

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

Juan Pablo Ruiz , Ignacio Grossmann*

Department of Chemical Engineering, Carnegie Mellon University
Pittsburgh, PA, USA 15213

March 2009

Abstract

This paper is concerned with global optimization of Bilinear and Concave Generalized Disjunctive Programs. A major objective is to propose a procedure to find relaxations that yield strong lower bounds. We first present a general framework for obtaining a hierarchy of linear relaxations for nonconvex Generalized Disjunctive Programs (GDP). This framework combines linear relaxation strategies proposed in the literature for nonconvex MINLPs with the results of the work by Sawaya & Grossmann (2008) for Linear GDPs. We further exploit the theory behind Disjunctive Programming to guide more efficiently the generation of relaxations by considering the particular structure of the problems. Finally, we show through a set of numerical examples how these new relaxations can strengthen substantially the lower bounds for the global optimum, often leading to a significant reduction of the number of nodes when used within a spatial branch and bound framework.

*Author to whom correspondence should be addressed

1- Introduction

Generalized Disjunctive Programming (GDP), developed by Raman and Grossmann (1994), has been proposed as a framework that facilitates the modeling of discrete-continuous optimization problems by allowing the use of algebraic and logical equations through disjunctions and logic propositions that are expressed in terms of Boolean and continuous variables. In order to take advantage of existing solvers (Westerlund & Pettersson, 1995; Viswanathan & Grossmann, 1990; Sahinidis, 1996; Leyffer, 2001; Kesavan et al, 2004; Bonami et. al, 2008), GDPs are often reformulated as MILP/MINLP problems by using either the Big-M (BM) (Nemhauser & Wolsey, 1988), or the Convex Hull (CH) (Lee & Grossmann, 2000) reformulation (See Appendix 1). It is important to note that GDP problems can always be reformulated as an MINLP. However, these reformulations are not unique and may have associated relaxations that are not very tight, consequently having an adverse effect on the efficiency of the algorithm that is used. In general, the tighter the relaxation of the reformulation and the fewer the number of variables and constraints, the smaller the computational effort is.

In the particular case of nonconvex GDP problems the direct application of traditional algorithms to solve the reformulated MINLPs such as Generalized Benders Decomposition (GBD) (Benders, 1962 and Geoffrion, 1972) or Outer Approximation (OA) (Duran & Grossmann, 1986), may fail to find the global optimum since the solution of the NLP subproblem may correspond to a local optimum and the cuts in the master problem may not be valid. Therefore, specialized algorithms should be used in order to find the global optimum (Horst & Tuy, 1996, Tawarmalani & Sahinidis, 2002 and Floudas, 2000). Nonconvex GDP problems with bilinear constraints are of particular interest since these arise in many applications, for instance, in the design of pooling problems (Meyer & Floudas, 2006), in the synthesis of integrated water treatment networks (Karuppiah & Grossmann, 2006), or generally, in the synthesis of process networks with multicomponent flows (Quesada & Grossmann, 1995). In addition, nonconvex GDP problems with concave constraints frequently arise when nonlinear investment cost functions are considered (Turkay & Grossmann, 1996). To tackle this problem, Lee and Grossmann (2003) proposed a global optimization method that first relaxes the bilinear terms by using the convex envelopes of McCormick (1976) and the

concave terms by using linear underestimators. The convex hull (Balas, 1985) is then applied to each disjunction. This formulation is then used within a spatial branch and bound technique in which the branching is first performed on the Boolean variables followed by the continuous variables. While the method proved to be effective in solving several problems, a major question is whether one might be able to obtain stronger lower bounds to enhance the efficiency for globally optimizing GDP problems.

Sawaya and Grossmann (2008) have recently established new connections between Linear GDP and the Disjunctive Programming theory by Balas (1979). As a result, a family of tighter reformulations has been identified. These are obtained by performing a sequence of basic steps on the original disjunctive set (i.e. each basic step is characterized by generating a new set of disjunctions by intersecting the former), bringing it to a form closer to the Disjunctive Normal Form (DNF), and hence tightening its discrete relaxation (Balas, 1985). It is important to note that each intersection usually creates new variables and constraints. Therefore, it is important to recognize when it may be useful to make these intersections. Some general rules are described in this work.

In this work we build on the work by Sawaya and Grossmann (2008) exploiting the newly discovered hierarchy of relaxations in order to solve more efficiently nonconvex GDP problems, particularly, with bilinearities and concave functions in their constraints, namely Bilinear GDP and Concave GDP.

This paper is organized as follows. In section 2 we present the general structure and particular properties of the problems for which we aim at finding better relaxations (i.e. Bilinear GDP and Concave GDP). In section 3 and 4, a general theoretical framework is proposed for obtaining tighter linear relaxations efficiently for nonconvex GDPs. The implementation of this framework is then illustrated in section 5 by finding a relaxation for two small examples, one of them formulated as a Bilinear GDP and the other as a Concave GDP. Section 6 outlines the implementation of the tighter reformulation within a spatial branch and bound procedure whose performance is compared with current methodologies (i.e. Lee & Grossmann, 2003) in section 7.

