Relaxations for Convex Nonlinear Generalized Disjunctive Programs and their Application to Nonconvex Problems

Ignacio Grossmann
Sangbum Lee, Juan Pablo Ruiz, Nick Sawaya

Center for Advanced Process Decision-making
Department of Chemical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213, U.S.A

OSE, Abo Akademi, Turku
December 8, 2011
Basic question:
How can we obtain strong relaxations for MINLP/GDP problems?

Outline

1. Overview of major relaxations for nonlinear GDP problems (*big-M and hull relaxation*)

2. Convex nonlinear GDP: hierarchy of relaxations
   - Concept of basic steps
   - Equivalent NLP formulation

3. Application to global Optimization of nonconvex GDP
   - Bilinear, concave and linear fractional functions
- **Global optimization techniques** find the global optimum by sequentially approximating the non-convex problem with a **convex relaxation**

- **Tighter formulations** lead to more efficient algorithms

**Finding strong relaxations** is a key element in
1. Global Optimization
2. Efficient solution of convex MINLP problems
MINLP

- Mixed-Integer Nonlinear Programming

\[
\min Z = f(x, y) \quad \text{Objective Function}
\]
\[
s.t. \quad g(x, y) \leq 0 \quad \text{Inequality Constraints}
\]
\[
x \in X, \ y \in Y
\]
\[
X = \{ x \mid x \in \mathbb{R}^n, x^L \leq x \leq x^U, Bx \leq b \}
\]
\[
Y = \{ y \mid y \in \{0,1\}^m, Ay \leq a \}
\]

- \(f(x,y)\) and \(g(x,y)\) - assumed to be **convex and bounded** over \(X\).
- \(f(x,y)\) and \(g(x,y)\) commonly **linear** in \(y\)
Mixed-integer Nonlinear Programming

Algorithms

Branch and Bound (BB) Ravindran and Gupta (1985), Stubbs, Mehrotra (1999), Leyffer (2001)

Generalized Benders Decomposition (GBD) Geoffrion (1972)


LP/NLP based Branch and Bound Quesada, Grossmann (1994), Bonami et al. (2008)

Extended Cutting Plane (ECP) Westerlund and Pettersson (1992)
Continuous MINLP Relaxation

- **Nonlinear Programming Problem**

\[
\begin{align*}
\min & \quad Z^L = f(x, y) \\
\text{s.t.} & \quad g(x, y) \leq 0 \\
& \quad x \in X, \ y \in Y
\end{align*}
\]

\[
X = \{x \mid x \in \mathbb{R}^n, x^L \leq x \leq x^U, Bx \leq b\}
\]

\[
Y = \{y \mid 0 \leq y \leq 1, Ay \leq a\}
\]

\[
Z^L = \text{Lower bound to optimal MINLP solution}
\]

Can we develop tighter lower bounds?
Generalized Disjunctive Programming (GDP)

- Raman and Grossmann (1994) (Extension Balas, 1979)
- Motivation: Facilitate modeling discrete/continuous problems

\[
\min \ Z = \sum_{k} c_k + f(x) \\
\text{s.t.} \quad \forall j \in J_k \quad \begin{bmatrix}
Y_{jk} \\
g_{jk}(x) \leq 0 \\
c_k = \gamma_{jk}
\end{bmatrix} \quad k \in K \\
\Omega(Y) = \text{true} \\
x \in \mathbb{R}^n, c_k \in \mathbb{R}^l \\
Y_{jk} \in \{\text{true, false}\}
\]

Objective Function
Common Constraints
Disjunction
Constraints
Fixed Charges
Logic Propositions
Continuous Variables
Boolean Variables

Properties: a) Every GDP can be transformed into an MINLP
b) Every bounded MINLP can be transformed into GDP
Process Network with fixed charges

**GDP model**

Min \( Z = c_1 + c_2 + c_3 + d^T x \)

s.t.

\[
\begin{align*}
  x_1 &= x_2 + x_4 \\
  x_6 &= x_3 + x_5 \\
  Y_{11} &= \gamma_1 \\
  x_3 &= p_1 x_2 \\
  c_1 &= 0 \\
  \vdots \\
  Y_{12} &= \gamma_2 \\
  x_5 &= p_2 x_4 \\
  c_2 &= 0 \\
  \vdots \\
  Y_{13} &= \gamma_3 \\
  x_7 &= p_3 x_6 \\
  c_3 &= 0 \\
  \vdots \\
  Y_{11} \lor Y_{21} &= 0 \\
  Y_{12} \lor Y_{22} &= 0 \\
  Y_{13} \lor Y_{23} &= 0 \\
  Y_{11} \lor Y_{12} \Rightarrow Y_{13} &= 0 \\
  Y_{13} \Rightarrow Y_{11} \lor Y_{12} &= 0 \\
  Y_{21} \lor Y_{22} &= 0 \\
  0 \leq x \leq x^U \\
  Y_{11}, Y_{21}, Y_{12}, Y_{22}, Y_{13}, Y_{23} \in \{True, False\} \\
  c_1, c_2, c_3 \in \mathbb{R}^1
\end{align*}
\]
Generalized Disjunctive Programming (GDP)

