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?

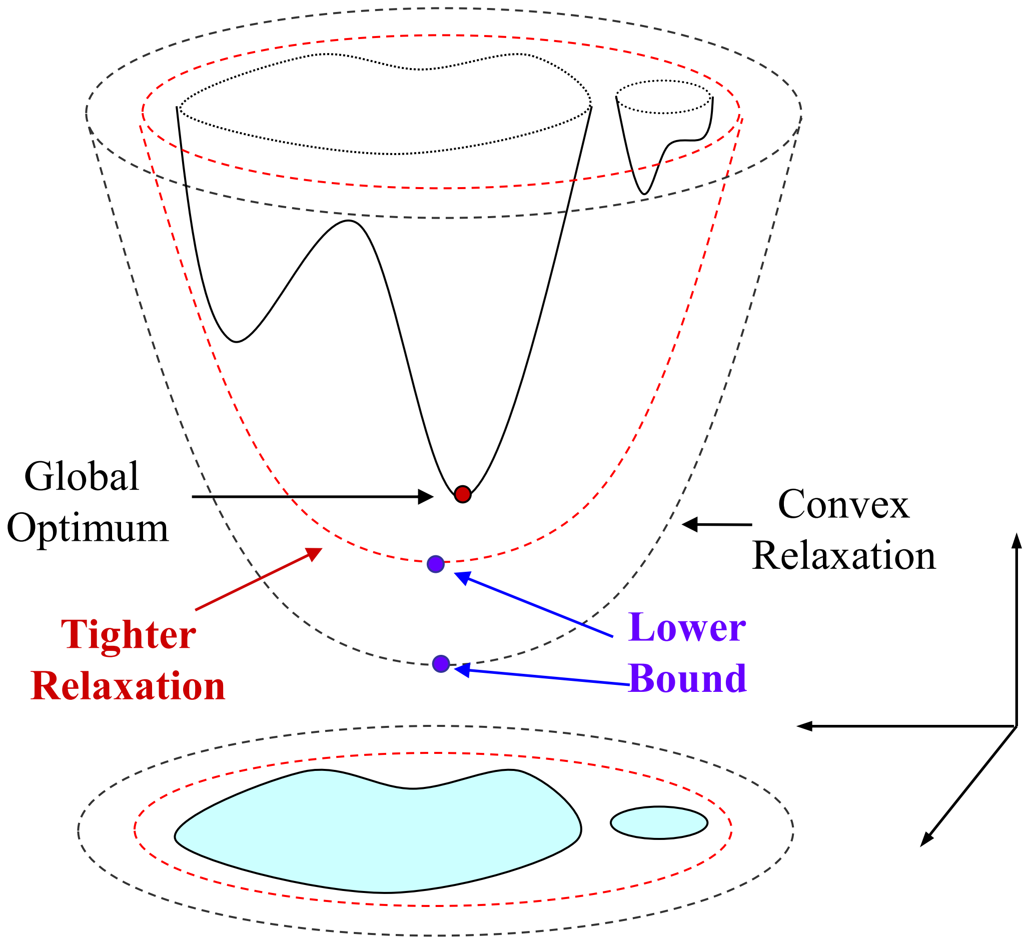
Outline

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 of MINLP



- **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**

$$\min Z = f(x, y)$$ *Objective Function*

$$s.t. \quad g(x, y) \le 0$$ *Inequality Constraints*

$$x \in X,\ y \in Y$$

$$X = \{x \mid x \in R^n, x^L \le x \le x^U, Bx \le b\}$$

$$Y = \{y \mid y \in \{0,1\}^m, Ay \le a\}$$

- ***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)***

Outer-Approximation (OA) ***Duran and Grossmann (1986), Fletcher and Leyffer (1994)***

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**

$$\min Z^L = f(x,y)$$

$$s.t. \quad g(x,y) \le 0$$

$$x \in X,\ y \in Y$$

$$X = \{x \mid x \in R^n, x^L \le x \le x^U, Bx \le b\}$$

$$Y = \{y \mid 0 \le y \le 1, Ay \le a\}$$

Z^L = *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*

$$\min \; Z = \sum_k c_k + f(x)$$ **Objective Function**

$$s.t. \quad r(x) \le 0$$ **Common Constraints**

OR operator →

$$\bigvee_{j \in J_k} \begin{bmatrix} Y_{jk} \\ g_{jk}(x) \le 0 \\ c_k = \gamma_{jk} \end{bmatrix} \quad k \in K$$

Disjunction

Constraints

Fixed Charges

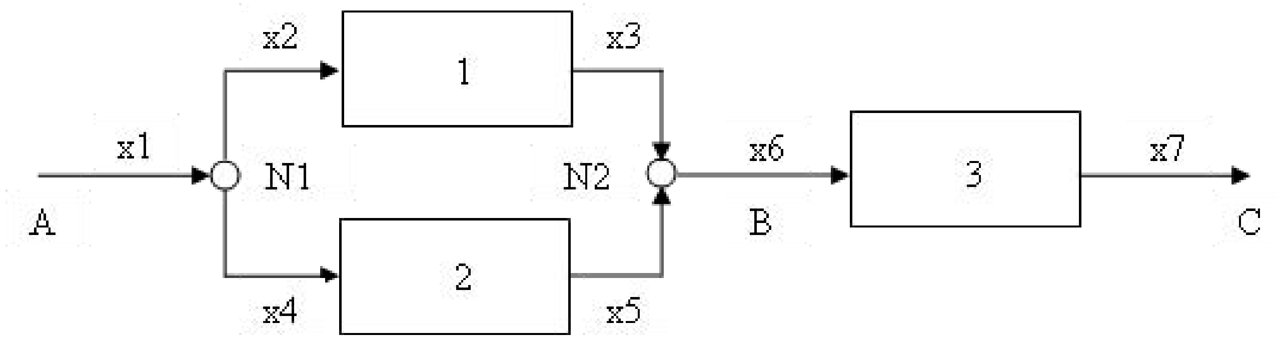
$$\Omega(Y) = true$$ **Logic Propositions**

$$x \in R^n, c_k \in R^1$$ **Continuous Variables**

$$Y_{jk} \in \{ true, false \}$$ **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

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

$$s.t.$$

$$x_1 = x_2 + x_4$$

$$x_6 = x_3 + x_5$$

$$\begin{bmatrix} Y_{11} \\ x_3 = p_1 x_2 \\ c_1 = \gamma_1 \end{bmatrix} \vee \begin{bmatrix} Y_{21} \\ x_3 = x_2 = 0 \\ c_1 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_{12} \\ x_5 = p_2 x_4 \\ c_2 = \gamma_2 \end{bmatrix} \vee \begin{bmatrix} Y_{22} \\ x_5 = x_4 = 0 \\ c_2 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_{13} \\ x_7 = p_3 x_6 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} Y_{23} \\ x_7 = x_6 = 0 \\ c_3 = 0 \end{bmatrix}$$

$$Y_{11} \veebar Y_{21}$$

$$Y_{12} \veebar Y_{22}$$

$$Y_{13} \veebar Y_{23}$$

$$Y_{11} \vee Y_{12} \Rightarrow Y_{13}$$

$$Y_{13} \Rightarrow Y_{11} \vee Y_{12}$$

$$Y_{21} \vee Y_{22}$$

$$0 \le x \le x^U$$

$$Y_{11}, Y_{21}, Y_{12}, Y_{22}, Y_{13}, Y_{23} \in \{True, False\}$$

$$c_1, c_2, c_3 \in \mathbf{R}^1$$

Generalized Disjunctive Programming (GDP)

- Raman and Grossmann (1994)

$$\min \; Z = \sum_k c_k + f(x)$$ **Objective Function**

$$s.t. \quad r(x) \le 0$$ **Common Constraints**

$$\bigvee_{j \in J_k} \begin{bmatrix} Y_{jk} \\ g_{jk}(x) \le 0 \\ c_k = \gamma_{jk} \end{bmatrix} \quad k \in K$$

Disjunction

Constraints

Fixed Charges

$$\Omega(Y) = \; true$$ **Logic Propositions**

$$x \in R^n, c_k \in R^1$$ **Continuous Variables**

$$Y_{jk} \in \{ \, true, \; false \, \}$$ **Boolean Variables**

Relaxation of GDP?

Big-M MINLP (BM)

- **MINLP reformulation of GDP**

$$
\begin{aligned}
\min Z \;&= \sum_{k\in K}\sum_{j\in J_k}\gamma_{jk}\lambda_{jk} + f(x)\\
s.t. \quad & r(x) \le 0\\
& g_{jk}(x) \le M_{jk}(1-\lambda_{jk}) \quad ,j\in J_k, k\in K\\
& \sum_{j\in J_k}\lambda_{jk} = 1,\;\; k\in K\\
& A\lambda \le a\\
& x \ge 0, \lambda_{jk}\in\{0,1\}
\end{aligned}
$$

Big-M Parameter

Logic constraints

Williams (1990)

NLP Relaxation $0 \le \lambda_{jk} \le 1$ **=> Lower bound to optimum of GDP**

Hull Relaxation Formulation

- Consider **Disjunction** $k \in K$

$$\bigvee_{j \in J_k} \begin{bmatrix} Y_{jk} \\ g_{jk}(x) \le 0 \\ c = \gamma_{jk} \end{bmatrix}$$

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

$$\{(x,c) \mid x = \sum_{j \in J_k} v^{jk}, \quad c = \sum_{j \in J} \lambda_{jk} \gamma_{jk},$$

$$0 \le v^{jk} \le \lambda_{jk} U^{k}_{jk},\ j \in J_k$$

=> Convex Constraints

$$\sum_{j \in J_k} \lambda_{jk} = 1, \quad 0 \le \lambda_{jk} \le 1,$$

$$\lambda_{jk} g_{jk}(v^{jk} / \lambda_{jk}) \le 0,\ j \in J_k \}$$

    - λ_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 $h(v, \lambda) = \lambda g(v/\lambda)$

   If g(x) is a bounded convex function,

   $h(v, \lambda)$ is a bounded convex function

   Hiriart-Urruty and Lemaréchal (1993)

   $h(v, 0) = 0$ for bounded g(x)

2. Replace $\lambda_{jk} g_{jk}(v_{jk}/\lambda_{jk}) \le 0$ where $0 \le v_{jk} \le U\lambda_{jk}$ by:

$$\boxed{((1-\varepsilon)\lambda_{jk} + \varepsilon)(g_{jk}(v_{jk}/((1-\varepsilon)\lambda_{jk} + \varepsilon))) - \varepsilon g_{jk}(0)(1-\lambda_{jk}) \le 0}$$

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 ε.

$$if\ \lambda_{jk} = 0, \Rightarrow (\varepsilon)(g_{jk}(0)) - \varepsilon g_{jk}(0) = 0 \le 0$$

$$if\ \lambda_{jk} = 1, \Rightarrow ((1)(g_{jk}(v_{jk}/(1)) - \varepsilon g_{jk}(0)(0) = (1) g_{jk}(v_{jk}/(1)) \le 0$$

c. The LHS of the new constraint is **convex**.