2- Nonconvex Generalized Disjunctive Programs

The general structure of a nonconvex GDP can be represented as follows (Raman & Grossmann, 1994, Turkay & Grossmann, 1996, Lee & Grossmann, 2000):

$$\begin{aligned}
 & \text{Min } Z = f(x) + \sum_{k \in K} c_k \\
 & \text{s.t. } g^l(x) \leq 0 \quad l \in L \\
 & \bigvee_{i \in D_k} \left[\begin{array}{l} Y_{ik} \\ r_{ik}^j(x) \leq 0 \quad j \in J_{ik} \\ c_k = \gamma_{ik}(x) \end{array} \right] \quad k \in K \quad (GDP_{NC})
 \end{aligned}$$

$$\Omega(Y) = \text{True}$$

$$x^{lo} \leq x \leq x^{up}$$

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

where $f : R^n \rightarrow R^1$ is a function of the continuous variables x in the objective function, $g^l : R^n \rightarrow R^1$, $l \in L$, belongs to the set of global constraints, the disjunctions $k \in K$, are composed of a number of terms $i \in D_k$, that are connected by the OR operator. In each term there is a Boolean variable Y_{jk} , a set of inequalities $r_{ik}^j(x) \leq 0$, $r_{ik}^j : R^n \rightarrow R^1$, $j \in J$ and a cost variable c_k . If Y_{jk} is true, then $r_{ik}^j(x) \leq 0$ and $c_k = \gamma_{ik}(x)$ are enforced; otherwise they are ignored. Also, $\Omega(Y) = \text{True}$ are logic propositions for the Boolean variables. As indicated in Sawaya & Grossmann (2008), we assume that the logic constraints $\bigvee_{j \in J} Y_{jk}$ are contained in $\Omega(Y) = \text{True}$. In a nonconvex GDP, f , r_{ik} , γ_{ik} and/or g^l are nonconvex functions.

Bilinear GDPs (BGDP) are the first class of nonconvex GDP problems that we address in this paper.

A BGDP is a nonconvex GDP where the functions in the constraints only contain bilinear and linear terms. In general we can represent a BGDP as:

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

$$\text{s.t. } x^T Q^l x + a^l x \leq b^l \quad l \in L$$

$$\bigvee_{i \in D_k} \left[\begin{array}{c} Y_{ik} \\ x^T Q_{ik}^j x + a_{ik}^j x \leq b_{ik}^j \quad j \in J_{ik} \\ c_k = \gamma_{ik} \end{array} \right] \quad k \in K \quad (GDP_B)$$

$$\Omega(Y) = \text{True}$$

$$x^{lo} \leq x \leq x^{up}$$

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

where some of the matrices Q^l, Q_{ik}^j are indefinite

REMARK: Note that if all matrices Q^l, Q_{ik}^j are positive semidefinite then the problem is convex and no global optimization methods are required to find the optimal solution.

In order to solve (GDP_B) with a spatial branch and bound method a *convex GDP relaxation* is required. A valid *Linear GDP relaxation* (See Proposition 1) can be obtained by finding suitable under and over estimating functions of the nonconvex constraints. Although this set of estimators is not unique, we propose to use the convex envelopes proposed by McCormick (1976) for bilinear terms (See also Al-Khayyal & Falk, 1983).

Defining $X = xx^T$ we can find a relaxation for each term $X_{ij} = x_i x_j$ as:

$$\begin{aligned} X_{ij} &\leq x_i x_j^{up} + x_j x_i^{lo} - x_j^{up} x_i^{lo} \\ X_{ij} &\leq x_i x_j^{lo} + x_j x_i^{up} - x_j^{lo} x_i^{up} \\ X_{ij} &\geq x_i x_j^{lo} + x_j x_i^{lo} - x_j^{lo} x_i^{lo} \\ X_{ij} &\geq x_i x_j^{up} + x_j x_i^{up} - x_j^{up} x_i^{up} \end{aligned} \quad i = 1, 2, \dots, n, \quad i < j < n+1$$

This leads us to the following Linear GDP,

$$\begin{aligned}
\text{Min } Z^L &= d^T x + \sum_{k \in K} c_k \\
\text{s.t. } Q^l \bullet X + a^l x &\leq b^l & l \in L \\
\bigvee_{i \in D_k} \left[\begin{array}{l} Y_{ik} \\ Q_{ik}^j \bullet X + a_{ik}^j x \leq b_{ik}^j \quad j \in J_{ik} \\ c_k = \gamma_{ik} \end{array} \right] & & k \in K & (GDP_{RB})
\end{aligned}$$

$$\begin{aligned}
X_{ij} &\leq x_i x_j^{up} + x_j x_i^{lo} - x_j^{up} x_i^{lo} \\
X_{ij} &\leq x_i x_j^{lo} + x_j x_i^{up} - x_j^{lo} x_i^{up} \\
X_{ij} &\geq x_i x_j^{lo} + x_j x_i^{lo} - x_j^{lo} x_i^{lo} \\
X_{ij} &\geq x_i x_j^{up} + x_j x_i^{up} - x_j^{up} x_i^{up} \\
& i = 1, 2, \dots, n, \quad i < j < n+1
\end{aligned}$$

$$\Omega(Y) = \text{True}$$

$$x^{lo} \leq x \leq x^{up}$$

$$X \in R^{n \times n}, x \in R^n, c_k \in R^1, Y_{ik} \in \{\text{True}, \text{False}\}$$

where \bullet represents the scalar product of matrices.

Traditionally, (GDP_{RB}) has been used to predict lower bounds in the spatial branch and bound method (Lee & Grossmann, 2003). In this work we will show that by the application of a systematic procedure described in sections 3 and 4, we can improve the strength of the continuous relaxation of (GDP_{RB}) , leading to stronger lower bound predictions for (GDP_B) .

The second class of nonconvex GDP problems we are interested in are Concave GDPs. These problems often arise when economies of scale are considered in the economic evaluation of potential designs (Turkay & Grossmann, 1996). A Concave GDP is a nonconvex GDP where the functions in the constraints are concave and linear.