- Raman and Grossmann (1994)

\[
\text{min } Z = \sum_{k} c_k + f(x)
\]

\[
s.t. \quad r(x) \leq 0
\]

\[
\bigvee_{j \in J_k} \begin{bmatrix}
Y_{jk} \\
g_{jk}(x) \leq 0
\end{bmatrix} \quad k \in K
\]

\[
\Omega(Y) = true
\]

\[
x \in R^n, c_k \in R^1
\]

\[
Y_{jk} \in \{ true, false \}
\]

**Objective Function**

**Common Constraints**

**Disjunction**

**Constraints**

**Fixed Charges**

**Logic Propositions**

**Continuous Variables**

**Boolean Variables**

**Relaxation of GDP?**
**Big-M MINLP (BM)**

- **MINLP reformulation of GDP**

\[
\begin{align*}
\min \ Z &= \sum_{k \in K} \sum_{j \in J_k} \gamma_{jk} \lambda_{jk} + f(x) \\
\text{s.t.} \quad r(x) &\leq 0 \\
g_{jk}(x) &\leq M_{jk}(1 - \lambda_{jk}) \quad \forall j \in J_k, k \in K \\
\sum_{j \in J_k} \lambda_{jk} &= 1, \quad k \in K \\
A\lambda &\leq a \\
x &\geq 0, \quad \lambda_{jk} \in \{0,1\}
\end{align*}
\]

**Big-M Parameter**

**Logic constraints**

*Williams (1990)*

**NLP Relaxation**

\[0 \leq \lambda_{jk} \leq 1 \quad \Rightarrow \quad \text{Lower bound to optimum of GDP}\]
Hull Relaxation Formulation

- Consider Disjunction \( k \in K \)

\[
\bigvee_{j \in J_k} Y_{jk} g_{jk}(x) \leq 0
\]

\[
c = \gamma_{jk}
\]

- **Theorem:** Convex Hull of Disjunction \( k \)  
  (Lee, Grossmann, 2000)
  
  - Disaggregated variables \( \nu^{jk} \)

\[
\{(x, c) | x = \sum_{j \in J_k} \nu^{jk}, c = \sum_{j \in J} \lambda_{jk} \gamma_{jk},
\]

\[
0 \leq \nu^{jk} \leq \lambda_{jk} U_{jk}, j \in J_k
\]

\[
\sum_{j \in J_k} \lambda_{jk} = 1, \quad 0 \leq \lambda_{jk} \leq 1,
\]

\[
\lambda_{jk} g_{jk}(\nu^{jk} / \lambda_{jk}) \leq 0, j \in J_k \}
\]

- \( \lambda_j \) - weights for linear combination
  - Generalization of Balas (1979)
  - Stubbs and Mehrotra (1999)

**Hull relaxation:** intersection of convex hull of each disjunction
Remarks

1. Perspective function \( h(\nu, \lambda) = \lambda g(\nu / \lambda) \)

   If \( g(x) \) is a bounded convex function, \( h(\nu, \lambda) \) is a bounded convex function

   \[ h(\nu,0) = 0 \] for bounded \( g(x) \)

2. Replace \( \lambda_{jk} g_{jk}(\nu_{jk} / \lambda_{jk}) \leq 0 \) where \( 0 \leq \nu_{jk} \leq U \lambda_{jk} \) by:

   \[
   ((1 - \varepsilon)\lambda_{jk} + \varepsilon)(g_{jk}(\nu_{jk} / ((1 - \varepsilon)\lambda_{jk} + \varepsilon))) - \varepsilon g_{jk}(0)(1 - \lambda_{jk}) \leq 0
   \]

   Furman, Sawaya & Grossmann (2009)

   a. Exact approximation of the original constraints as \( \varepsilon \to 0 \).

   b. The constraints are exact at \( \lambda_{jk} = 0 \) and at \( \lambda_{jk} = 1 \) regardless of value of \( \varepsilon \).

   \[ \text{if } \lambda_{jk} = 0, \implies (\varepsilon)(g_{jk}(0)) - \varepsilon g_{jk}(0) = 0 \leq 0 \]

   \[ \text{if } \lambda_{jk} = 1, \implies ((1)(g_{jk}(\nu_{jk} / (1)) - \varepsilon g_{jk}(0)(0) = (1)g_{jk}(\nu_{jk} / (1)) \leq 0 \]

   c. The LHS of the new constraint is \textbf{convex}.
Hull Relaxation Problem (HRP)

**HRP:**

$$\text{min } Z = \sum_{k \in K} \sum_{j \in J_k} \gamma_{jk} \lambda_{jk} + f(x)$$

s.t. $$\quad r(x) \leq 0$$

$$x = \sum_{j \in J_k} \nu_{jk}, \ k \in K$$

$$0 \leq \nu_{jk} \leq \lambda_{jk} U_{jk}, \ j \in J_k, \ k \in K$$

$$\sum_{j \in J_k} \lambda_{jk} = 1, \ k \in K$$

$$\lambda_{jk} g_{jk}(\nu_{jk} / \lambda_{jk}) \leq 0, \ j \in J_k, \ k \in K$$

$$A\lambda \leq a$$

$$x, \nu_{jk} \geq 0, \ 0 \leq \lambda_{jk} \leq 1, \ j \in J_k, \ k \in K$$

*Property:* The NLP (HRP) yields a lower bound to optimum of (GDP).
**Strength Lower Bounds**