Hull Relaxation Problem (HRP)

HRP:

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

$$s.t. \qquad r(x)\le 0$$

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

$$0 \le \nu^{jk} \le \lambda_{jk} U_{jk}, j\in J_k, k \in K$$

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

$$\lambda_{jk} g_{jk}(\nu^{jk}/\lambda_{jk}) \le 0, \quad j\in J_k, k \in K$$

Convex Hull each disjunction

$$A\lambda \le a$$

Logic constraints

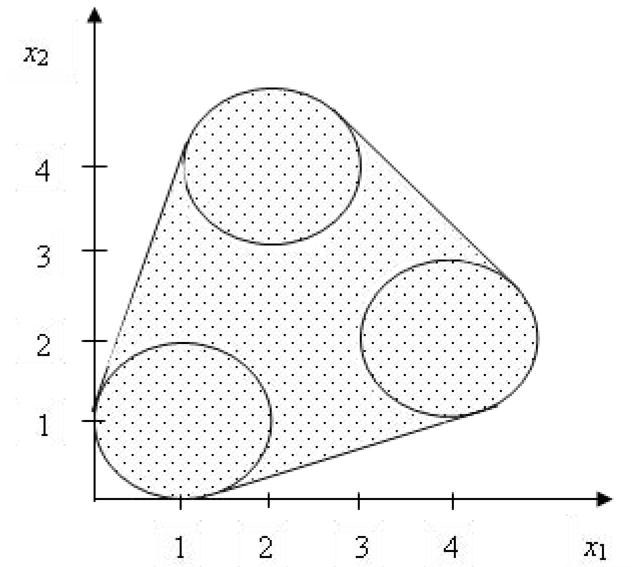
$$x, \nu^{jk} \ge 0,\; 0\le \lambda_{jk} \le 1,\; j\in J_k, k \in K$$

- <u>**Property**</u>: *The NLP (HRP) yields a <u>lower bound</u> to optimum of (GDP).*

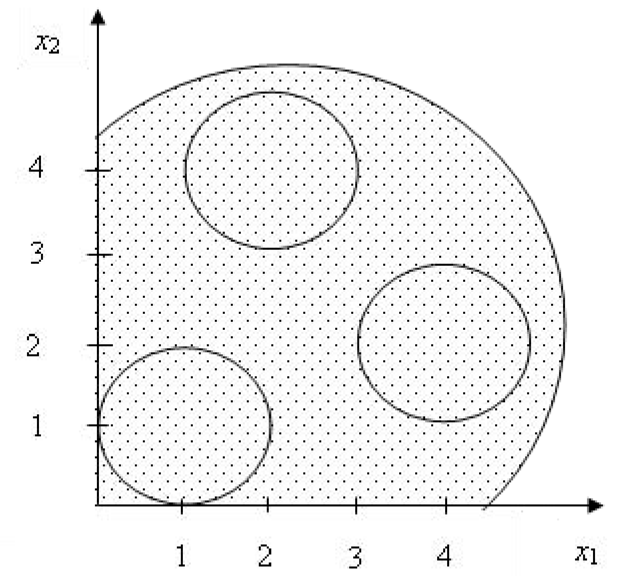
Strength Lower Bounds

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

Grossmann, Lee (2003)



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

GDP

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**

GDP Example

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

$$s.t.$$

$$\begin{bmatrix} Y_1 \\ (x_1)^2 + (x_2)^2 - 1 \le 0 \\ c = 2 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ (x_1 - 4)^2 + (x_2 - 1)^2 - 1 \le 0 \\ c = 1 \end{bmatrix} \vee \begin{bmatrix} Y_3 \\ (x_1 - 2)^2 + (x_2 - 4)^2 - 1 \le 0 \\ c = 3 \end{bmatrix}$$

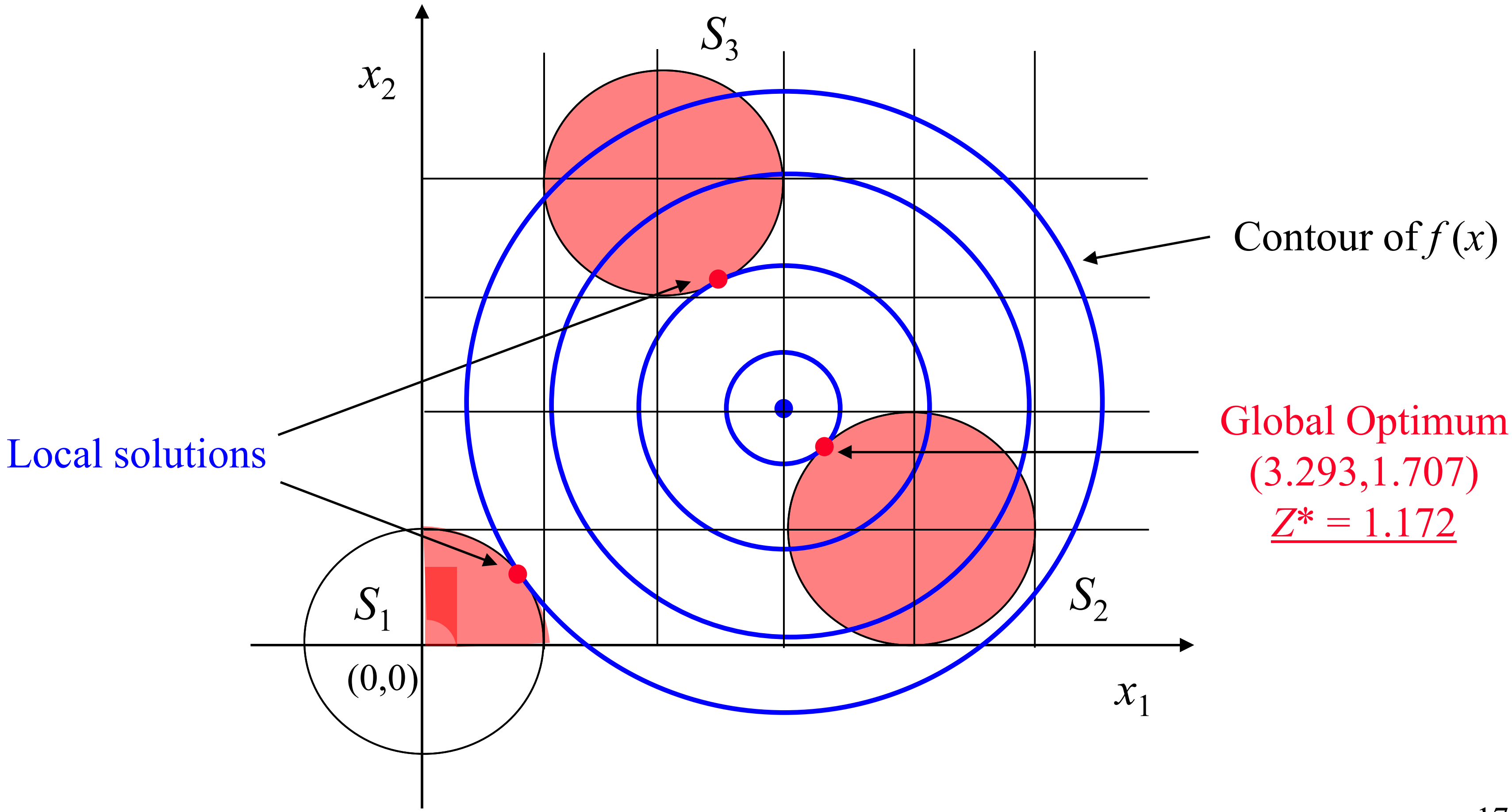
$$Y_1 \underline{\vee} Y_2 \underline{\vee} Y_3 = True$$

$$0 \le x_1, x_2 \le 8, c \ge 0, \qquad Y_j \in \{true, false\}, j = 1,2,3.$$

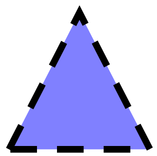
GDP Example

- **Find $x \geq 0$, $(x \in S_1) \vee (x \in S_2) \vee (x \in S_3)$**

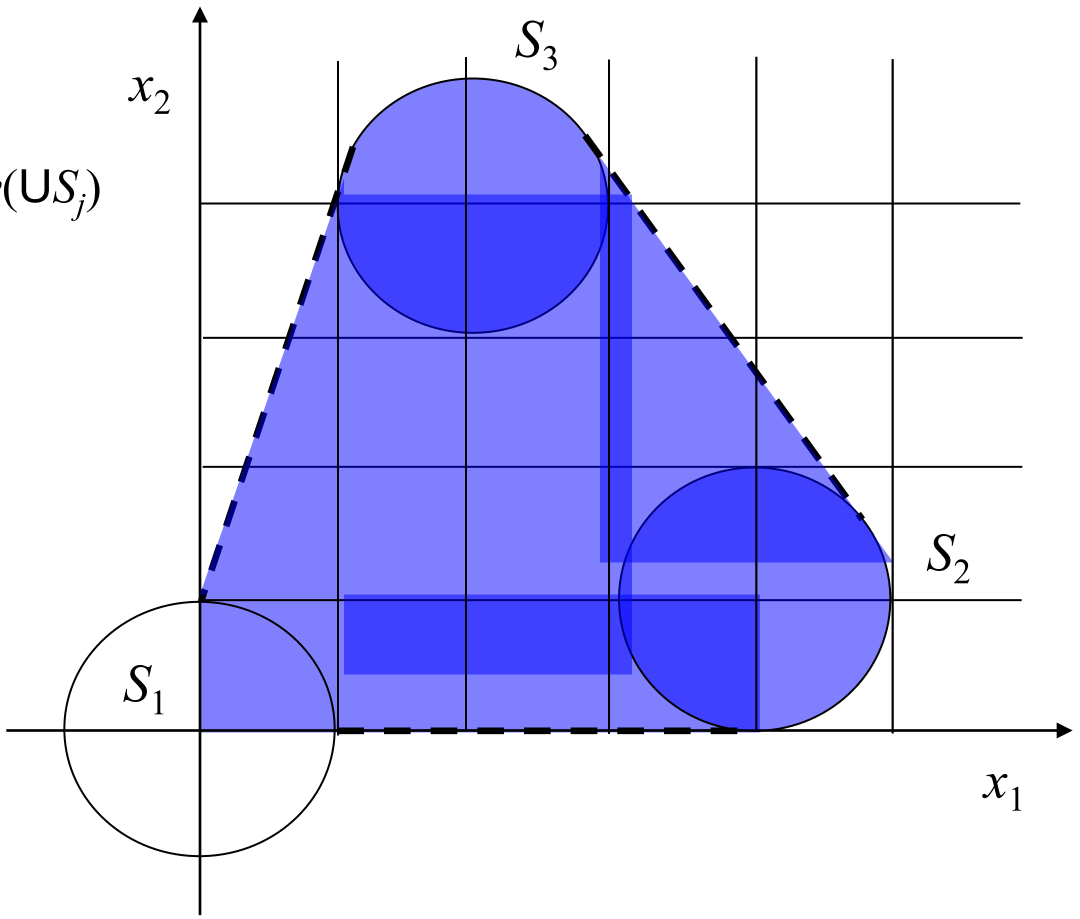
  to minimize $Z = (x_1 - 3)^2 + (x_2 - 2)^2 + c$