The general representation of these problems is as follows:

$$\begin{aligned}
\text{Min } Z &= f(x) + \sum_{k \in K} c_k \\
\text{s.t. } g^l(x) &\leq 0 & l \in L \\
\bigvee_{i \in D_k} \left[\begin{array}{l} Y_{ik} \\ r_{ik}^j(x) \leq 0 \quad j \in J_{ik} \\ c_k = \gamma_{ik}(x) \end{array} \right] & & k \in K & (GDP_{Co})
\end{aligned}$$

$\Omega(Y) = True$

$x^{lo} \leq x \leq x^{up}$

$x \in R^n, c_k \in R^1, Y_{ik} \in \{True, False\}$

where r_{ik}, γ_{ik} and/or g^l are concave or linear.

REMARK: Note that the effect of the economies of scale are often defined in the equality $c_k = \gamma_{ik}(x)$. Frequently, the function $\gamma_{ik}(x)$ is concave univariate.

A valid Linear GDP relaxation of the concave GDP (GDP_{RCO}) can be obtained by finding suitable convex under and over linear estimators for the concave functions. As in the case of BGDP, this set of estimators is not unique. However, in this work we propose making use of the framework proposed by Tawarmalani and Sahinidis (Theorem 2.15, 2.16 and Corollary 2.18) to obtain valid under estimators and polyhedral outer approximations to obtain valid over estimators.

For the particular case in which $\gamma_{ik}(x)$ is a concave univariate function defined in the domain $[x^{lo}, x^{up}]$, a valid linear under estimator is given by the secant

$$\gamma_{ik}^* = \frac{\gamma_{ik}(x^{up}) - \gamma_{ik}(x^{lo})}{x^{up} - x^{lo}}(x - x^{lo}) + \gamma_{ik}(x^{lo})$$

As in the case of (GDP_{RB}) we will show that the framework we propose in section 3 and 4 can also strengthen the relaxation given by GDP_{RCO}

REMARK: It is important to note that sometimes concave functions and bilinear terms arise in the same GDP formulation (See Appendix 4 – Example 3 and Example 6). The method we propose can be readily extended to solve this class of problems.

3 - A hierarchy of linear relaxations for nonconvex GDP

In this section we present a general framework to obtain a hierarchy of linear relaxations for the nonconvex GDP problem (GDP_{NC}) that can serve as a basis to predict tight lower bounds to the global optimum. This framework will be presented for general nonconvex

GDP problems, although it will later be only applied to the bilinear and concave case (i.e. GDP_B and GDP_{Co})

Without loss of generality, we can consider in (GDP_{NC}) f to be a linear function of x . Hence, we can represent it as $d^T x$.

Let us define the following sets:

$$L_c := \{ l \in L \mid g^l(x) \text{ is convex} \}$$

$$L_{nc} := \{ l \in L \mid g^l(x) \text{ is nonconvex} \}$$

$$J_{cik} := \{ j \in J_{ik}, i \in D_k, k \in K \mid r_{ik}^j(x) \text{ is convex} \}$$

$$J_{ncik} := \{ j \in J_{ik}, i \in D_k, k \in K \mid r_{ik}^j(x) \text{ is nonconvex} \}$$

$$D_{cik} := \{ i \in D_k, k \in K \mid \gamma_{ik}(x) \text{ is convex} \}$$

$$D_{ncik} := \{ i \in D_k, k \in K \mid \gamma_{ik}(x) \text{ is nonconvex} \}$$

In the first step of this approach we replace each nonconvex function with a valid linear under/over-estimator as was done with (GDP_{RB}) and (GDP_{RCo}). Generalizing,

$$g^l(x) \leq 0 \text{ is replaced by } A_g^l x \leq b_g^l, l \in L_{nc},$$

$$r_{ik}^j(x) \leq 0 \text{ is replaced by } A_{rik}^j x \leq b_{rik}^j, j \in J_{ncik}, i \in D_k, k \in K,$$

$$c_k = \gamma_{ik}(x), \text{ is replaced by } A_{\gamma ik}(x, c_k) \leq b_{\gamma ik}, i \in D_{ncik}, k \in K$$

Note that the dimensions of the matrices, $A_g^l, A_{rik}^j, A_{\gamma ik}$, and the right-hand side vectors $b_{rik}^j, b_g^l, b_{\gamma ik}$, depend on the particular under/over estimators that are chosen as was discussed in section 2.

Similarly, we replace each convex inequality with a valid linear outer-approximation. Namely,

$$g^l(x) \leq 0 \text{ is replaced by } A_g^l x \leq b_g^l, l \in L_c,$$

$$r_{ik}^j(x) \leq 0 \text{ is replaced by } A_{rik}^j x \leq b_{rik}^j, j \in J_{cik}, i \in D_k, k \in K,$$

$c_k = \gamma_{ik}(x)$, is replaced by $A_{\gamma_{ik}}(x, c_k) \leq b_{\gamma_{ik}}$, $i \in D_{cik}$, $k \in K$

Note that the dimensions of the matrices $A_g^l, A_{rik}^l, A_{\gamma_{ik}}$ and the vectors $b_{rik}^j, b_g^l, b_{\gamma_{ik}}$ depend on the polyhedral outer-approximation technique that is chosen (Tawarmalani & Sahinidis, 2002, Gruber & Kenderov, 1982).

Replacing the nonconvex and concave functions in (GDP_{NC}) by the corresponding under/over-estimators and outer-approximations, leads to the following Linear GDP:

$$\begin{aligned}
 \text{Min } Z^L &= d^T x + \sum_{k \in K} c_k \\
 \text{s.t. } & A_g^l x \leq b_g^l \quad l \in L \\
 & \bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ A_{rik}^j x \leq b_{rik}^j \quad j \in J_{ik} \\ A_{\gamma_{ik}}(x, c_k) \leq b_{\gamma_{ik}} \end{bmatrix} \quad k \in K \quad (GDP_{RLP})
 \end{aligned}$$

$$\Omega(Y) = \text{True}$$

$$x^{lo} \leq x \leq x^{up}$$

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

Note that (GDP_{RB}) and (GDP_{RCo}) are particular cases of GDP_{RLP} .

