Global Optimization of Bilinear Generalized Disjunctive Programs

Juan Pablo Ruiz

Ignacio E. Grossmann

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

Non-Convex Discrete/Continuous Optimization Models

Mixed Integer Program (MIP)

- Most common non-linear discrete/continuous optimization model.
- Purely equation-based.
- If some functions in MIP are non convex → **non-convex MINLP**.

Generalized Disjunctive Program (GDP)

- Developed by Raman & Grossmann (1994)
- Combination of algebraic equations, disjunctions and logic propositions.
- Natural representation of engineering problems.
- If some functions in GDP are non convex → **non-convex GDP**.
- If all non linear functions are given by bilinear terms → **Bilinear GDP**.

Goal:
Global Optimization of Bilinear GDP with improved relaxations

Bilinear Generalized Disjunctive Programs

$$Min\ \ Z = \boxed{f(x)} + \sum_{k \in K} c_k$$ **Objective Function**

$$s.t.\ \ \boxed{g(x)} \le 0$$ **Global Constraints**

$$\bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ \boxed{r_{ik}(x)} \le 0 \\ c_k = \gamma_{ik} \end{bmatrix}$$ **Disjunctions** $k \in K$

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

Bilinearities

$$x \in R^n, c_k \in R, Y_{ik} \in \{True, False\} \qquad i \in D_k\ ,\ k \in K$$

Process engineering models. (e.g. HENS, Water Treatment Networks, Process Networks in general)

Bilinearities may lead to multiple local minima → Global Optimization techniques are required

Global optimization method by Lee & Grossmann (2001) → *Is their relaxation the tightest?*

Relaxation of Bilinear terms using McCormick envelopes leads to a LGDP → *Improved relaxations for Linear GDP has recently been obtained (Sawaya & Grossmann, 2007)*

LGDP to Disjunctive Programming Reformulation

Sawaya N.W. and Grossmann I.E. (2006)

LGDP

$$\min Z = \sum_{k \in K} c_k + d^T x$$

$$\text{s.t.} \quad Bx \ge b$$

$$\bigvee_{j \in J_k} \begin{bmatrix} \boxed{Y_{jk}} \\ A^{jk} x \ge a^{jk} \\ c_k = \gamma_{jk} \end{bmatrix} \quad k \in K$$

$$\boxed{\begin{array}{l} \underset{j \in J_k}{\veebar} Y_{jk} \\ \Omega(Y) = True \end{array}} \quad k \in K$$

$$x^L \le x \le x^U$$

$$\boxed{Y_{jk} \in \{True, False\}} \quad j \in J_k, k \in K$$

$$c_k \in \mathbf{R}^1 \quad k \in K$$

DP

$$\min Z = \sum_{k \in K} c_k + d^T x$$

$$\text{s.t.} \quad Bx \ge b$$

$$\bigvee_{j \in J_k} \begin{bmatrix} \boxed{\lambda_{jk} = 1} \\ A^{jk} x \ge a^{jk} \\ c_k = \gamma_{jk} \end{bmatrix} \quad k \in K$$

$$\boxed{\begin{array}{l} \sum_{j \in J_k} \lambda_{jk} = 1 \\ H\lambda \ge h \end{array}} \quad k \in K$$

$$x^L \le x \le x^U$$

$$\boxed{0 \le \lambda_{jk} \le 1} \quad j \in J_k, k \in K$$

$$c_k \in \mathbf{R}^1 \quad k \in K$$

Integrality λ guaranteed

Equivalent Disjunctive Programs

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

Thus the RF is:

$$F = \bigcap_{t \in T} S_t$$

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

There exists many forms between CNF and DNF that are equivalent

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

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

$$S_{rs} = \bigcup_{\substack{i \in Q_r \\ t \in Q_s}} (P_i \cap P_t).$$

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

A Hierarchy of Relaxations for DP

Hull Relaxation (Balas, 1985):

Let us take the following disjunctive set:

$$F = \bigcap_{j \in T} S_j$$

Then the hull-relaxation corresponds to:

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

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

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

ii) F_t *is in DNF;*

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

Then,

$$P_0 = h-rel\ F_0 \supseteq h-rel\ F_1 \supseteq \cdots \supseteq h-rel\ F_t = clconv\ F_t.$$ ***(true convex hull)***