Example : convex hull



Convex hull = $conv(\cup S_j)$



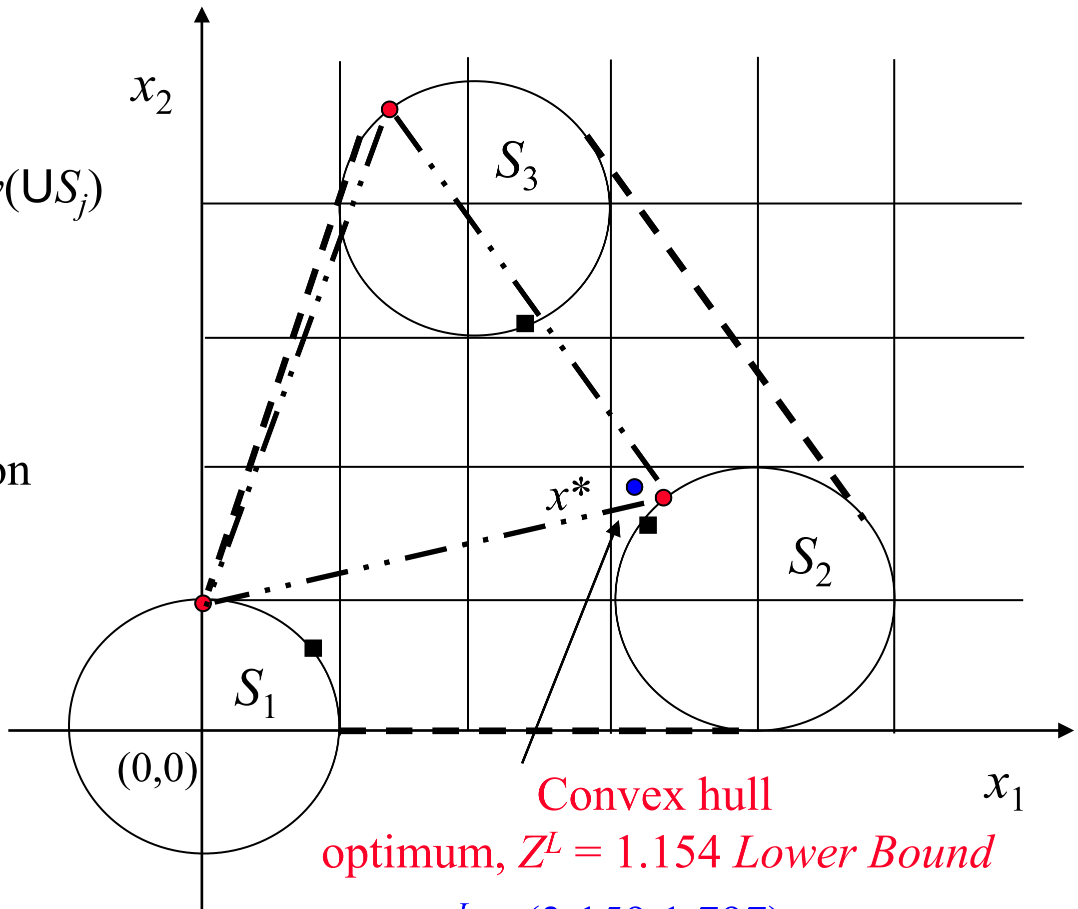
Example: CRP solution

Convex hull = $conv(\cup S_j)$

Convex combination of z_j

$z_j = v^j/\lambda_j$

Local solution point



$\lambda_1 = 0.016$
$\lambda_2 =$ **0.955**
$\lambda_3 = 0.029$

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

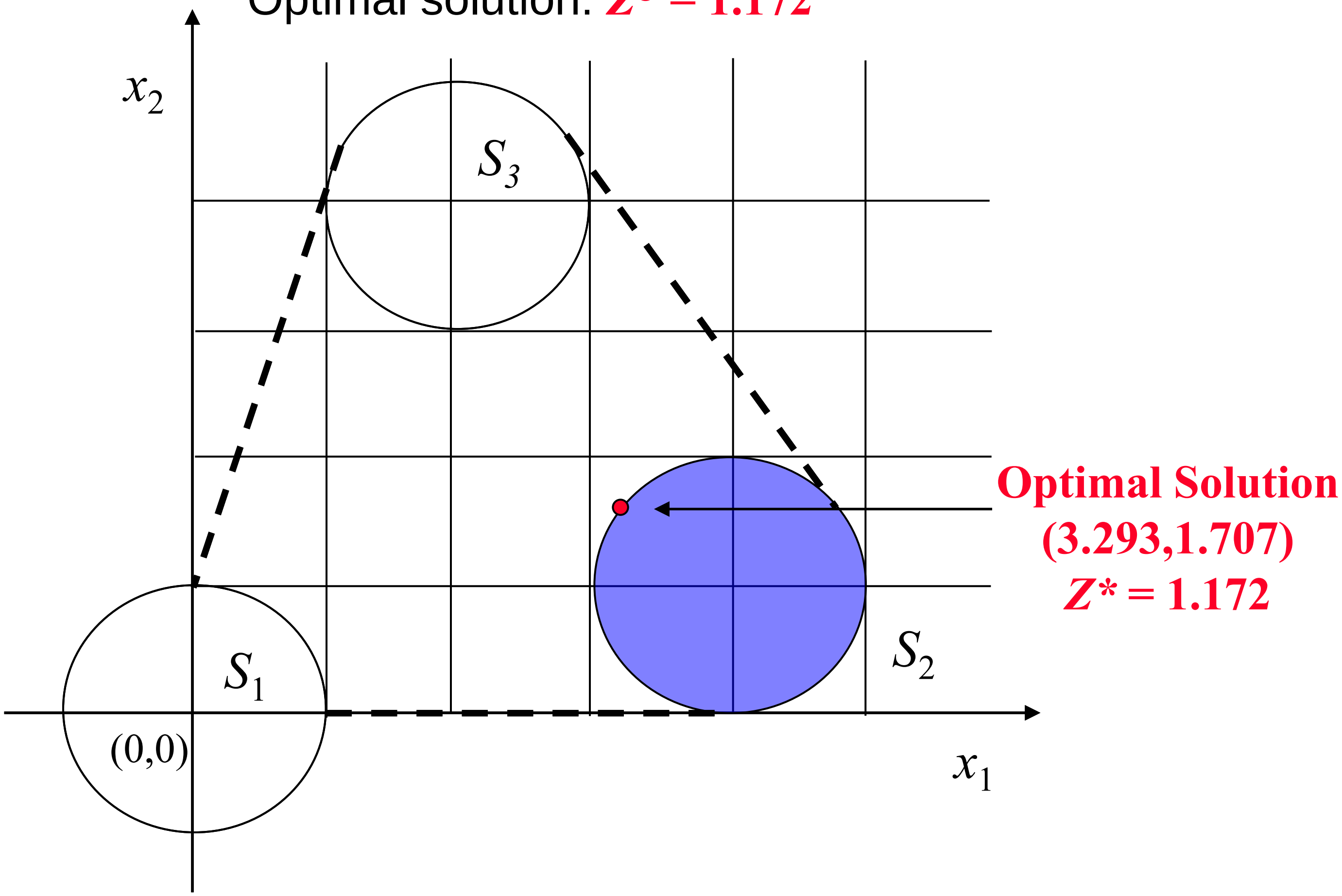
$x^L = (3.159, 1.797)$

Infeasible for GDP

Example : branch and bound

First Node: S_2

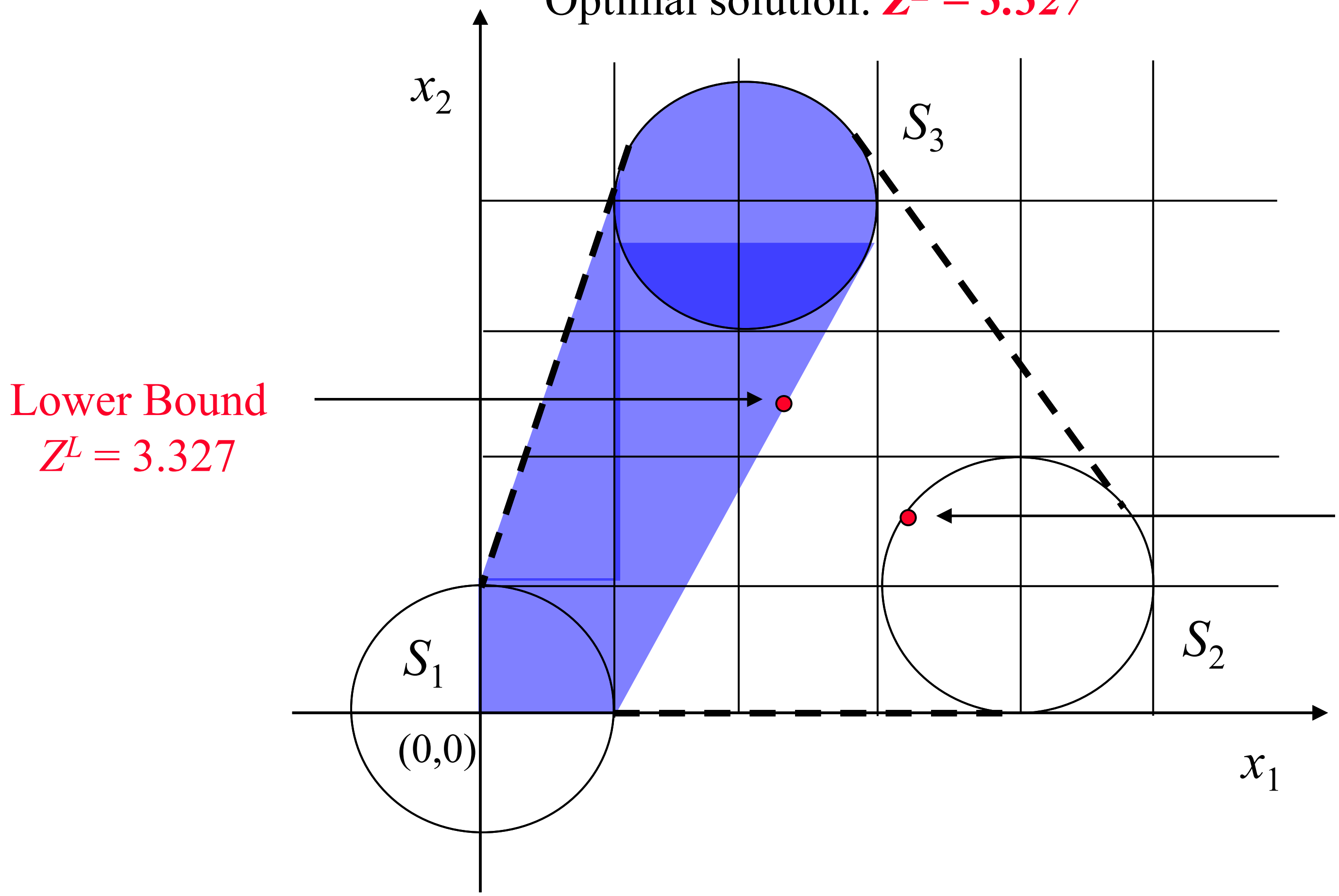
Optimal solution: $Z^U = 1.172$



Example : branch and bound

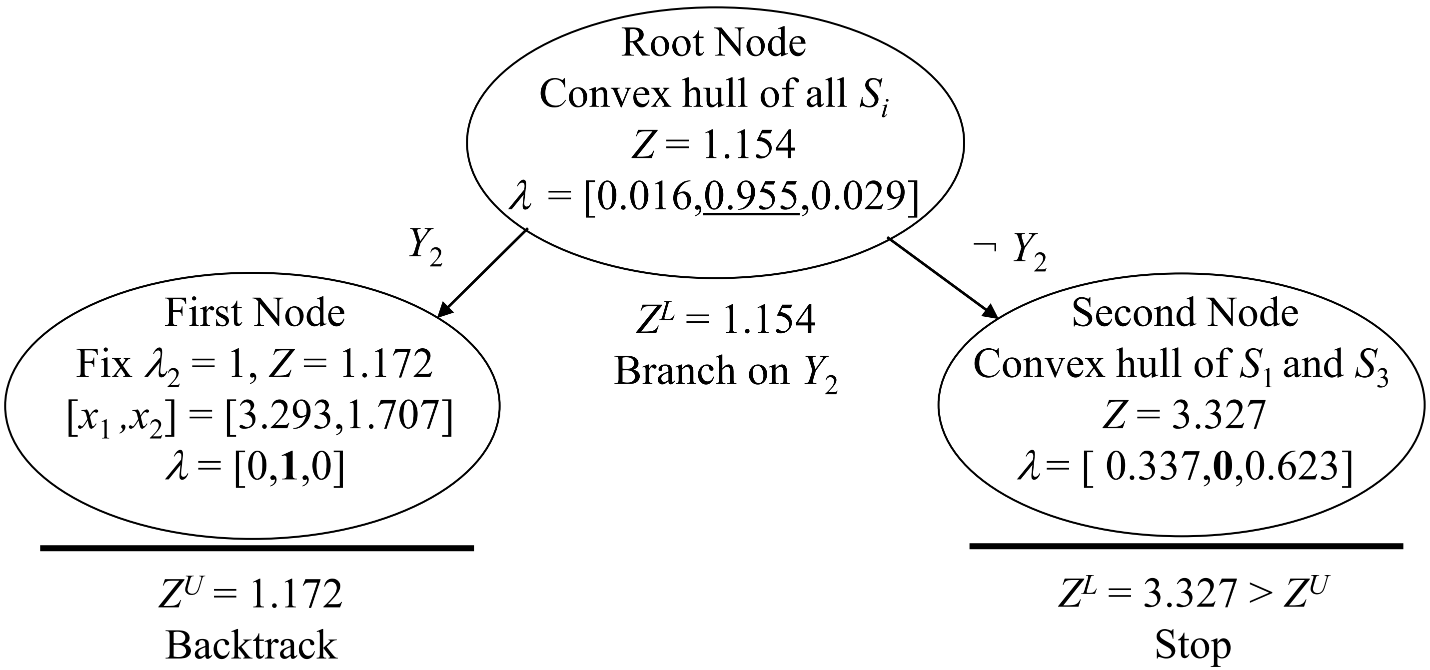
Second Node: ***conv(S_1 U S_3)***

Optimal solution: ***$Z^L = 3.327$***



Example: Search Tree

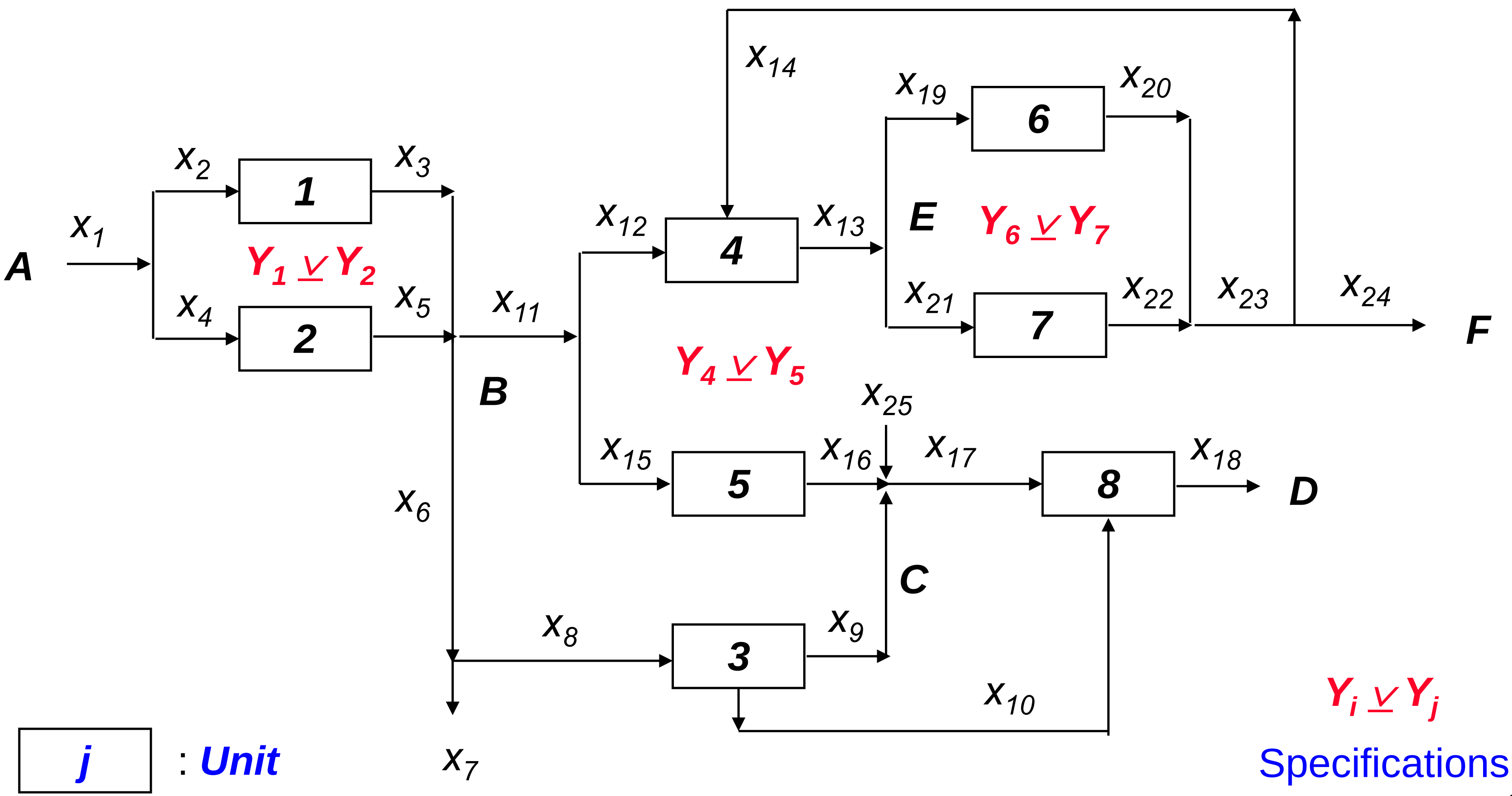
- **Branching Rule**: λ_j - the "weight" of disjunction
  - Fix Y_j as true: fix λ_j as 1.



Process Network with Fixed Charges

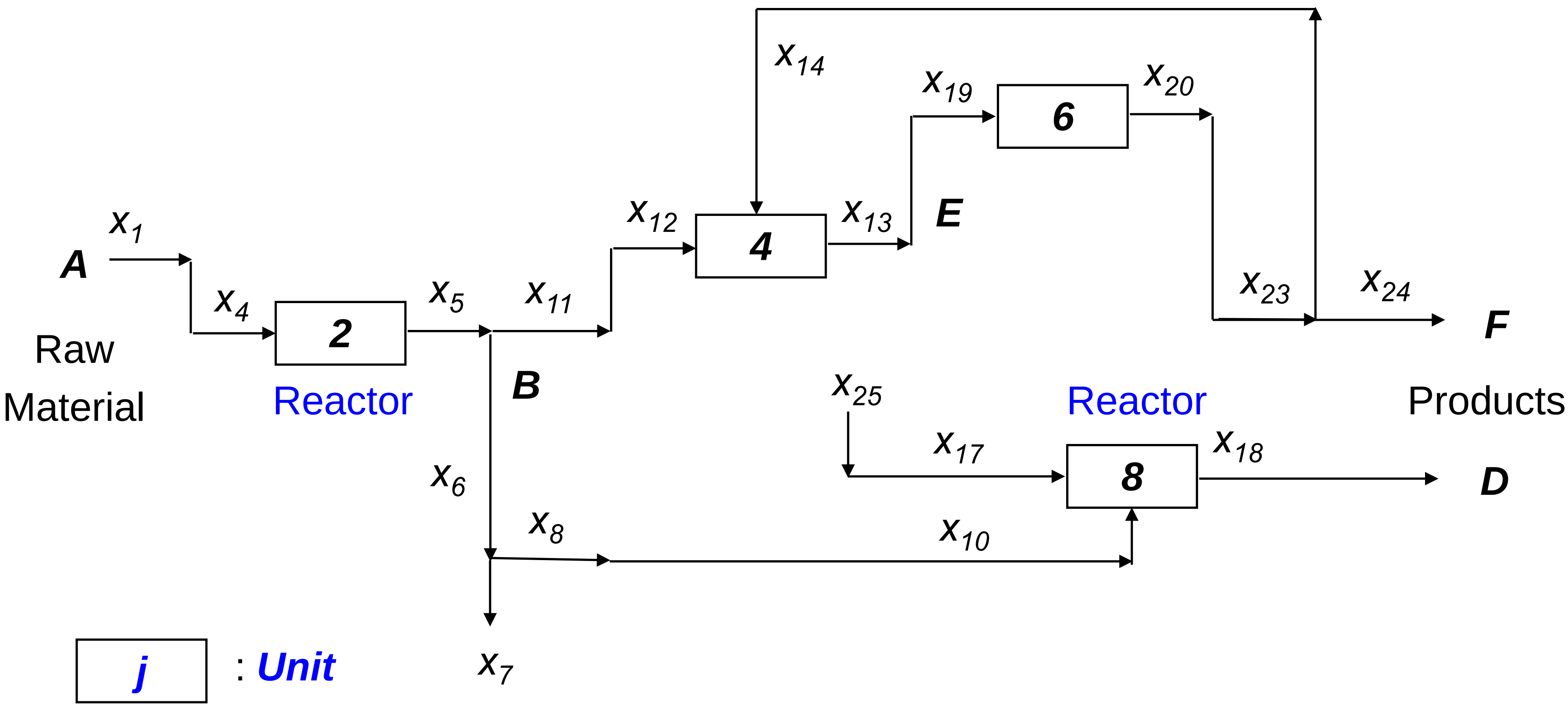
- *Türkay and Grossmann (1997)*

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

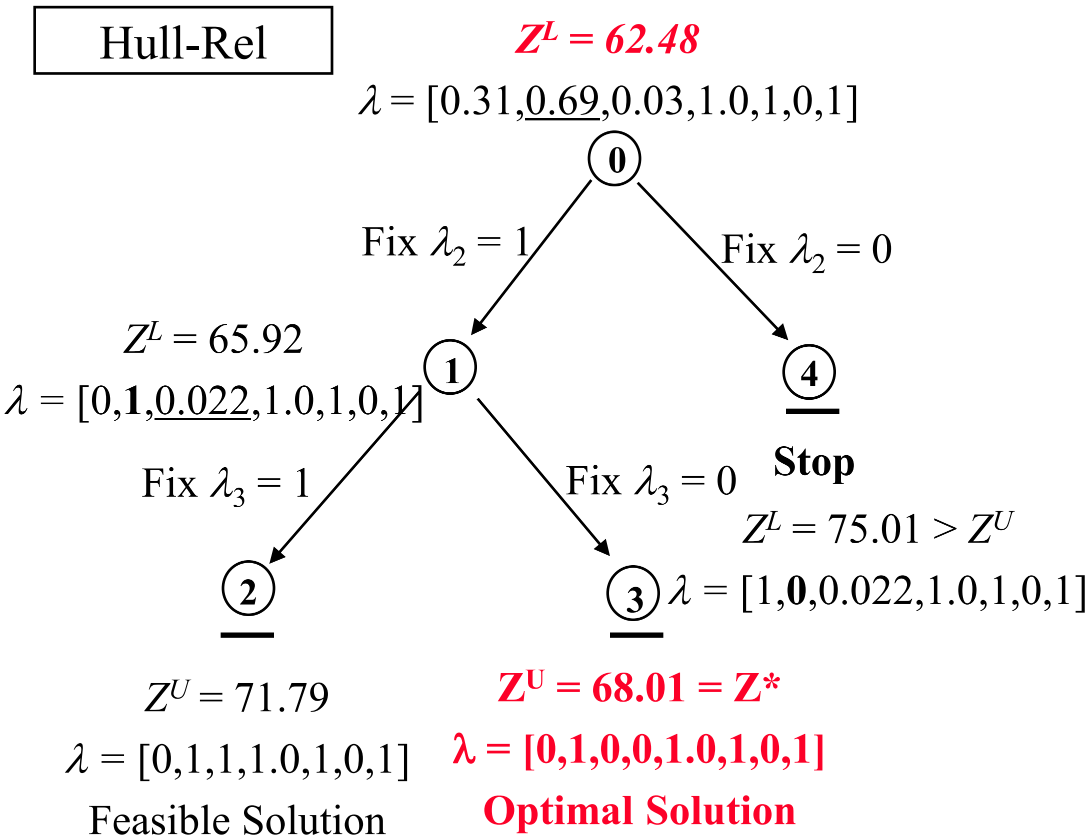


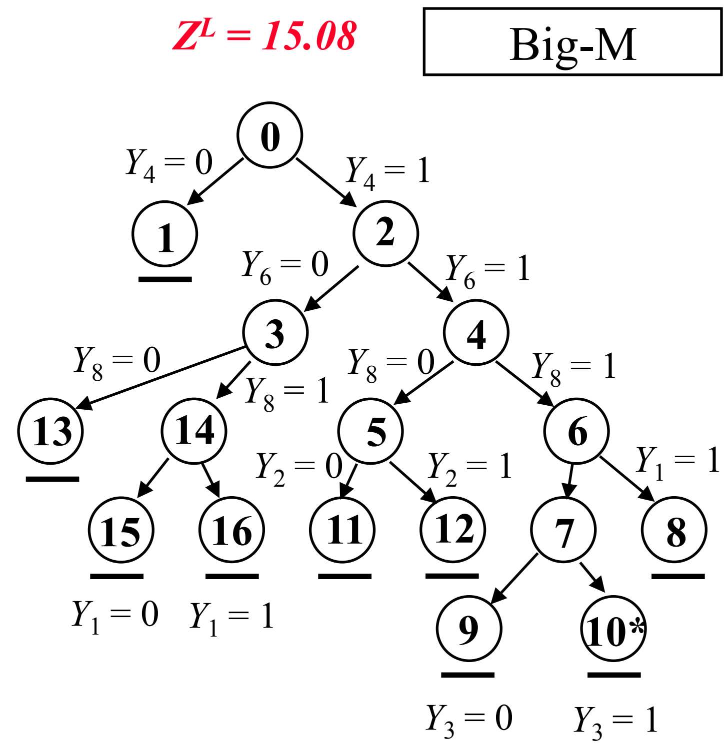
