{GDP}_\text{RLP} \text{ is a valid Linear GDP relaxation for GDP}_\text{NC}\]
Proposed Framework (Phase 2)
Valid Hierarchy of Relaxations

From theorem 2.1 (Balas, 1985)

\[ \text{GDP}_{RLP_0} \sim \text{GDP}_{RLP_1} \sim \ldots \text{GDP}_{RLP_i} \sim \text{GDP}_{RLP_t} \]

\( \text{GDP}_{RLP_i} \) is obtained by the application of a basic step on \( \text{GDP}_{RLP_{i-1}} \)

From Proposition 1 (Ruiz & Grossmann, 2008)

\[ \text{GDP}_{NC} \subseteq \text{GDP}_{RLP_0} \sim \text{GDP}_{RLP_t} \]

Hence applying Theorem 4.3 (Balas, 1979)

\[ h - \text{rel} \text{GDP}_{RLP_0} \supseteq h - \text{rel} \text{GDP}_{RLP_1} \supseteq \ldots h - \text{rel} \text{GDP}_{RLP_i} \supseteq \ldots \supseteq h - \text{rel} \text{GDP}_{RLP_t} \supseteq \text{GDP}_{RLP_t} \supseteq \text{GDP}_{NC} \]

Valid Hierarchy of Relaxations for \( \text{GDP}_{NC} \)
Illustrative Example: Optimal reactor selection I

Min Z = \(-qFX + gF + CP\)

s.t.

\[
\begin{align*}
FX & \leq d \\
0 & \leq X \leq X_{UP} \\
CP & = C_p
\end{align*}
\]

\[
\begin{align*}
Y_{11} & = \begin{cases} 1 & \text{if } F = \alpha_1X + \beta_1 \\
0 & \text{otherwise} \end{cases} \\
Y_{21} & = \begin{cases} 1 & \text{if } F = \alpha_2X + \beta_2 \\
0 & \text{otherwise} \end{cases}
\end{align*}
\]

\[
\begin{align*}
Y_{11} \vee Y_{21} & = \text{True} \\
CP, X, F & \in R \\
F_{LO} & \leq F \leq F_{UP} \\
Y_{11}, Y_{21} & \in \{\text{True, False}\}
\end{align*}
\]

Optimum Z* = -1.01
Illustrative Example: Optimal reactor selection I
Lee & Grossmann (2003) Relaxation

Lower bound $Z^* = -1.28 < -1.01$
Illustrative Example: Optimal reactor selection I

 Proposed Relaxation

**Relaxation (Basic Steps)**

**Bilinear Terms**

\[
\begin{align*}
\text{Min } Z &= -F \cdot X + F_{\text{UP}} \cdot \epsilon P \\
\text{s.t. } F R &\leq d
\end{align*}
\]

Convex Envelopes

Basic Step

Illustrative Example: Optimal reactor selection I

Proposed Relaxation

Lower bound \( Z^* = -1.1 < -1.01 \) and tighter than \(-1.28!\)
The application of **basic step** prior to the discrete relaxation leads to a **tighter** relaxed feasible region = > **stronger lower bounds**
Rules to apply basic steps

**Proposition 4.**

Let $S_1$ and $S_2$ be two disjunctions in which the variables appearing in $S_1$ do not appear in $S_2$ and vice versa. Then the system satisfies $\text{cl conv}(S_1 \cap S_2) = (\text{cl conv } S_1) \cap (\text{cl conv } S_2)$

**RULE 1:** Apply basic steps between disjunctions with at least one variable in common.

**RULE 2:** Apply basic steps between disjunctions with more variables in common.
Proposition 5.

Let $S = P_1 \cup P_2$ where $P_1$ and $P_2$ are polyhedra defined in the $x$ real space, and let $H$ be a half space defined as $ax + b \leq 0$ and $H^*$ be a facet of $H$. If $P_1$ is a point such that $P_1$ is in $H^*$ then $\text{cl conv}(H \cap S) = \text{cl conv}(H) \cap \text{cl conv}(S)$

