Review of Mixed-Integer Nonlinear Programming and Generalized Disjunctive Programs

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

Zhejiang University
Hangzhou
January 17, 2014

Outline

1. **Review of MINLP methods**

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

3. **Convex nonlinear GDP: hierarchy of relaxations**
   Concept of basic steps
   Equivalent NLP formulation

4. **Application to global Optimization of nonconvex GDP**
   Bilinear, concave and linear fractional functions

5. **Algorithm reformulating GDP to MI(N)LP using basic steps**
   Convex linear/nonlinear GDP

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**

Solution Algorithms

- **Branch and Bound method (BB)**

Ravindran and Gupta (1985) Leyffer and Fletcher (2001)

Branch and cut: *Stubbs and Mehrotra (1999)*

- **Generalized Benders Decomposition (GBD)**

Geoffrion (1972)

- **Outer-Approximation (OA)**

Duran & Grossmann (1986), Yuan et al. (1988), Fletcher & Leyffer (1994)

- **LP/NLP based Branch and Bound**

Quesada and Grossmann (1992)

- **Extended Cutting Plane (ECP)**

Westerlund and Pettersson (1995)

Basic NLP subproblems

a) NLP Relaxation *Lower bound*

$$\begin{aligned} &\min Z_{LB}^k = f(x, y) \\ s.t.\ \ & g_j(x, y) \le 0 \quad j \in J \\ & x \in X,\ y \in Y_R \\ & y_i \le \alpha_i^k \ \ i \in I_{FL}^k \\ & y_i \ge \beta_i^k \ \ i \in I_{FU}^k \end{aligned} \qquad \text{(NLP1)}$$

b) NLP Fixed y^k *Upper bound*

$$\begin{aligned} &\min Z_U^k = f(x, y^k) \\ s.t.\ \ & g_j(x, y^k) \le 0 \quad j \in J \\ & x \in X \end{aligned} \qquad \text{(NLP2)}$$

c) Feasibility subproblem for fixed y^k.

$$\begin{aligned} &\min u \\ s.t.\ \ & g_j(x, y^k) \le u \quad j \in J \\ & x \in X,\ u \in R^1 \end{aligned} \qquad \text{(NLPF)}$$

Infinity-norm

Cutting plane MILP master

(Duran and Grossmann, 1986)

Based on solution of K subproblems (x^k, y^k) $k=1,...K$

Lower Bound

M-MIP

$$\min \; Z_L^K = \alpha$$

$$\left.\begin{array}{l} st \;\; \alpha \geq f(x^k, y^k) + \nabla f(x^k, y^k)^T \begin{bmatrix} x - x^k \\ y - y^k \end{bmatrix} \\ \qquad g_j(x^k, y^k) + \nabla g_j(x^k, y^k)^T \begin{bmatrix} x - x^k \\ y - y^k \end{bmatrix} \leq 0 \;\; j \in J \end{array}\right\} \; k = 1, \ldots K$$

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

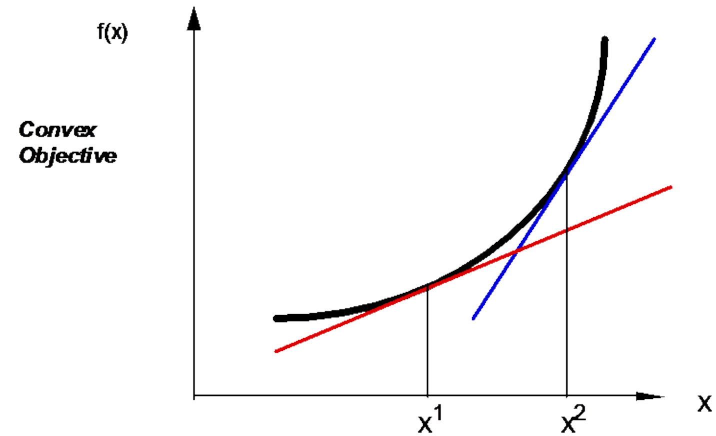
Notes:

a) Point (x^k, y^k) $k=1,...K$ normally from NLP2

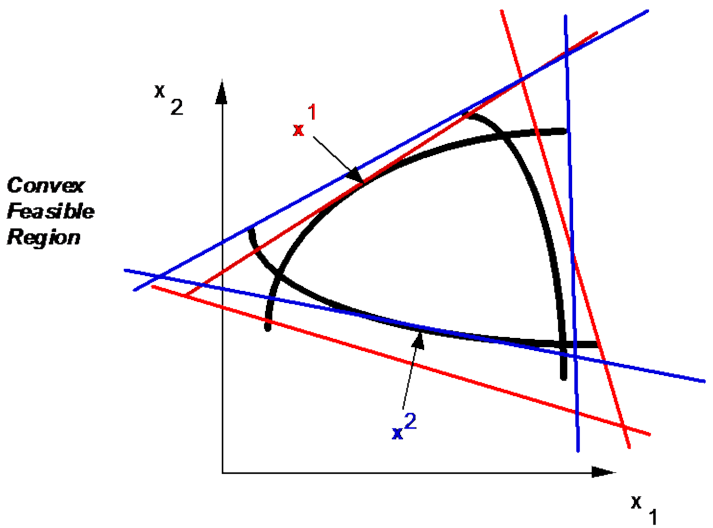
b) Linearizations *accumulated* as iterations K increase

c) Non-decreasing sequence lower bounds

Linearizations and Cutting Planes



Underestimate Objective Function



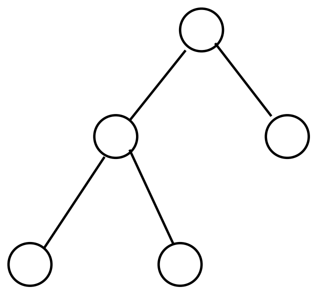
Overestimate Feasible Region

Branch and Bound

NLP1:

$$\min \; Z_{LB}^{k} = f(x,y)$$
$$s.t. \quad g_j(x,y) \le 0 \qquad j \in J$$
$$x \in X, \quad y \in Y_R$$
$$y_i \le \alpha_i^k \qquad i \in I_{FL}^k$$
$$y_i \ge \beta_i^k \qquad i \in I_{FU}^k$$

Tree Enumeration



Successive solution of NLP1 subproblems

Advantage:
Tight formulation may require one NLP1 *($I_{FL}=I_{FU}=\emptyset$)*

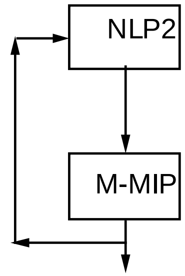
Disadvantage:
Potentially many NLP subproblems

Convergence global optimum:
Uniqueness solution NLP1 *(sufficient condition)*

Less stringent than other methods

Outer-Approximation

Alternate solution of NLP and MIP problems:



Upper bound

Lower bound

NLP2:

$$\min Z_U^k = f(x, y^k)$$
$$s.t.\ \ g_j(x, y^k) \le 0 \quad j \in J$$
$$x \in X$$

M-MIP:

$$\min \quad Z_L^K = \alpha$$
$$\left.\begin{aligned} st \quad \alpha &\ge f(x^k, y^k) + \nabla f(x^k, y^k)^T \begin{bmatrix} x - x^k \\ y - y^k \end{bmatrix} \\ g_j(x^k, y^k) &+ \nabla g_j(x^k, y^k)^T \begin{bmatrix} x - x^k \\ y - y^k \end{bmatrix} \le 0 \quad j \in J^k \end{aligned}\right\} \quad k = 1, \ldots K$$
$$x \in X, \quad y \in Y$$

Property. Trivially converges in one iteration if *f(x,y)* and *g(x,y)* are linear

- If infeasible NLP solution of feasibility NLP-F required to guarantee convergence.

Carnegie Mellon

Generalized Benders Decomposition

Benders (1962), Geoffrion (1972)

Particular case of Outer-Approximation as applied to (P1)

1. Consider Outer-Approximation at (x^k, y^k)

$$\begin{aligned} &\alpha \geq f(x^k,y^k) + \nabla f(x^k,y^k)^T \begin{bmatrix} x-x^k \\ y-y^k \end{bmatrix} \\ &g(x^k,y^k) + \nabla g_j(x^k,y^k)^T \begin{bmatrix} x-x^k \\ y-y^k \end{bmatrix} \leq 0 \quad j \in J^k \end{aligned} \qquad (1)$$

2. Obtain linear combination of (1) using Karush-Kuhn-Tucker multipliers μ^k and eliminating x variables

$$\begin{aligned} \alpha \geq f(x^k,y^k) &+ \nabla_y f(x^k,y^k)^T (y-y^k) \Big\} \\ &+ (\mu^k)^T \left[g(x^k,y^k) + \nabla_y g(x^k,y^k)^T (y-y^k) \right] \end{aligned} \qquad (2)$$

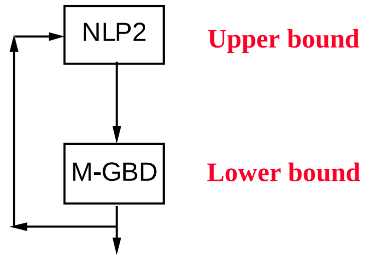
Lagrangian cut

Remark. Cut for infeasible subproblems can be derived in a similar way.

$$(\lambda^k)^T \left[g(x^k,y^k) + \nabla_y g(x^k,y^k)^T (y-y^k) \right] \leq 0$$

Generalized Benders Decomposition

Alternate solution of NLP and MIP problems:



NLP2:

$$\begin{aligned} &\min Z_U^k = f(x, y^k) \\ &s.t.\ \ g_j(x, y^k) \le 0 \quad j \in J \\ &x \in X \end{aligned}$$

M-GBD:

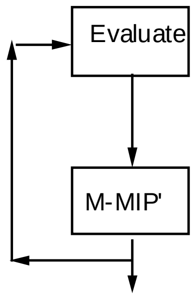
$$\begin{aligned} \min\ \ & Z_L^K = \alpha \\ st\ \ & \alpha \ \ge\ f(x^k, y^k) + \nabla_y f(x^k, y^k)^T \left(y - y^k\right) \\ & + \left(\mu^k\right)^T \left[g(x^k, y^k) + \nabla_y g\ (x^k, y^k)^T \left(y - y^k\right)\right] \quad k \in KFS \\ & \left(\lambda^k\right)^T \left[g(x^k, y^k) + \nabla_y g\ (x^k, y^k)^T \left(y - y^k\right)\right] \le 0 \quad k \in KIS \\ & y \in Y,\ \ \alpha \in R^1 \end{aligned}$$

Property 1. If problem (P1) has zero integrality gap, Generalized Benders Decomposition converges in one iteration when optimal (x^k, y^k) are found. ***Sahinidis, Grossmann (1991)***

=> *Also applies to Outer-Approximation*

Extended Cutting Plane

Westerlund and Pettersson *(1995)*



No NLP !

Add linearization most violated constraint to M-MIP

$$J^k = \{ \hat{j} \in \arg\{ \max_{j \in J} g_j(x^k, y^k) \} \}$$

Remarks.

- Can also add full set of linearizations for M-MIP

- Successive M-MIP's produce non-decreasing sequence lower bounds

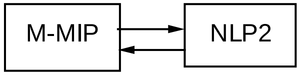
- Simultaneously optimize x^k, y^k with M-MIP

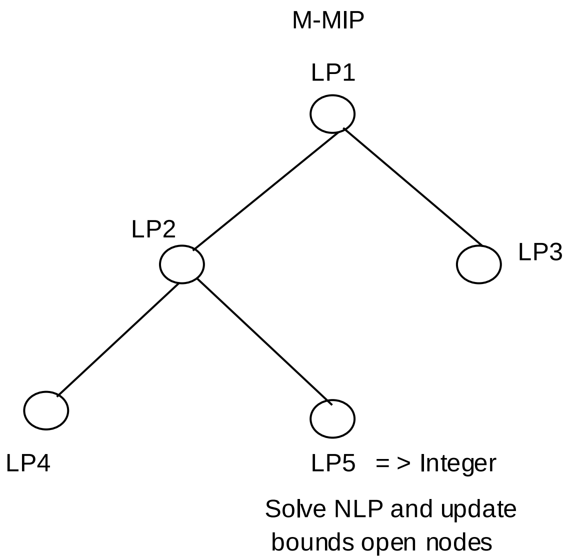
= > Convergence may be slow

LP/NLP Based Branch and Bound *(Branch & Cut)*

Quesada and Grossmann (1992)

Integrate NLP and M-MIP problems





Remark.

Fewer number branch and bound nodes for LP subproblems

May increase number of NLP subproblems

Carnegie Mellon

Numerical Example

$$\min Z = y_1 + 1.5y_2 + 0.5y_3 + x_1^2 + x_2^2$$

$$\text{s.t.} \quad (x_1 - 2)^2 - x_2 \le 0$$

$$x_1 - 2y_1 \ge 0$$

$$x_1 - x_2 - 4(1-y_2) \le 0$$

$$x_1 - (1 - y_1) \ge 0$$

$$x_2 - y_2 \ge 0 \qquad \text{(MIP-EX)}$$

$$x_1 + x_2 \ge 3y_3$$

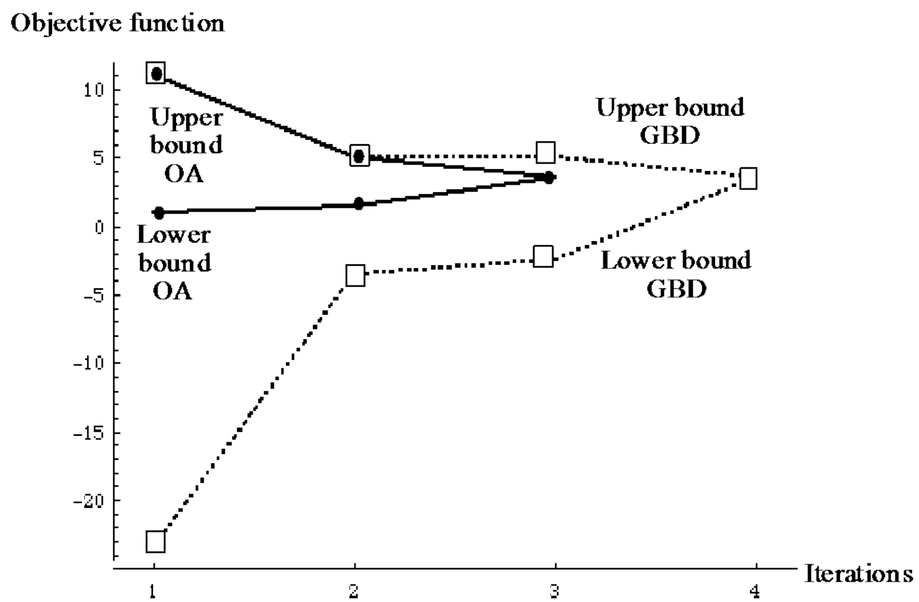
$$y_1 + y_2 + y_3 \ge 1$$

$$0 \le x_1 \le 4, \quad 0 \le x_2 \le 4$$

$$y_1, y_2, y_3 = 0, 1$$

Optimum solution: $y_1=0$, $y_2 = 1$, $y_3 = 0$, $x_1 = 1$, $x_2 = 1$, $Z = 3.5$.

Starting point $y_1 = y_2 = y_3 = 1$.