In the following proposition (PROPOSITION 1) we prove that GDP_{RLP} is indeed a valid Linear GDP relaxation for (GDP_{NC})

PROPOSITION 1 : Let us define the set S_{NC} as the set of points $(x, c, Y) \in R^{n+|K|} \times B_{k \in K}^{\sum |D_k|}$ such that (x, c, Y) belongs to the feasible region defined in GDP_{NC} . Similarly, let us define the set S_{RLP} as the set of points $(x, c, Y) \in R^{n+|K|} \times B_{k \in K}^{\sum |D_k|}$ such that (x, c, Y) belongs to the feasible region defined in GDP_{RLP} . Then $S_{NC} \subseteq S_{RLP}$

PROOF: Defining

$$NC_g = \{(x, c, Y) \in R^{n+|K|} \times B_{k \in K}^{\sum |D_k|} \mid g^l(x) \leq 0 \quad \forall l \in L\}$$

$$NC_{rik}^j = \{(x, c, Y) \in R^{n+|K|} \times B_{k \in K}^{\sum |D_k|} \mid r_{ik}^j \leq 0 \quad \forall i \in D_k, k \in K, j \in J_{ik}\}$$

$$NC_{\gamma_{ik}} = \{(x, c, Y) \in R^{n+|K|} \times B_{k \in K}^{\sum |D_k|} \mid c_k = \gamma_{ik}(x) \quad \forall i \in D_k, k \in K\}$$

$$C_g = \{(x, c, Y) \in R^{n+|K|} \times B^{\sum_{k \in K} |D_k|} \mid A_g^l x \leq b_g^l \quad \forall l \in L\}$$

$$C_{rik}^j = \{(x, c, Y) \in R^{n+|K|} \times B^{\sum_{k \in K} |D_k|} \mid A_{rik}^j x \leq b_{rik}^j \quad \forall i \in D_k, k \in K, j \in J_{ik}\}$$

$$C_{\gamma ik} = \{(x, c, Y) \in R^{n+|K|} \times B^{\sum_{k \in K} |D_k|} \mid A_{\gamma ik}(x, c_k) \leq b_{\gamma ik} \quad \forall i \in D_k, k \in K\}$$

$$G = \{(x, c, Y) \in R^{n+|K|} \times B^{\sum_{k \in K} |D_k|} \mid \Omega(Y) = true, \quad x^{lo} \leq x \leq x^{up}\}$$

The following relations clearly hold,

$$S_{NC} = NC_g \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} \bigcap_{j \in J_{ik}} NC_{rik}^j \right\} \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} NC_{\gamma ik} \right\} \cap G$$

$$S_{RLP} = C_g \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} \bigcap_{j \in J_{ik}} C_{rik}^j \right\} \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} C_{\gamma ik} \right\} \cap G$$

Since $NC_g \subseteq C_g$, $NC_{rik}^j \subseteq C_{rik}^j$ and $NC_{\gamma ik} \subseteq C_{\gamma ik}$ by definition of the linear over/under-estimators or polyhedral outer-approximations (Tawarmalani & Sahinidis 2002) it follows that,

$$S_{NC} = NC_g \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} \bigcap_{j \in J_{ik}} NC_{rik}^j \right\} \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} NC_{\gamma ik} \right\} \cap G \subseteq$$

$$C_g \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} \bigcap_{j \in J_{ik}} NC_{rik}^j \right\} \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} NC_{\gamma ik} \right\} \cap G \subseteq C_g \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} \bigcap_{j \in J_{ik}} C_{rik}^j \right\} \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} NC_{\gamma ik} \right\} \cap G \subseteq$$

$$C_g \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} \bigcap_{j \in J_{ik}} C_{rik}^j \right\} \cap \left\{ \bigcup_{k \in K} \bigcap_{i \in D_k} C_{\gamma ik} \right\} \cap G = S_{RLP}$$

Hence, $S_{NC} \subseteq S_{RLP}$. □

COROLLARY 1: *The optimal solution of (GDP_{RLP}) yields a valid lower bound Z^L to the global optimum of problem (GDP_{NC})*

Since the objective function for (GDP_{RLP}) and (GDP_{NC}) are the same, the proof follows trivially from PROPOSITON 1.

In order to predict strong lower bounds for the global optimum of (GDP_{NC}) , we consider the hierarchy of relaxations for (GDP_{RLP}) from the work of Sawaya & Grossmann (2008). These authors proved that any Linear Generalized Disjunctive Program (LGDP) that involves Boolean and continuous variables can be equivalently formulated as a Disjunctive Program (DP), that only involves continuous variables (See

Appendix 2). This means that we are able to exploit the wealth of theory behind DP from Balas (1979,1985) in order to solve LGDP.

One of the properties of disjunctive sets is that they can be expressed in many different equivalent forms. Among these forms, two extreme ones are the Conjunctive Normal Form (CNF), which is expressed as the intersection of elementary sets (i.e. sets that are the union of half spaces), and the Disjunctive Normal Form (DNF), which is expressed as the union of polyhedra. One important result in Disjunctive Programming Theory, as presented in the work of Balas (1985), is that we can systematically generate a set of equivalent DP formulations going from the CNF to the DNF by performing an operation called “basic step” that preserves regularity. A basic step is defined by the following theorem.

THEOREM 2.1. (Balas, 1985) *Let F be the disjunctive set in RF given by $F = \bigcap_{j \in T} S_j$ where $S_j = \bigcup_{i \in Q_j} P_i$, P_i a polyhedron, $i \in Q_j$. Then F can be brought to DNF by $|T|-1$ applications of the following basic steps, which preserve regularity:*

For some $k, l \in T, k \neq l$, bring $S_k \cap S_l$ to DNF by replacing it with $S_{kl} = \bigcup_{\substack{i \in Q_k \\ j \in Q_l}} (P_i \cap P_j)$

Note that from the above theorem, a basic step involves intersecting a given pair of disjunctions S_k and S_l .