A Hierarchy of Relaxations for GDP

Proposition 4 (Sawaya & Grossmann, 2006) *For* $i = 1,2,\ldots|T|+|K|-1$ *let* F_{GDP_i} *be a sequence of regular forms of the disjunctive set:*

$$F = \left\{ z := (x,\lambda,c) \in \mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|} : \bigcap_{i\in\hat{T}} \bar{b}^i z \geq \bar{b}_0^{\,i} \bigcap_{k\in\tilde{K}} \bigcup_{j\in J_k} (\tilde{A}^{jk} z \geq \tilde{a}^{jk}) \bigcap_{n\in\hat{K}} \bigcup_{m\in J_n} (\hat{A}^{mn} z \geq \hat{a}^{mn}) \right\}, \text{ such that}$$

i) F_{GDP_0} *corresponds to the disjunctive form:*

$$F = \left\{ z := (x,\lambda,c) \in \mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|} : \bigcap_{i\in\bar{T}} \bar{b}^i z \geq \bar{b}_0^{\,i} \bigcap_{k\in K} \bigcup_{j\in J_k} (\bar{A}^{jk} z \geq \bar{a}^{jk}) \right\};$$

ii) $F_{GDP_{|\bar{T}|+|K|-1}} := F_t$ *is in DNF;*

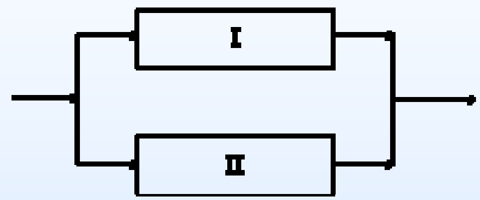
iii) for $i = 1,2,\ldots|T|+|K|-1$, F_{GDP_i} *is obtained from* $F_{GDP_{i-1}}$ *by a basic step.*

Then,

$$h-rel\ F_{GDP_0} \supseteq h-rel\ F_{GDP_1} \supseteq \cdots \supseteq h-rel\ F_{GDP_{|\bar{T}|+|K|-1}} = clconv\ F_{GDP_{|\bar{T}|+|K|-1}} = clconv\ F_t.$$ ***(true convex hull)***