- **Theorem:** *The relaxation of (HRP) yields a lower bound that is greater than or equal to the lower bound that is obtained from the relaxation of problem (BM)*
  

Convex hull relaxation

Big-M relaxation

Convex Hull of a set of disjunctions is smallest convex set that includes set of disjunctions. Projected relaxation of (CH) onto the space of (BM) is as tight or tighter than that of (BM)
Methods Generalized Disjunctive Programming

- Logic based methods
  - Branch and bound
    (Lee & Grossmann, 2000)
  - Decomposition
    Outer-Approximation
    Generalized Benders
    (Turkay & Grossmann, 1997)

- Reformulation MINLP
  - Branch and Bound
  - Outer-Approximation
  - Generalized Benders
  - Extended Cutting Plane

- Hull relaxation
- Big-M relaxation
GDP Example

\[
\min Z = (x_1 - 3)^2 + (x_2 - 2)^2 + c
\]

s.t.
\[
\begin{bmatrix}
Y_1 \\
(x_1)^2 + (x_2)^2 - 1 \leq 0 \\
c = 2
\end{bmatrix}
\lor
\begin{bmatrix}
Y_2 \\
(x_1 - 4)^2 + (x_2 - 1)^2 - 1 \leq 0 \\
c = 1
\end{bmatrix}
\lor
\begin{bmatrix}
Y_3 \\
(x_1 - 2)^2 + (x_2 - 4)^2 - 1 \leq 0 \\
c = 3
\end{bmatrix}
\]

\[Y_1 \lor Y_2 \lor Y_3 = True\]

\[0 \leq x_1, x_2 \leq 8, c \geq 0, \quad Y_j \in \{true, false\}, j = 1,2,3.\]
GDP Example

- Find \( x \geq 0, (x \in S_1) \lor (x \in S_2) \lor (x \in S_3) \)

to minimize \( Z = (x_1 - 3)^2 + (x_2 - 2)^2 + c \)
Example: convex hull

Convex hull = \text{conv}(U S_j)
Example: CRP solution

Convex hull = \text{conv}(\bigcup S_j)

Convex combination of $z_j$

$z_j = \nu / \lambda_j$

Local solution point

Convex hull optimum, $Z^L = 1.154$ Lower Bound

$x^L = (3.159, 1.797)$

Infeasible for GDP

$\lambda_1 = 0.016$

$\lambda_2 = 0.955$

$\lambda_3 = 0.029$
Example: branch and bound

First Node: $S_2$
Optimal solution: $Z^U = 1.172$

Optimal Solution
(3.293, 1.707)
$Z^* = 1.172$
Example: branch and bound

Second Node: \( \text{conv}(S_1 \cup S_3) \)
Optimal solution: \( Z^L = 3.327 \)

Lower Bound \( Z^L = 3.327 \)
Upper Bound \( Z^U = 1.172 \)
Example: Search Tree

- **Branching Rule**: $\lambda_j$ - the “weight” of disjunction
  - Fix $Y_j$ as true: fix $\lambda_j$ as 1.
Process Network with Fixed Charges

- Türkay and Grossmann (1997)

8 Boolean variables, 25 continuous, 31 constraints (8 disjunctions, 5 nonlinear)
Optimal solution

- Minimum Cost: $68.01M/year
MINLP - Branch and Bound Method

**Hull-Rel**

\[ Z^L = 62.48 \]
\[ \lambda = [0.31, 0.69, 0.03, 1.0, 1, 0, 1] \]

Fix \( \lambda_2 = 0 \)

\[ Z^L = 65.92 \]
\[ \lambda = [0.1, 0.022, 1.0, 1, 0, 1] \]

Fix \( \lambda_3 = 0 \)

Fix \( \lambda_2 = 1 \)

\[ Z^L = 75.01 > Z^U \]

\[ \lambda = [1.0, 0.022, 1.0, 1, 0, 1] \]

Feasible Solution

**Optimal Solution**

\[ Z^U = 71.79 \]
\[ \lambda = [0.1, 1.0, 1, 0, 1, 0, 1] \]

\[ Z^U = 68.01 = Z^* \]
\[ \lambda = [0.1, 0.0, 1.0, 1, 0, 1] \]

\[ Z^L = 15.08 \]

**Big-M**

\[ Y_4 = 0 \]
\[ Y_4 = 1 \]
\[ Y_6 = 0 \]
\[ Y_6 = 1 \]
\[ Y_8 = 0 \]
\[ Y_8 = 1 \]
\[ Y_2 = 0 \]
\[ Y_2 = 1 \]
\[ Y_1 = 0 \]
\[ Y_1 = 1 \]
\[ Y_3 = 0 \]
\[ Y_3 = 1 \]

\[ Y_{10*} \]

\[ 5 \text{ nodes vs. 17 nodes of Big-M (lower bound = 15.08)} \]
Question

Can we obtain stronger relaxations than with Hull-Relaxation?