REMARK: A particular case arises when for a given $k \in T$, S_k is given by a single polyhedron. In this case we denote S_k , as an improper disjunction (otherwise, we denote it as a proper disjunction). For example, the set of global constraints $A_g^l x \leq b_g^l$ $l \in L$ that appear in (GDP_{RLP}) are improper disjunctions. From the practical point of view, one important property of improper disjunctions is the fact that when a basic step is applied between one of them and a proper disjunction (i.e. intersecting them), the number of polyhedra in the resulting disjunctive set is not increased. This will become important later when the implementation of basic steps are discussed (See section 4).

Although the formulations obtained after the application of basic steps on the disjunctive sets are equivalent, their *continuous relaxations* are not. We denote the continuous relaxation of a disjunctive set $F = \bigcap_{j \in T} S_j$ in regular form where each S_j is a union of

polyhedra, as the *hull-relaxation* of F (or *h-rel* F). Where $h-rel F := \bigcap_{j \in T} cl conv S_j$ and $cl conv S_j$ denotes the closure of the convex hull of S_j . That is, if $S_j = \bigcup_{i \in Q_j} P_i$, $P_i = \{x \in R^n, A^i x \leq b^i\}$, then $cl conv S_j$ is given by, $x = \sum_{i \in Q_j} v^i, \lambda_i \geq 0, \sum_{i \in Q_j} \lambda_i = 1, A^i v^i \leq b^i \lambda_i, i \in Q_j$. Note that the convex hull of F is in general different from its hull-relaxation. As described in Theorem 4.3., the application of a basic step on a disjunctive set leads to a new disjunctive set whose relaxation is at least as tight, if not tighter, as the former.

THEOREM 4.3 (Balas, 1985) *For $i=0,1,\dots,t$ 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 $i=1,\dots,t$, F_i is obtained from F_{i-1} by a basic step*
 Then $h-rel F_0 \supseteq h-rel F_1 \supseteq \dots \supseteq h-rel F_t$.

Consider now the linear relaxation of (GDP_{NC}) , namely (GDP_{RLP0}) , which is equivalent to (GDP_{RLP}) . We introduce the subscript i to indicate the number of basic steps that has been applied to the initial Linear GDP relaxation (GDP_{RLP0}) that is obtained by the under/over-estimation of the nonconvex terms from the (GDP_{NC}) . Clearly, (GDP_{RLP0}) represents a disjunctive set that is between the CNF and the DNF. We denote (GDP_{RLPi}) as the DNF form of the Linear GDP relaxation.

By THEOREM 2.1 (Balas, 1985) we can write:

$$GDP_{RLP0} \sim GDP_{RLP1} \sim \dots \sim GDP_{RLPi} \dots \sim GDP_{RLPt}$$

where “ \sim ” means equivalent disjunctive sets and where GDP_{RLPi} can be obtained from GDP_{RLPi-1} by the application of a basic step.

From PROPOSITION 1 we know that $GDP_{RLP0} \supseteq GDP_{NC}$ then $GDP_{RLP0} \sim GDP_{RLPi} \supseteq GDP_{NC}$. Hence, the hierarchy of continuous linear relaxations for the nonconvex GDP_{NC} can be described as follows:

$$h-rel GDP_{RLP0} \supseteq h-rel GDP_{RLP1} \supseteq \dots \supseteq h-rel GDP_{RLPi} \dots \supseteq h-rel GDP_{RLPt} \supseteq GDP_{NC}$$

We can then establish the following proposition.

PROPOSITION 2: *The lower bounds for the global optimum obey the following relationship: $Z_{RLP0} \leq Z_{RLP1} \leq \dots \leq Z_{RLPi} \dots \leq Z_{RLPi} \leq Z_{NC}$ where $Z_{RLP0}, Z_{RLP1}, \dots, Z_{RLPi} \dots, Z_{RLPi}$ are the optimal solutions of the hull-relaxations of problems $GDP_{RLP0}, GDP_{RLP1}, \dots, GDP_{RLPi} \dots, GDP_{RLPi}$ respectively and Z_{NC} is the optimal solution of GDP_{NC} .*

PROOF: Since $Z_{RLP0}, Z_{RLP1}, \dots, Z_{RLPi} \dots, Z_{RLPi}, Z_{NC}$, are defined by the same objective function and $h-rel GDP_{RLP0} \supseteq \dots \supseteq h-rel GDP_{RLPi} \dots \supseteq h-rel GDP_{RLPi} \supseteq GDP_{NC}$, then the proof follows trivially. \square

Note that $h-rel GDP_{RLP0}$ is equivalent to the relaxation proposed by Lee and Grossmann (2003) when the under/over-estimation functions used on GDP_{NC} are linear.

REMARK: It is important to note that every time a basic step is applied, there might be an increase in the size of the problem. Hence, some rules to guide the implementation of this operation are necessary. These rules should consider two aspects. Firstly, the effect on the tightening of the relaxation given by a basic step, and secondly, the effect on the increase in the size of the formulation.

4- Rules to implement the basic steps on GDP

In order to make good use of the hierarchy of relaxations described in the previous section, one important aspect is to understand which basic steps will lead to an improvement in the tightness of the relaxation, and hence in a potential increase in the lower bound of the global optimum. In other words, we need to be able to differentiate among the basic steps that will lead to a strict inclusion with those that will keep the relaxation unchanged. With this objective in mind, Balas(1985) proposed the following theorem which can be readily extended to the case of linear GDP.