Optimal solution

- **Minimum Cost: $ 68.01M/year**



MINLP- Branch and Bound Method





- **5 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

Balas (1974, 1979, 1985, 1988)

Equivalence between GDP and DP

Sawaya (2007)

GDP

$$\begin{aligned}
\min \quad & Z = f(x) + \textstyle\sum_{k\in K} c_k \\
s.t. \quad & r(x) \le 0 \\
& \bigvee_{i\in D_k} \begin{bmatrix} \boxed{Y_{ik}} \\ g_{ik}(x) \le 0 \\ c_k = \gamma_{ik} \end{bmatrix} \qquad k \in K \\
& \boxed{\Omega(Y) = True} \\
& x^{lo} \le x \le x^{up} \\
& x \in R^n,\ c_k \in R^1,\ Y_{ik} \in \{True, False\}
\end{aligned}$$

DP

$$\begin{aligned}
\min \quad & Z = f(x) + \textstyle\sum_{k\in K} c_k \\
s.t. \quad & r(x) \le 0 \\
& \bigvee_{i\in D_k} \begin{bmatrix} \boxed{\lambda_{ik} = 1} \\ g_{ik}(x) \le 0 \\ c_k = \gamma_{ik} \end{bmatrix} \qquad k \in K \\
& \boxed{A\lambda \ge a} \qquad \sum_{i\in D} \lambda_i = 1, \\
& x^{lo} \le x \le x^{up} \\
& x \in R^n,\ c_k \in R^1, \boxed{\lambda_{ik} \ge 0}
\end{aligned}$$

The integrality of λ is guaranteed

Proposition:
Discrete/continuous GDP and continuous DP have equivalent solutions.