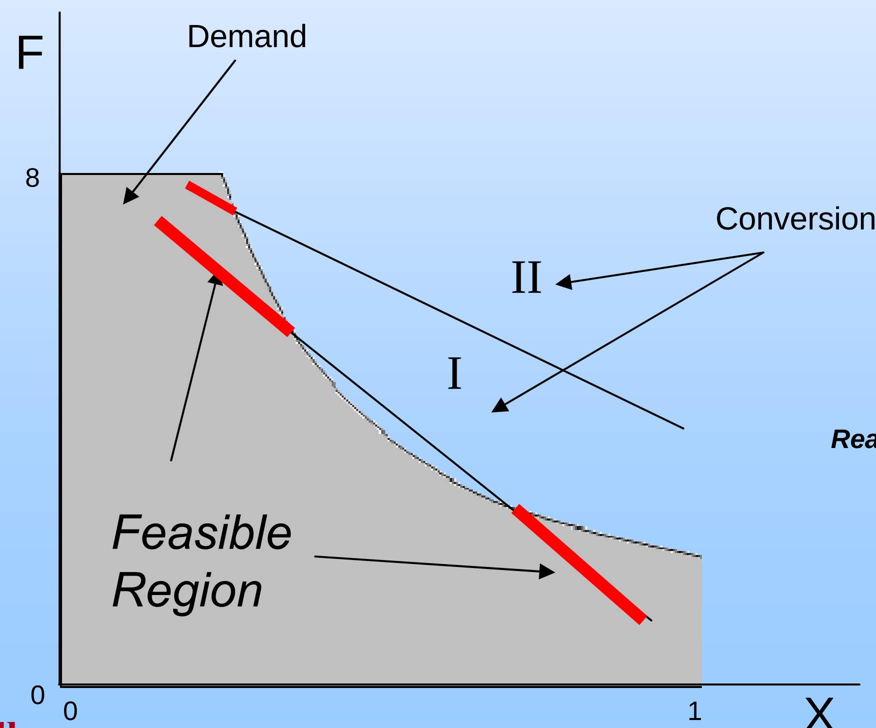
Bilinear GDP Relaxation

Illustrative Example: *Optimal reactor selection*



F: Flow
X: Conversion

Feasible region projected onto the FX space



GDP Formulation

Objective Function

Demand constraint

$$Max\ Z = \theta FX - \gamma F - CP$$

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

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

Reactor Curves

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

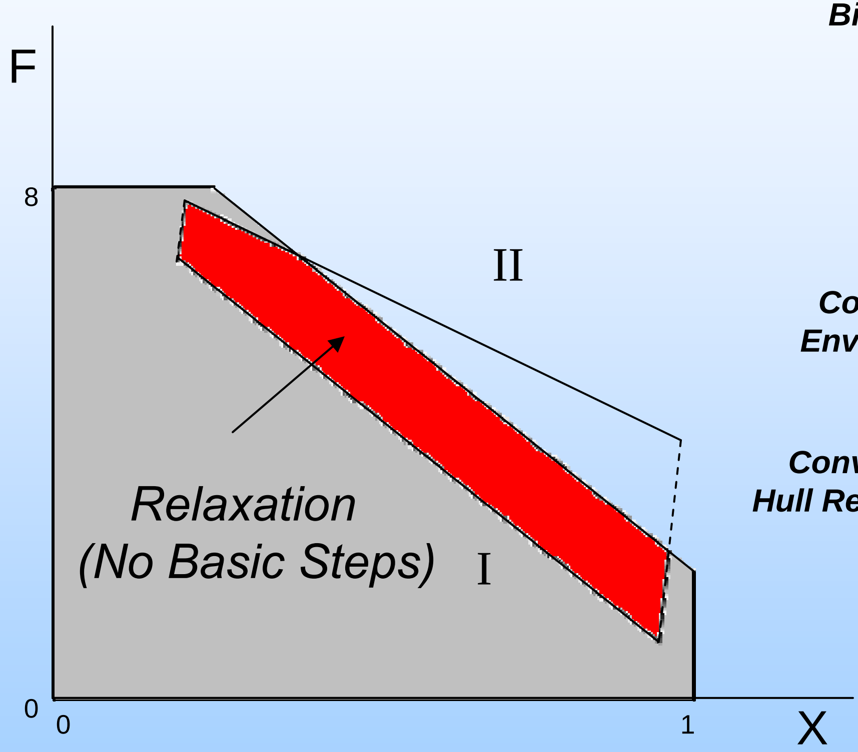
$CP, X, F \in R$

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

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

Bilinear GDP Relaxation

(Lee & Grossmann 2001)