**RULE 3:** Do not apply basic steps between disjunctions with the properties presented above
Proposition 6.

When a basic step is applied between an improper disjunction with a proper disjunction the number of polyhedra in the resulting disjunction is not increased

Intersection proper disjunction 2 polyhedra with one improper disjunction 1 polyhedron

\[ F = S_1 \cap S_2 \]  
\[ S_1 = P_1 \cup P_2 \]  
\[ S_2 = P_3 \]

Number of polyhedra after the intersection of proper disjunctions: 2

\[ F = S_{12} = ((P_1 \cap P_3) \cup (P_2 \cap P_3)) \]

**RULE 4: Apply basic steps between improper and proper disjunctions**
Summary of “practical” rules to apply basic steps

- Apply basic steps between those **disjunctions** with at least one **variable in common**.

- The **more variables in common** two disjunctions have the **more** the **tightening** expected.

- A basic step between a half space and a disjunctions with two disjuncts one of which is a point contained in the facet of the half space **will not tighten the relaxation**.

- A **smaller increase in the size** of the formulation is expected when **basic steps** are applied between **improper** disjunctions and **proper** disjunctions.
Example 1: Optimal Reactor selection I

**Bilinear GDP**

**Superstructure of Reactor Network**

No. of cont. vars. : 3
No. of disc. vars. : 2
No. of bilinear terms: 1
No. of concave functions: 0

**Generalized Disjunctive Program**

\[
\begin{align*}
\text{Min } Z &= -(\theta FX - \gamma F - CP) \\
\text{s.t. } & FX \leq d \\
& \begin{bmatrix} Y_{11} \\ F = \alpha_1 X + \beta_1 \\ X_{l1}^{\alpha} \leq X \leq X_{u1}^{\beta} \\ CP = C_{p1} \end{bmatrix} \lor \begin{bmatrix} Y_{21} \\ F = \alpha_2 X + \beta_2 \\ X_{l2}^{\alpha} \leq X \leq X_{u2}^{\beta} \\ CP = C_{p2} \end{bmatrix} \\
& Y_{11} \lor Y_{21} \\
& X, F, CP \in R^1, F_{lo} \leq F \leq F_{up}, Y_{11}, Y_{21} \in \{\text{True, False}\}
\end{align*}
\]

\[Z^* = -1.01\]
Example 2: Optimal Reactor selection II

Superstructure of Reactor Network

Generalized Disjunctive Program

\[ \text{Min } Z = C_p \]
\[ \text{s.t. } F_a = F_{c_1} + F_{c_2} \]
\[ Y_1 \leq Y_2 \]

\[ \begin{align*}
Y_1 & = \alpha_1 F_b + \beta_1 \\
C_p & = \gamma_1 F_a^{0.7} + \varepsilon_1 \\
F_a^{LO_1} & \leq F_a \leq F_a^{UP_1}
\end{align*} \lor \begin{align*}
Y_2 & = \alpha_2 F_b + \beta_2 \\
C_p & = \gamma_2 F_a^{0.7} + \varepsilon_2 \\
F_a^{LO_2} & \leq F_a \leq F_a^{UP_2}
\end{align*} \]

\[ D_{c_1}^{LO} \leq F_{c_1} \leq D_{c_1}^{UP} \]
\[ D_{c_2}^{LO} \leq F_{c_2} \leq D_{c_2}^{UP} \]
\[ F_a^{LO} \leq F_a \leq F_a^{UP} \]

No. of cont. vars. : 5
No. of disc. vars. : 2
No. of bilinear terms: 0
No. of concave functions: 2

Optimal selection

\[ Z^* = 6.31 \]
Example 3: Heat exchanger network design

**Bilinear/Concave GDP**

**Heat Exchanger Network**

No. of cont. vars.: 8  
No. of disc. vars.: 9  
No. of bilinear terms: 4  
No. of concave functions: 9

**Optimal Objective**

\[ Z^* = 114384.78 \]