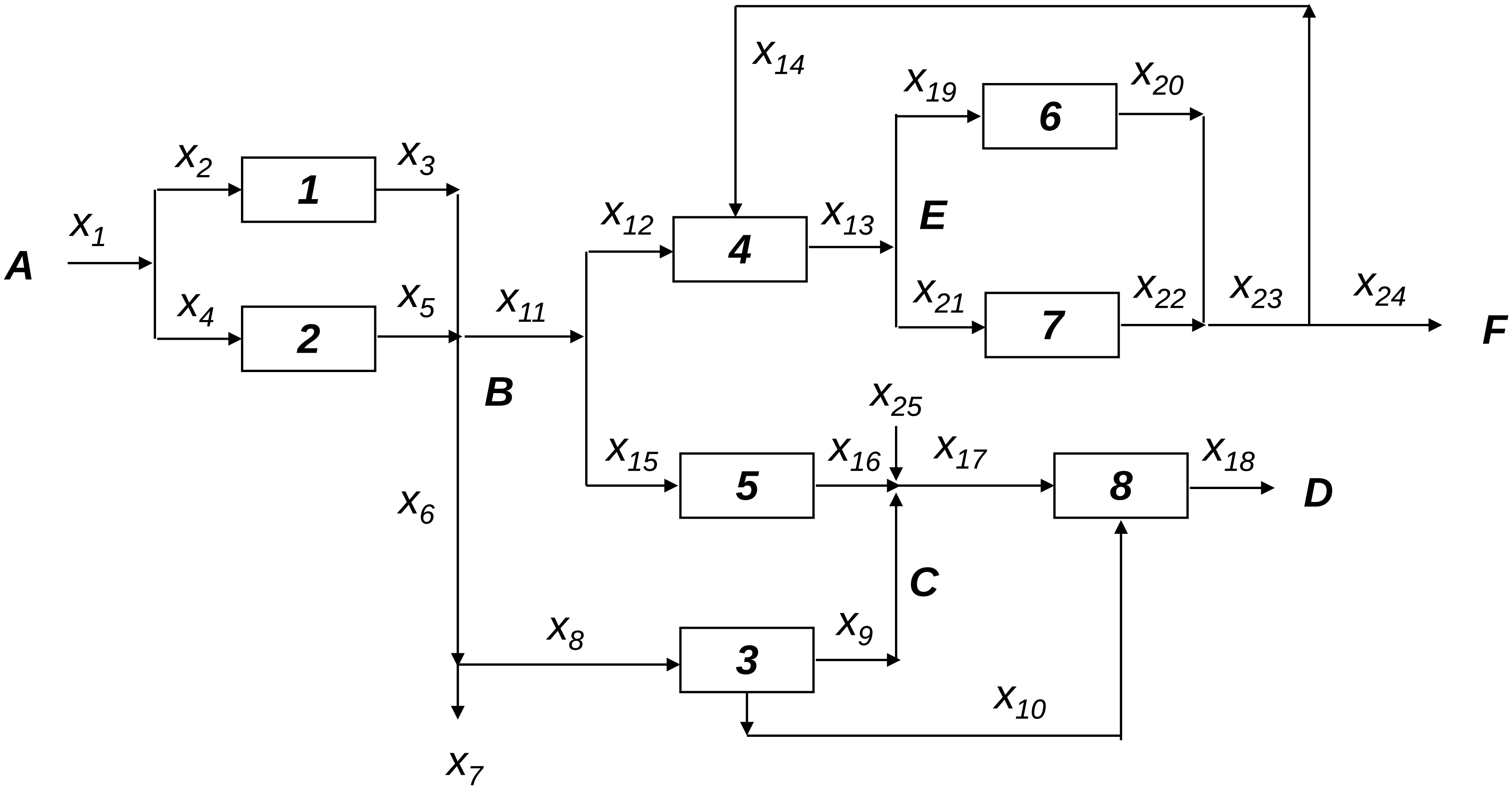


Summary of Computational Results

Method	Subproblems	Master problems (LP's solved)
BB	5 (NLP1)	
OA	3 (NLP2)	3 (M-MIP) (19 LP's)
GBD	4 (NLP2)	4 (M-GBD) (10 LP's)
ECP	-	5 (M-MIP) (18 LP's)

Example: Process Network with Fixed Charges

- *Duran and Grossmann (1986)*
  - Network superstructure



Example *(Duran and Grossmann, 1986)*

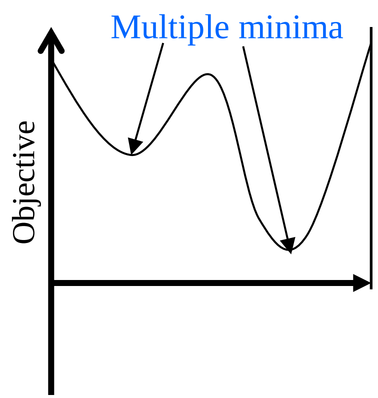
Algebraic MINLP: *linear in y, convex in x*

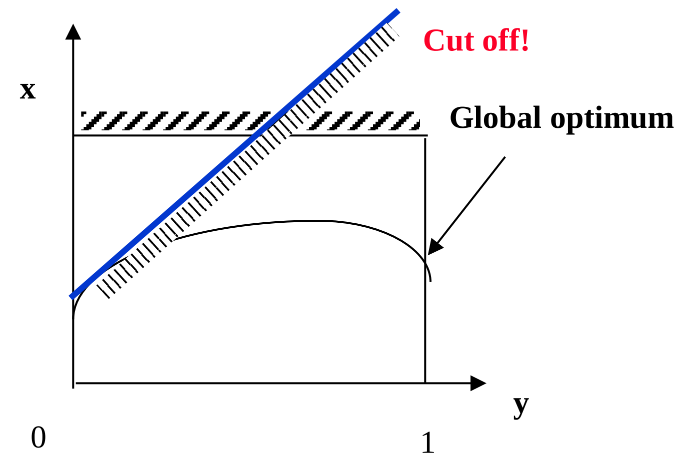
8 0-1 variables, 25 continuous, 31 constraints (5 nonlinear)

	NLP	MIP
Branch and Bound *(F-L)*	20	-
Outer-Approximation:	3	3
Generalized-Benders	10	10
Extended Cutting Plane	-	15
LP/NLP based	3	7 LP's vs 13 LP's OA

Effects of Nonconvexities

1. NLP supbroblems may have local optima
2. MILP master may cut-off global optimum





Handling of Nonconvexities

1. **Rigorous approach (global optimization):**
   Replace nonconvex terms by underestimtors/convex envelopes
   Solve convex MINLP within spatial branch and bound

2. **Heuristic approach:**
   Add slacks to linearizations
   Search until no imprvement in NLP

Handling nonlinear equations

h(x,y) = 0

1. **In branch and bound no special provision-simply add to NLPs**

2. **In GBD no special provision- cancels in Lagrangian cut**

3. **In OA equality relaxation**

$$T^k \nabla h(x^k, y^k)^T \begin{bmatrix} x - x^k \\ y - y^k \end{bmatrix} \le 0$$

$$T^k = [t_{ii}^k], \quad t_{ii}^k = \begin{cases} 1 \ \ if \ \lambda_i^k > 0 \\ -1 \ if \ \lambda_i^k < 0 \\ 0 \ \ if \ \lambda_i^k = 0 \end{cases}$$

Lower bounds may not be valid

Rigorous if eqtn relaxes as h(x,y) ≤ 0 h(x,y) is convex

MIP-Master Augmented Penalty

Viswanathan and Grossmann, 1990

Slacks: p^k, q^k with weights w^k

$$\min \quad Z^K = \alpha + \sum_{k=1}^{K} \left[w_p^k p^k + w_q^k q^k \right] \qquad \text{(M-APER)}$$

$$
\left.
\begin{aligned}
s.t. \quad & \alpha \geq f(x^k, y^k) + \nabla f(x^k, y^k)^T \begin{bmatrix} x - x^k \\ y - y^k \end{bmatrix} \\
& T^k \nabla h(x^k, y^k)^T \begin{bmatrix} x - x^k \\ y - y^k \end{bmatrix} \leq p^k \\
& g(x^k, y^k) + \nabla g(x^k, y^k)^T \begin{bmatrix} x - x^k \\ y - y^k \end{bmatrix} \leq q^k
\end{aligned}
\right\} \; k = 1, \ldots K
$$

$$\sum_{i \in B^k} y_i - \sum_{i \in N^k} y_i \leq \left| B^k \right| - 1 \quad k = 1, \ldots K$$

$$x \in X, \; y \in Y, \; \alpha \in \boldsymbol{R}^1, \; p^k, q^k \geq 0$$

If convex MINLP then slacks take value of zero
=> reduces to OA/ER

Basis DICOPT (nonconvex version)

1. Solve relaxed MINLP

2. Iterate between MIP-APER and NLP subproblem
   until no improvement in NLP

Chemical ENGINEERING

CAPD CENTER

Carnegie Mellon

Mixed-integer Nonlinear Programming

MINLP:

Algorithms

Branch and Bound (BB) *Leyffer (2001), Bussieck, Drud (2003)*
Generalized Benders Decomposition (GBD) *Geoffrion (1972)*
Outer-Approximation (OA) *Duran and Grossmann (1986)*
Extended Cutting Plane(ECP) *Westerlund and Pettersson (1992)*

Codes:

SBB *GAMS simple B&B*
MINLP-BB (AMPL)*Fletcher and Leyffer (1999)*

Bonmin (COIN-OR) *Bonami et al (2006)*
FilMINT Linderoth and Leyffer (2006)

DICOPT (GAMS) *Viswanathan and Grossman (1990)*
AOA (AIMSS)

α–ECP *Westerlund and Peterssson (1996)*
MINOPT *Schweiger and Floudas (1998)*

BARON *Sahinidis et al. (1998)* ***Global***
Couenne *Belotti & Margot (2008)* ***Global***
SCIP *ZIB (2012)* ***Global***
GLOMIQO ***Floudas and Meisner (2011)***

Note: MIQPs can be solved with CPLEX

http://www.minlp.org

IBM | Carnegie Mellon

MINLP
CYBER-INFRASTRUCTURE

Username
Password
Login

Home | Goals | Participants | Instructions | Contribute Problems | Open Problems | Libraries of Problems | Forum | Resources

CMU-IBM Cyber-Infrastructure for MINLP collaborative site

This collaborative site has as a major goal to promote the optimization of linear and nonlinear models with one or several alternative model formulations involving discrete and continuous variables through mixed-integer nonlinear programming (MINLP), or generalized disjunctive programming (GDP). Three major objectives are:

- Create a library of optimization problems that can be generally formulated as MINLP/GDP models.
- Provide high level descriptions of the problems with one or several model formulations with corresponding input files for one or several instances.
- Allow users to pose open problems that are unsolved and with unknown or tentative formulations

$$\begin{aligned} &\min Z = f(x,y) \\ &s.t. \quad g(x,y) \le 0 \\ &\qquad x \in X,\ y \in Y \end{aligned}$$

We invite researchers and practitioners to contribute to the library of problems and models, and to participate in the discussions on these problems. We look forward to collaborating with you!

About us

- Goals of our project
- Participants of the project

Contribute

- Create an account
- Learn how to contribute problems
- Contribute solved problems, models, and instanes to our library
- Post open unsolved problems

Our library

- View our library of problems
- Discuss problems in the forums

Resources

- Conferences
- Lectures and Tutorials

Carnegie Mellon

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**

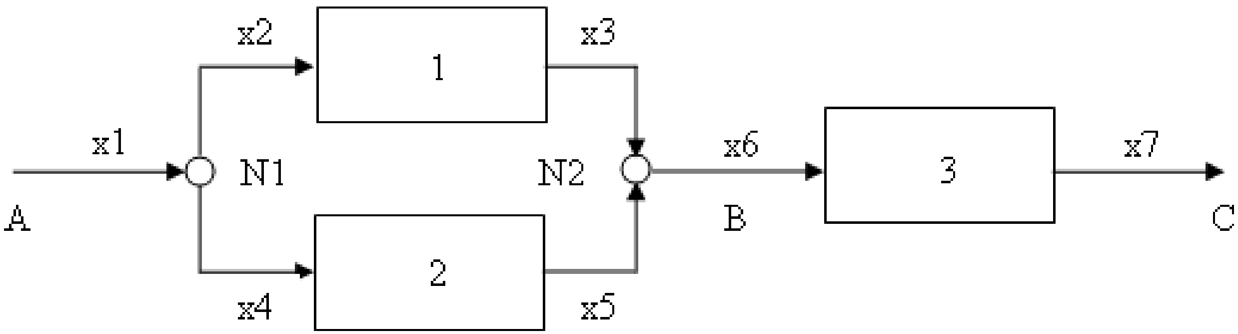
$$x \in R^n, c_k \in R^1$$ **Continuous Variables**

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

Lee, Grossmann (2000)

Big-M MINLP (BM)

- **MINLP reformulation of GDP**

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

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

$$g_{jk}(x) \le M_{jk}(1 - \lambda_{jk}) \quad ,j \in J_k, k \in K$$

Big-M Parameter

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

$$A\lambda \le a$$

Logic constraints

Williams (1990)

$$x \ge 0, \lambda_{jk} \in \{0,1\}$$

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
    - Stubbs and Mehrotra (1999)
    - Generalization of Balas (1979)

Hull relaxation: intersection of convex hull of each disjunction

Remarks

1. Perspective function $\boldsymbol{h(v, \lambda) = \lambda\, g(v/\lambda)}$

If g(x) is a bounded convex function,

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

Hiriart-Urruty and Lemaréchal (1993)

$\boldsymbol{h(\nu, 0) = 0}$ for bounded g(x)

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

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

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.

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

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

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

Remark

For linear disjunctions

$$\bigvee_{j \in J_k} \left[A_{jk} x \le b_{jk} \right]$$

Convex-hull *set* $g_{jk}(x) = A_{jk} x - b_{jk}$

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

$$A_{jk} \nu_{jk} \le b_{jk} \lambda_{jk} \quad j \in J_k$$

Balas (1985)

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

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

Hull Relaxation Problem (HRP)

HRP:

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

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

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

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

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

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

Convex Hull each disjunction

$$A\lambda \le a$$

Logic constraints

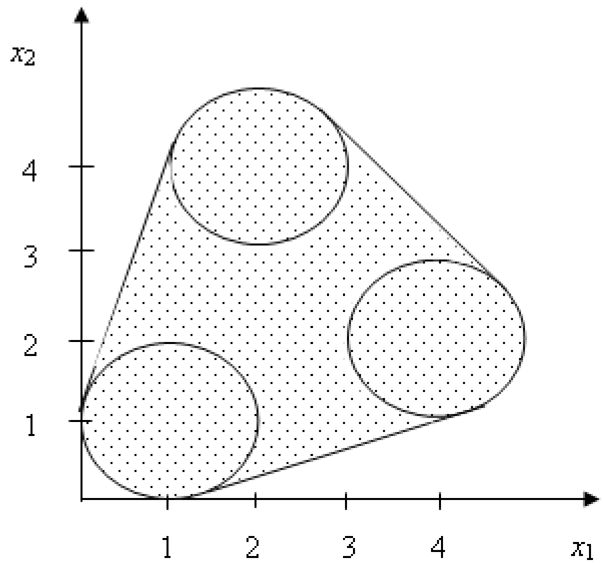
$$x, \nu^{jk} \ge 0, \; 0 \le \lambda_{jk} \le 1, \; j \in J_k, k \in K$$