Bilinear Terms

Max Z = $\theta \boxed{FX} - \gamma F - CP$

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

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

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

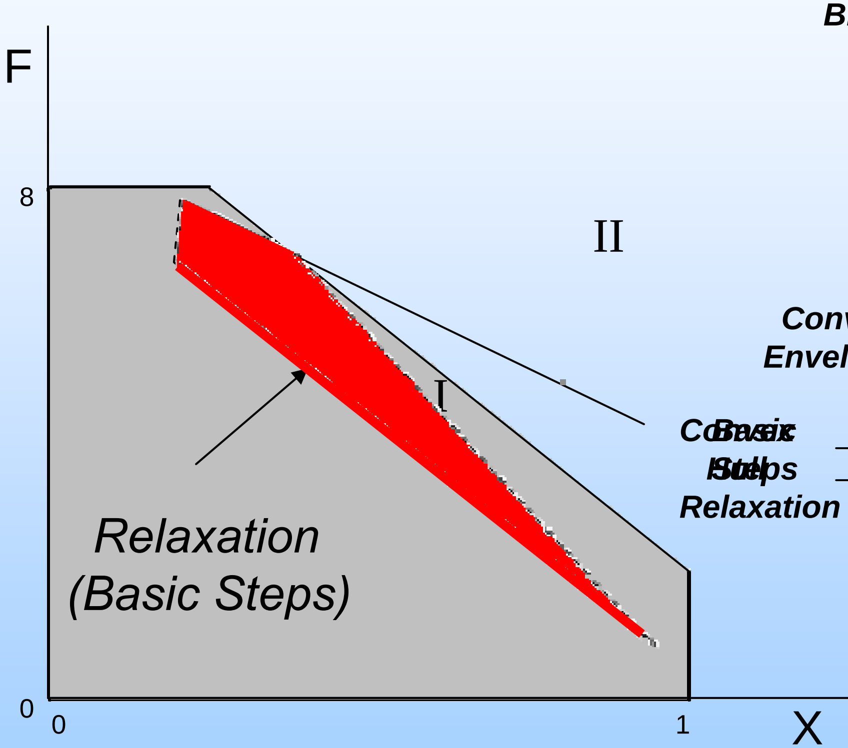
$Y_{11} \vee Y_{21}$ = *True*

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

Carnegie Mellon

Bilinear GDP Relaxation

Proposed Relaxation



Bilinear Terms

Convex Envelopes

Convex Hull

Basic Steps Relaxation

$$Max\ Z = \theta P - \gamma F - CP$$

$$s.t.\ \ P \le d$$

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

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

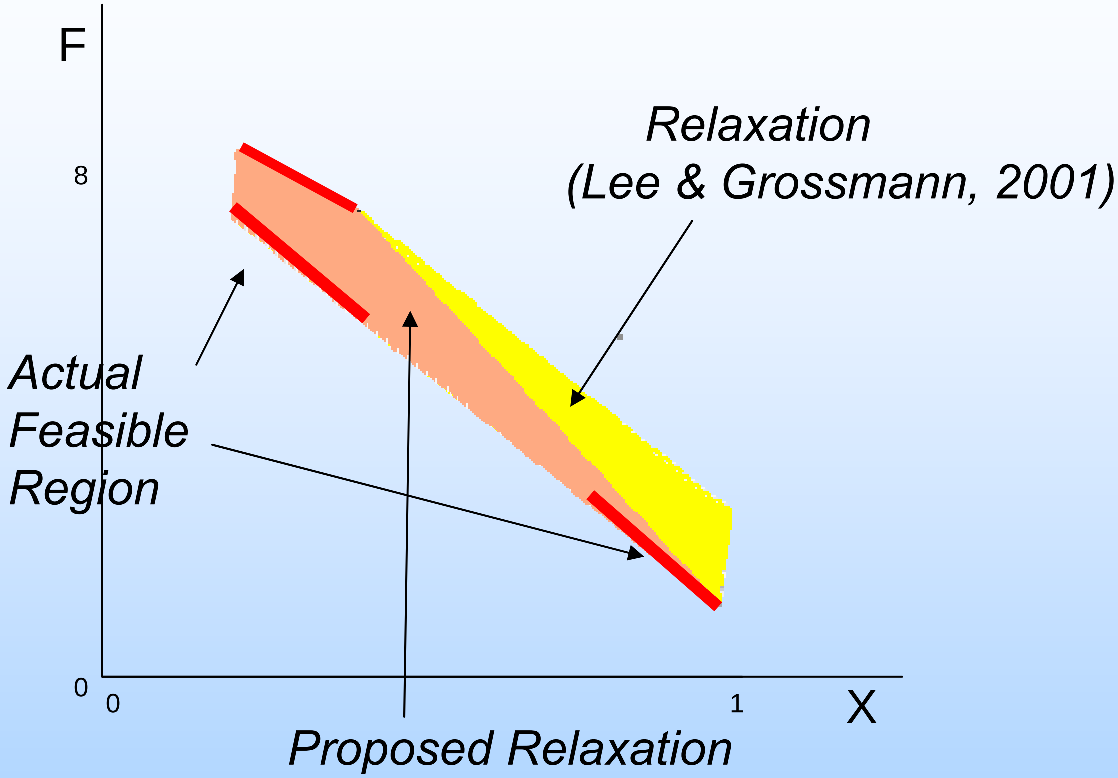
$$CP, X, F \in R$$

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

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

Bilinear GDP Relaxation

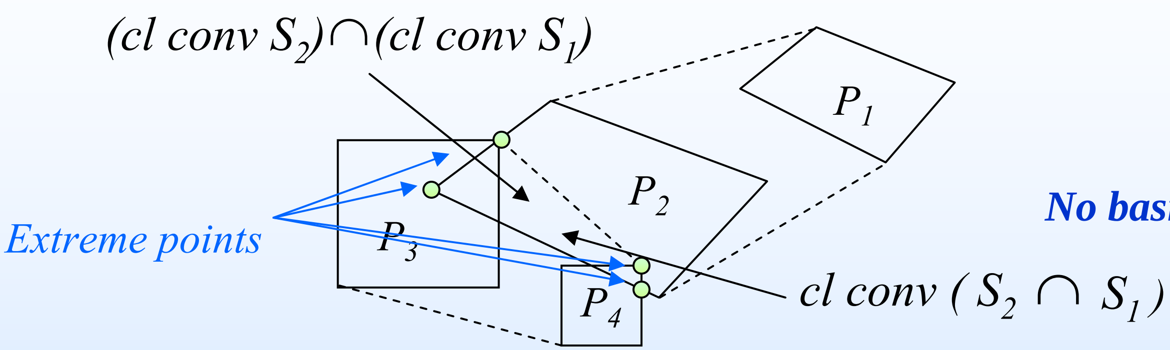
Comparison