**Generalized Disjunctive Program**

\[
\begin{align*}
\text{Min } Z &= \sum_i CP_i + FCP_h (T_1 - T_{out_{th}}) C_{cu} + FCP_c (T_{out_{c}} - T_2) C_{hu} \\
\text{s.t.} \\
FCP_h (T_{in_{h}} - T_1) &= A_1 U_1 \left( \frac{T_{in_{h}} - T_2}{T_1 - T_{inc}} \right) \\
FCP_h (T_{in_{h}} - T_1) &= A_2 U_2 \left( \frac{T_{incw} + \frac{2}{2} T_{out_{cw}}} + (T_1 - T_{out_{cw}}) \right) \\
FCP_c (T_{out_{c}} - T_2) &= A_3 U_3 \left( \frac{T_{out_{c}} - T_2}{2} + (T_{out_{c}} - T_{out_{c}}) \right) \\
FCP_h (T_{in_{h}} - T_1) &= FCP_c (T_2 - T_{inc}) \\
\end{align*}
\]

\[
\begin{bmatrix}
Y_{1i} \\
A_{i1}^{lo} \leq A_i \leq A_{i1}^{up}
\end{bmatrix} \vee \begin{bmatrix}
Y_{2i} \\
A_{i2}^{lo} \leq A_i \leq A_{i2}^{up}
\end{bmatrix} \vee \begin{bmatrix}
Y_{3i} \\
A_{i3}^{lo} \leq A_i \leq A_{i3}^{up}
\end{bmatrix}
\]

\[ Y_{1i} \vee Y_{2i} \vee Y_{3i} = \text{True} \quad i = 1, 2, 3 \]

\[ T_{11}^{lo} \leq T_1 \leq T_{11}^{up} \]

\[ T_{21}^{lo} \leq T_2 \leq T_{21}^{up} \]
Example 4: Water Treatment Network design

**Bilinear GDP (bilinear outside disjunction)**

**Process superstructure**

**Optimal structure**

**Generalized Disjunctive Program**

Min \( Z = \sum_{k \in PU} CP_k \)

s.t.

\[ f_k^j = \sum_{i \in M_k} f_i^j \quad \forall j \quad k \in MU \]

\[ \sum_{i \in S_k} f_i^j = f_k^j \quad \forall j \quad k \in SU \]

\[ \sum_{i \in S_k} \zeta_i^k = 1 \quad k \in SU \]

\[ f_i^j = \zeta_i^k f_k^j \quad \forall j \quad i \in S_k \quad k \in SU \]

\[ F_k = \sum_{j} f_i^j, i \in OPU_k \quad k \in PU \]

\[ CP_k = \delta_{ik} F_k \]

\[ 0 \leq \zeta_i^k \leq 1 \quad \forall j, k \]

\[ 0 \leq f_i^j, f_k^j \quad \forall i, j, k \]

\[ 0 \leq CP_k \quad \forall k \]

\[ YP_k^h \in \{true, false\} \quad \forall h \in D_k \quad \forall k \in PU \]

No. of cont. vars. : 114
No. of disc. vars. : 9
No. of bilinear terms: 36

\[ Z^* = 1214 \]
**Example 5: Pooling network design**

**Bilinear GDP (bilinear inside disjunction)**

**Generalized Disjunctive Program**

Min \( Z = \sum_{j \in J} CP_j + \sum_{i \in I} CST_i + \sum_{j \in J} \sum_{i \in I} c_{ij} \sum_{w \in W} f_{jw} - \sum_{k \in K} d_k \sum_{j \in J} \sum_{w \in W} f_{jkw} \)

s.t.
\[
\sum_{i \in I} \sum_{w \in W} f_{jw} = \sum_{k \in K} f_{jkw} \quad \forall j \in J
\]
\[
\sum_{j \in J} f_{jkw} - S_k = 0 \quad \forall k \in K
\]
\[
f_{jw} = \lambda_{iw} \sum_{w' \in W} f_{jw'} \quad \forall i \in I, \forall j \in J, \forall w \in W
\]
\[
\sum_{j \in J} f_{jkw} - Z_{kw} \sum_{j \in J} \sum_{w \in W} f_{jkw} = 0 \quad \forall k \in K, \forall w \in W
\]