Extend Disjunctive Programming Theory to Nonlinear Convex Sets

DP: Linear programming with disjunctions

Equivalence between GDP and DP

Sawaya (2007)

\begin{align*}
\text{GDP} & \quad \min \quad Z = f(x) + \sum_{k \in K} c_k \\
\text{s.t.} & \quad r(x) \leq 0 \\
& \quad \bigvee_{i \in D_k} \begin{cases} 
Y_{ik} \\
\frac{g_{ik}(x)}{c_k} \leq 0 \\
Y_{ik} = \gamma_{ik} \\
\Omega(Y) = True
\end{cases} \\
& \quad x_{lo} \leq x \leq x_{up} \\
& \quad x \in \mathbb{R}^n, \ c_k \in \mathbb{R}_1, \ Y_{ik} \in \{True, False\}
\end{align*}

\begin{align*}
\text{DP} & \quad \min \quad Z = f(x) + \sum_{k \in K} c_k \\
\text{s.t.} & \quad r(x) \leq 0 \\
\text{subject to} & \quad \bigvee_{i \in D_k} \begin{cases} 
\lambda_{ik} = 1 \\
\frac{g_{ik}(x)}{c_k} \leq 0 \\
\gamma_{ik} \\
\Omega(Y) = True \\
A \lambda \geq a \\
\sum_{i \in D} \lambda_i = 1, \\
x_{lo} \leq x \leq x_{up} \\
x \in \mathbb{R}^n, \ c_k \in \mathbb{R}_1, \ \lambda_{ik} \geq 0
\end{cases} \\
& \quad x \in \mathbb{R}^n, \ c_k \in \mathbb{R}_1, \ \lambda_{ik} \geq 0
\end{align*}

The integrality of \( \lambda \) is guaranteed

Proposition: Discrete/continuous GDP and continuous DP have equivalent solutions.
Regular Form: Form represented by the intersection of the union of convex sets

\[ F = \bigcap_{k \in K} S_k, \quad k \in K, \quad S_k = \bigcup_{i \in D_k} P_i \]

\( P_i \) a convex set for \( i \in D_k \)

\[ F \text{ is in regular form} \]

**Theorem 2.1.** Let \( F \) be a disjunctive set in regular form. Then \( F \) can be brought to DNF by \( |K| - 1 \) recursive applications of the following basic step which preserves regularity:

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

\[ S_{rs} = \bigcup_{i \in D_r, j \in D_s} (P_i \cap P_j) \]

Balas (1985)
Illustrative Example: Basic Steps

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

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

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

\[ F = S_{12} \cap S_3 \quad \text{as} \quad F = S_{123} \quad \text{which is its equivalent DNF} \]
Theorem 2.4. For \( i = 1, 2, \ldots, k \) let \( F_i = \bigcap_{k \in K} S_k \) be a sequence of regular forms of a disjunctive set such that \( F_i \) is obtained from \( F_{i-1} \) by the application of a basic step, then:

\[
h-\text{rel}(F_i) \subseteq h-\text{rel}(F_{i-1})
\]

Illustration: \( F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22}) \)
Theorem 2.4. For $i = 1, 2, \ldots, k$ let $F_i = \bigcap_{k \in K} S_k$ be a sequence of regular forms of a disjunctive set such that $F_i$ is obtained from $F_{i-1}$ by the application of a basic step, then:

$$h\text{-}rel(F_i) \subseteq h\text{-}rel(F_{i-1})$$

Illustration: $F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22})$

$clconv(P_{11} \cup P_{12})$

$clconv(P_{21} \cup P_{22})$
Hierarchy of Relaxations for Convex Disjunctive Programs

**Theorem 2.4.** For \( i = 1, 2, \ldots, k \) let \( F_i = \bigcap_{k \in K} S_k \) be a sequence of regular forms of a disjunctive set such that \( F_i \) is obtained from \( F_{i-1} \) by the application of a basic step, then:

\[
h\text{-rel}(F_i) \subseteq h\text{-rel}(F_{i-1})
\]

**Illustration:** \( F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22}) \)

No Basic Step Applied => HR
Theorem 2.4. For \( i = 1, 2, \ldots, k \) let \( F_i = \bigcap_{k \in K} S_k \) be a sequence of regular forms of a disjunctive set such that \( F_i \) is obtained from \( F_{i-1} \) by the application of a basic step, then: 

\[
h\text{-rel}(F_i) \subseteq h\text{-rel}(F_{i-1})
\]

Illustration: 

\[
F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22}) \quad F_1 = (P_{11} \cap P_{21}) \cup (P_{11} \cap P_{22}) \cup (P_{12} \cap P_{21}) \cup (P_{12} \cap P_{22})
\]

No Basic Step Applied => HR

Basic Step Applied
Hierarchy of Relaxations for Convex Disjunctive Programs

Theorem 2.4. For $i = 1, 2, \ldots, k$ let $F_i = \bigcap_{k \in K} S_k$ be a sequence of regular forms of a disjunctive set such that $F_i$ is obtained from $F_{i-1}$ by the application of a basic step, then:

$$h\text{-rel}(F_i) \subseteq h\text{-rel}(F_{i-1})$$

**Illustration:**

$F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22})$

$F_1 = (P_{11} \cap P_{21}) \cup (P_{11} \cap P_{22}) \cup (P_{12} \cap P_{21}) \cup (P_{21} \cap P_{22})$

No Basic Step Applied $\Rightarrow$ HR

Basic Step Applied $\Rightarrow$ CH

Tighter relaxation!
Convex nonlinear program equivalent to a convex disjunctive program