The application of basic steps prior to the discrete relaxation leads to a ***tighter*** *relaxed feasible region*

Rules to apply basic steps. When they make a difference.

Motivating example:



$S_1 = P_1 \cup P_2$

$S_2 = P_3 \cup P_4$

No basic steps are necessary!

Theorem 4.5. (Balas, 1985)

For j=1,2 let

$$S_j = \bigcup_{i \in Q_j} P_i \, ,$$

Where each $P_i, i \in Q, j = 1,2,$ *is a polyhedron. Then*

$$cl\ conv(S_1 \cap S_2) = (cl\ conv\ S_1) \cap (cl\ conv\ S_2)$$

If and only if every extreme point (extreme direction) of $(cl\ conv\ S_1) \cap (cl\ conv\ S_2)$ *is an extreme point (extreme direction) of* $P_i \cap P_k$ *for some* $(i,k) \in Q_1 \times Q_2$

In general the hypothesis of the theorem is not easy to verify

To gain insight, the structure of the program should be exploited

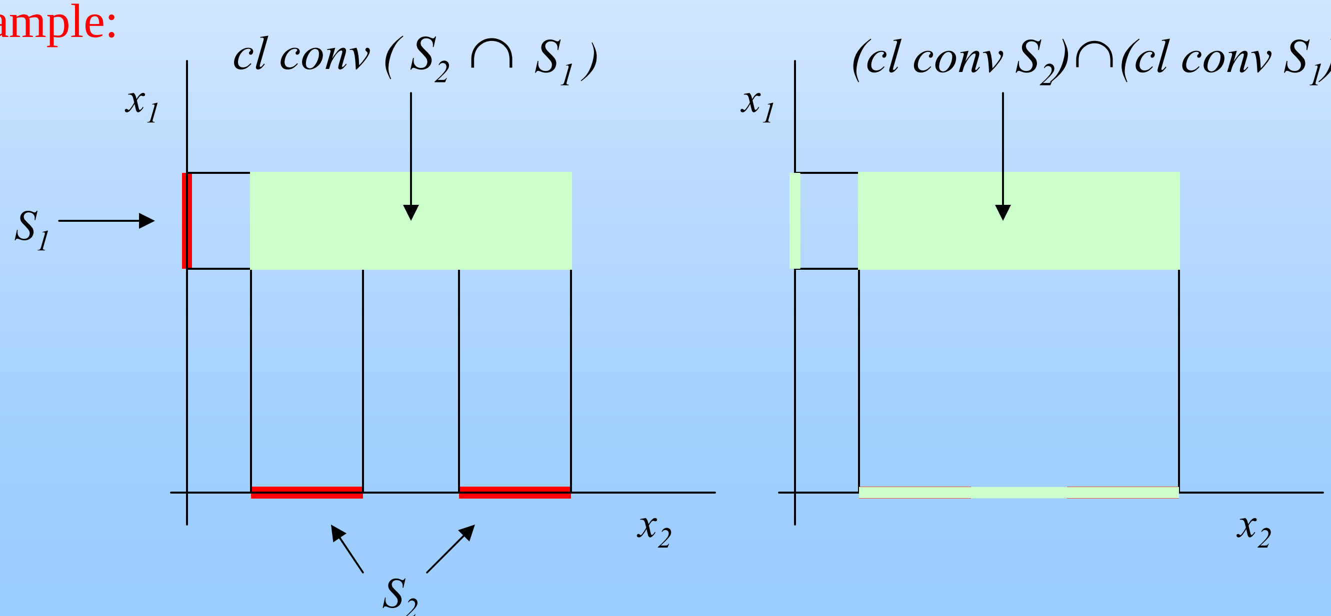
Rules to apply basic steps. When they make a difference.