\[
\begin{bmatrix}
YST_i \\
f^{lo} \leq \sum_{i \in I} \sum_{w' \in W} f_{jw'} \\
CST_i = \alpha_i
\end{bmatrix}
\begin{bmatrix}
\neg YST_i \\
f^{lo} = 0 \\
f_{jw} = 0
\end{bmatrix}
\quad \forall i \in I
\]

\[
\begin{bmatrix}
YP_j \\
f^{lo} \leq \sum_{i \in I} \sum_{w' \in W} f_{jw'} \\
\sum_{k \in K} f_{jkw} = \sum_{i \in I} f_{jw}, \forall w \in W \\
f_{jkw} = \zeta^k_j \sum_{i \in I} f_{jw} \quad \forall w \in W, k \in K
\end{bmatrix}
\begin{bmatrix}
\neg YP_j \\
f_{jw} = 0, \forall i \in I, w \in W \\
f_{jkw} = 0, \forall k \in K, w \in W \\
CP_j = \gamma_j
\end{bmatrix}
\quad \forall j \in J
\]

\[0 \leq \zeta^k_j \leq 1; 0 \leq f_{jkw}, f_{jw} \leq f^{up}\]
\[0 \leq CST_i, CP_j, YST_i, YP_j \in \{true, false\}\]
### Dimensions of Test Problems

<table>
<thead>
<tr>
<th>Example</th>
<th>Bilinear Terms</th>
<th>Concave Functions</th>
<th>Discrete Variables</th>
<th>Continuous Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Example 2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Example 3</td>
<td>4</td>
<td>9</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Example 4</td>
<td>36</td>
<td>0</td>
<td>9</td>
<td>114</td>
</tr>
<tr>
<td>Example 5</td>
<td>24</td>
<td>0</td>
<td>9</td>
<td>76</td>
</tr>
</tbody>
</table>

**Examples**

1- Optimal Reactor selection I
2- Optimal Reactor selection II
3- HEN with investment cost - multiple size Regions (Turkay & Grossmann, 1996)
4- Water Treatment Network Design problem (Galan & Grossmann, 1998)
5- Pooling Network Design problem (Lee & Grossmann, 2003)
## Relaxation Results

<table>
<thead>
<tr>
<th>Example</th>
<th>Global Optimum</th>
<th>Lower Bound (Lee &amp; Grossmann Relaxation)</th>
<th>Lower Bound (Proposed Relaxation)</th>
<th>DNF Lower Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>-1.01</td>
<td>-1.28</td>
<td>-1.10</td>
<td>-1.10</td>
</tr>
<tr>
<td>Example 2</td>
<td>6.31</td>
<td>5.65</td>
<td>6.08</td>
<td>6.08</td>
</tr>
<tr>
<td>Example 3</td>
<td>114384.78</td>
<td>91671.18</td>
<td>94925.77</td>
<td>97858.86</td>
</tr>
<tr>
<td>Example 4</td>
<td>1214.87</td>
<td>400.66</td>
<td>431.90</td>
<td>431.90</td>
</tr>
<tr>
<td>Example 5</td>
<td>-4640</td>
<td>-5515</td>
<td>-5468</td>
<td>-5241</td>
</tr>
</tbody>
</table>

**Remarks**

- Proposed methodology leads to improvements in the lower bounds
  => 31% gap reduction

- The lower bound of the DNF is the best lower bound attainable for a given LGDP.