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

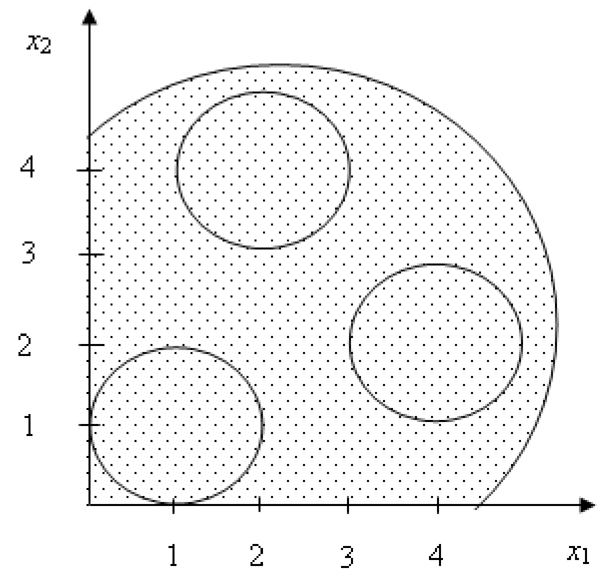
Strength Lower Bounds

- **<u>Theorem</u>:** ***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

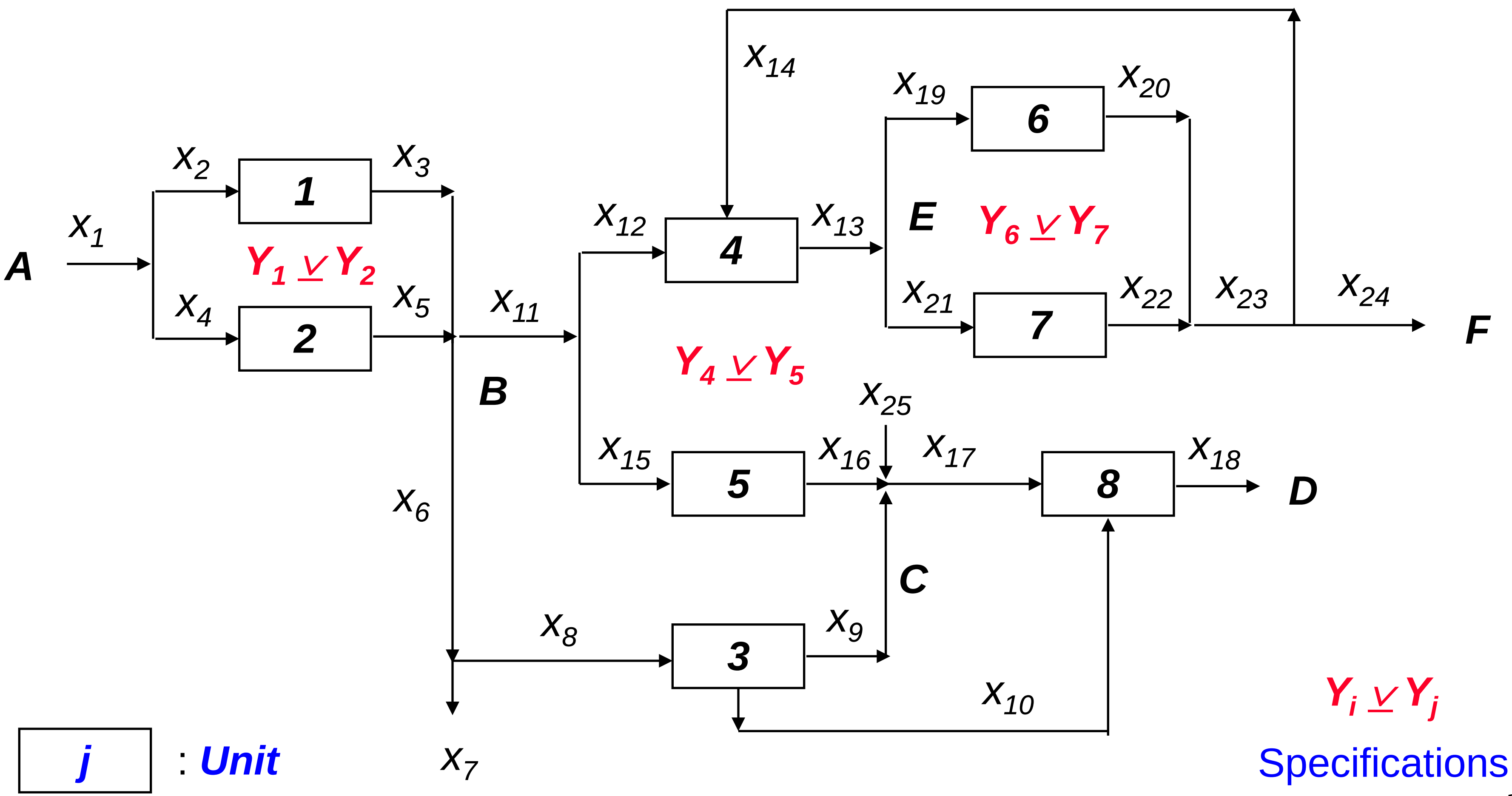
GDP

- **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**

Process Network with Fixed Charges

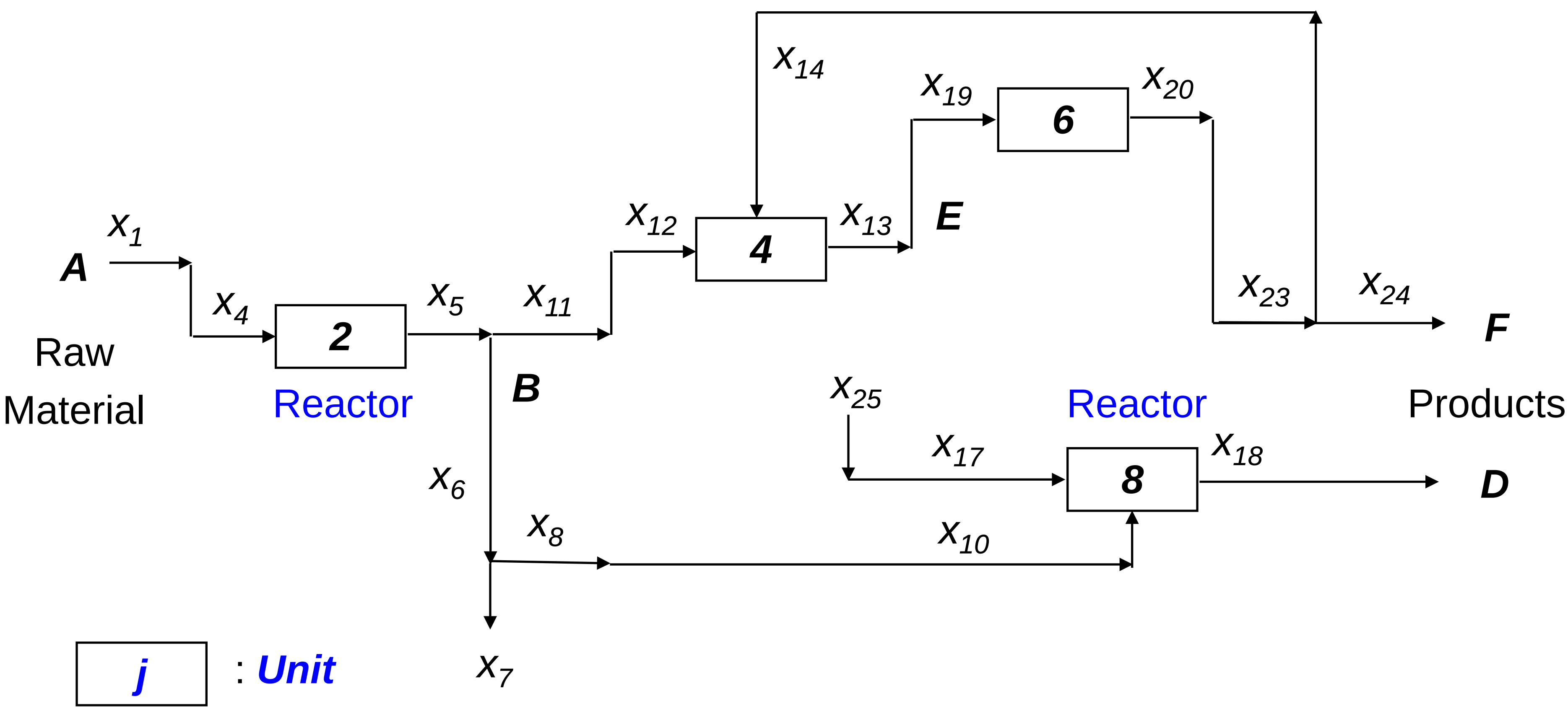
- *Türkay and Grossmann (1997)*

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

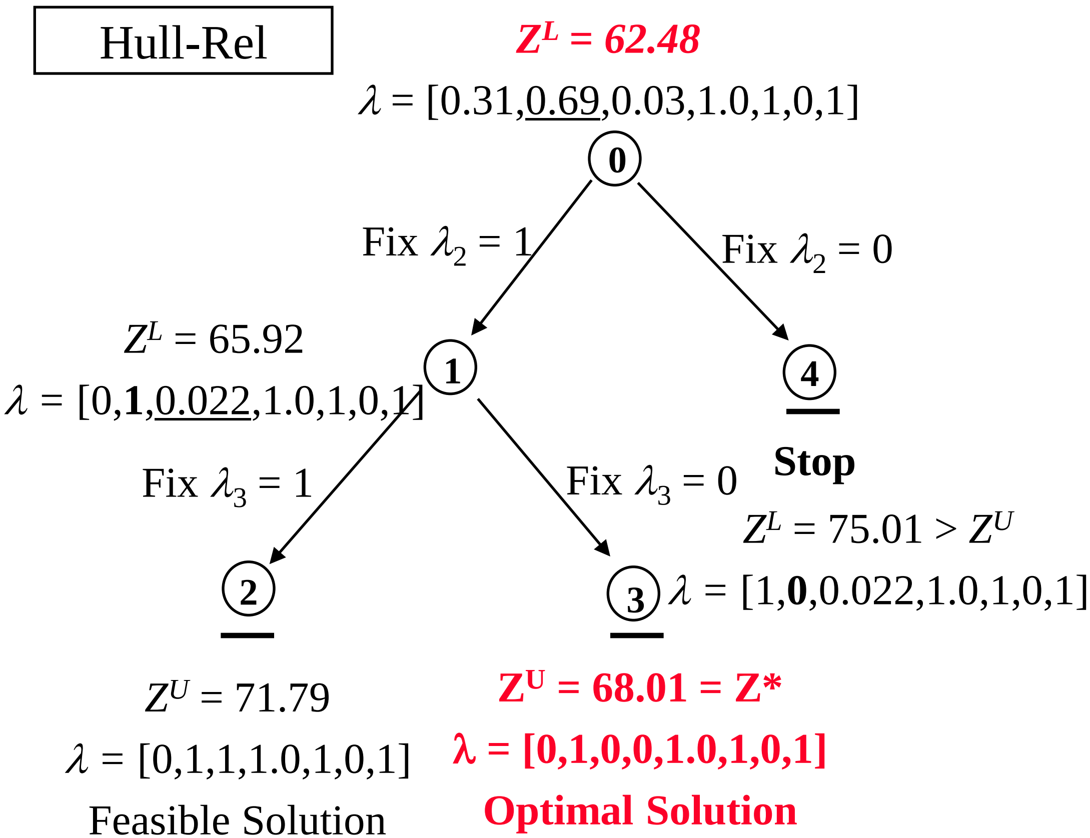


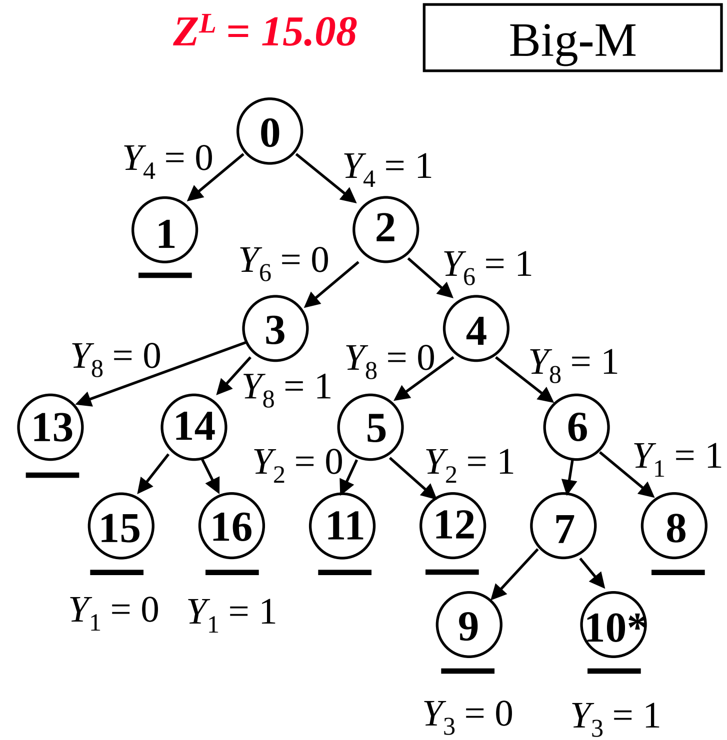
Optimal solution

- **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, Grossmann (2012)

GDP

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

DP

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

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

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

→ ***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:

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

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

We can then rewrite

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

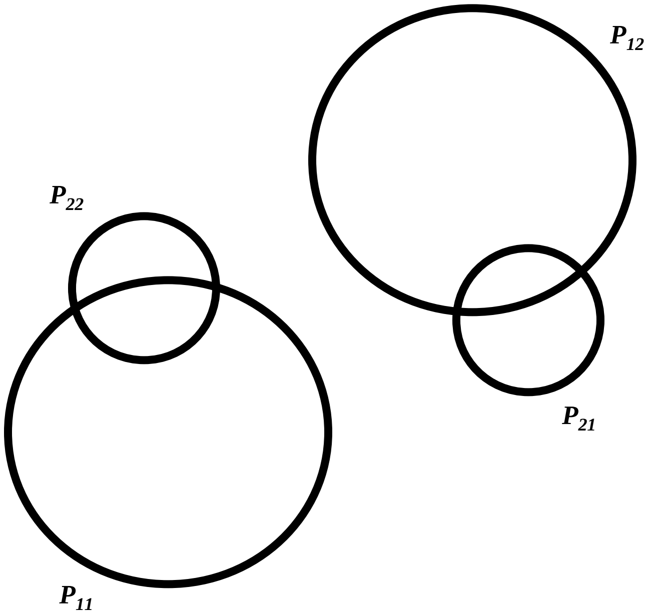
which is its equivalent DNF

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

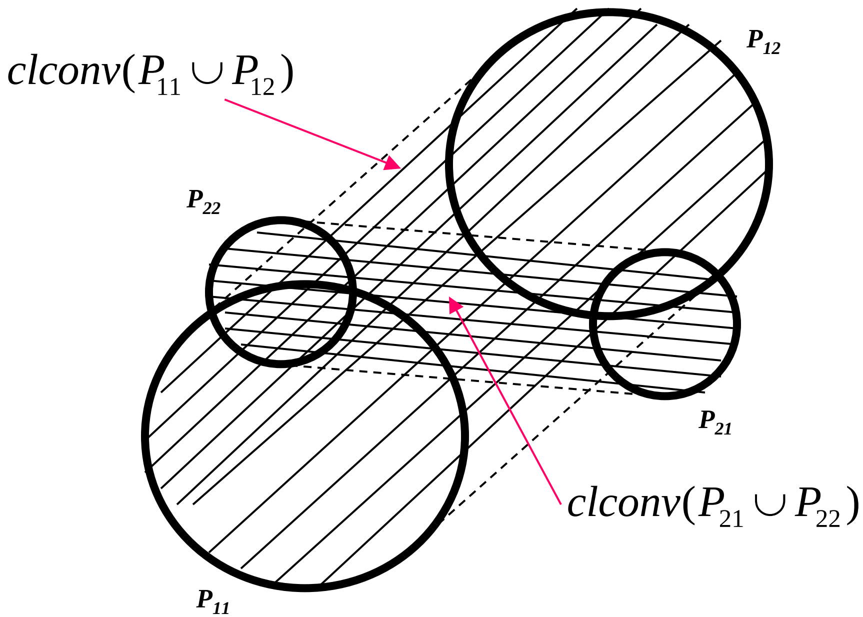


Hierarchy of Relaxations for Convex Disjunctive Programs

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

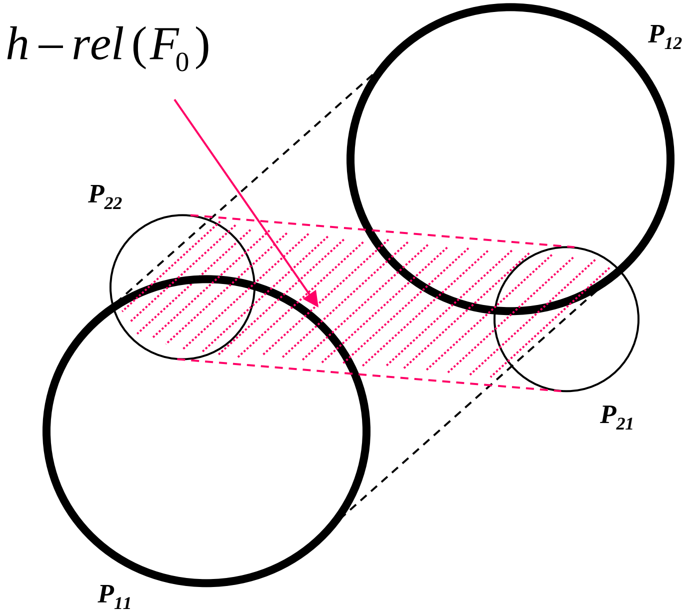


Hierarchy of Relaxations for Convex Disjunctive Programs

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



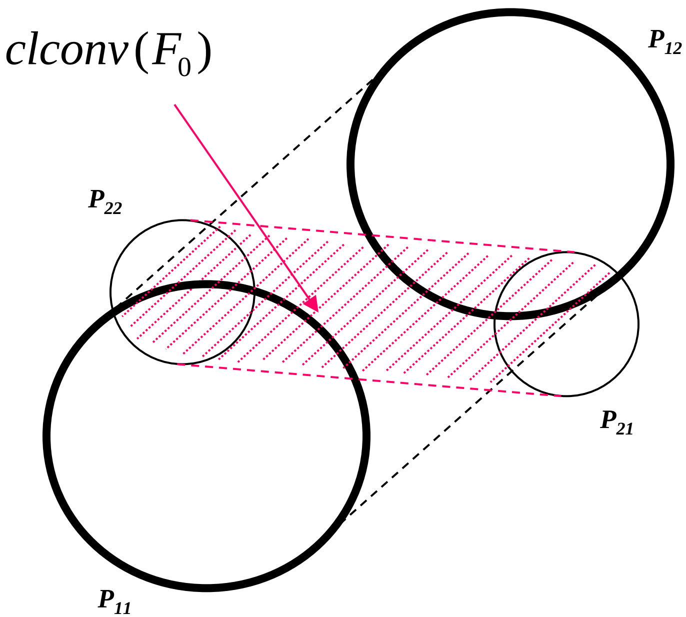
No Basic Step Applied => HR

Hierarchy of Relaxations for Convex Disjunctive Programs

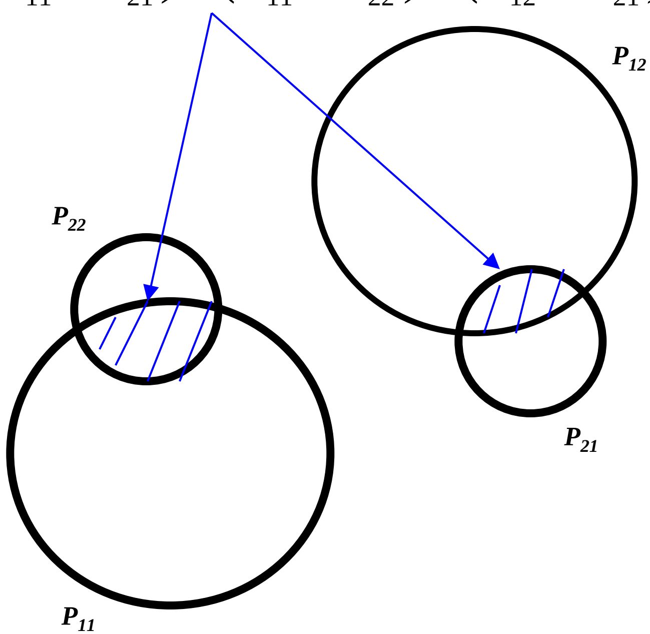
Theorem 2.4. *For* $i = 1, 2....k$ *let* $F_i = \bigcap_{k \in K} S_k$ *be a sequence of regular forms of a disjunctive set such that* F_i *is obtained from* F_{i-1} *by the application of a basic step, then:*