Theorem 2.8. Let $Z = \min \{ f(x) | x \in S \}$ be a convex disjunctive program where $S$ is a convex disjunctive set in DNF such that $S = \bigcup_{i \in D} P_i$ and $P_i = \{ x \in \mathbb{R}^n, g_i(x) \leq 0 \}$ where $P_i \neq \emptyset$ and that $x$ and $f(x)$ are bounded below and above by a large number $L$. Then, the following nonlinear program has at least one solution that is also solution of the disjunctive program:

Objective as constraint

**NLPDP:**

$$
\begin{align*}
\min & \quad \alpha \\
\text{s.t.} & \quad \alpha = \sum_{i \in D} \nu^i \\
& \quad x = \sum_{i \in D} \nu^i \\
& \quad \lambda_i g_i(\nu^i / \lambda_i) \leq 0, \quad i \in D \\
& \quad \lambda_i f(\nu^i / \lambda_i) \leq \nu^i, \quad i \in D \\
& \quad \sum_{i \in D} \lambda_i = 1, \quad \lambda_i \geq 0, \quad i \in D \\
& \quad |\nu^i| \leq L \lambda_i, \quad i \in D \\
& \quad |\nu^i_\alpha| \leq L \lambda_i, \quad i \in D
\end{align*}
$$

The solution of the NLP relaxation leads to the solution of the DP!

Generalizes convexification results of MILPs

Lovazc & Schrijver (1989), Sherali & Adams (1990), Balas, Ceria, Cornuejols (1993)
Convex nonlinear program equivalent to a convex disjunctive program

**Illustrative Example**

**Disjunctive Program**

\[
\begin{align*}
\text{min} & \quad Z = (x_1 - 3)^2 + (x_2 - 2)^2 + 1 \\
\text{s.t.} & \quad [x_1^2 + x_2^2 \leq 1] \lor [(x_1 - 4)^2 + (x_2 - 1)^2 \leq 1] \\
& \lor [(x_1 - 2)^2 + (x_2 - 4)^2 \leq 1] \\
& |x_i| \leq 5 \quad i \in 1, 2
\end{align*}
\]

Solution of the relaxed program is different from solution of the disjunctive program
Convex nonlinear program equivalent to a convex disjunctive program

Illustrative Example

Disjunctive Program
Place objective as constraint and intersect with disjunction

\[
\begin{align*}
\min & \quad Z = \alpha \\
n & \quad s.t.
\left[
\begin{array}{c}
\frac{x_1^2 + x_2^2}{(x_1 - 3)^2 + (x_2 - 2)^2 + 1} \leq 1 \\
\frac{(x_1 - 4)^2 + (x_2 - 1)^2}{(x_1 - 3)^2 + (x_2 - 2)^2 + 1} \leq 1 \\
\frac{(x_1 - 2)^2 + (x_2 - 4)^2}{(x_1 - 3)^2 + (x_2 - 2)^2 + 1} \leq 1
\end{array}
\right]
\end{align*}
\]

\[
|x_i| \leq 5 \quad i \in 1, 2
\]

Solution of the program and its relaxation

\[
Z^* = 1.172
\]

(3.293,1.707)

Solution of the hull relaxation of DNF is the same as the solution of the disjunctive program (Theorem 2.8)
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.
MINLP formulation of convex disjunctive program after several basic steps

\[ \min \quad f(x) \]
\[ \text{s.t.} \]
\[ x = \sum_{t \in Q_j} \nu^t, \quad j \in T_n \]
\[ \lambda_i G^t(\nu^t / \lambda_t) \leq 0, \quad t \in Q_j, \quad j \in T_n \]
\[ \sum_{t \in Q_j} \lambda_t = 1, \quad \lambda_t \geq 0, \quad t \in Q_j, \quad j \in T_n \]
\[ \sum_{t \in Q_j | s \in M^t_j} \lambda_t = \delta^r_s, \quad s \in Q_r, \quad r \in T_d, \quad j \in T_n \]
\[ \sum_{s \in Q_r} \delta^r_s = 1, \quad r \in T_d \]
\[ |\nu^t| \leq L \lambda_t, \quad \delta^r_s \in \{0, 1\}, \quad t \in Q_j, \quad j \in T_n \]
\[ t \in Q_j, \quad j \in T_n \]

No new binary variables are created (Balas, 1985)

No additional 0-1 variables are required!
We can obtain a tighter relaxation by applying basic steps between the improper disjunctions and the proper disjunctions.

Optimal Solution $Z_{rel} = 68.0097$ obtained from Hull Relaxation with basic steps.

Solves as an NLP!
# Sizes of Convex GDP Formulations

<table>
<thead>
<tr>
<th>Example</th>
<th>BM Approach</th>
<th>HR Approach</th>
<th>Proposed Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bin</td>
<td>Con</td>
<td>Const</td>
</tr>
<tr>
<td>Circles2D3</td>
<td>3</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Circles2D36</td>
<td>36</td>
<td>39</td>
<td>38</td>
</tr>
<tr>
<td>Circles3D36</td>
<td>36</td>
<td>40</td>
<td>38</td>
</tr>
<tr>
<td>Proc8</td>
<td>8</td>
<td>42</td>
<td>97</td>
</tr>
<tr>
<td>Proc10</td>
<td>10</td>
<td>51</td>
<td>98</td>
</tr>
<tr>
<td>Proc12</td>
<td>12</td>
<td>57</td>
<td>114</td>
</tr>
<tr>
<td>Flay02</td>
<td>4</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>Flay03</td>
<td>12</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td>Flay04</td>
<td>24</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>Clay0203</td>
<td>18</td>
<td>31</td>
<td>55</td>
</tr>
<tr>
<td>Clay0303</td>
<td>21</td>
<td>34</td>
<td>67</td>
</tr>
<tr>
<td>Clay0204</td>
<td>32</td>
<td>53</td>
<td>91</td>
</tr>
</tbody>
</table>
Numerical Results