- Note in examples 1, 2 and 4 the lower bound is the same as the lower bound of the DNF
  Often, it is not necessary to reach the DNF form to have good lower bounds.
Global Optimization Methodology

**GDP reformulation**
Apply basic steps following the rules presented

**Bound Contraction**
(Zamora & Grossmann, 1999)

**Spatial Branch and Bound**
(Lee & Grossmann, 2001)

- **Disjunctive B&B**
- **Feasible discrete**
- **Spatial B&B**
## Computational Performance

<table>
<thead>
<tr>
<th></th>
<th>Global Optimization Technique using Lee &amp; Grossmann Relaxation</th>
<th>Global Optimization Technique using Proposed Relaxation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Global Optimum</td>
<td>Nodes</td>
</tr>
<tr>
<td>Example 1</td>
<td>-1.01</td>
<td>5</td>
</tr>
<tr>
<td>Example 2</td>
<td>6.31</td>
<td>1</td>
</tr>
<tr>
<td>Example 3</td>
<td>114384.78</td>
<td>13</td>
</tr>
<tr>
<td>Example 4</td>
<td>1214.87</td>
<td>450</td>
</tr>
<tr>
<td>Example 5</td>
<td>-4640</td>
<td>502</td>
</tr>
</tbody>
</table>

**Remarks**

- Proposed relaxation led to a significant bound contraction at the root node.
- 44% reduction number of nodes, 23% reduction CPU time
  tighter relaxation but increased size of proposed relaxation

<table>
<thead>
<tr>
<th></th>
<th>Size of the LP Relaxation (Lee &amp; Grossmann)</th>
<th>Size of the LP Relaxation (Proposed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constraints</td>
<td>Variables</td>
</tr>
<tr>
<td>Example 1</td>
<td>23</td>
<td>15</td>
</tr>
<tr>
<td>Example 2</td>
<td>24</td>
<td>14</td>
</tr>
<tr>
<td>Example 3</td>
<td>87</td>
<td>52</td>
</tr>
<tr>
<td>Example 4</td>
<td>544</td>
<td>346</td>
</tr>
<tr>
<td>Example 5</td>
<td>3336</td>
<td>1777</td>
</tr>
</tbody>
</table>
Process Networks

- The optimization of process networks is one of the most frequent problems that is addressed in process systems engineering (e.g. optimization of pooling networks, water treatment networks)

- Mass balances are the common denominator of these systems and are often represented through equations with bilinear terms.

Goal: Propose a methodology to find stronger relaxations for the global optimization of process networks.
Basic equations

Building block of process networks

General mass balance formulation

**Bilinear:** \[ \sum_{i \in I_n} (F_{in} P_{in}^j) - F_{on} P_{on}^j = 0 \quad \forall n \in N, \forall j \in J \]

\[ \sum_{i \in I_n} F_{in} - F_{on} = 0 \quad \forall n \in N \]

**(MPB)**

**Sets:**

- \( N \): Nodes in the network
- \( I_n \): Streams entering node \( n \) and
- \( J \): Property type
Vectorial Representation

For a given node $n$ and property $j$ we define the *vectors*:

$$v_F = (F_1, F_2, \ldots, F_{|I|}, F_o) \quad v_E = (1, 1, \ldots, 1, -1) \quad v_P = (P_1, P_2, \ldots, P_{|I|}, -P_o)$$

*(MBP)* can be represented as:

$$v_P \cdot v_F = 0$$
$$v_F \cdot v_E = 0$$

Or equivalently, in *vectorial form*:

$3$-Vector Representation

$$(VMPB)$$

The *relation* between the vector spaces $v_F, v_P$ and the vector $v_E$ is *exposed* in the 3-Vector Representation.
We define a *minimal set*, the set composed by three elements (i.e. $|I|+1 = 3$)

**Lemma 1:**

Any system of the form *(VMBP)* can be decomposed as the intersection of $|I|-1$ 3-Vector Representation of *minimal sets*

**Illustration (|I| = 4):**

3 minimal sets
Properties of the minimal set

**Lemma 2**: The property vectors \((v_P)\) and flow vectors \((v_F)\) in a minimal set are related as follows:

\[
v_P \perp v_F \land v_F \perp v_E \Rightarrow v_P \times v_E \parallel v_F
\]

Or equivalently

\[
v_P \cdot v_F = 0, v_F \cdot v_E = 0 \Rightarrow v_P \times v_E = \alpha v_F
\]

where

\[
\alpha = \frac{\|v_P\| \sqrt{3} \sin \theta}{\|v_F\|}, \quad 0 \leq \theta \leq 2\pi
\]

**Illustration:**

The cross product between \(v_P\) and \(v_E\) is parallel to \(v_F\)
Properties of the minimal set (cont.)