Proposition :

Let S_1 and S_2 be two disjunctions in which the variables restricted in S_i are not restricted in the S_j ($i=1,2$; $j=1,2$ and $i \neq j$)

Then the system satisfies the hypothesis of theorem 4.2.

Example:



Summary of “practical” rules to apply basic steps

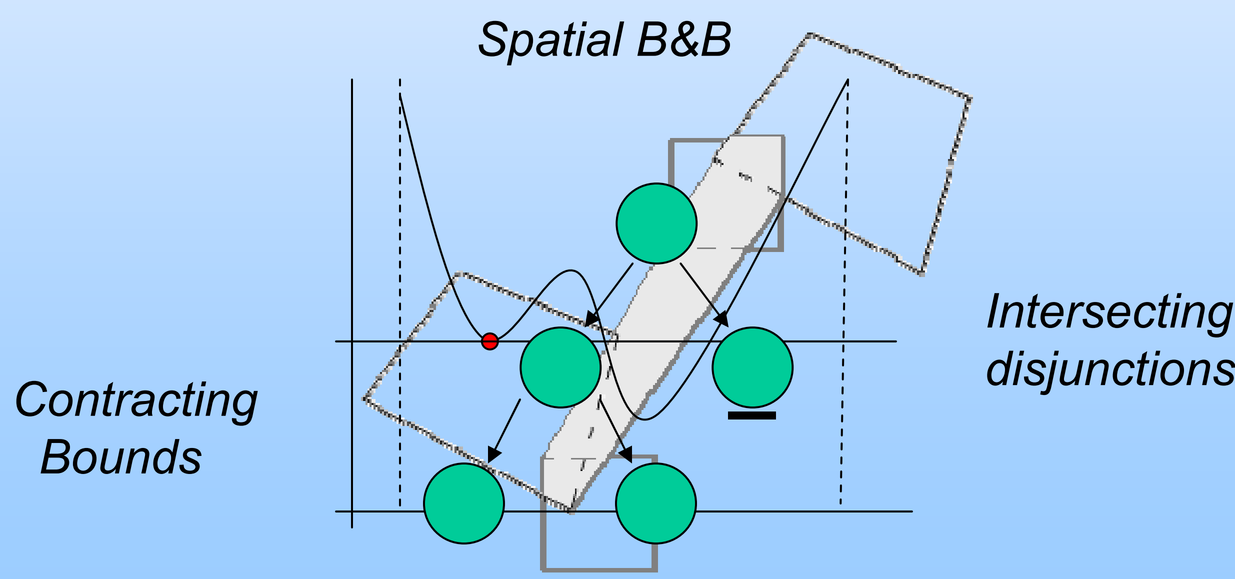
- Apply basic steps between those disjunctions with at least one variable in common.
- The more variables in common two disjunctions have the more the tightening expected
- If bilinearities are outside the disjunctions apply basic steps by introducing them in the disjunctions previous to the relaxation.
- If bilinearities are inside the disjunctions a less tightening effect is expected.
- A less increase in the size of the formulation is expected when basic steps are applied between improper disjunctions and proper disjunctions.