All problems were solved using NLP branch-and-bound SBB/CONOPT 3.14 (GAMS)

Table: Performance using different reformulation strategies

<table>
<thead>
<tr>
<th>Example</th>
<th>Opt.</th>
<th>BM Approach</th>
<th>LB</th>
<th>Nds</th>
<th>T(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circles2D3</td>
<td>1.17</td>
<td>LB</td>
<td>0.00</td>
<td>4</td>
<td>1.0</td>
</tr>
<tr>
<td>Circles2D36</td>
<td>2.25</td>
<td>LB</td>
<td>0.00</td>
<td>70</td>
<td>7.8</td>
</tr>
<tr>
<td>Circles3D36</td>
<td>15.77</td>
<td>LB</td>
<td>0.44</td>
<td>70</td>
<td>7.3</td>
</tr>
<tr>
<td>Proc8</td>
<td>68.01</td>
<td>LB</td>
<td>-829.0</td>
<td>34</td>
<td>4.6</td>
</tr>
<tr>
<td>Proc10</td>
<td>-73.51</td>
<td>LB</td>
<td>-1,108.88</td>
<td>197</td>
<td>21.7</td>
</tr>
<tr>
<td>Proc12</td>
<td>-69.51</td>
<td>LB</td>
<td>-1,108.88</td>
<td>234</td>
<td>27.7</td>
</tr>
<tr>
<td>Flay02</td>
<td>37.95</td>
<td>LB</td>
<td>28.28</td>
<td>6</td>
<td>1.0</td>
</tr>
<tr>
<td>Flay03</td>
<td>48.99</td>
<td>LB</td>
<td>30.98</td>
<td>104</td>
<td>10.7</td>
</tr>
<tr>
<td>Flay04</td>
<td>54.40</td>
<td>LB</td>
<td>30.98</td>
<td>2,415</td>
<td>234.0</td>
</tr>
<tr>
<td>Clay0203</td>
<td>41,573.30</td>
<td>LB</td>
<td>0.00</td>
<td>323</td>
<td>32.7</td>
</tr>
<tr>
<td>Clay0303</td>
<td>26,670.00</td>
<td>LB</td>
<td>0.00</td>
<td>380</td>
<td>42.0</td>
</tr>
<tr>
<td>Clay0204</td>
<td>6,545.00</td>
<td>LB</td>
<td>0.00</td>
<td>2,265</td>
<td>229.0</td>
</tr>
</tbody>
</table>

Poor lower bounds
Numerical Results

All problems were solved using NLP branch-and-bound SBB/CONOPT 3.14 (GAMS)

Table: Performance using different reformulation strategies

<table>
<thead>
<tr>
<th>Example</th>
<th>Opt.</th>
<th>BM Approach</th>
<th>HR Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LB</td>
<td>Nds</td>
</tr>
<tr>
<td>Circles2D3</td>
<td>1.17</td>
<td>0.00</td>
<td>4</td>
</tr>
<tr>
<td>Circles2D36</td>
<td>2.25</td>
<td>0.00</td>
<td>70</td>
</tr>
<tr>
<td>Circles3D36</td>
<td>15.77</td>
<td>0.44</td>
<td>70</td>
</tr>
<tr>
<td>Proc8</td>
<td>68.01</td>
<td>-829.0</td>
<td>34</td>
</tr>
<tr>
<td>Proc10</td>
<td>-73.51</td>
<td>-1,108.88</td>
<td>197</td>
</tr>
<tr>
<td>Proc12</td>
<td>-69.51</td>
<td>-1,108.88</td>
<td>234</td>
</tr>
<tr>
<td>Flay02</td>
<td>37.95</td>
<td>28.28</td>
<td>6</td>
</tr>
<tr>
<td>Flay03</td>
<td>48.99</td>
<td>30.98</td>
<td>104</td>
</tr>
<tr>
<td>Flay04</td>
<td>54.40</td>
<td>30.98</td>
<td>2,415</td>
</tr>
<tr>
<td>Clay0203</td>
<td>41,573.30</td>
<td>0.00</td>
<td>323</td>
</tr>
<tr>
<td>Clay0303</td>
<td>26,670.00</td>
<td>0.00</td>
<td>380</td>
</tr>
<tr>
<td>Clay0204</td>
<td>6,545.00</td>
<td>0.00</td>
<td>2,265</td>
</tr>
</tbody>
</table>

- Improved lower bounds 50%probs
Numerical Results

All problems were solved using NLP branch-and-bound SBB/CONOPT 3.14 (GAMS)