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

$$clconv(S_1 \cap S_2) = (clconv S_1) \cap (clconv S_2)$$

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

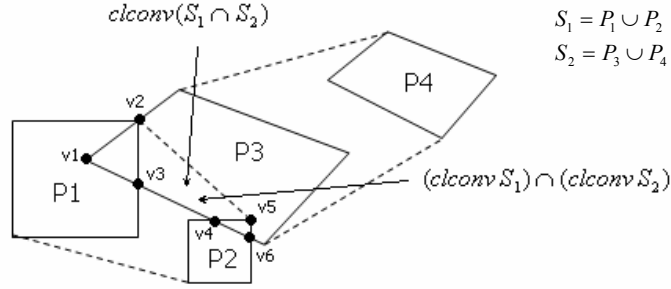


Fig. 1 Illustrative example of Theorem 4.5.

Fig. 1 illustrates the basic idea behind this theorem. Clearly, $clconv(S_1 \cap S_2) = (clconv S_1) \cap (clconv S_2)$. This is in agreement with Theorem 4.5 considering that the set of extreme points of $(clconv S_1) \cap (clconv S_2)$ given by $\{v1, v2, v5, v6\}$ is contained in the set of extreme points of $\{P_1 \cap P_3\} \cup \{P_2 \cap P_3\}$ given by $\{v1, v2, v3, v4, v5, v6\}$

Although to the best of our knowledge it is not easy to find in general a systematic and computationally efficient way to check the hypothesis of Theorem 4.5, we can still make use of it by exploiting the structure of particular problems. Some of the cases that are commonly found in GDP formulations are described below.

Case 1:

PROPOSITION 3 : Let $S = \bigcup_{i \in Q} P_i$ where $P_i, i \in Q = \{1,2\}$ are polyhedra defined in the

space, $x \in R^n$, H is a half space defined by $\alpha x + \beta \leq 0$ and H^* is a facet of H (i.e.

$\alpha x + \beta = 0$). If P_1 is a point in R^n such that $P_1 \subseteq H^*$, then

$$clconv(H \cap S) = clconv(H) \cap clconv(S).$$

PROOF: The system trivially satisfies the hypothesis of Theorem 4.5. □

Example: A particular case is found in process systems when the unit operation decisions are represented as follows:

$$\begin{bmatrix} Y \\ Bx \leq b \\ \alpha^T x \leq 0 \\ cp = s \end{bmatrix} \vee \begin{bmatrix} -Y \\ x = 0 \\ cp = 0 \end{bmatrix}$$

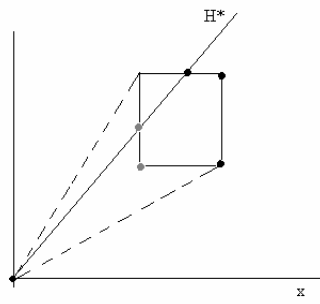


Fig 2: Illustration of Proposition 2

Clearly the hull relaxation of the above disjunction is the same as the hull relaxation of the following system (See Fig. 2),

$$\begin{bmatrix} Y \\ Bx \leq b \\ cp = s \end{bmatrix} \vee \begin{bmatrix} -Y \\ x = 0 \\ cp = 0 \end{bmatrix}$$

This particular case was already noted by Vecchiotti and Grossmann (2003)

Case 2:

PROPOSITION 4: Let $S_1 = \bigcup_{i \in Q_1} P_i^1$ and $S_2 = \bigcup_{i \in Q_2} P_i^2$ be two disjunctive sets defined in \mathbb{R}^n . If

the set of variables constrained in S_1 are not constrained in S_2 , and the set of variables constrained in S_2 are not constrained in S_1 , then $clconv(S_1 \cap S_2) = clconv(S_1) \cap clconv(S_2)$

.

PROOF: The system trivially satisfies the hypothesis of Theorem 4.5. □

Example:

$$S_1 = [a \leq x_1 \leq b] \quad , \quad S_2 = [c \leq x_2 \leq d] \cup [e \leq x_2 \leq f]$$

As can be seen from Fig 3, $clconv(S_1 \cap S_2) = clconv(S_1) \cap clconv(S_2)$

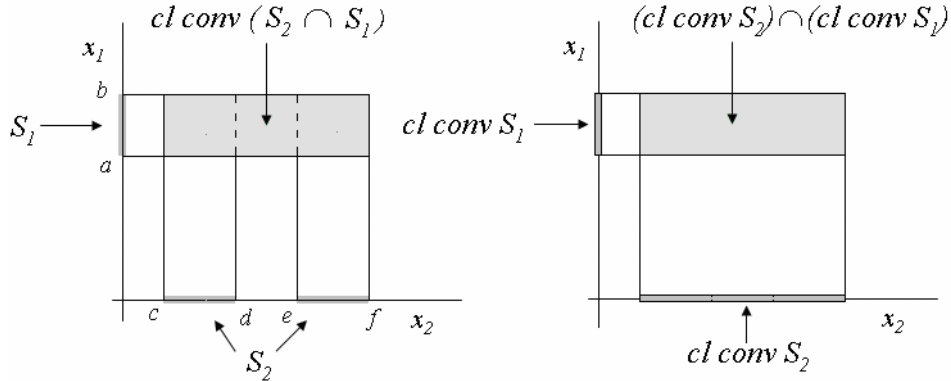


Fig 3 Illustration of Proposition 3

Another important aspect to consider is the effect that a particular basic step has in the increasing of the size of the formulation. In this respect we can differentiate two types of basic steps. Firstly, the ones that are implemented between two proper disjunctions, and second, the ones that are implemented between a proper and an improper disjunction. In this work we propose to use the latter approach. Note that in this case, parallel basic steps (i.e. simultaneous intersection of each improper disjunction with all proper disjunctions) will not lead to an increase in the number of polyhedra in the disjunctive set, keeping the size of the formulation smaller.