Equivalent Convex Disjunctive Programs

Regular Form: Form represented by the intersection of the union of convex sets

$$F = \bigcap_{k \in K} S_k \,, k \in K,\; S_k = \bigcup_{i \in D_k} P_i$$

$\longrightarrow$ ***F is in regular form***

P_i a convex set for $i \in D_k$

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

$$F = S_1 \cap S_2 \cap S_3$$

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

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

$$F = S_1 \cap S_2 \cap S_3 \qquad \text{as } F = S_{12} \cap S_3$$

Apply Basic Step to:

$$S_{12} \cap S_3 = ((P_{11} \cap P_{12}) \cup (P_{11} \cap P_{22}) \cup (P_{21} \cap P_{12}) \cup (P_{21} \cap P_{22})) \cap (P_{13} \cup P_{23})$$

$$S_{123} = \begin{pmatrix} (P_{11} \cap P_{12} \cap P_{13}) \cup (P_{11} \cap P_{22} \cap P_{13}) \cup (P_{21} \cap P_{12} \cap P_{13}) \cup (P_{21} \cap P_{22} \cap P_{13}) \\ \cup (P_{11} \cap P_{12} \cap P_{23}) \cup (P_{11} \cap P_{22} \cap P_{23}) \cup (P_{21} \cap P_{12} \cap P_{23}) \cup (P_{21} \cap P_{22} \cap P_{23}) \end{pmatrix}$$

We can then rewrite

$$F = S_{12} \cap S_3 \qquad \text{as } F = S_{123}$$

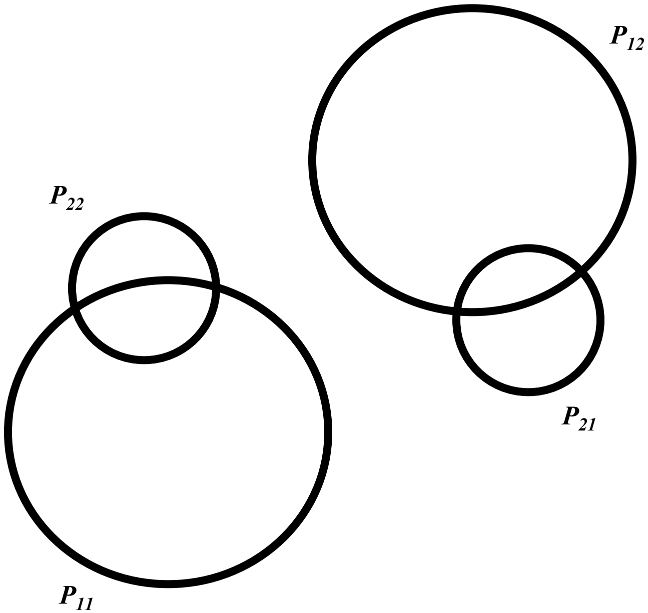
which is its equivalent DNF

Hierarchy of Relaxations for Convex Disjunctive Programs

Theorem 2.4. *For $i = 1, 2....k$ 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: $F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22})$

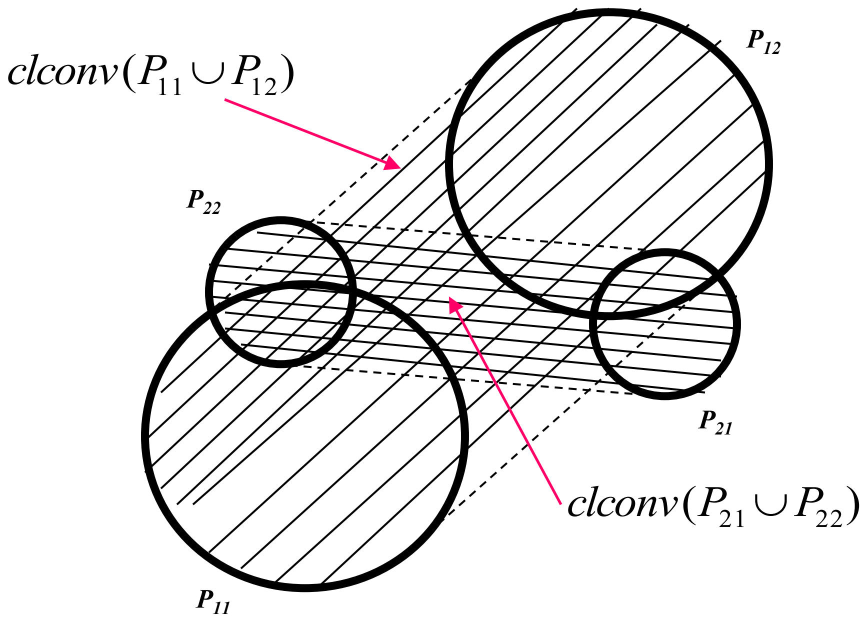


Hierarchy of Relaxations for Convex Disjunctive Programs

Theorem 2.4. *For $i = 1, 2 \dots k$ 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: $F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22})$

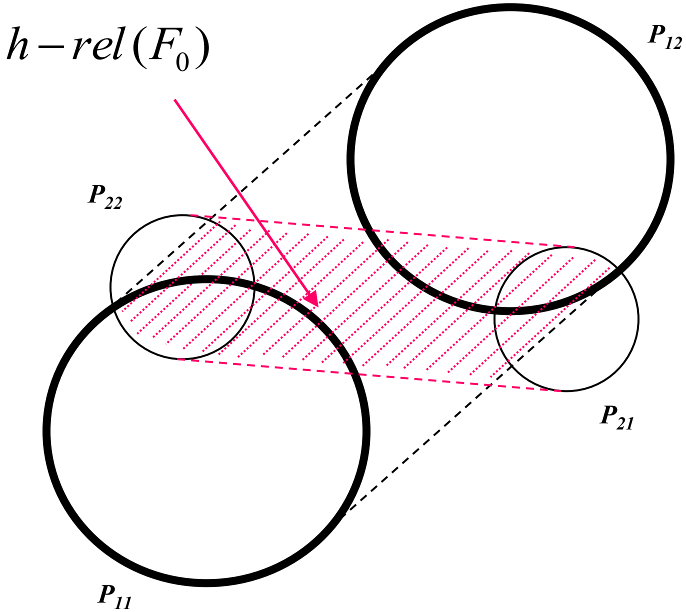


Hierarchy of Relaxations for Convex Disjunctive Programs

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



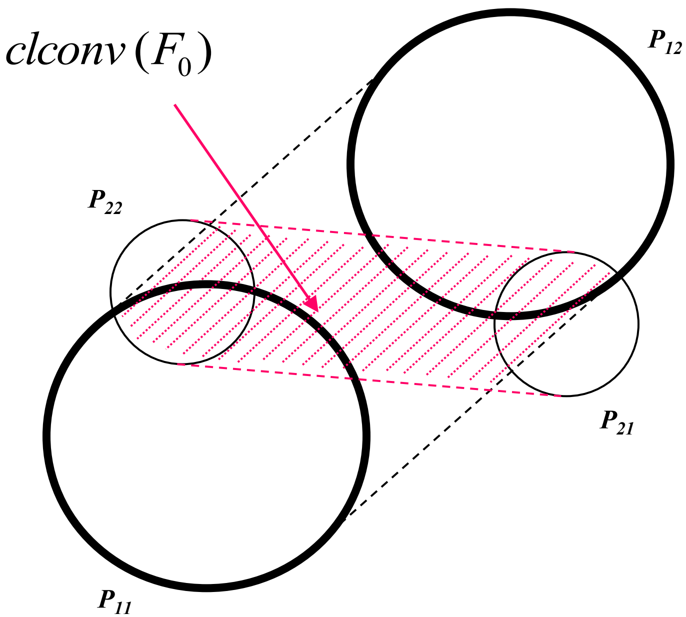
No Basic Step Applied => HR

Hierarchy of Relaxations for Convex Disjunctive Programs

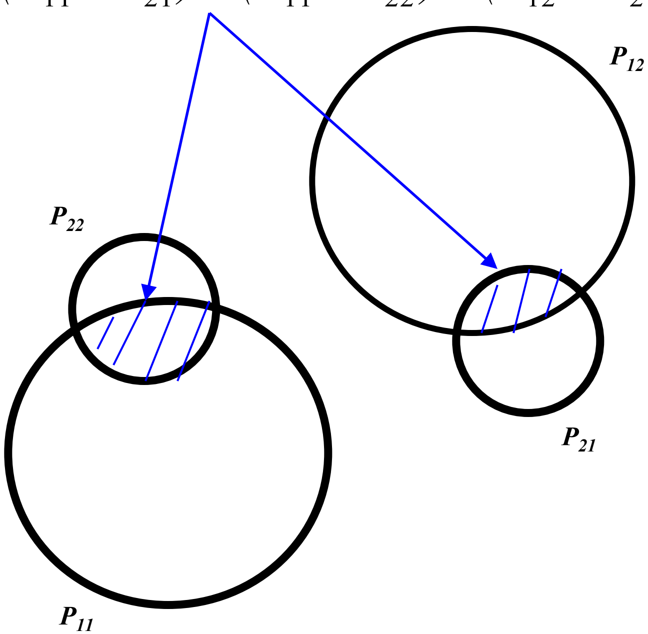
Theorem 2.4. *For* $i = 1, 2 \ldots k$ *let* $F_i = \bigcap_{k \in K} S_k$ *be a sequence of regular forms of a disjunctive set such that* F_i *is obtained from* F_{i-1} *by the application of a basic step, then:*

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

Illustration: $F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22})$ $F_1 = (P_{11} \cap P_{21}) \cup (P_{11} \cap P_{22}) \cup (P_{12} \cap P_{21}) \cup (P_{12} \cap P_{22})$