Methodology

Step 1: GDP reformulation (Apply basic steps following the rules presented)

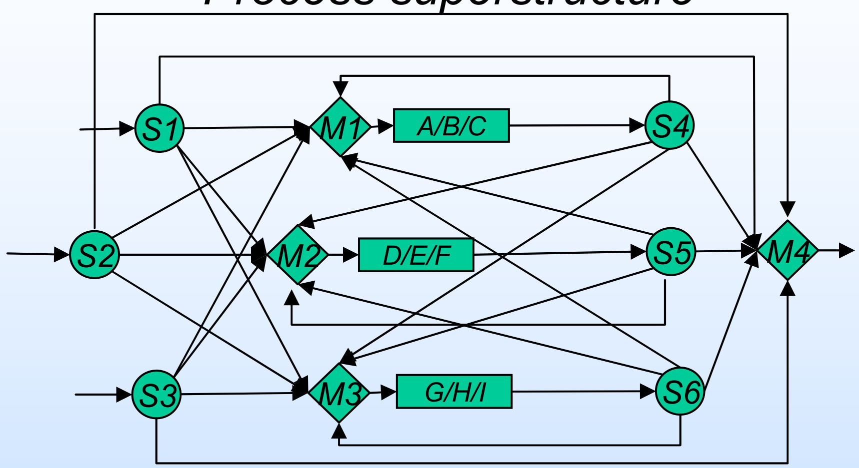
Step 2: Bound Contraction (Zamora & Grossmann, 1999)

Step 3: Branch and Bound Procedure (Lee & Grossmann, 2001)



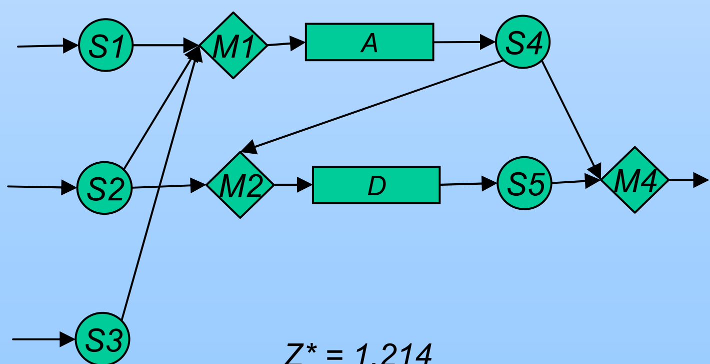
Case Study I: Water treatment network design

Process superstructure



N of cont. vars. : 114
N of disc. vars. : 9
N of bilinear terms: 36

Optimal structure



Z = 1.214*

Generalized Disjunctive Program

$$Min\ Z = \sum_{k \in PU} CP_k$$

$s.t.$

$$f_k^{\,j} = \sum_{i \in M_k} f_i^{\,j} \qquad \forall j \quad k \in MU$$

$$\sum_{i \in S_k} f_i^{\,j} = f_k^{\,j} \qquad \forall j \quad k \in SU$$

$$\sum_{i \in S_k} \zeta_i^{\,k} = 1 \qquad k \in SU$$

$$\boxed{f_i^{\,j} = \zeta_i^{\,k} f_k^{\,j}} \qquad \forall j \quad i \in S_k \qquad k \in SU$$

$$\bigvee_{h \in D_k} \left[\begin{array}{c} YP_k^{\,h} \\ f_i^{\,j} = \beta_k^{\,jh} f_{i'}^{\,j}, i \in OPU_k, i' \in IPU_k, \forall j \\ F_k = \sum_j f_i^{\,j}, i \in OPU_k \\ CP_k = \partial_{ik} F_k \end{array} \right] \qquad k \in PU$$

$$0 \le \zeta_i^{\,k} \le 1 \qquad \forall j,k$$

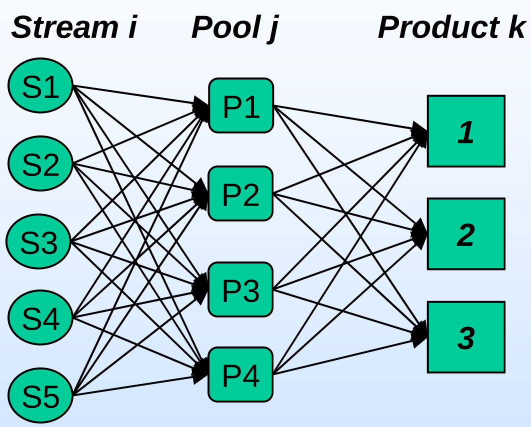
$$0 \le f_i^{\,j}, f_k^{\,j} \qquad \forall i,j,k$$

$$0 \le CP_k \qquad \forall k$$

$$YP_k^{\,h} \in \{true, false\} \qquad \forall h \in D_k \qquad \forall k \in PU$$