5- Illustrative Examples

In this section our aim is to illustrate how improved lower bounds can be obtained with the proposed framework in two simple examples. One of them is described by a BGDP, and the other described by a Concave GDP.

5-1- Example 1: Simple Bilinear GDP problem

In this section we illustrate how to obtain a strong relaxation in a simple BGDP. Fig. 4 shows a small superstructure consisting of two reactors, each characterized by a flow-conversion curve, a conversion range for which it can be designed, and its corresponding cost as can be seen in Table 1. The problem consists in choosing the reactor and

conversion that maximize the profit from sales of the product considering that there is a limit on the demand.

Table 1: Data for the reactors

Reactor	Curve*		Range		Cost
	α	β	X^{lo}	X^{up}	Cp
I	-8	9	0.2	0.95	2.5
II	-10	15	0.7	0.99	1.5

The characteristic curve is defined as $F = \alpha X + \beta$ in the range of conversions $[X^{lo}, X^{up}]$ where F and X are the flow of raw material and conversion respectively.

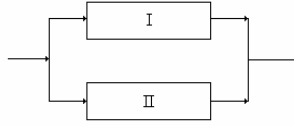


Fig 4. Two reactor network

The bilinear GDP model, which maximizes the profit, can be stated as follows:

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

$$\text{s.t. } FX \leq d$$

$$\left[\begin{array}{c} Y_{11} \\ F = \alpha_1 X + \beta_1 \\ X_1^{lo} \leq X \leq X_1^{up} \\ CP = Cp_1 \end{array} \right] \vee \left[\begin{array}{c} Y_{21} \\ F = \alpha_2 X + \beta_2 \\ X_2^{lo} \leq X \leq X_2^{up} \\ CP = Cp_2 \end{array} \right] \quad (GDP1_{NC})$$

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

$$X, F, CP \in R^1, F^{lo} \leq F \leq F^{up}, Y_{11}, Y_{21} \in \{\text{True}, \text{False}\}$$

The associated Linear GDP relaxation is obtained by replacing the bilinear term, FX , using the McCormick convex envelopes:

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

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

$$P \leq FX^{lo} + F^{up}X - F^{up}X^{lo} ; P \leq FX^{up} + F^{lo}X - F^{lo}X^{up}$$

$$P \geq FX^{lo} + F^{lo}X - F^{lo}X^{lo} ; P \geq FX^{up} + F^{up}X - F^{up}X^{up}$$

$$\left[\begin{array}{c} Y_{11} \\ F = \alpha_1 X + \beta_1 \\ X_1^{lo} \leq X \leq X_1^{up} \\ CP = Cp_1 \end{array} \right] \vee \left[\begin{array}{c} Y_{21} \\ F = \alpha_2 X + \beta_2 \\ X_2^{lo} \leq X \leq X_2^{up} \\ CP = Cp_2 \end{array} \right] \quad (GDP1_{RLPC})$$

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

$$X, F, CP \in R^1, F^{lo} \leq F \leq F^{up}, Y_{11}, Y_{21} \in \{\text{True}, \text{False}\}$$

Intersecting the improper disjunctions given by the inequalities of the relaxed bilinear term with the only proper disjunction (i.e. by applying five basic steps), we obtain the following GDP formulation,

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

s.t.

$$\left[\begin{array}{c} Y_{11} \\ P \leq d \\ P \leq FX^{up} + F^{lo}X - F^{lo}X^{up} \\ P \leq FX^{lo} + F^{up}X - F^{up}X^{lo} \\ P \geq FX^{lo} + F^{lo}X - F^{lo}X^{lo} \\ P \geq FX^{up} + F^{up}X - F^{up}X^{up} \\ F = \alpha_1 X + \beta_1 \\ X_1^{lo} \leq X \leq X_1^{up} \\ CP = Cp_1 \end{array} \right] \vee \left[\begin{array}{c} Y_{21} \\ P \leq d \\ P \leq FX^{up} + F^{lo}X - F^{lo}X^{up} \\ P \leq FX^{lo} + F^{up}X - F^{up}X^{lo} \\ P \geq FX^{lo} + F^{lo}X - F^{lo}X^{lo} \\ P \geq FX^{up} + F^{up}X - F^{up}X^{up} \\ F = \alpha_2 X + \beta_2 \\ X_2^{lo} \leq X \leq X_2^{up} \\ CP = Cp_2 \end{array} \right] \quad (GDPI_{RLPI})$$

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

$$X, F, CP \in R^1, F^{lo} \leq F \leq F^{up}, Y_{11}, Y_{21} \in \{\text{True}, \text{False}\}$$

Fig. 5 shows the actual feasible region of $(GDPI_{NC})$ and the projection onto the F-X space of the hull relaxations of $(GDPI_{RLP0})$ and $(GDPI_{RLPI})$. Notice that in this case the choice of reactor II is infeasible.