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

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



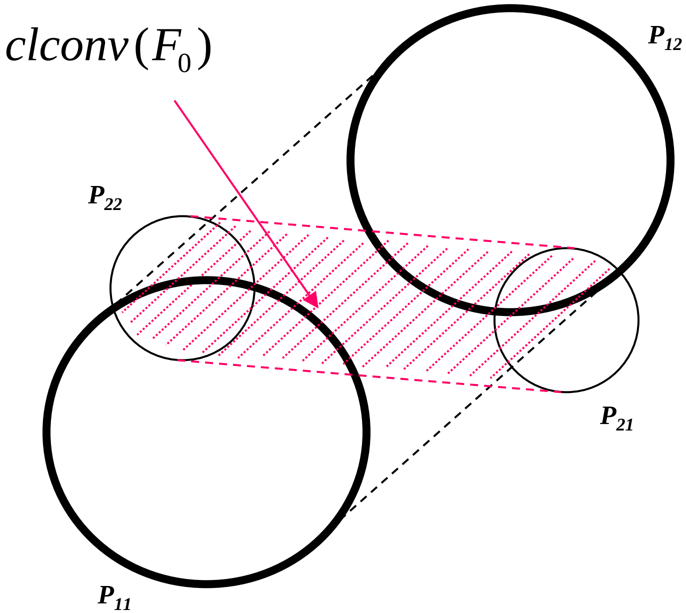
Basic Step Applied

Hierarchy of Relaxations for Convex Disjunctive Programs

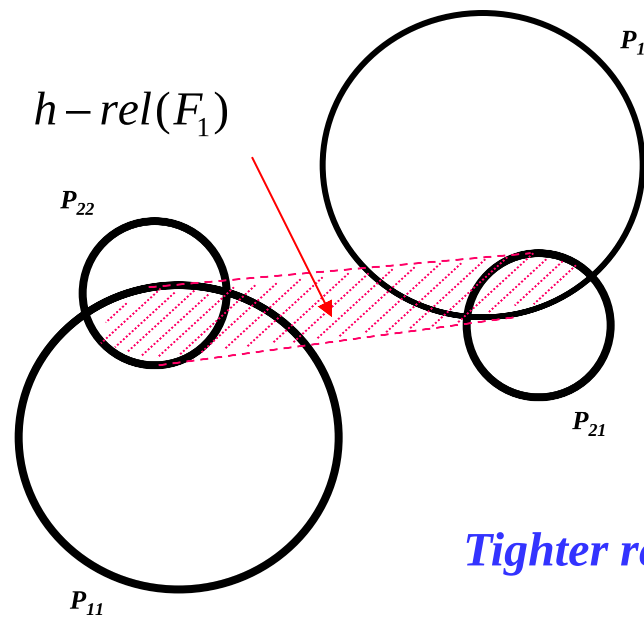
Theorem 2.4. *For* $i = 1, 2....k$ *let* $F_i = \bigcap_{k \in K} S_k$ *be a sequence of regular forms of a disjunctive set such that* F_i *is obtained from* F_{i-1} *by the application of a basic step, then:*

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

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



No Basic Step Applied => HR



Tighter relaxation!

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!

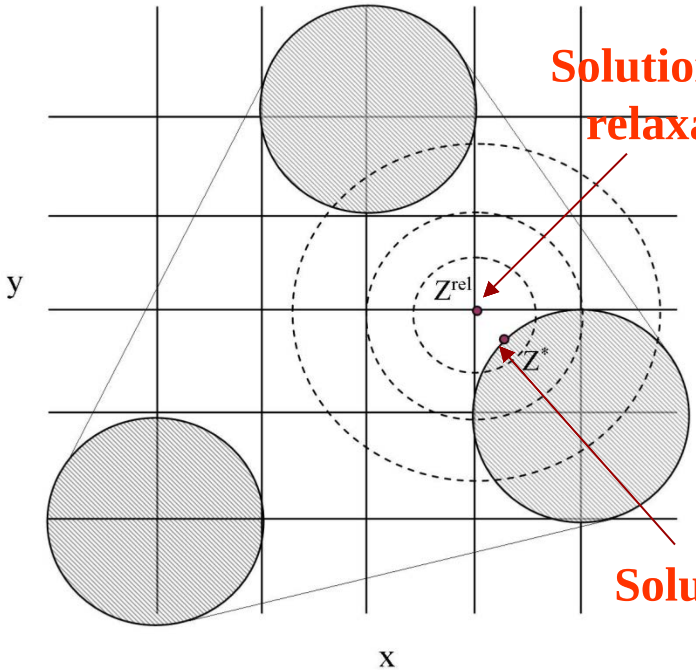
Similar to convexification of MILPs

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

For DP/MINLP: ***Soares, Ceria (1999);*** *Implicit in Stubbs and Mehrotra (1999)*

Convex nonlinear program equivalent to a convex disjunctive program

Illustrative Example



Disjunctive Program

$$\min \quad Z = (x_1 - 3)^2 + (x_2 - 2)^2 + 1$$

$$s.t.$$

$$[x_1^2 + x_2^2 \leq 1] \vee [(x_1 - 4)^2 + (x_2 - 1)^2 \leq 1]$$

$$\vee [(x_1 - 2)^2 + (x_2 - 4)^2 \leq 1]$$

$$|x_i| \leq 5 \quad i \in 1, 2$$

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

Carnegie Mellon

Convex nonlinear program equivalent to a convex disjunctive program

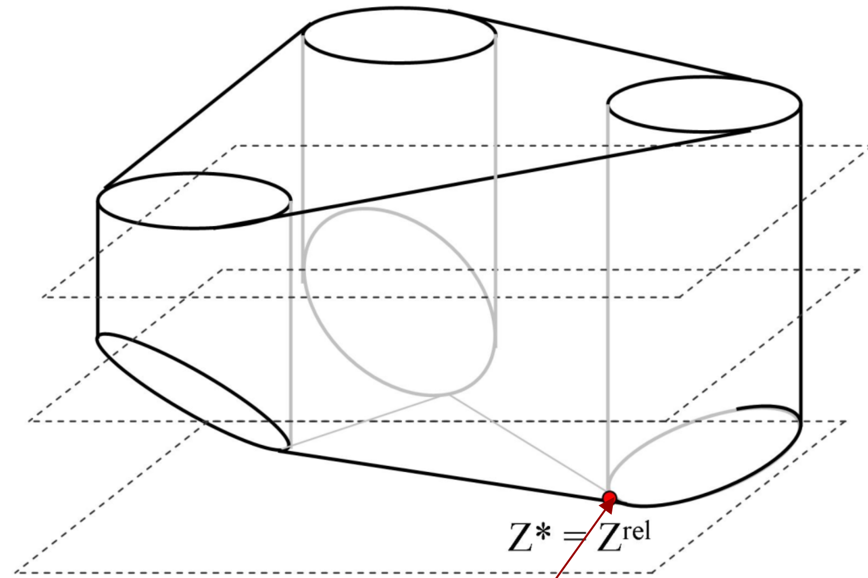
Illustrative Example

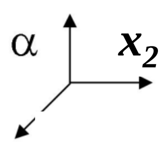
Disjunctive Program

Place objective as constraint and intersect with disjunction

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

DNF!





Solution of the program and its relaxation

$Z = 1.172$

(3.293,1.707)

Solution of the hull relaxation of DNF (NLP) is the <u>same</u> 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 disjunction 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

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

Constraints after basic steps →

Set of disjunctions after basic steps ← T_n

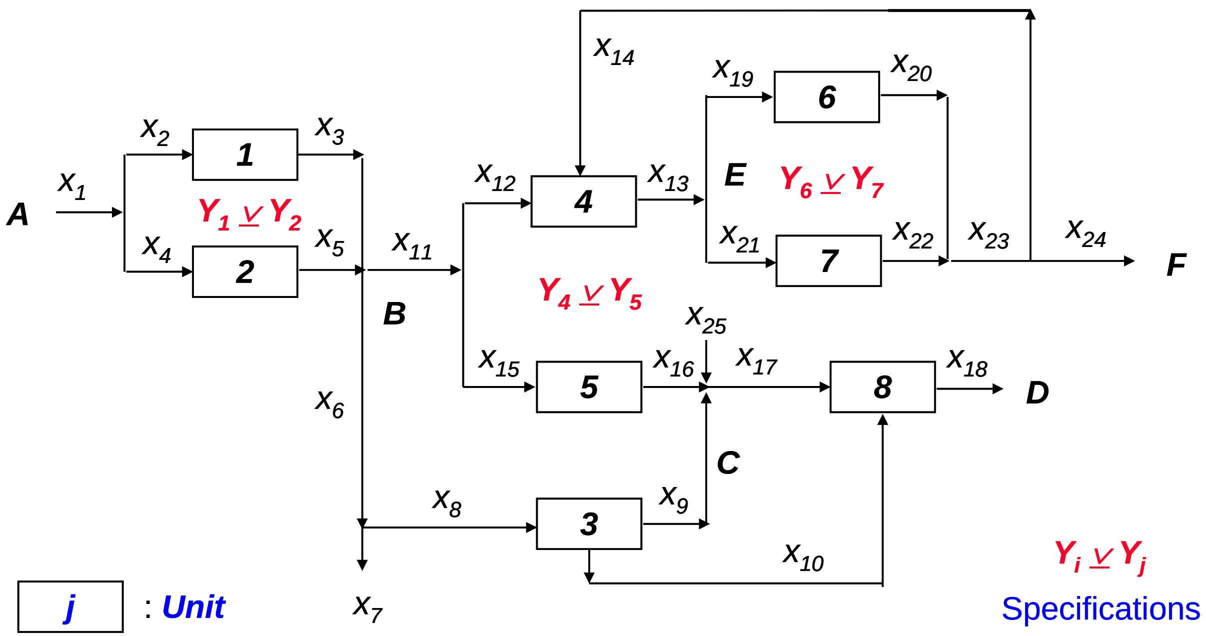
No new binary variables are created **(Balas, 1985)** →

Set of disjunctions before basic steps ← T_d

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	109	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			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

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			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 50%probs

Numerical Results

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

Table: Performance using different reformulation strategies

Example	Opt.	BM Approach LB	BM Nds	BM T(s)	HR Approach LB	HR Nds	HR T(s)	Proposed Approach LB	Proposed Nds	Proposed 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

Cutting Planes for Linear Generalized Disjunctive Programming

GDP Model:

Sawaya, Grossmann (2004)

$$Min\ Z = \sum_{k \in K} c_k + h^T x$$

Objective Function

$$s.t. \quad Bx \le b$$

Common Constraints

OR Operator →

$$\bigvee_{j \in J_k} \begin{bmatrix} Y_{jk} \\ A_{jk} x \le a_{jk} \\ c_k = \gamma_{jk} \end{bmatrix} \quad k \in K$$

Disjunctive Constraints

$$\Omega(Y) = True$$

Logic Constraints

$$x \in R^n,\ Y_{jk} \in \{True, False\},\ c_k \in R$$

$$j \in J_k,\ k \in K$$

Boolean Variables

Reformulations as MILP

Big-M

$$\min Z = \sum_{k\in K}\sum_{j\in J_k}\gamma_{jk}\lambda_{jk} + h^T x$$

Big-M parameters

$$\begin{aligned}
s.t.\quad & Bx \le b \\
& A_{jk}x - a_{jk} \le M_{jk}(1-\lambda_{jk}) && j\in J_k,\ k\in K \\
& \sum_{j\in J_k}\lambda_{jk} = 1 && k\in K \\
& D\lambda \le d \\
& x\in R^n,\ \lambda_{jk}\in\{0,1\} && j\in J_k,\ k\in K
\end{aligned}$$

(BM)

Convex Hull

$$\min Z = \sum_{k\in K}\sum_{j\in J_k}\gamma_{jk}\lambda_{jk} + h^T x$$

Disaggregated variables

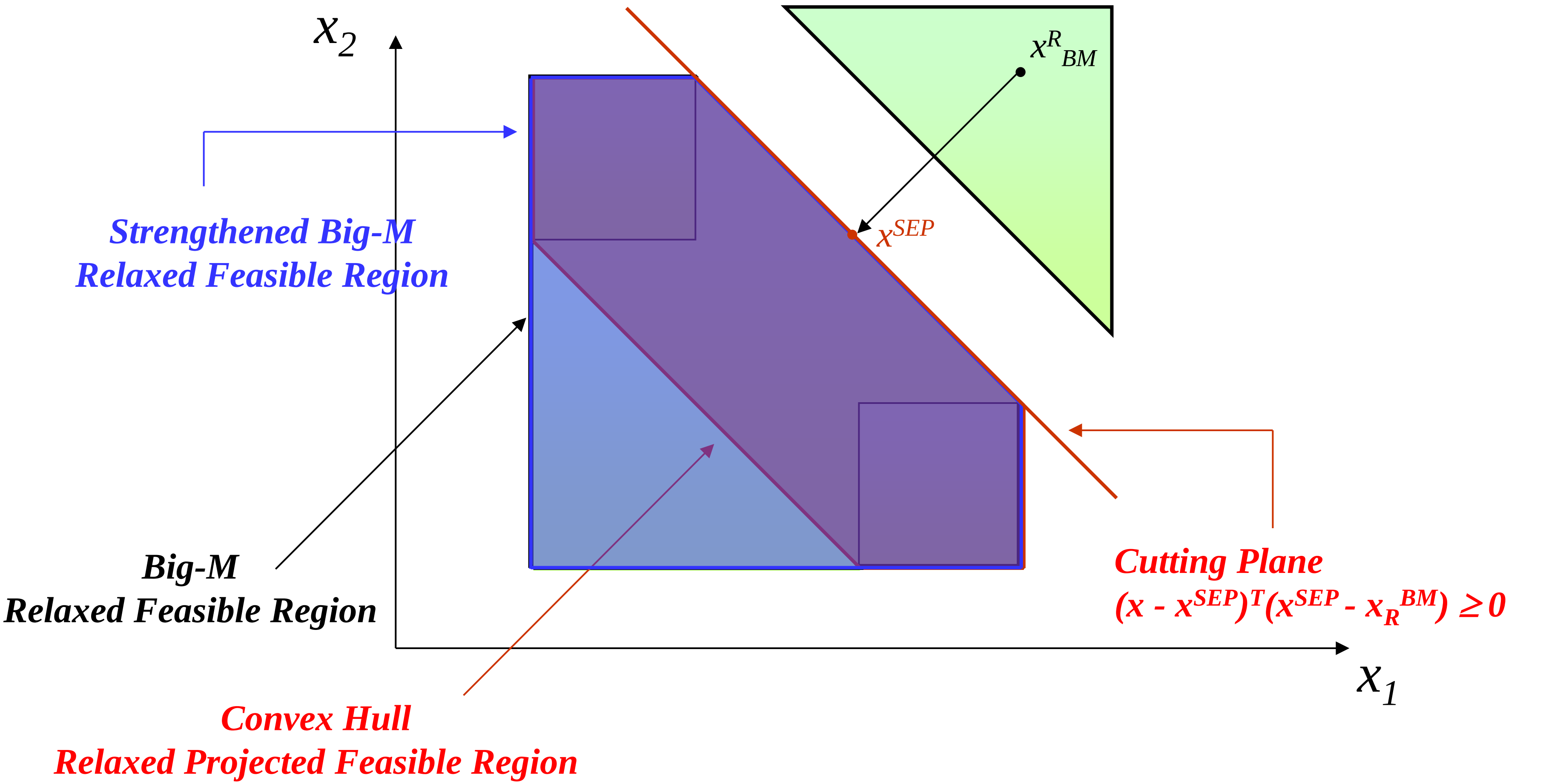
$$\begin{aligned}
s.t.\quad & Bx \le b \\
& A_{jk}\nu^{jk} - a_{jk}\lambda_{jk} \le 0 && j\in J_k,\ k\in K \\
& x = \sum_{j\in J_k}\nu^{jk} && k\in K \\
& 0 \le \nu^{jk} \le \lambda_{jk}U_{jk} && j\in J_k,\ k\in K \\
& \sum_{j\in J_k}\lambda_{jk} = 1 && k\in K \\
& D\lambda \le d \\
& x\in R^n,\ \nu^{jk}\in R^n_+,\ \lambda_{jk}\in\{0,1\} && j\in J_k,\ k\in K
\end{aligned}$$