No Basic Step Applied => HR



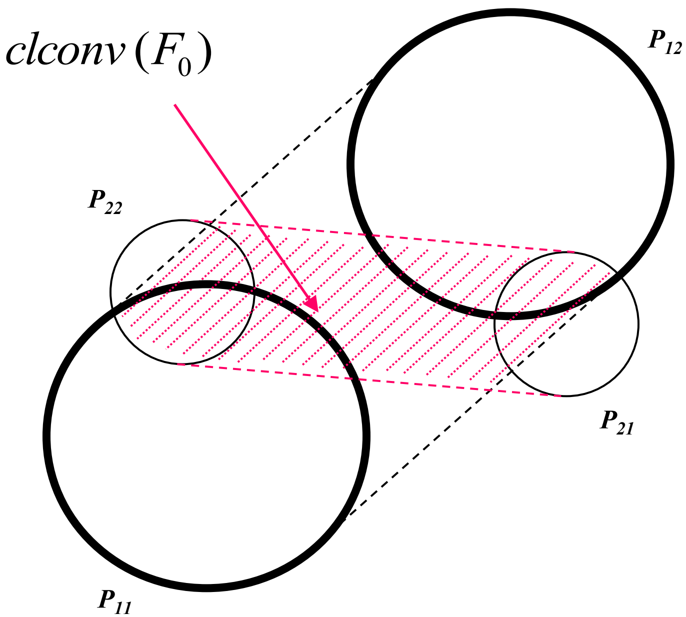
Basic Step Applied

Hierarchy of Relaxations for Convex Disjunctive Programs

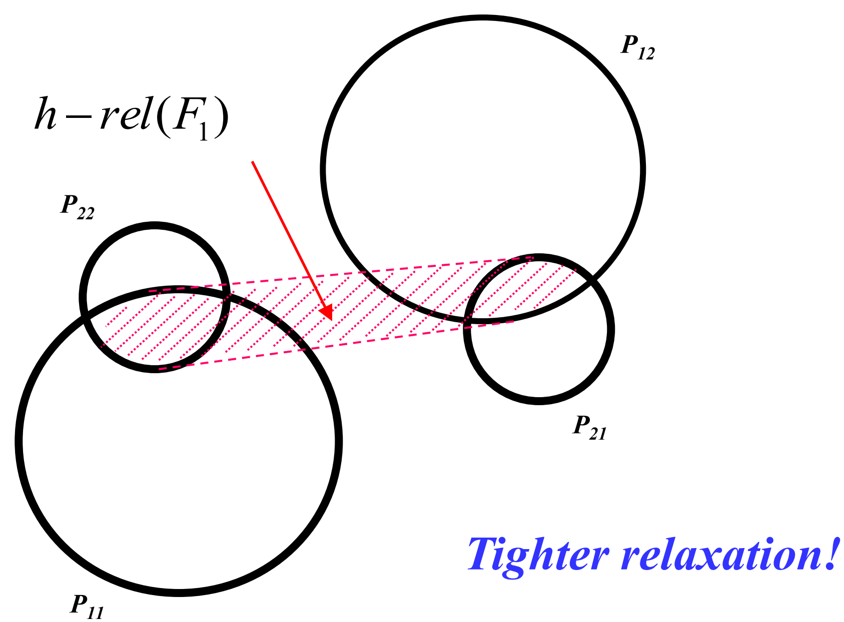
Theorem 2.4. *For $i = 1, 2....k$ let $F_i = \bigcap_{k \in K} S_k$ be a sequence of regular forms of a disjunctive set such that F_i is obtained from F_{i-1} by the application of a basic step, then:*

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

Illustration: $F_0 = (P_{11} \cup P_{12}) \cap (P_{21} \cup P_{22})$ $\qquad$ $F_1 = (P_{11} \cap P_{21}) \cup (P_{11} \cap P_{22}) \cup (P_{12} \cap P_{21}) \cup (P_{21} \cap P_{22})$



No Basic Step Applied => HR



Basic Step Applied => CH

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 $P_i = \{x \in R^n, g_i(x) \leq 0\}$ 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:*

NLPDP:

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

Objective as constraint

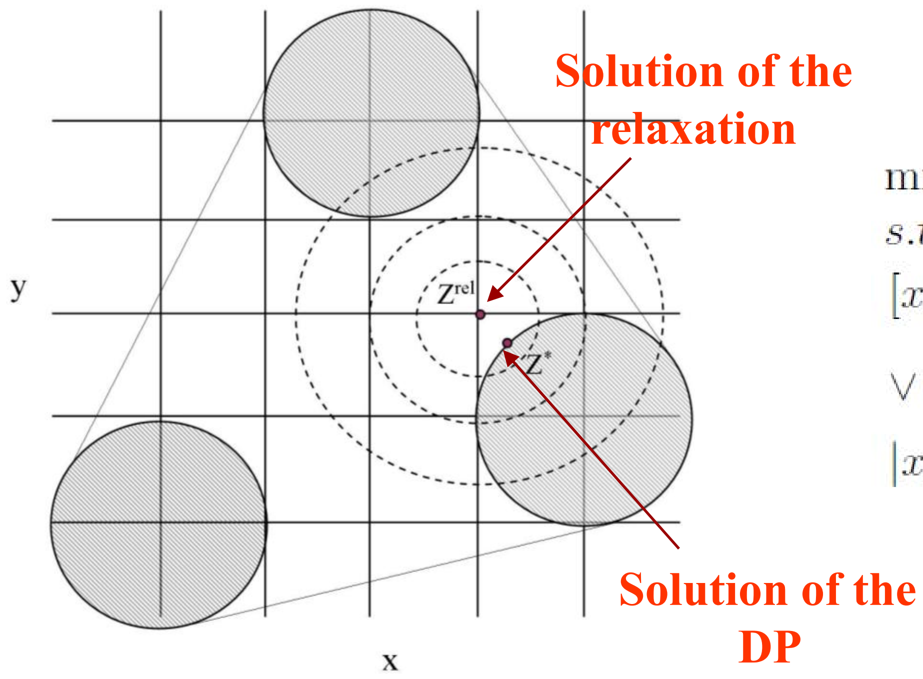
The solution of the NLP relaxation leads to the solution of the DP!

Generalizes convexification results of MILPs

Lovacz & Schrijver (1989), Sherali & Adams (1990), Balas, Ceria, Cornuejols (1993)

Convex nonlinear program equivalent to a convex disjunctive program

Illustrative Example



Disjunctive Program

$$\min \quad Z = (x_1 - 3)^2 + (x_2 - 2)^2 + 1$$
$$s.t.$$
$$[x_1^2 + x_2^2 \le 1] \vee [(x_1 - 4)^2 + (x_2 - 1)^2 \le 1]$$
$$\vee [(x_1 - 2)^2 + (x_2 - 4)^2 \le 1]$$
$$|x_i| \le 5 \quad i \in 1, 2$$

Solution of the relaxed program is different from solution of the disjunctive program

Convex nonlinear program equivalent to a convex disjunctive program

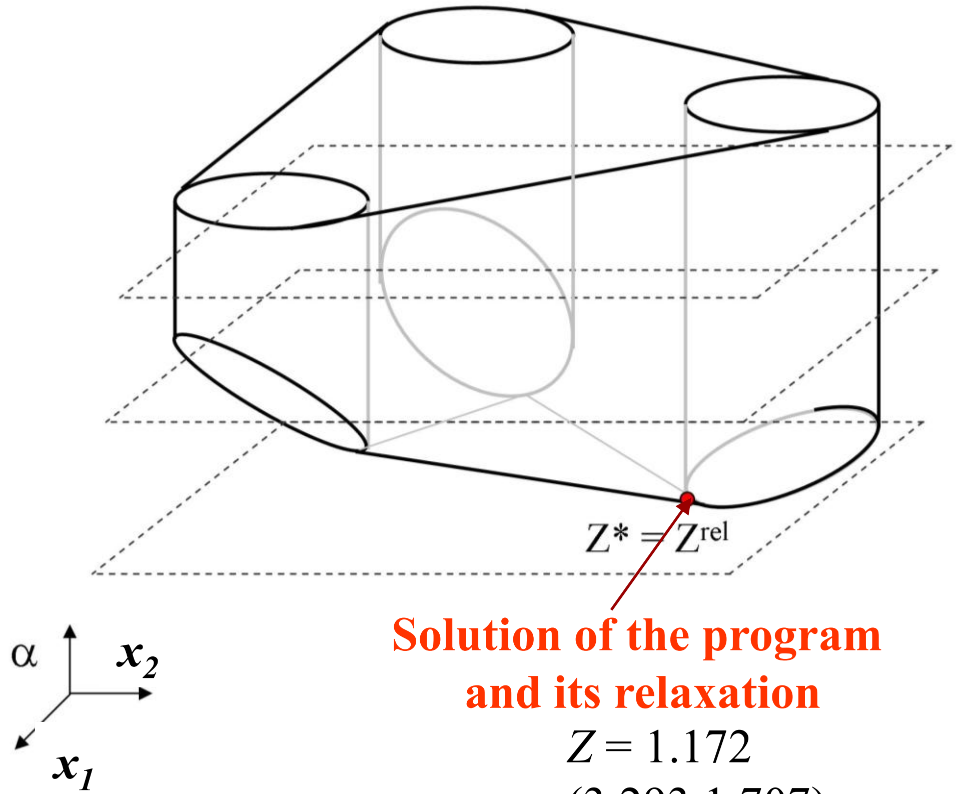
Illustrative Example

Disjunctive Program

Place objective as constraint and intersect with disjunction

$$
\begin{aligned}
&\min \quad Z = \alpha \\
&s.t. \\
&\begin{bmatrix} x_1^2 + x_2^2 \le 1 \\ (x_1-3)^2 + (x_2-2)^2 + 1 \le \alpha \end{bmatrix} \\
\vee &\begin{bmatrix} (x_1-4)^2 + (x_2-1)^2 \le 1 \\ (x_1-3)^2 + (x_2-2)^2 + 1 \le \alpha \end{bmatrix} \\
\vee &\begin{bmatrix} (x_1-2)^2 + (x_2-4)^2 \le 1 \\ (x_1-3)^2 + (x_2-2)^2 + 1 \le \alpha \end{bmatrix} \\
&|x_i| \le 5 \quad i \in 1,2
\end{aligned}
$$

DNF!



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

$$
\begin{aligned}
\min\quad & f(x) \\
s.t.\quad & \\
& x = \sum_{t \in Q_j} \nu^t, && j \in \boxed{T_n} \\
& \lambda_i G^t(\nu^t/\lambda_t) \le 0, && t \in Q_j, \quad j \in T_n \\
& \sum_{t \in Q_j} \lambda_t = 1, \ \lambda_t \ge 0, && t \in Q_j, \quad j \in T_n \\
& \sum_{t \in Q_j | s \in M_j^t} \lambda_t = \delta_s^r, && s \in Q_r, \quad r \in T_d, j \in T_n \\
& \sum_{s \in Q_r} \delta_s^r = 1, && r \in \boxed{T_d} \\
& |\nu^t| \le L\lambda_t, && t \in Q_j, \quad j \in T_n \\
& \delta_s^r \in \{0, 1\}, && t \in Q_j, \quad j \in T_n
\end{aligned}
$$

Set of disjunctions after basic steps

Constraints after basic steps

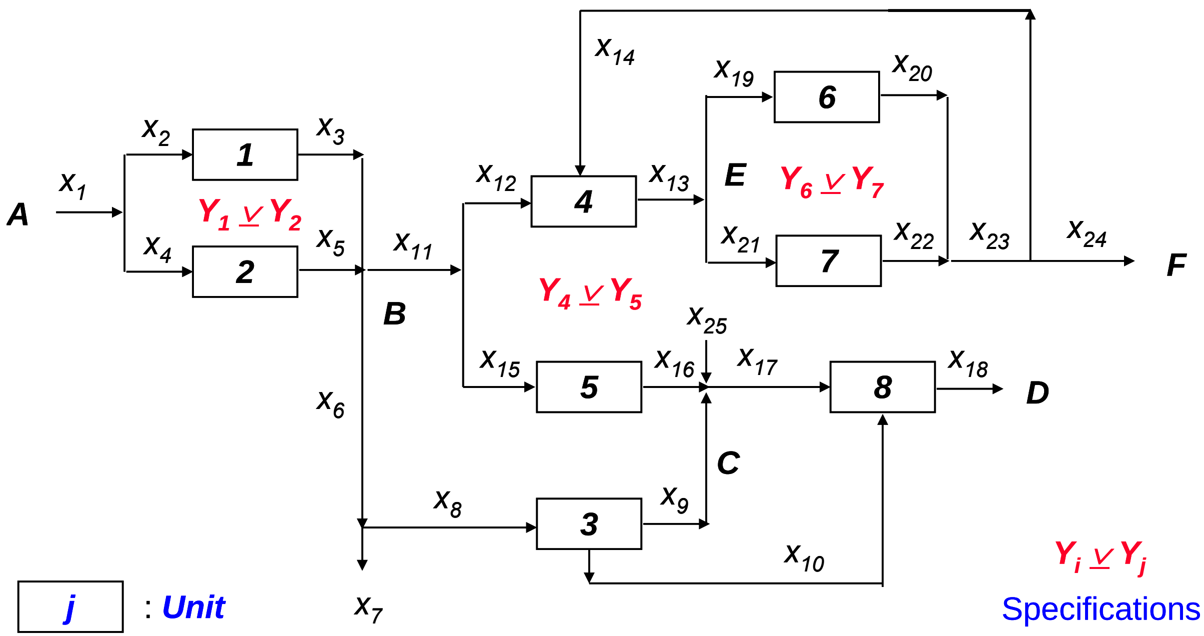
No new binary variables are created (Balas, 1985)

Set of disjunctions before basic steps

No additional 0-1 variables are required!