Table: Performance using different reformulation strategies

<table>
<thead>
<tr>
<th>Example</th>
<th>BM Approach</th>
<th>HR Approach</th>
<th>Proposed Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Opt.</td>
<td>LB</td>
<td>Nds</td>
</tr>
<tr>
<td>Circles2D3</td>
<td>1.17</td>
<td>0.00</td>
<td>4</td>
</tr>
<tr>
<td>Circles2D36</td>
<td>2.25</td>
<td>0.00</td>
<td>70</td>
</tr>
<tr>
<td>Circles3D36</td>
<td>15.77</td>
<td>0.44</td>
<td>70</td>
</tr>
<tr>
<td>Proc8</td>
<td>68.01</td>
<td>-829.0</td>
<td>34</td>
</tr>
<tr>
<td>Proc10</td>
<td>-73.51</td>
<td>-1,108.88</td>
<td>197</td>
</tr>
<tr>
<td>Proc12</td>
<td>-69.51</td>
<td>-1,108.88</td>
<td>234</td>
</tr>
<tr>
<td>Flay02</td>
<td>37.95</td>
<td>28.28</td>
<td>6</td>
</tr>
<tr>
<td>Flay03</td>
<td>48.99</td>
<td>30.98</td>
<td>104</td>
</tr>
<tr>
<td>Flay04</td>
<td>54.40</td>
<td>30.98</td>
<td>2,415</td>
</tr>
<tr>
<td>Clay0203</td>
<td>41,573.30</td>
<td>0.00</td>
<td>323</td>
</tr>
<tr>
<td>Clay0303</td>
<td>26,670.00</td>
<td>0.00</td>
<td>380</td>
</tr>
<tr>
<td>Clay0204</td>
<td>6,545.00</td>
<td>0.00</td>
<td>2,265</td>
</tr>
</tbody>
</table>

Proposed vs BM: faster 10 out of 12
Proposed vs HR: faster 8 out of 12

Improved lower bounds 100%probs
Extension to Nonconvex GDP

Basic idea: strengthen lower bound of global optimum

Phase 1

Nonconvex GDP

Convex GDP

Initial lower bound

Phase 2

Tight Convex GDP

Stronger lower bound

Relaxation

Under/over estimating functions
Convex envelopes

Strengthen relaxation
Apply basic steps

Remarks
1. Since transformation to DNF impractical special rules are applied to identify promising basic steps
2. Stronger relaxation can also be used to infer tighter bounds for variables
Illustrative Example: Optimal reactor selection I

\[
\begin{align*}
F: \text{Flow} \\
X: \text{Conversion}
\end{align*}
\]

**Objective Function (Profit)**

\[
\min Z = -qFX + gF + CP
\]

**Demand constraint**

\[
FX \leq d
\]

**Conversion**

\[
F = \alpha_i X + \beta_i \quad X_{11}^{LO} \leq X \leq X_{11}^{UP} \quad CP = C_{p1}
\]

\[
F = \alpha_2 X + \beta_2 \quad X_{21}^{LO} \leq X \leq X_{21}^{UP} \quad CP = C_{p2}
\]

\[
Y_{I1} \lor Y_{21} = \text{True}
\]

\[
CP, X, F \in \mathbb{R} \\
F_{LO} \leq F \leq F_{UP} \\
Y_{I1}, Y_{21} \in \{\text{True, False}\}
\]

**Selection Reactor**

**Feasible Region**

**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

Min Z = \(-FX + \gamma F + CP\)

s.t. \(FR \leq d\)

\[ Y_{P} \leq F \cdot X_{LO} + F \cdot X_{LO} \cdot X - F \cdot X_{LO} \cdot X \]
\[ P \leq d \]
\[ P \leq F \cdot X_{LO} + F \cdot X_{LO} \cdot X - F \cdot X_{LO} \cdot X \]
\[ P \leq F \cdot X_{LO} + F \cdot X_{LO} \cdot X \]
\[ P \geq F \cdot X_{LO} + F \cdot X_{LO} \cdot X \]
\[ F \geq X_{LO} \cdot X \]
\[ X_{LO} \leq F \cdot X_{LO} + F \cdot X_{LO} \cdot X \]
\[ CP = C_{p_{1}} \]

\[ Y_{11} \vee Y_{21} = \text{True} \]

\[ CP, X_{F} \in R \]
\[ F_{LO} \leq F \leq F_{UP} \]
\[ Y_{11}, Y_{21} \in \{ \text{True, False} \} \]

Lower bound \(Z^{*} = -1.1 < -1.01\) and tighter than \(-1.28!\)
Illustrative Example: Optimal reactor selection I
Comparison of Relaxations

The application of \textit{basic step} prior to the discrete relaxation leads to a \textit{tighter} relaxed feasible region $\Rightarrow$ \textit{stronger lower bounds}
Dimensions of Test Problems
Bilinear/Concave

<table>
<thead>
<tr>
<th></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)

Strong linear relaxations exist for bilinear and concave functions
Dimension of Case Studies
Linear Fractional, Posynomial, Exponential

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PROC1</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>PROC2</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>RXN1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>RXN2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>HEN1</td>
<td>18</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>21</td>
</tr>
</tbody>
</table>

Reference

PROC1, PROC2 : Optimal Process Network Problem
RXN1, RXN2 : Optimal Reactor Network Problem
HEN1 : Optimal Heat Exchanger Network Problem

Strong nonlinear relaxations exist for linear fractional and posynomial functions
Heat Exchanger Network Problem