Lemma 3: The space defined by the minimal set is nonconvex

Illustration:

Given two points in the set

\begin{align*}
  v^1_F &= \{2,1,3\} & v^2_F &= \{1,1,2\} \\
  v^1_P &= \{1,1,-1\} & v^2_P &= \{3,1,-2\} & \in & \text{minimal set} \\
  v^1_E &= \{1,1,-1\} & v^2_E &= \{1,1,-1\}
\end{align*}

The following point, which is a convex combination, is not in the set

\begin{align*}
  0.5v^1_F + 0.5v^2_F &= v^{12}_F = \{1.5,1,2.5\} & \notin & \text{minimal set} \\
  0.5v^1_P + 0.5v^2_P &= v^{12}_P = \{2,1,-1.5\} \\
  0.5v^1_E + 0.5v^2_E &= v^{12}_E = \{1,1,-1\} & v_F \cdot v_P & \neq 0
\end{align*}
Convex relaxation of minimal set
(Traditional Approach)

A traditional relaxation of (MPB) is given by replacing the bilinear terms with the McCormick convex envelopes.

\[ \sum_{i \in \{1,2\}} (F_i P_i) - F_o P_o = 0 \]
\[ \sum_{i \in \{1,2\}} F_i - F_o = 0 \]
\[ F_P \leq F_i P_i^{up} + F_i^{lo} P_i - F_i^{lo} P_i^{up} \]
\[ F_P \leq F_i P_i^{lo} + F_i^{up} P_i - F_i^{up} P_i^{lo} \]
\[ F_P \geq F_i P_i^{up} + F_i^{up} P_i - F_i^{up} P_i^{up} \]
\[ F_P \geq F_i P_i^{lo} + F_i^{lo} P_i - F_i^{lo} P_i^{lo} \]

\[ v_F \times v_P = \alpha v_F \] implicitly defines the orthogonality between \( v_P \) and \( v_F \)

The orthogonality between \( v_P \) and \( v_F \) is lost!
Valid cuts from cross product

Based on **Lemma 2** the following is a valid cut

\[ v_F \times v_P = \alpha v_F \]

which in algebraic form reads

\[ -P_2 + P_o = \alpha F_1 \]

**Nonconvex!**

\[ P_1 - P_o = \alpha F_2 \]

\[ P_1 - P_2 = \alpha F_o \]

From where the following **linear cuts** are derived:

\[ \beta_i \leq \alpha F_i^{up} + \alpha^{lo} F_i - \alpha^{lo} F_i^{up} \]

\[ \beta_i \leq \alpha F_i^{lo} + \alpha^{up} F_i - \alpha^{up} F_i^{lo} \quad i = o, 1, 2 \]

\[ \beta_i \geq \alpha F_i^{up} + \alpha^{up} F_i - \alpha^{up} F_i^{up} \]

\[ \beta_i \geq \alpha F_i^{lo} + \alpha^{lo} F_i - \alpha^{lo} F_i^{lo} \]

where \( \beta_1 = P_o - P_2 \), \( \beta_2 = P_1 - P_o \) and \( \beta_o = P_1 - P_2 \)
Bounds for $\alpha$

From the definition of cross product:

$$\alpha^{up} = -\alpha^{lo} = \max \left( \frac{\|v_p\| \sqrt{3} \sin \theta}{\|v_F\|} \right) = \frac{\|v_p\|_{\max} \sqrt{3}}{\|v_F\|_{\min}} = \frac{\|v_p^{up}\| \sqrt{3}}{\|v_F^{lo}\|}$$