(CH)

Motivation for Cutting Plane Method

Proposition: The projected relaxation of (CH) onto the space of (BM) is always as tight or tighter than that of (BM) ***(Grossmann I.E. , S. Lee, 2003)***

Trade-off: Big-M fewer vars/weaker relaxation vs Convex-Hull tighter relaxation/more vars



Cutting Plane Method

1. **Solve relaxed Big-M MILP x^R_{BM}**

2. **Solve separation problem: find point x^{SEP} closest to x^R_{BM}**
 Feasible region corresponds to relaxed Convex Hull.

$$
\begin{aligned}
&\textit{Min } Z = \Phi(x) && \text{(SEP)}\\
&\text{s.t.} \quad Bx \le b &&\\
&\quad A_{jk}\nu^{jk} - a_{jk}\lambda_{jk} \le 0 && j \in J_k,\ k \in K\\
&\quad x = \sum_{j \in J_k} \nu^{jk} && k \in K\\
&\quad 0 \le \nu^{jk} \le \lambda_{jk} U_{jk} && j \in J_k,\ k \in K\\
&\quad \sum_{j \in J_k} \lambda_{jk} = 1 && k \in K\\
&\quad D\lambda \le d &&\\
&\quad x \in R^n,\ \nu^{jk} \in R^n_+,\ 0 \le \lambda_{jk} \le 1 && j \in J_k,\ k \in K
\end{aligned}
$$

Note: $\Phi(x)$ can be represented by either the Euclidean norm (? $x - x_R^{BM}$?) (NLP) or the Infinity norm ($\max_i |x_i - x_{iR}^{BM}|$) (LP).

3. **Cutting plane is generated and added to relaxed big-M MILP.**

4. **Solve strengthened relaxed Big-M MILP. Go to 2.**

Cutting Plane Method: Different Cuts

Proposition: There exists a vector ξ such that

$$\xi^T (z^{SEP} - z^{BM}) \geq 0$$

is a valid linear inequality, where ξ is a subgradient of $\Phi(z)$ at z^{SEP}.

Note: $z=(x,\lambda)$

Proposition: (1) Let $\Phi(z) \equiv \| z - z^{BM} \|_2 \equiv (z - z^{BM})^T (z - z^{BM})$. Then,

$$\xi \equiv \nabla \Phi = (z - z^{BM})$$

(2) Let $\Phi(z) \equiv \| z - z^{BM} \|_\infty \equiv max_i | z_i - z_i^{BM} |$. Then,

$$\xi \equiv (\mu^+ - \mu^-)$$

$$\begin{aligned} & Min\ u & & \text{Lagrange Multipliers} \\ & s.t\quad u \geq z_i - z_i^{BM} \quad i \in I & \leftarrow & \quad \mu^+ \\ & \qquad u \geq -z_i + z_i^{BM} \quad i \in I & \leftarrow & \quad \mu^- \\ & \qquad \text{Feasible region of (SEP)} & & \end{aligned}$$

(3) Let $\Phi(z) \equiv \| z - z^{BM} \|_1 \equiv \Sigma | z_i - z_i^{BM} |$. Then,

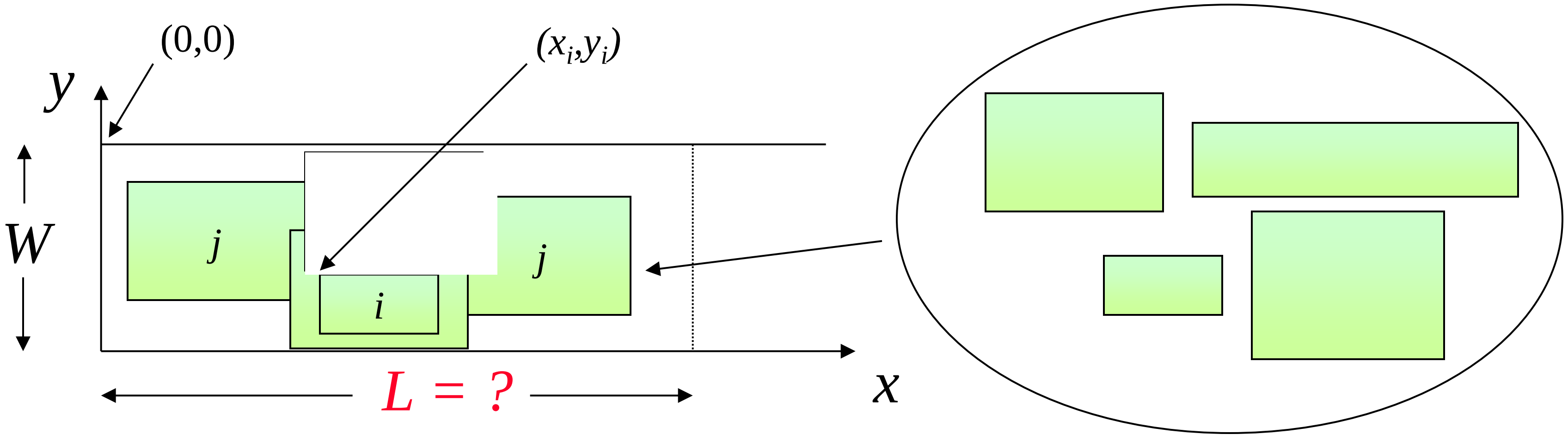
$$\xi \equiv (\mu^+ - \mu^-)$$

$$\begin{aligned} & Min\ \Sigma u_i & & \text{Lagrange Multipliers} \\ & s.t\quad u_i \geq z_i - z_i^{BM} \quad i \in I & \leftarrow & \quad \mu^+ \\ & \qquad u_i \geq -z_i + z_i^{BM} \quad i \in I & \leftarrow & \quad \mu^- \\ & \qquad \text{Feasible region of (SEP)} & & \end{aligned}$$

Strip-packing Problem

Problem statement: *Hifi M. (1998)*

- We need to fit a set of small rectangles with width w_i and length l_i onto a large rectangular strip of fixed width W and unknown length L. The objective is to fit all small rectangles onto the strip without overlap and rotation while minimizing length L of the strip.



Set of small rectangles

GDP Model For Strip-packing Problem

$$Min\ Z = L \qquad \text{(SP-GDP)}$$

$$s.t.\quad L \geq x_i + l_i \qquad i \in N$$

$$\begin{bmatrix} Y^1_{ij} \\ x_i + l_i \leq x_j \end{bmatrix} \vee \begin{bmatrix} Y^2_{ij} \\ x_j + l_j \leq x_i \end{bmatrix} \vee \begin{bmatrix} Y^3_{ij} \\ y_i - h_i \geq y_j \end{bmatrix} \vee \begin{bmatrix} Y^4_{ij} \\ y_j - h_j \geq y_i \end{bmatrix} \qquad i,j \in N,\ i < j$$

$$0 \leq x_i \leq U_i - l_i \qquad i \in N$$

$$h_i \leq y_i \leq W \qquad i \in N$$

$$x_i, y_i \in R \qquad i \in N$$

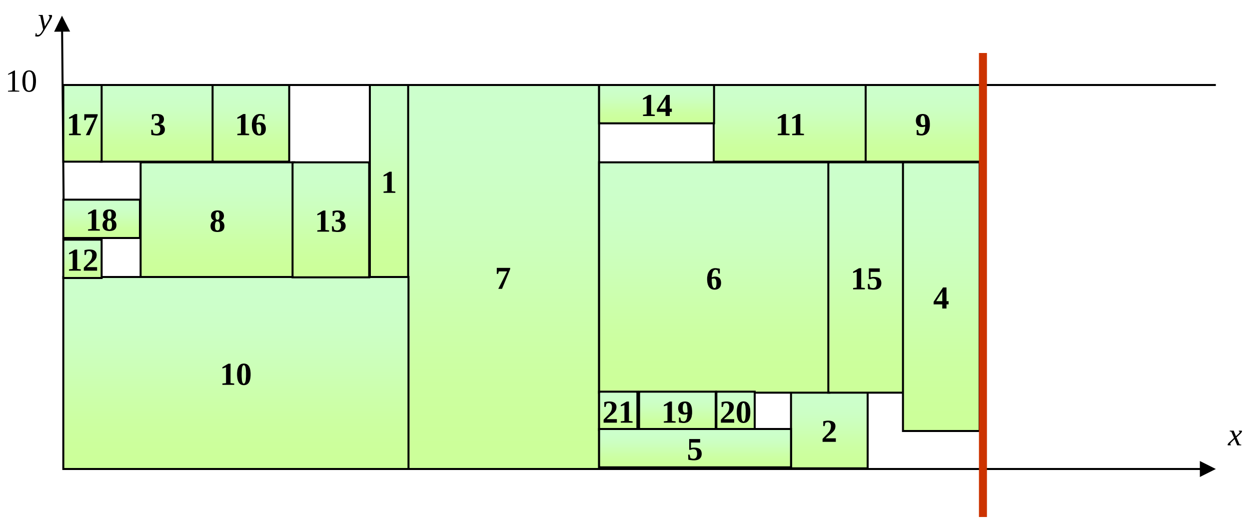
$$Y^1_{ij}, Y^2_{ij}, Y^3_{ij}, Y^4_{ij} \in \{True, False\} \qquad i,j \in N,\ i < j$$

21-rectangle Strip-packing Problem

Problem Size

	Total number of constraints	Total number of variables	Number of discrete variables
Convex Hull	5272	4244	840
Big-M	1072	884	840

Solution



Optimal Length: 24

Numerical Results

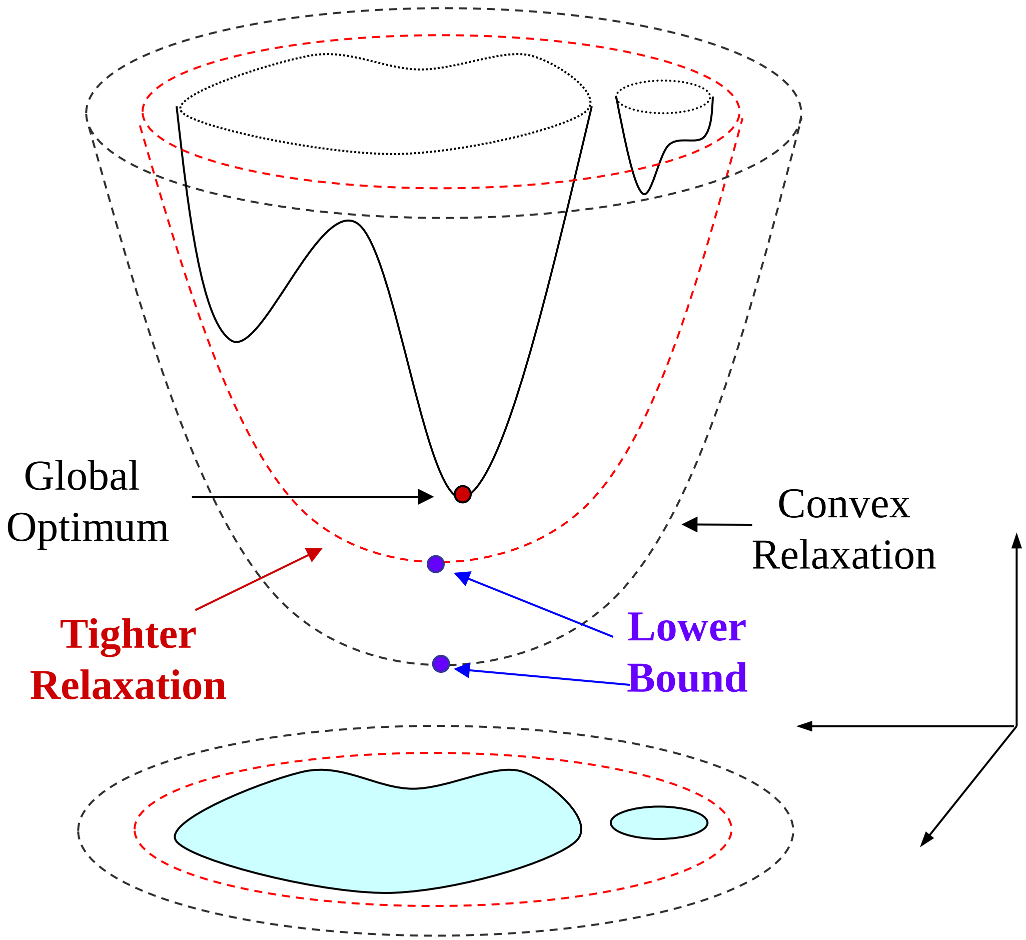
(CPLEX v. 8.1, default MIP options turned on)

	Relaxation	Optimal Solution	Gap (%)	Total Nodes in MIP	Solution Time for Cut Generation (sec)	*Total Solution Time (sec)
Convex Hull	9.1786	---	---	968 652	0	>10 800
Big-M	9	24	62.5	1 416 137	0	4 093.39
Big-M + 20 cuts	9.1786	24	61.75	306 029	3.74	917.79
Big-M + 40 cuts	9.1786	24	61.75	547 828	7.48	1 063.51
Big-M + 60 cuts	9.1786	24	61.75	28 611	11.22	79.44
Big-M + 62 cuts	9.1786	24	61.75	32 185	11.59	91.4

* Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP

Results also for retrofit, scheduling problems

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

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

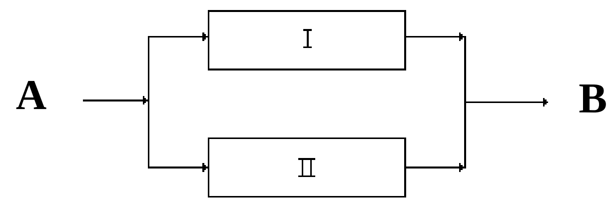
Strengthen relaxation

Apply basic steps

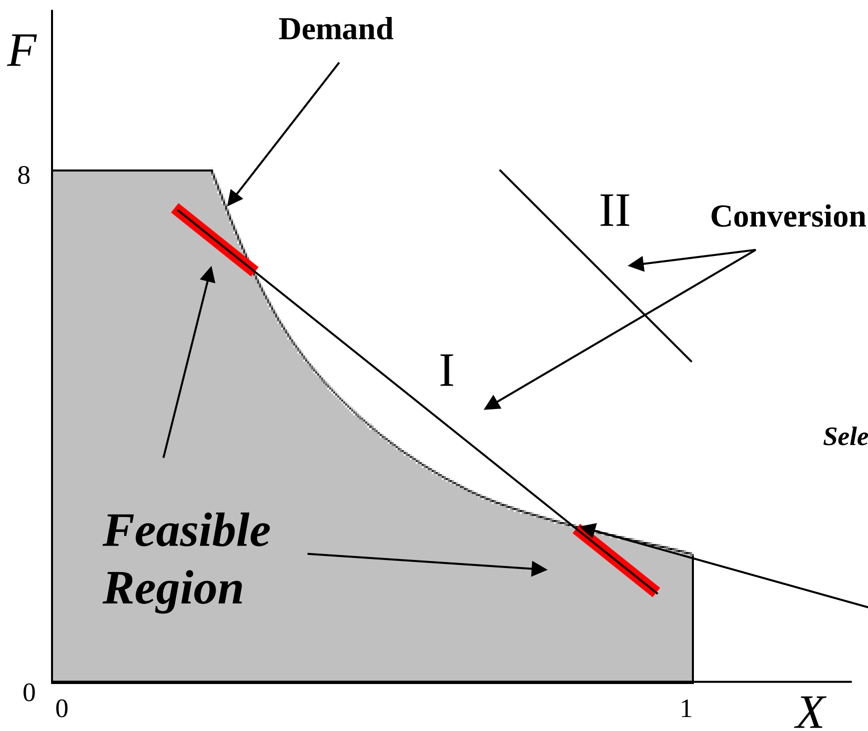
Remarks

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)

$$Min\ Z = -qFX + gF + CP$$

Demand constraint

$$s.t.\quad FX \le d$$

Selection Reactor

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

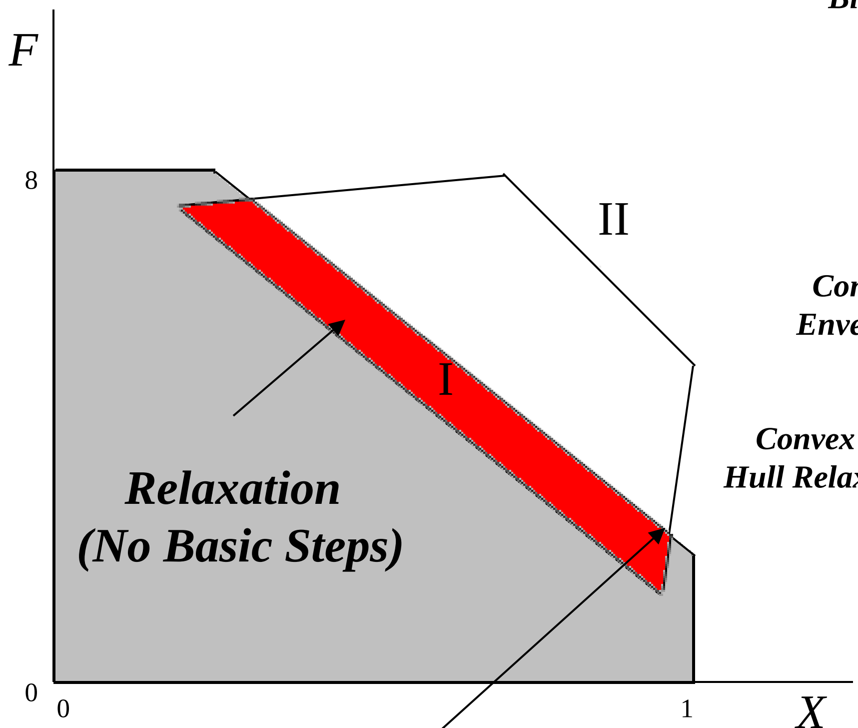
$$Y_{11} \vee Y_{21} = True$$

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

Optimum Z* = -1.01

Illustrative Example: Optimal reactor selection I

Lee & Grossmann (2003) Relaxation



Lower bound Z* = -1.28 < -1.01

Bilinear Terms

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

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

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

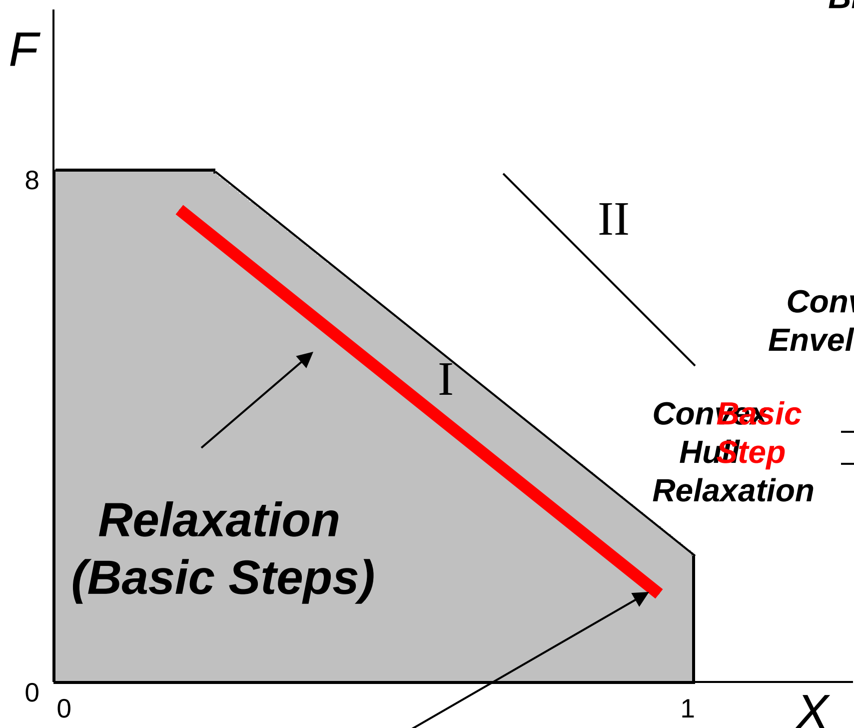
$$\begin{bmatrix} Y_{11} \\ F = \alpha_1 X + \beta_1 \\ X_1^{LO} \le X \le X_1^{UP} \\ CP = Cp_1 \end{bmatrix} \vee \begin{bmatrix} Y_{21} \\ F = \alpha_2 X + \beta_2 \\ X_2^{LO} \le X \le X_2^{UP} \\ CP = Cp_2 \end{bmatrix}$$

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

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

Illustrative Example: Optimal reactor selection I

Proposed Relaxation



Convex Hull Relaxation

Basic Step

Convex Envelopes

Bilinear Terms

$\min Z = -\theta FX - \gamma F + CP$

s.t. $FX \le d$

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

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

$CP, X, F \in R$

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

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

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

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

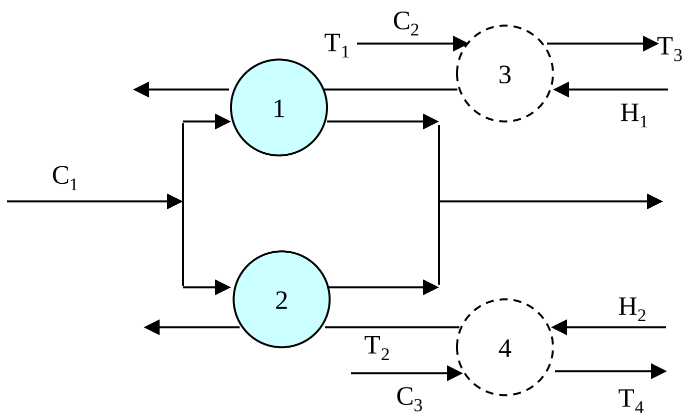
Reference

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.

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

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

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

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

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

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

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

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

Prediction of Lower Bounds Global Optimum

		Global Optimum	Lower Bound Hull Relaxation	Lower Bound Basic Steps	DNF Lower Bound
Bilinear Concave	*React 1*	-1.01	-1.28	-1.10	-1.10
	React 2	6.31	5.65	6.08	6.08
	HEN	114384.78	91671.18	94925.77	97858.86
	Water	1214.87	400.66	431.90	431.90
	Pool	-4640	-5515	-5468	-5241
Linear Fractional, Posynomial, Exponential	*Process 1*	18.61	11.85	16.01	16.01
	Process 2	19.48	12.38	17.07	17.07
	RXN 1	42.89	-337.5	-320.0	-320.0
	RXN 2	76.47	22.5	40.0	40.0
	HEN 1	48531	38729.3	48230	48531

Lower bounds improved in all cases ***Ave. increase 22%***

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

Global Optimization Methodology

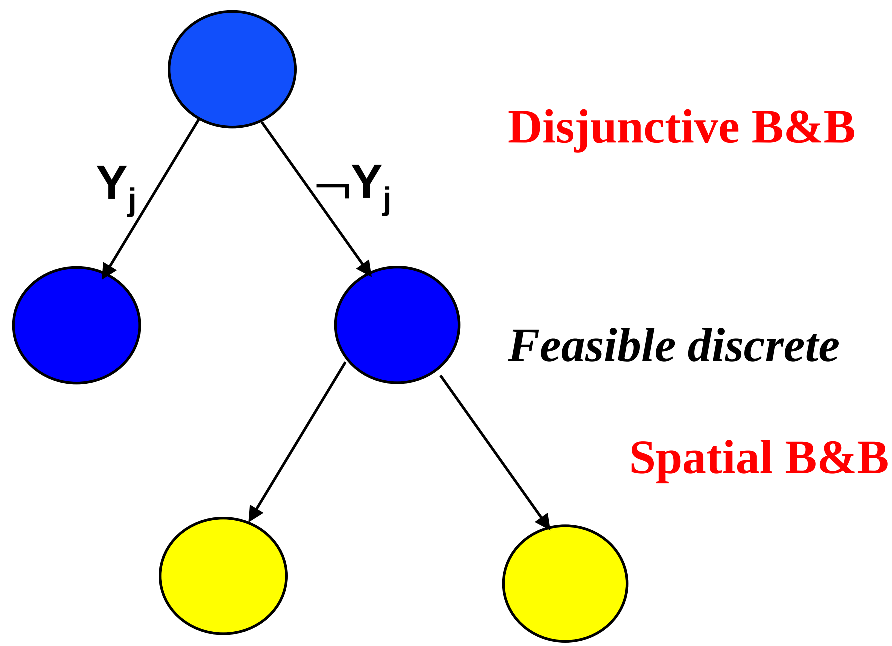
GDP reformulation
Apply basic steps following the rules presented

↓

Bound Contraction
(Zamora & Grossmann, 1999)

↓

Spatial Branch and Bound
(Lee & Grossmann, 2001)



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

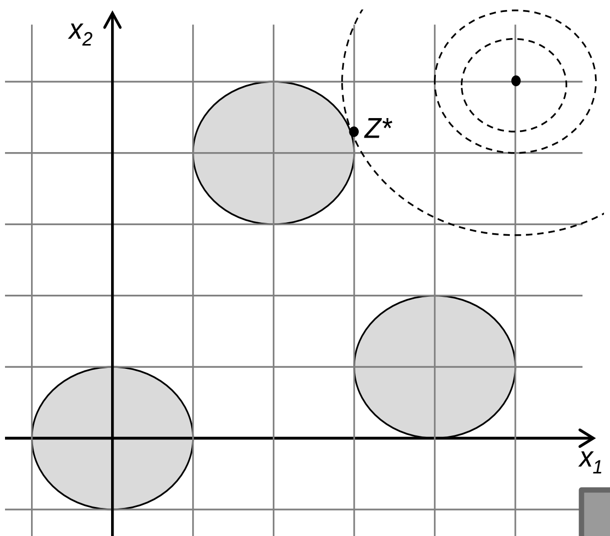
	Size of the LP Relaxation (Hull Relaxation)		Size of the LP Relaxation (Proposed)	
	Constraints	Variables	Constraints	Variables
Example 1	23	15	28	15
Example 2	24	14	31	18
Example 3	87	52	206	106
Example 4	544	346	3424	1210
Example 5	3336	1777	4237	1777

Software Implementation GDP

Extended Mathematical Programming

(GAMS-EMP) syntax (big-M or HR)

Example



$$\min \ Z = (x_1 - 5)^2 + (x_2 - 5)^2$$

$st.$

$$\begin{bmatrix} Y_1 \\ x_1^2 + x_2^2 \leq 1 \end{bmatrix} \veebar \begin{bmatrix} Y_2 \\ (x_1 - 4)^2 + (x_2 - 1)^2 \leq 1 \end{bmatrix}$$

$$\veebar \begin{bmatrix} Y_3 \\ (x_1 - 2)^2 + (x_2 - 4)^2 \leq 1 \end{bmatrix}$$

$$-5 \leq x_1, x_2 \leq 5$$

$$Y_1, Y_2, Y_3 \in \{True, False\}$$

Model definition: variables and equations

Writes EMP model,

Write disjunctions:
First term: "disjunction"
Last term: "else"
All others: "elseif"

Equations not written here are global

Binary variables are automatically assigned

Finish writing emp model and solve

Syntax 1: All default = (HR)

```
variables x1,x2,z;

equation obj, f1, f2, f3;

obj.. z =e= sqr(x1-5) + sqr(x2-5);
f1.. sqr(x1) + sqr(x2) =l= 1;
f2.. sqr(x1-4) + sqr(x2-1) =l= 1;
f3.. sqr(x1-2) + sqr(x2-4) =l= 1;

x1.lo = -5; x2.lo = -5; x1.up = 5; x2.up = 5;

model circles /all/;
```

```
file emp / '%emp.info%' /
```

```
put emp / "disjunction * " f1
        / "elseif * " f2
        / "else " f3;
```

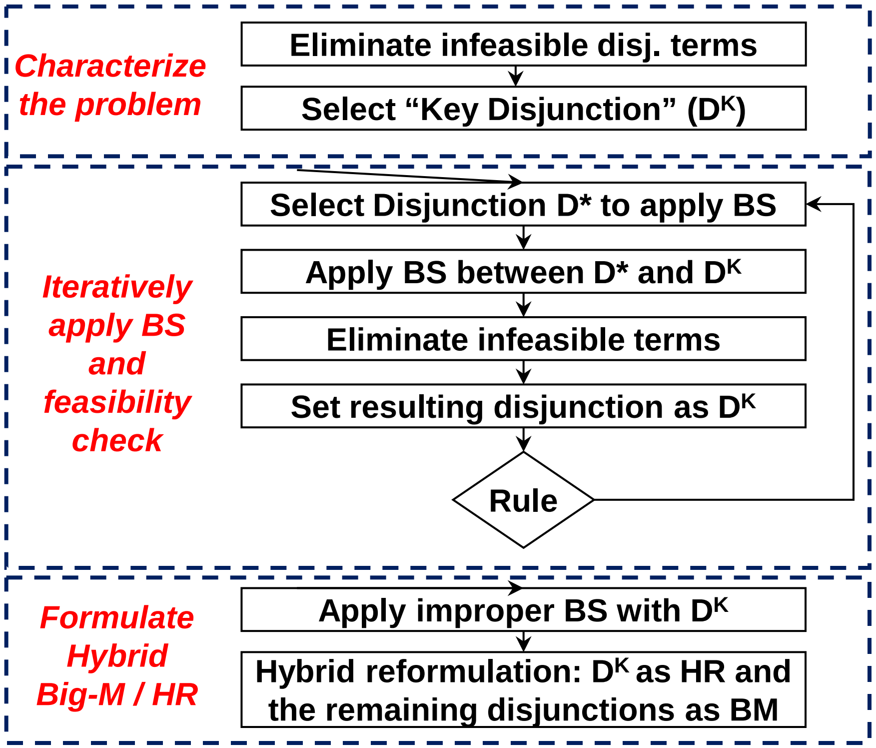
```
putclose emp;

solve circles using emp min z;
```

Reformulation algorithm from GDP to MI(N)LP

Algorithm consists of 3 stages

Algorithm

(Trespalacios, Grossmann, 2013)



Reduction of terms through a preprocessing

Preprocessing allows us to reduce problem size and identify better bounds

Preprocess: HR when each disjunctive term is TRUE

10.0

$$min\ 20 - x_1 - x_2$$

s.t..