Heat Exchanger Network

Generalized Disjunctive Program

\[
\min Z = c_1 A_1 + c_2 A_2 + c_3 A_3 + c_4 A_4
\]

s.t.

\[
A_1 = \frac{Q_1}{U_1 \Delta T_1}, \quad A_2 = \frac{Q_2}{U_2 \Delta T_2},
\]

\[
A_3 = \begin{cases}
\frac{Q_3}{U_3 \Delta T_3}, & \gamma_3 = 0 \\
0, & \gamma_3 = 0
\end{cases}
\]

\[
A_4 = \begin{cases}
\frac{Q_4}{U_4 \Delta T_4}, & \gamma_4 = 0 \\
0, & \gamma_4 = 0
\end{cases}
\]

\[
Q_1 = FCP_{H1}(T_1 - T_{H1, out}),
Q_2 = FCP_{H2}(T_2 - T_{H2, out}),
Q_3 = FCP_{C2}(T_3 - T_{C2, in}),
Q_4 = FCP_{H2}(T_{H2, in} - T_2),
\]

\[
T_1 \geq T_{C1, in} + EMAT, \quad T_2 \geq T_{C1, in} + EMAT
\]

\[
Q_1 + Q_2 = Q_{total}
\]

\[
\Delta T_1 = \frac{(T_1 - T_{C1, out}) + (T_{H1, in} - T_{C1, in})}{2}, \quad \Delta T_2 = \frac{(T_2 - T_{C1, out}) + (T_{H2, in} - T_{C1, in})}{2}
\]

\[
\Delta T_3 = \frac{(T_3 - T_{C2, in}) + (T_{H1, out} - T_3)}{2}, \quad \Delta T_4 = \frac{(T_4 - T_{C3, in}) + (T_{H2, out} - T_4)}{2}
\]

\[
T_{H1, out} \leq T_1 \leq T_{H1, in}, \ T_{H2, out} \leq T_2 \leq T_{H2, in}
\]

\[
T_{C2, in} \leq T_3, \ T_{C3, in} \leq T_4
\]

\[
Q_i \geq 0, \Delta T_i \geq EMAT \ I = 1, \ldots, 4
\]
Prediction of Lower Bounds Global Optimum

<table>
<thead>
<tr>
<th></th>
<th>Global Optimum</th>
<th>Lower Bound Hull Relaxation</th>
</tr>
</thead>
<tbody>
<tr>
<td>React 1</td>
<td>-1.01</td>
<td>-1.28</td>
</tr>
<tr>
<td>React 2</td>
<td>6.31</td>
<td>5.65</td>
</tr>
<tr>
<td>HEN</td>
<td>114384.78</td>
<td>91671.18</td>
</tr>
<tr>
<td>Water</td>
<td>1214.87</td>
<td>400.66</td>
</tr>
<tr>
<td>Pool</td>
<td>-4640</td>
<td>-5515</td>
</tr>
<tr>
<td>Process 1</td>
<td>18.61</td>
<td>11.85</td>
</tr>
<tr>
<td>Process 2</td>
<td>19.48</td>
<td>12.38</td>
</tr>
<tr>
<td>RXN 1</td>
<td>42.89</td>
<td>-337.5</td>
</tr>
<tr>
<td>RXN 2</td>
<td>76.47</td>
<td>22.5</td>
</tr>
<tr>
<td>HEN 1</td>
<td>48531</td>
<td>38729.3</td>
</tr>
</tbody>
</table>

Lower bounds improved in all cases  Ave. increase 22%

8 out of 10 achieved theoretically best lower bound (DNF)!
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
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>Global Optimization Technique using Hull 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>
## Computational Performance - Nonlinear

<table>
<thead>
<tr>
<th></th>
<th>Global Optimization Technique using Lee &amp; Grossmann Relaxation</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Global Optimum</td>
<td>Nodes</td>
<td>Bound contract. (%) Avg</td>
</tr>
<tr>
<td><strong>PROC1</strong></td>
<td>18.61</td>
<td>3</td>
<td>51.3</td>
</tr>
<tr>
<td><strong>PROC2</strong></td>
<td>19.48</td>
<td>2</td>
<td>40.5</td>
</tr>
<tr>
<td><strong>RXN1</strong></td>
<td>42.89</td>
<td>2</td>
<td>51.0</td>
</tr>
<tr>
<td><strong>RXN2</strong></td>
<td>76.46</td>
<td>2</td>
<td>51.0</td>
</tr>
<tr>
<td><strong>HEN1</strong></td>
<td>48531</td>
<td>3</td>
<td>13.8</td>
</tr>
</tbody>
</table>

**Remarks**

- Proposed relaxation led to a **significant bound contraction** at the root node.
- The **reduced number of nodes** is a further indication of tighter relaxation
- **Modest savings** compared to bilinear/concave due to small size
Conclusions

- Proposed an extension of disjunctive programming theory to nonlinear convex sets that yields hierarchy of relaxations (concept basic steps)

- Tightest of these relaxations allows in theory the solution of the DP as an NLP

- Applied the proposed framework to several instance obtaining significant improvements in the performance (tighter lower bounds)

- Proposed framework can be applied to nonconvex GDP problems yielding tighter lower bounds on global optimum (bilinear, concave, linear fractional) and can be extended to nonlinear convex envelopes