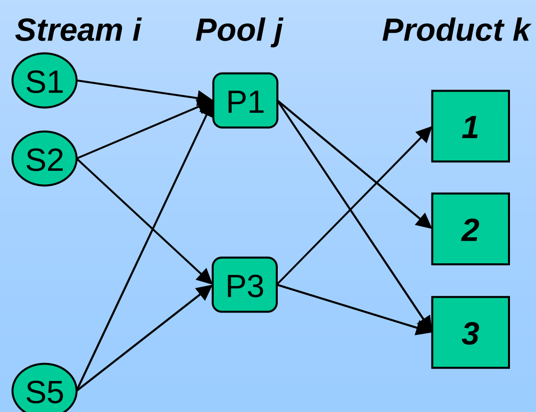
Case Study II: Pooling network design

Process superstructure



N of cont. vars. : 76
N of disc. vars. : 9
N of bilinear terms: 24

Optimal structure



Z = -4.640*

Generalized Disjunctive Program

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

$s.t.$

$$\sum_{i\in I}\sum_{w\in W} f_{ijw} = \sum_{k\in K}\sum_{w\in W} f_{jkw} \qquad \forall j\in J$$

$$\sum_{j\in J}\sum_{w\in W} f_{jkw} - S_k = 0 \qquad \forall k\in K$$

$$f_{ijw} = \lambda_{iw} \sum_{w'\in W} f_{ijw'} \qquad \forall i\in I, \forall j\in J, \forall w\in W$$

$$\sum_{j\in J} f_{jkw} - Z_{kw} \sum_{j\in J}\sum_{w'\in W} f_{jkw'} = 0 \qquad \forall k\in K, \forall w\in W$$

$$\begin{bmatrix} YST_i \\ f^{lo} \le \sum_{j\in J}\sum_{w\in W} f_{ijw} \\ CST_i = \alpha_i \end{bmatrix} \vee \begin{bmatrix} \neg YST_i \\ f_{ijw} = 0 \\ CST_i = 0 \end{bmatrix} \qquad \forall i\in I$$

$$\begin{bmatrix} YP_j \\ f^{lo} \le \sum_{i\in I}\sum_{w\in W} f_{ijw} \\ \sum_{k\in K} f_{jkw} = \sum_{i\in I} f_{ijw}, \forall w\in W \\ \boxed{f_{jkw} = \zeta_j^{\,k} \sum_{i\in I} f_{ijw}}, \forall w\in W, k\in K \\ \sum_{k\in K} \zeta_j^{\,k} = 1 \\ CP_j = \gamma_j \end{bmatrix} \vee \begin{bmatrix} \neg YP_j \\ f_{ijw} = 0, \forall i\in I, w\in W \\ f_{jkw} = 0, \forall k\in K, w\in W \\ CP_j = 0 \end{bmatrix} \qquad \forall j\in J$$

$$0 \le \zeta_j^{\,k} \le 1; 0 \le f_{jkw}, f_{ijw} \le f^{up}$$

$$0 \le CST_i, CP_j; YST_i, YP_j \in \{true, false\}$$

Performance

		Global Optimization Technique using Lee & Grossmann relaxation	*Global Optimization Technique using proposed relaxation*	*Relative Improvement*
Example 1	**Initial Lower Bound**	**400.66**	**499.86**	**24.90%**
	Bound contraction			**99.7%**
	Nodes	**399**	**204**	**51%**

		Global Optimization Technique using Lee & Grossmann relaxation	*Global Optimization Technique using proposed relaxation*	*Relative Improvement*
Example 2	**Initial Lower Bound**	**-5515**	**-5468**	**0.90%**
	Bound contraction			**8%**
	Nodes	**748**	**683**	**9%**

Conclusions and Remarks

- Tighter reformulation of bilinear GDPs.
- Proposed general rules to implement basic steps.
- Proposed methodology to solve Bilinear GDPs.
- Application in two cases showed improved performance.