Tighter lower and upper bounds can be obtained by using (CPC):

$$\alpha^{up} = \min(\max \frac{-P_2 + P_o}{F_1}, \max \frac{-P_1 + P_o}{F_2}, \max \frac{P_2 - P_1}{F_o})$$

$$\alpha^{lo} = \max(\min \frac{-P_2 + P_o}{F_1}, \min \frac{P_1 - P_o}{F_2}, \min \frac{-P_2 + P_1}{F_o})$$
Proposed vs Traditional Approach

Proposition

The proposed cuts are not dominated by the McCormick convex envelopes

Illustration

Given the minimal set:

\[ F_1 P_1 + F_2 P_2 = F_3 P_3 \]
\[ F_1 + F_2 = F_3 \]

where:

\[ 0.5 \leq F_1 \leq 2, \ 1.5 \leq F_2 \leq 2.5, \ 2 \leq F_3 \leq 4.5 \]
\[ 0.5 \leq P_1 \leq 1.5, \ 0 \leq P_2 \leq 2, \ 0 \leq P_3 \leq 2 \]

the region in the space with fixed \( F_1 = 0.5, \ F_2 = 2.3, \ F_3 = 2.8, \ P_1 = 1.2, \ P_2 = 0.1 \) is

\( P_3 = [0.19-0.43] \) using the McCormick envelopes
\( P_3 = [0.23-0.36] \) using the proposed cuts

Carnegie Mellon
Test Problems
(Data Reconciliation)

Problem Statement:
Find the set of values of flows and composition that minimize the
squared error when compared with the measurements.

System representation (Instances 1-2):

Formulation:
\[
\begin{align*}
\text{Min } Z &= w_1(F_1 - FI_1)^2 + w_2(F_2 - FI_2)^2 \\
&\quad + w_3(F_3 - FI_3)^2 + w_4(C_1 - CI_1)^2 \\
&\quad + w_5(C_2 - CI_2)^2 + w_6(C_3 - CI_3)^2 \\
F_1C_1 + F_2C_2 &= F_3C_3 \\
F_1 + F_2 &= F_3
\end{align*}
\]
Nonconvex set

\[
\begin{align*}
F_1^{lo} &\leq F_1 \leq F_1^{up} & F_2^{lo} &\leq F_2 \leq F_2^{up} \\
C_1^{lo} &\leq C_1 \leq C_1^{up} & C_2^{lo} &\leq C_2 \leq C_2^{up}
\end{align*}
\]
Numerical Results

System representation (Instances 3-4):
## Numerical Results

On average, the proposed approach led to:

- **46% gap reduction**,  
- **70% of nodes** necessary to find the solution  
- **48% the computational time**.

<table>
<thead>
<tr>
<th>Instance</th>
<th>GO</th>
<th>LB</th>
<th>Nodes</th>
<th>Time(s)</th>
<th>Proposed Approach</th>
<th>LB</th>
<th>Nodes</th>
<th>Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>82.78</td>
<td>78.25</td>
<td>11</td>
<td>20</td>
<td>81.6</td>
<td>4</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>5.26</td>
<td>4.89</td>
<td>80</td>
<td>109</td>
<td>5.01</td>
<td>9</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>13.14</td>
<td>10.45</td>
<td>428</td>
<td>518</td>
<td>11.32</td>
<td>242</td>
<td>283</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>17.19</td>
<td>17.08</td>
<td>7</td>
<td>30</td>
<td>17.13</td>
<td>1</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>
Conclusions

• Tighter reformulations of bilinear/concave GDPs by a 2 Phase procedure.
  - Linear GDP relaxation in Phase 1
  - Proposed general rules to implement basic steps in Phase 2

• Proposed global optimization algorithm to solve bilinear/concave GDPs

• Proposed vector space properties to strengthen the relaxation in bilinear process networks

• Application in several test problems showed improved performance.

Acknowledgements: The authors would like to acknowledge financial support from the National Science Foundation under grant NSF OCI-0750826