True *True* *True*

$$\begin{bmatrix} x_2 = 8 + x_1 \\ x_2 = 12 - x_1 \end{bmatrix} \veebar \begin{bmatrix} x_1 \le 5 \\ x_2 \ge 6 \\ x_2 \le x_1 + 5 \end{bmatrix} \veebar \begin{bmatrix} x_1 \ge 9 \\ x_2 \le 5 \\ x_2 \ge x_1 - 8 \end{bmatrix}$$

$$\begin{bmatrix} 4 \le x_1 \le 7 \\ 7 \le x_2 \le 8 \end{bmatrix} \veebar \begin{bmatrix} 7 \le x_1 \le 11 \\ 2 \le x_2 \le 4 \end{bmatrix}$$

Hull Reformulation Relaxation

Solution: 9.5
(BM) lower bound: 4.0
(HR) lower bound: 8.5

Characteristic value of each disjunction

Lower bound

	Min obj
$[Infeas] \veebar [9.5] \veebar [10]$	9.5
$[8.5] \veebar [10.5]$	8.5

After preprocessing

$$min\ 20 - x_1 - x_2$$

s.t..

$$\begin{bmatrix} x_1 \le 5 \\ x_2 \ge 6 \\ x_2 \le x_1 + 5 \end{bmatrix} \veebar \begin{bmatrix} x_1 \ge 9 \\ x_2 \le 5 \\ x_2 \ge x_1 - 8 \end{bmatrix}$$

$$\begin{bmatrix} 4 \le x_1 \le 7 \\ 7 \le x_2 \le 8 \end{bmatrix} \veebar \begin{bmatrix} 7 \le x_1 \le 11 \\ 2 \le x_2 \le 4 \end{bmatrix}$$

New lower bound: **9.5**

vs. initial H-Ref: **8.5**

We eliminate one disjunctive term and obtain a stronger lower bound!

With hybrid GDP reformulation it is possible to exploit advantages of Hull-R and Big-M

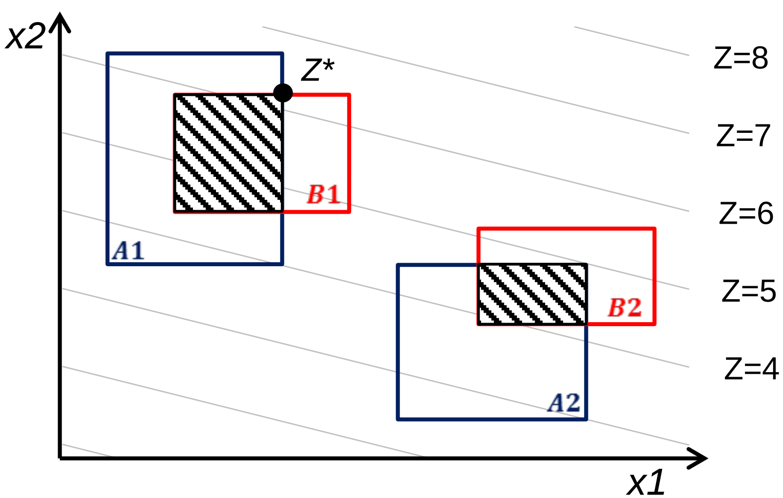
Maximization example

Max $Z = x1 + 4x2$

s.t

$[A1] \veebar [A2]$

$[B1] \veebar [B2]$



Different reformulations

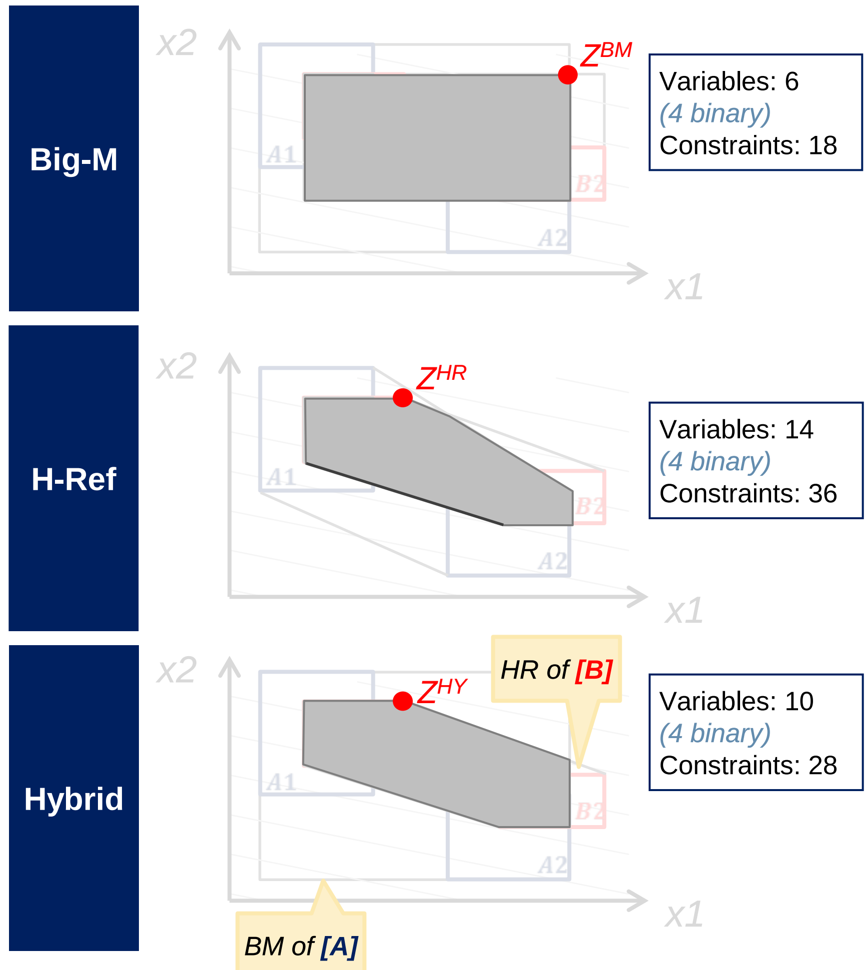
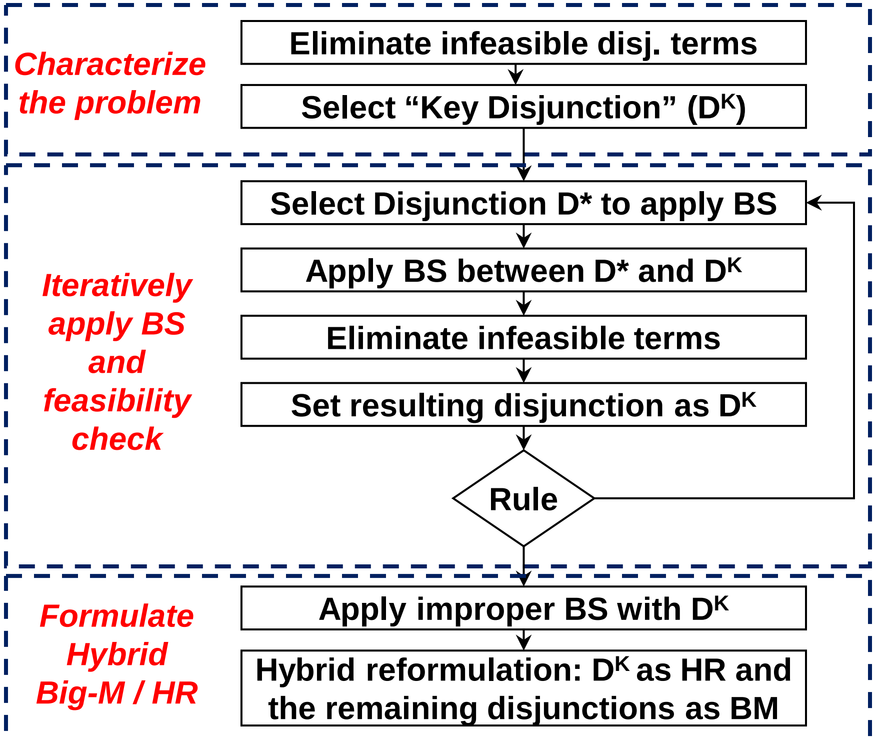


Illustration 3 stages of Algorithm

GDP example shows the application of these, and improvement in relaxation

Algorithm



[Rule]. In this example: H-Ref improves after 2 BS, then continue iterating, else stop. Other rules, such as limiting the # of terms could be used

Example (Ex5)

$\min lt$

s.t.

$lt \geq x_4 + 3$

$$\begin{bmatrix} lt \geq x_1 + 6 \\ lt \geq x_2 + 5 \\ lt \geq x_3 + 4 \\ x_1 + 6 \leq x_2 \\ x_1 + 6 \leq x_3 \\ x_2 + 5 \leq x_3 \end{bmatrix} \veebar \begin{bmatrix} lt \geq x_1 + 6 \\ lt \geq x_2 + 5 \\ lt \geq x_3 + 4 \\ x_1 + 6 \leq x_2 \\ x_1 + 6 \leq x_3 \\ x_3 + 4 \leq x_2 \end{bmatrix} \veebar \begin{bmatrix} lt \geq x_1 + 6 \\ lt \geq x_2 + 5 \\ lt \geq x_3 + 4 \\ x_2 + 5 \leq x_1 \\ x_3 + 4 \leq x_1 \\ x_2 + 5 \leq x_3 \end{bmatrix} \veebar \begin{bmatrix} lt \geq x_1 + 6 \\ lt \geq x_2 + 5 \\ lt \geq x_3 + 4 \\ x_2 + 5 \leq x_1 \\ x_3 + 4 \leq x_1 \\ x_3 + 4 \leq x_2 \end{bmatrix}$$

$$[x_1 + 6 \leq x_4] \veebar [x_4 + 3 \leq x_1] \veebar [y_1 - 6 \geq y_4] \veebar [y_4 - 3 \geq y_1]$$

$$[x_2 + 5 \leq x_4] \veebar [x_4 + 3 \leq x_2] \veebar [y_2 - 7 \geq y_4] \veebar [y_4 - 3 \geq y_2]$$

$$[x_3 + 4 \leq x_4] \veebar [x_4 + 3 \leq x_3] \veebar [y_3 - 5 \geq y_4] \veebar [y_4 - 3 \geq y_3]$$

$0 \leq lt \leq 18$
$0 \leq x_1 \leq 12$
$0 \leq x_2 \leq 13$
$0 \leq x_3 \leq 14$
$0 \leq x_4 \leq 15$
$6 \leq y_1 \leq 10$
$7 \leq y_2 \leq 10$
$5 \leq y_3 \leq 10$
$3 \leq y_4 \leq 10$

Solution: 15.0
HR lower bound: 8.3
Big-M lower bound: 6.0

Algorithm lower bound: 15.0

Same as optimal solution!

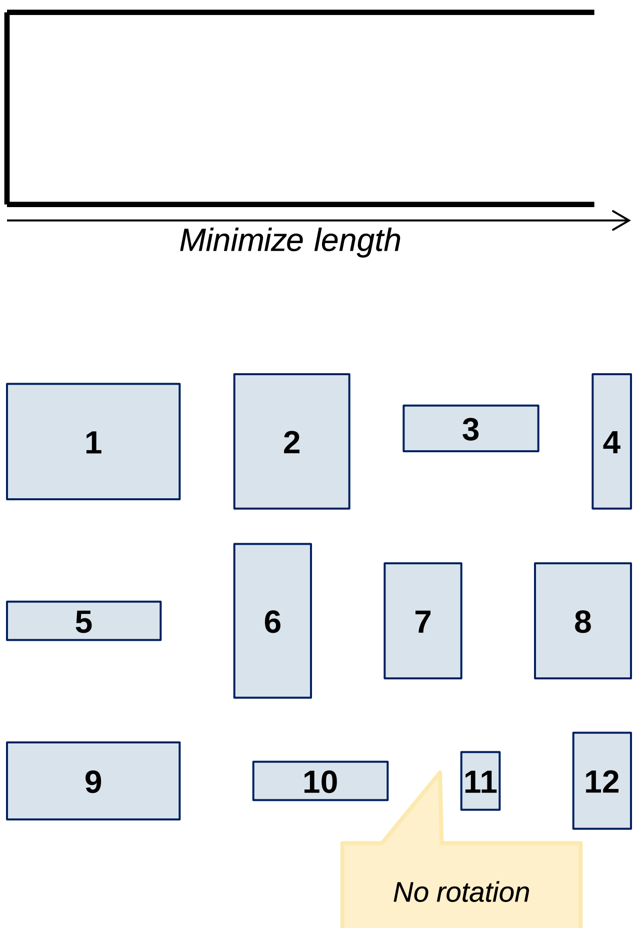
Selection of D^K and D* is based on three key concepts

Consequence of Theorem 4.5[1]	***A Basic Step between two disjunctions that do not share variables in common will not improve the tightness of the formulation***
Growth in proper basic step	***# of terms in resulting disjunction = (# of terms in D1)*(# of terms in D2)*** ***If we are applying several BS over the same disjunction this growth is even more important***
Characteristic value of disjunctions (Pre-analysis)	***The disjunction with highest characteristic value is expected to provide tightest relaxation when the Basic Step is applied***

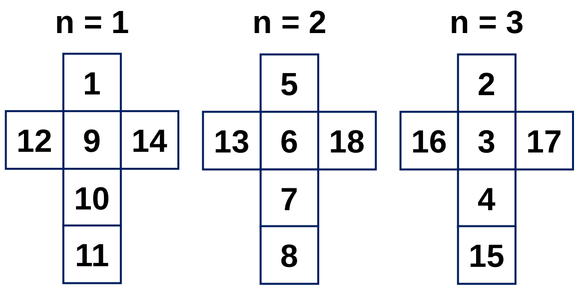
1: Balas E., Disjunctive Programming and a hierarchy of relaxations for discrete optimization problems, SIAM J. Alg. Disc. Meth., 6, 466-486, 1985

Algorithm was tested with several convex problems (I/III)

MILP: Strip packing (Stpck)



MILP: Nontransitive dice (Dice)



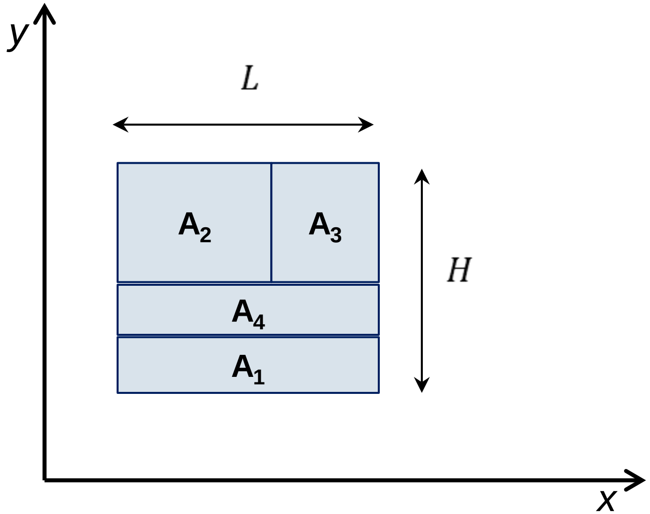
- Dice 1 beats dice 2 in 21 of 36 possible outcomes
- Dice 2 beats dice 3 in 21 of 36 possible outcomes
- Dice 3 beats dice 21in 21 of 36 possible outcomes

Algorithm was tested with several convex problems (II/III)

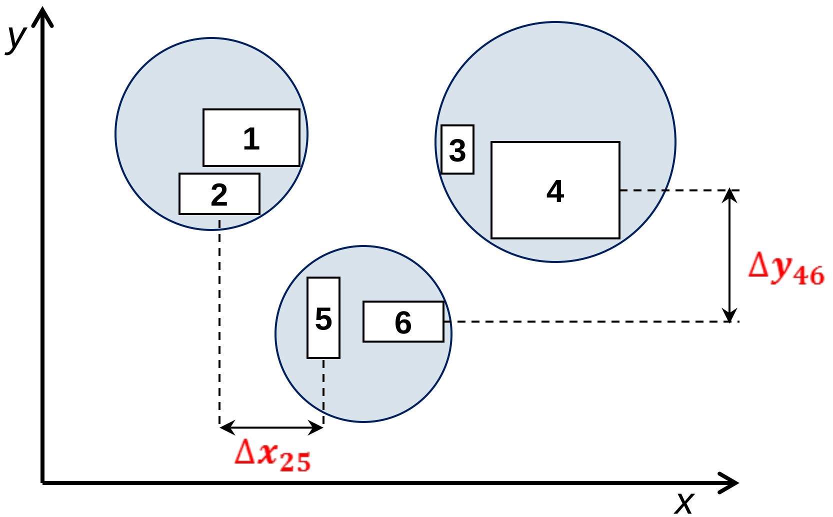
MINLP: Farm Layout (Flay)