Process Network Revisited

Illustrative Example



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

	BM Approach			HR Approach			Proposed Approach		
Example	Bin	Con	Const	Bin	Con	Const	Bin	Con	Const
Circles2D3	3	8	12	3	16	20	3	20	27
Circles2D36	36	39	38	36	111	112	36	147	184
Circles3D36	36	40	38	36	148	149	36	184	221
Proc8	8	42	97	8	98	152	8	444	843
Proc10	10	51	98	10	124	158	10	638	1181
Proc12	12	57	114	12	137	184	12	805	1462
Flay02	4	15	12	4	47	52	4	55	80
Flay03	12	27	25	12	123	145	12	195	334
Flay04	24	43	43	24	235	283	24	511	865
Clay0203	18	31	55	15	88	130	15	160	316
Clay0303	21	34	67	21	100	424	21	268	571
Clay0204	32	53	91	32	165	235	32	641	1503

Numerical Results

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

Table: Performance using different reformulation strategies

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

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

		BM Approach			HR Approach		
Example	Opt.	LB	Nds	T(s)	LB	Nds	T(s)
Circles2D3	1.17	0.00	4	1.0	1.15	4	1.04
Circles2D36	2.25	0.00	70	7.8	0.00	70	15.40
Circles3D36	15.77	0.44	70	7.3	12.04	70	18.50
Proc8	68.01	-829.0	34	4.6	67.12	2	1.0
Proc10	-73.51	-1,108.88	197	21.7	-78.81	4	1.0
Proc12	-69.51	-1,108.88	234	27.7	-74.81	8	1.0
Flay02	37.95	28.28	6	1.0	28.28	6	1.7
Flay03	48.99	30.98	104	10.7	30.98	108	12.1
Flay04	54.40	30.98	2,415	234.0	30.98	2,887	288.0
Clay0203	41,573.30	0.00	323	32.7	0.00	216	22.0
Clay0303	26,670.00	0.00	380	42.0	0.00	879	99.0
Clay0204	6,545.00	0.00	2,265	229.0	0.00	2,835	507.0

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

		BM Approach			HR Approach			Proposed Approach		
Example	Opt.	LB	Nds	T(s)	LB	Nds	T(s)	LB	Nds	T(s)
Circles2D3	1.17	0.00	4	1.0	1.15	4	1.04	1.17	0	0.7
Circles2D36	2.25	0.00	70	7.8	0.00	70	15.40	2.24	29	4.9
Circles3D36	15.77	0.44	70	7.3	12.04	70	18.50	15.72	34	10.8
Proc8	68.01	-829.0	34	4.6	67.12	2	1.0	68.01	0	1.7
Proc10	-73.51	-1,108.88	197	21.7	-78.81	4	1.0	-73.56	2	1.9
Proc12	-69.51	-1,108.88	234	27.7	-74.81	8	1.0	-69.51	2	2.9
Flay02	37.95	28.28	6	1.0	28.28	6	1.7	37.95	3	1.0
Flay03	48.99	30.98	104	10.7	30.98	108	12.1	41.94	30	9.0
Flay04	54.40	30.98	2,415	234.0	30.98	2,887	288.0	41.69	52	48.0
Clay0203	41,573.30	0.00	323	32.7	0.00	216	22.0	3,010.00	206	28.7
Clay0303	26,670.00	0.00	380	42.0	0.00	879	99.0	3,103.00	331	69.0
Clay0204	6,545.00	0.00	2,265	229.0	0.00	2,835	507.0	4,760.00	546	157.0

Improved lower bounds 100%probs

Proposed vs BM: faster 10 out of 12

Proposed vs HR: faster 8 out of 12

Extension to Nonconvex GDP

Basic idea: strengthen lower bound of global optimum

Phase 1 *Phase 2*

Nonconvex GDP	→	Convex GDP *Initial lower bound*	→	Tight Convex GDP *Stronger lower bound*

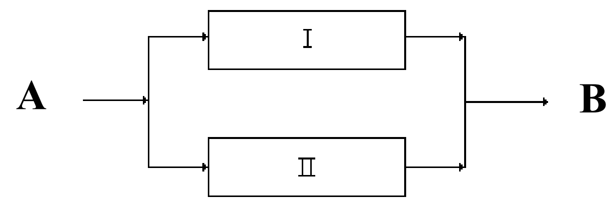
Relaxation
Under/over estimating functions
Convex envelopes

Strengthen relaxation
Apply basic steps

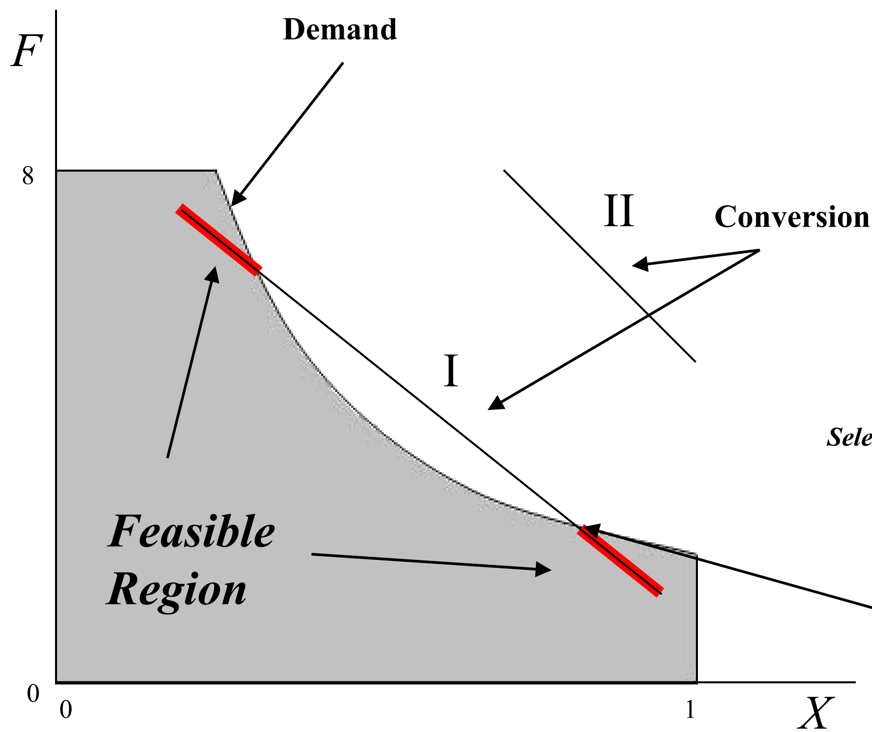
Remarks

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



F: Flow
X: Conversion



GDP Formulation

Objective Function (- Profit)

$$\text{Min } Z = -qFX + gF + CP$$

Demand constraint

$$\text{s.t.} \quad \boxed{FX \le d}$$

Selection Reactor

$$\boxed{\begin{bmatrix} Y_{11} \\ F = \alpha_1 X + \beta_1 \\ X_1^{LO} \le X \le X_1^{UP} \\ CP = Cp_1 \end{bmatrix} \vee \begin{bmatrix} Y_{21} \\ F = \alpha_2 X + \beta_2 \\ X_2^{LO} \le X \le X_2^{UP} \\ CP = Cp_2 \end{bmatrix}}$$

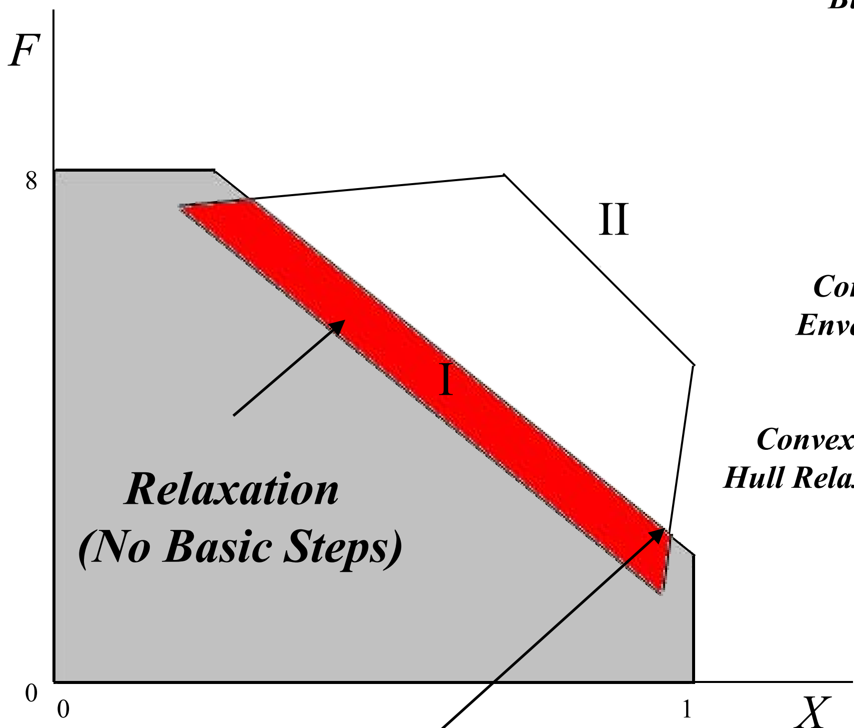
$$Y_{11} \vee Y_{21} = True$$

$$CP, X, F \in R$$
$$F^{LO} \le F \le F^{UP}$$
$$Y_{11}, Y_{21} \in \{True, False\}$$

Optimum Z* = -1.01

Illustrative Example: Optimal reactor selection I

Lee & Grossmann (2003) Relaxation



Lower bound Z* = -1.28 < -1.01

Bilinear Terms

$$Max\ Z = -qP + gF + CP$$

$$s.t.\ \ \boxed{P} \le d$$

Convex Envelopes

$$P \le F.X^{LO} + F^{UP}.X - F^{UP}.X^{LO}$$
$$P \le F.X^{UP} + F^{LO}.X - F^{LO}.X^{UP}$$
$$P \ge F.X^{LO} + F^{LO}.X - F^{LO}.X^{LO}$$
$$P \ge F.X^{UP} + F^{UP}.X - F^{UP}.X^{UP}$$

Convex Hull Relaxation

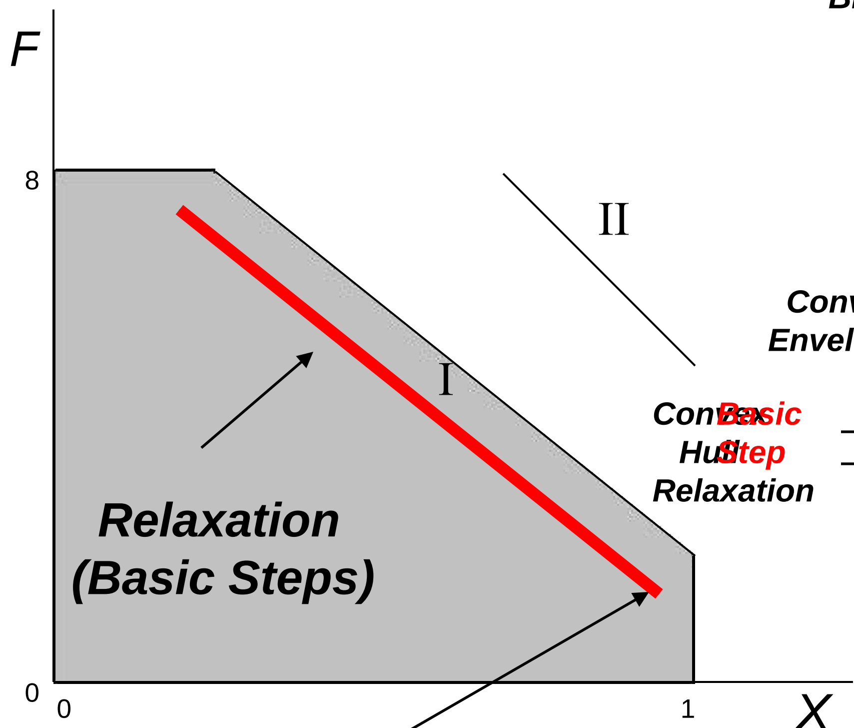
$$\begin{bmatrix} Y_{11} \\ F = \alpha_1 X + \beta_1 \\ X_1^{LO} \le X \le X_1^{UP} \\ CP = Cp_1 \end{bmatrix} \vee \begin{bmatrix} Y_{21} \\ F = \alpha_2 X + \beta_2 \\ X_2^{LO} \le X \le X_2^{UP} \\ CP = Cp_2 \end{bmatrix}$$

$$Y_{11} \vee Y_{21} = True$$

$CP, X, F \in R$
$F^{LO} \le F \le F^{UP}$
$Y_{11}, Y_{21} \in \{True, False\}$

Illustrative Example: Optimal reactor selection I

Proposed Relaxation



Bilinear Terms

Convex Envelopes

Convex Hull Relaxation

Basic Step

$$\text{Min } Z = -\theta P + \gamma F + CP$$

$$\text{s.t. } P \le d$$

$$\begin{bmatrix} Y_{11} \\ P \le d \\ P \le F.X^{UP} + F^{LO}.X - F^{LO}.X^{UP} \\ P \le F.X^{LO} + F^{UP}.X - F^{UP}.X^{LO} \\ P \ge F.X^{LO} + F^{LO}.X - F^{LO}.X^{LO} \\ P \ge F.X^{UP} + F^{UP}.X - F^{UP}.X^{UP} \\ F = \alpha_1 X + \beta_1 \\ X_1^{LO} \le X \le X_1^{UP} \\ CP = Cp_1 \end{bmatrix} \vee \begin{bmatrix} Y_{21} \\ P \le d \\ P \le F.X^{UP} + F^{LO}.X - F^{LO}.X^{UP} \\ P \le F.X^{LO} + F^{UP}.X - F^{UP}.X^{LO} \\ P \ge F.X^{LO} + F^{LO}.X - F^{LO}.X^{LO} \\ P \ge F.X^{UP} + F^{UP}.X - F^{UP}.X^{UP} \\ F = \alpha_2 X + \beta_2 \\ X_2^{LO} \le X \le X_2^{UP} \\ CP = Cp_2 \end{bmatrix}$$

$$Y_{11} \vee Y_{21} = True$$

$CP, X, F \in R$

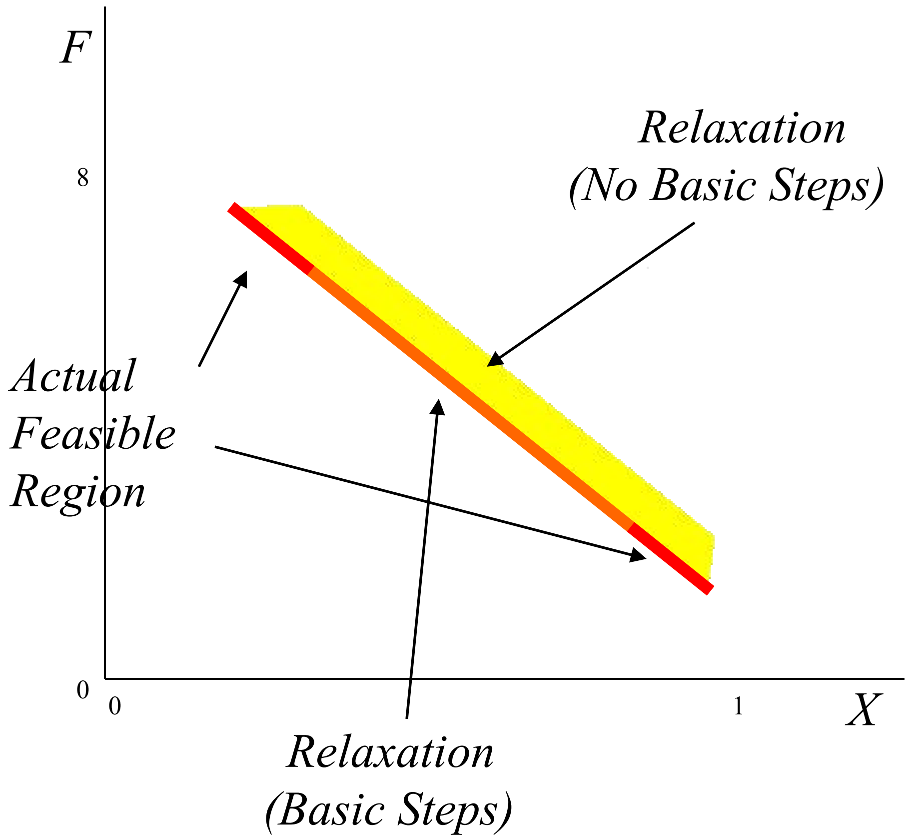
$F^{LO} \le F \le F^{UP}$

$Y_{11}, Y_{21} \in \{True, False\}$

Lower bound Z* = -1.1 < -1.01 and tighter than -1.28!

Illustrative Example: Optimal reactor selection I

Comparison of Relaxations



*The application of **basic step** prior to the discrete relaxation leads to a **tighter** relaxed feasible region = > **stronger lower bounds***

Dimensions of Test Problems
Bilinear/Concave

	Bilinear Terms	Concave Functions	Discrete Variables	Continuous Variables
Example 1	1	0	2	3
Example 2	0	2	2	5
Example 3	4	9	9	8
Example 4	36	0	9	114
Example 5	24	0	9	76

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

	Cont. Vars.	Boolean Vars.	Logic Const.	Disj. Const.	Global Const.
PROC1	5	2	1	1	3
PROC2	5	2	1	1	3
RXN1	4	2	1	1	6
RXN2	4	2	1	1	6
HEN1	18	2	2	2	21

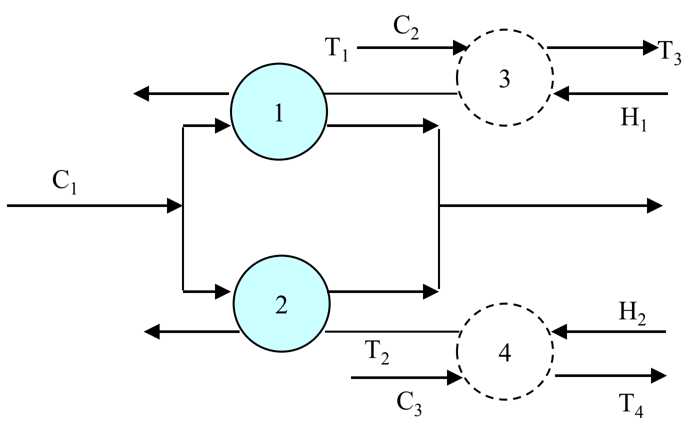
Reference

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



Linear Fractional Terms in constraints

Generalized Disjunctive Program

$$\min\ Z = c_1A_1 + c_1A_1 + c_1A_1 + c_1A_1 + C_3 + C_4$$

$s.t.$

$$A_1 = \boxed{\frac{Q_1}{U_1\Delta T_1}},\ A_2 = \boxed{\frac{Q_2}{U_2\Delta T_2}}$$

$$\begin{bmatrix} Y_3 \\ A_3 = \boxed{\dfrac{Q_3}{U_3\Delta T_3}} \\ C_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \neg Y_3 \\ Q_3 = 0 \\ A_3 = 0 \\ C_3 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_4 \\ A_3 = \boxed{\dfrac{Q_3}{U_3\Delta T_3}} \\ C_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \neg Y_4 \\ Q_3 = 0 \\ A_3 = 0 \\ C_3 = 0 \end{bmatrix}$$

$$Q_1 = FCP_{H1}(T_1 - T_{H1,out}), Q_2 = FCP_{H2}(T_2 - T_{H2,out})$$
$$Q_3 = FCP_{C2}(T_3 - T_{C2,in}), Q_3 = FCP_{H1}(T_{H1,in} - T_1)$$
$$Q_4 = FCP_{C3}(T_4 - T_{C3,in}), Q_4 = FCP_{H2}(T_{H2,in} - T_2)$$

$$T_1 \ge T_{C1,in} + EMAT,\ \ T_2 \ge T_{C1,in} + EMAT$$
$$Q_1 + Q_2 = Q_{total}$$

$$\Delta T_1 = \frac{(T_1 - T_{C1,out}) + (T_{H1,out} - T_{C1,in})}{2}, \Delta T_2 = \frac{(T_2 - T_{C1,out}) + (T_{H2,out} - T_{C1,in})}{2}$$
$$\Delta T_3 = \frac{(T_1 - T_{C2,in}) + (T_{H1,in} - T_3)}{2}, \Delta T_4 = \frac{(T_2 - T_{C3,in}) + (T_{H2,in} - T_4)}{2}$$

$$T_{H1,out} \le T_1 \le T_{H1,in}, T_{H2,out} \le T_4 \le T_{H2,in}$$
$$T_{C2,in} \le T_3\ ,\ T_{C3,in} \le T_4$$

$$Q_i \ge 0, \Delta T_i \ge EMAT\ I = 1,\ldots,4$$

Prediction of Lower Bounds Global Optimum

		Global Optimum	Lower Bound Hull Relaxation
Bilinear Concave	***React 1***	**-1.01**	**-1.28**
	React 2	**6.31**	**5.65**
	HEN	**114384.78**	**91671.18**
	Water	**1214.87**	**400.66**
	Pool	**-4640**	**-5515**
Linear Fractional, Posynomial, Exponential	***Process 1***	**18.61**	**11.85**
	Process 2	**19.48**	**12.38**
	RXN 1	**42.89**	**-337.5**
	RXN 2	**76.47**	**22.5**
	HEN 1	**48531**	**38729.3**

Lower bounds improved in all cases   Ave. increase 22%

8 out of 10 achieved theoretically best lower bound (DNF)!

Global Optimization Methodology

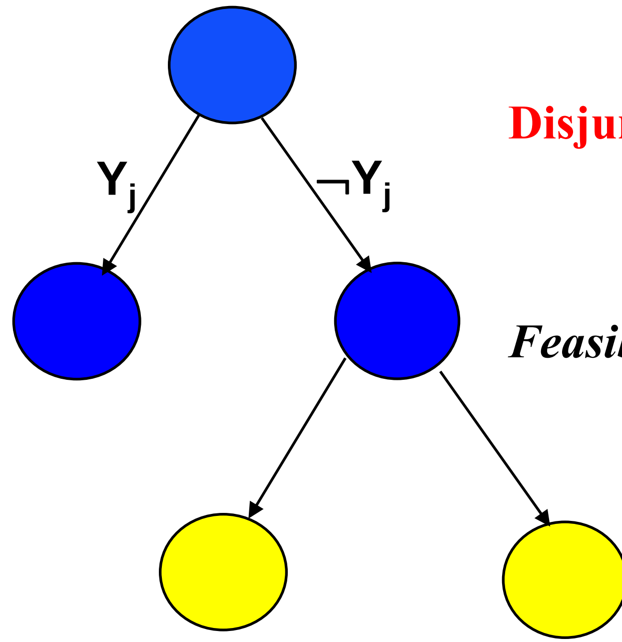
GDP reformulation
Apply basic steps following the rules presented

⇩

Bound Contraction
(Zamora & Grossmann, 1999)

⇩

Spatial Branch and Bound
(Lee & Grossmann, 2001)



Disjunctive B&B

Feasible discrete

Spatial B&B

Computational Performance- *Bilinear/Concave*

	Global Optimization Technique using Hull Relaxation			
	Global Optimum	Nodes	Bound contract. (% Avg)	CPU Time (sec)
Example 1	-1.01	5	35	2.1
Example 2	6.31	1	33	1.0
Example 3	114384.78	13	85	11.0
Example 4	1214.87	450	8	217
Example 5	-4640	502	1	268

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

	Size of the LP Relaxation (Hull Relaxation)		Size of the LP Relaxation (Proposed)	
	Constraints	Variables	Constraints	Variables
Example 1	23	15	28	15
Example 2	24	14	31	18
Example 3	87	52	206	106
Example 4	544	346	3424	1210
Example 5	3336	1777	4237	1777

Computational Performance- *Nonlinear*

		Global Optimization Technique using Lee & Grossmann Relaxation		
	Global Optimum	Nodes	Bound contract. (% Avg)	CPU Time (sec)
PROCN1	**18.61**	**3**	**51.3**	**6**
PROCN2	**19.48**	**2**	**40.5**	**4**
RXN1	**42.89**	**2**	**51.0**	**7**
RXN2	**76.46**	**2**	**51.0**	**6**
HEN1	**48531**	**3**	**13.8**	**15**

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/concve 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**