A_1 = 40 m^2
A_2 = 50 m^2
A_3 = 60 m^2
A_4 = 35 m^2

$$\min P = 2 * (H + L)$$



MINLP: Constrained Layout (Clay)

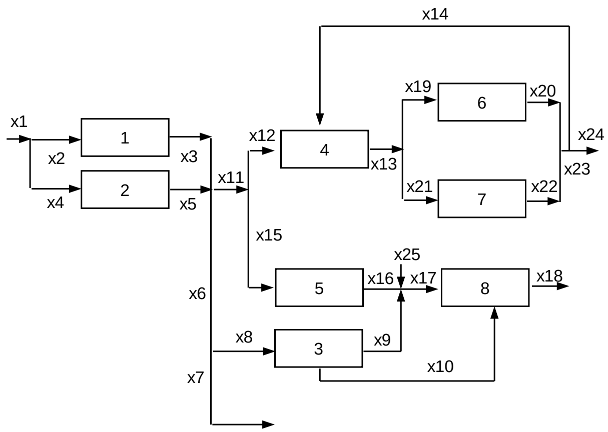


$$\min Q = \sum_i \sum_j c_{ij} (\Delta x_{ij} + \Delta y_{ij})$$

Algorithm was tested with several convex problems (III/III)

MINLP: Process flowsheet (Proc)

proc8:



$$\min \sum_{i=18} c_i + \sum_{j=1}^{25} p_j x_j + \gamma$$

st.

$$\sum_{j=1}^{25} r_{jn} x_j \le 0 \qquad \forall n \in \mathrm{N}$$

$$\begin{bmatrix} Y_i \\ \sum_{j=1}^{25} d_{ij}\left(e^{x_j/t_{ij}} - 1\right) - \sum_{j=1}^{25} s_{ij} x_j \le 0 \\ c_i = \gamma_i \end{bmatrix} \veebar \begin{bmatrix} \neg Y_i \\ x_j = 0; j \in J^i \\ c_i = 0 \end{bmatrix} \quad i = 1,8$$

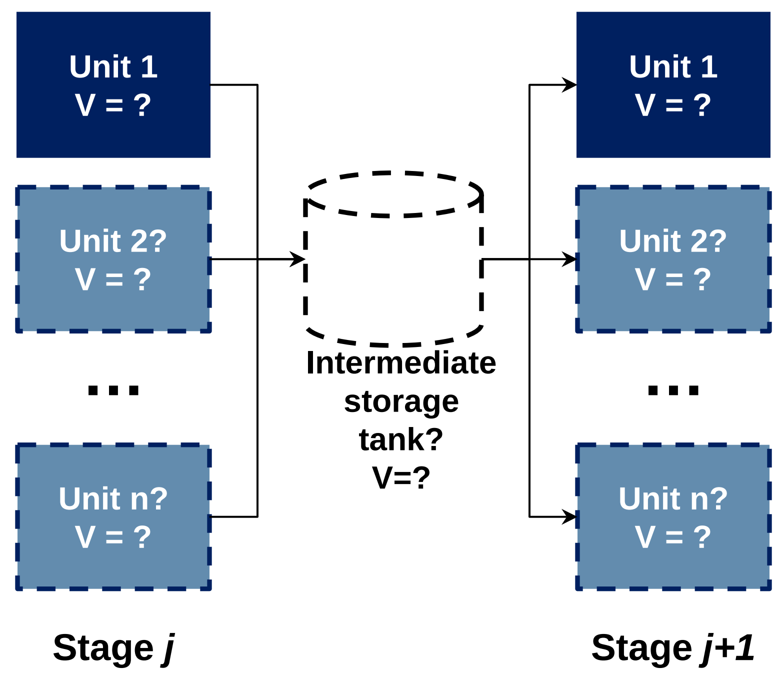
$$\Omega(Y) = True$$

$$x_j, c_i \ge 0 \qquad i = 1, \ldots, 8$$

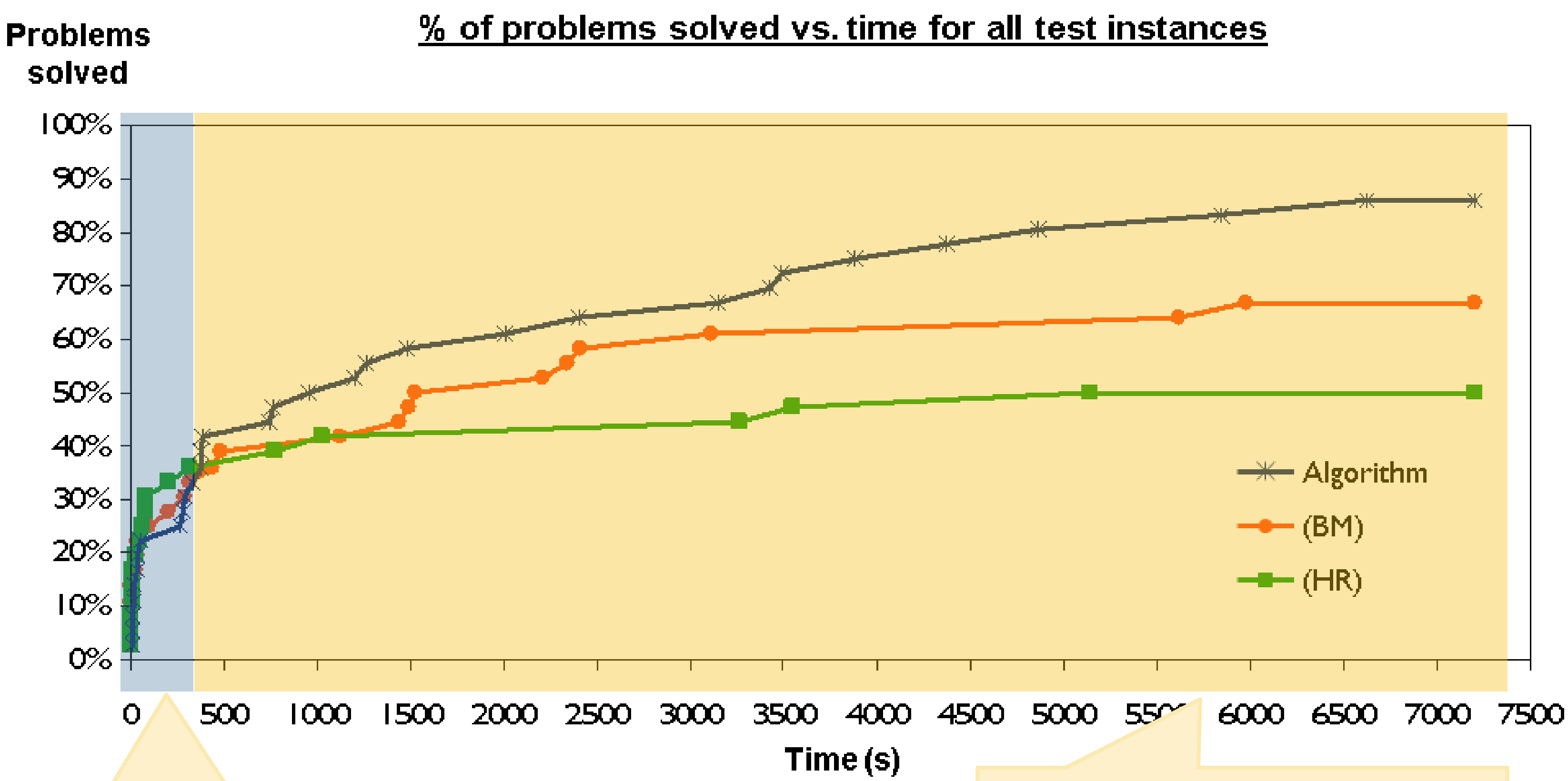
$$Y_i \in \{True, False\}$$

Non linear term

MINLP: Multiproduct batch (batch)



Results: The algorithm solves GDP generally faster for the 36 instances in which it was tested

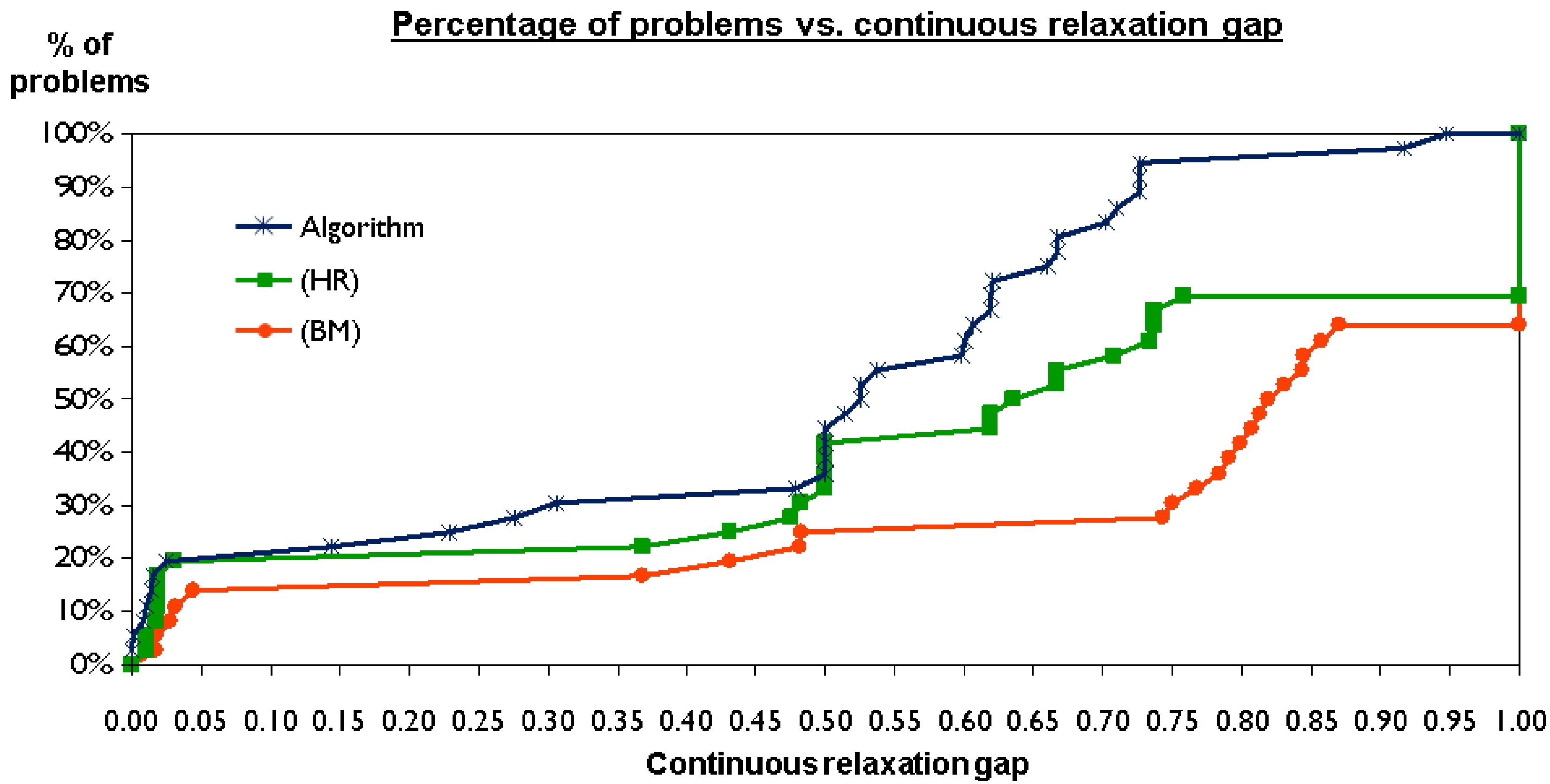


In smaller instances the presolve can take as much time as solving the MINLP, so the algorithm is slower than direct (BM) or (HR)

In larger instances the algorithm generally performs better than (BM) and (HR)

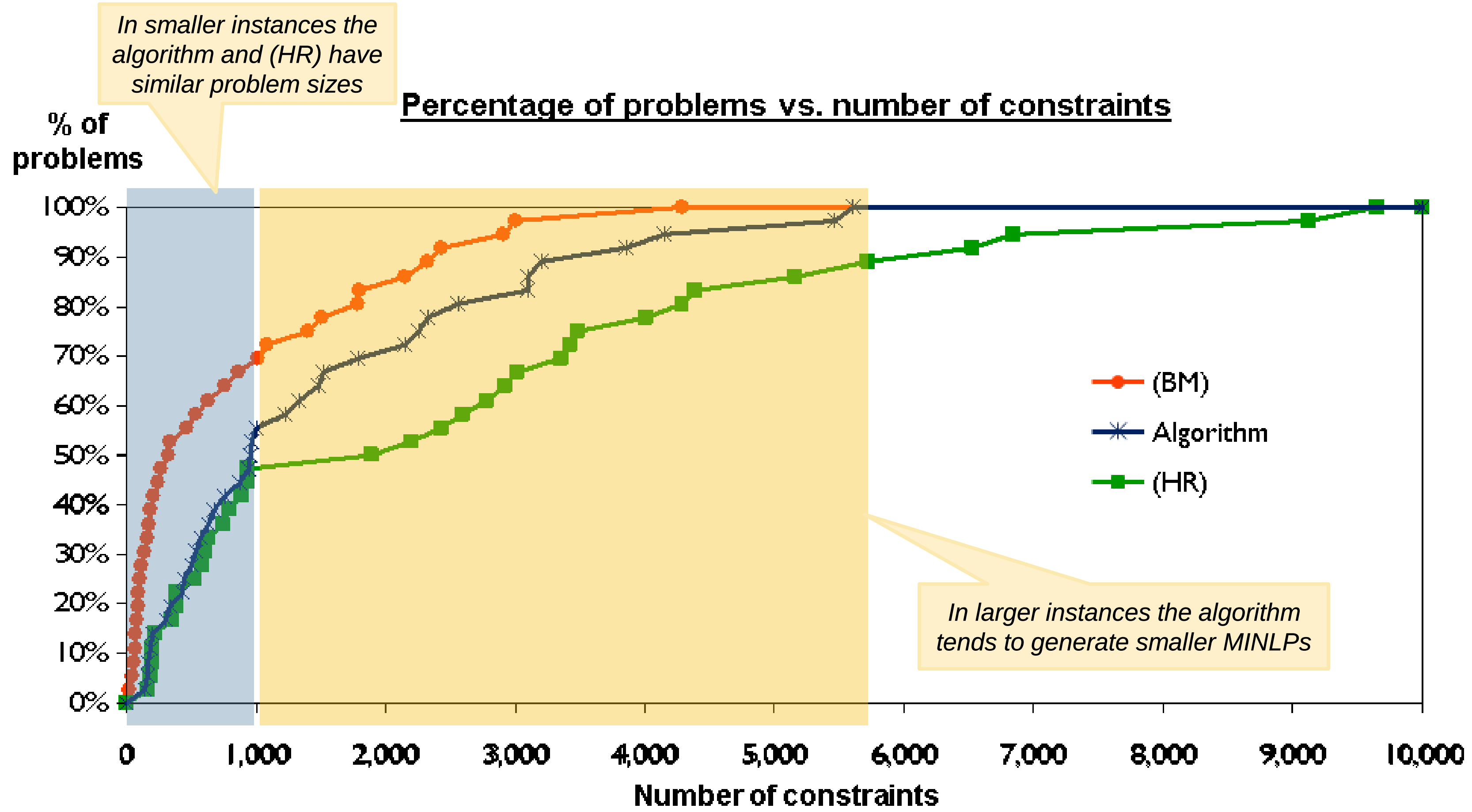
Note: MINLP problems solved with SBB/CONOPT. MILP solved with Gurobi (Pre-solve and cuts deactivated in solver for comparison purpose), in a 2.93 GHz Processor, Intel® Core™ i7. 4GB of RAM.

Results: MINLP after algorithm provides stronger relaxations



Note: MINLP problems solved with SBB/CONOPT. MILP solved with Gurobi (Pre-solve and cuts deactivated in solver for comparison purpose), in a 2.93 GHz Processor, Intel® Core™ i7. 4GB of RAM.

Results: MINLP size generally smaller than (HR)



Note: MINLP problems solved with SBB/CONOPT. MILP solved with Gurobi (Pre-solve and cuts deactivated in solver for comparison purpose), in a 2.93 GHz Processor, Intel® Core™ i7. 4GB of RAM.

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 convex 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

-Work currently underway to automate reformulation of convex GDP problems into MI(N)LP using concept basic steps