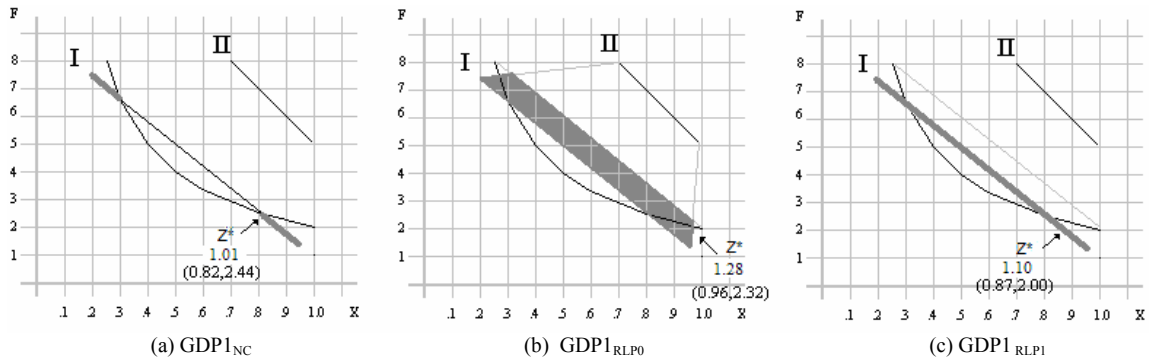


Fig. 5 a) Projected feasible region of $GDP1_{NC}$, b) Projected feasible region of relaxed $GDP1_{RLP0}$
c) Projected feasible region of relaxed $GDP1_{RLPI}$

From Fig 5 it is clear that the relaxed feasible region of $(GDPI_{RLPI})$ is contained in $(GDPI_{RLP0})$. This implies that if we solve the relaxed problem $(GDPI_{RLPI})$ we obtain an upper bound of the objective Z of 1.1 that is closer to the global optimal solution of 1.01. In contrast GDP_{RLP0} predicts a weaker upper bound of 1.28.

5-2- Example 2: Concave GDP problems

The following is a similar simple example to illustrate the proposed framework in Concave GDP problems. The problem consists in selecting one of the two reactors and its size with the objective of minimizing the loss while satisfying two demands, each specified within a given range (i.e. Range Demand 1 : $[Dc_1^{lo}, Dc_1^{up}]$, Range Demand 2 : $[Dc_2^{lo}, Dc_2^{up}]$) and a selling price (i.e. Price Demand 1: p_1 , Price Demand 2: p_2). Note that this problem can also be stated as a maximization of the profit.

Table 2: Data for the reactors

Reactor	Curve*		Range		Cost	
	α	β	Fb^{lo}	Fb^{up}	γ	δ
I	1.5	0	1.0	3.0	4.0	1.0
II	1.0	0	2.0	4.0	1.5	5.0

* The characteristic curve is defined as $Fb = \alpha Fa + \beta$ in the range of flows $[Fb^{LO}, Fb^{UP}]$ where Fb and Fa are the flow of final product and raw material, respectively. The processing cost is given by $Cp = \gamma Fb^{0.7} + \delta$

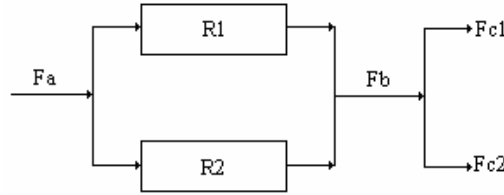


Fig 6 Superstructure for two reactor system

The concave GDP model, which minimizes the loss, can be stated as follows:

$$\text{Min } Z = Cp - p_1 Fc_1 - p_2 Fc_2$$

$$\text{s.t. } Fa = Fc_1 + Fc_2$$

$$\left[\begin{array}{c} Y_1 \\ Fb = \alpha_1 Fa + \beta_1 \\ Cp = \gamma_1 Fb^{0.7} + \delta_1 \\ Fb^{LO_1} \leq Fb \leq Fb^{UP_1} \end{array} \right] \vee \left[\begin{array}{c} Y_2 \\ Fb = \alpha_2 Fa + \beta_2 \\ Cp = \gamma_2 Fb^{0.7} + \delta_2 \\ Fb^{LO_2} \leq Fb \leq Fb^{UP_2} \end{array} \right] \quad (GDP2_{NC})$$

$$Y_1 \vee Y_2$$

$$Dc_1^{LO} \leq Fc_1 \leq Dc_1^{UP}$$

$$Dc_2^{LO} \leq Fc_2 \leq Dc_2^{UP}$$

$$Fb^{LO} \leq Fb \leq Fb^{UP}$$

The associated relaxed GDP program is obtained by replacing the concave terms using the secant as an under estimator. Note that in this particular case no over estimators are necessary since the solution of the relaxation will always be active in the inequality for the underestimator (i.e. secant in section 2). (For the general case, as noted in section 3, supporting hyperplanes can be used as over-estimators):

$$\begin{aligned}
& \text{Min } Z = Cp - p_1Fc_1 - p_2Fc_2 \\
& \text{s.t. } Fb = Fc_1 + Fc_2 \tag{GDP2_{RLP0}}
\end{aligned}$$

$$\left[\begin{array}{c} Y_1 \\ Fb = \alpha_1 Fa + \beta_1 \\ Cp = \gamma_1 Fb^* + \delta_1 \\ Fb^* \geq \frac{(Fb^{up})^{0.7} - (Fb^{lo})^{0.7}}{Fb^{up} - Fb^{lo}} (Fb - Fb^{lo}) + (Fb^{lo})^{0.7} \\ Fb^{lo1} \leq Fb \leq Fb^{up1} \end{array} \right] \vee \left[\begin{array}{c} Y_2 \\ Fb = \alpha_2 Fa + \beta_2 \\ Cp = \gamma_2 Fb^* + \delta_2 \\ Fb^* \geq \frac{(Fb^{up})^{0.7} - (Fb^{lo})^{0.7}}{Fb^{up} - Fb^{lo}} (Fb - Fb^{lo}) + (Fb^{lo})^{0.7} \\ Fb^{lo2} \leq Fb \leq Fb^{up2} \end{array} \right]$$

$$\begin{aligned}
& Y_1 \vee Y_2 \\
& Dc_1^{LO} \leq Fc_1 \leq Dc_1^{UP} \\
& Dc_2^{LO} \leq Fc_2 \leq Dc_2^{UP} \\
& Fb^{LO} \leq Fb \leq Fb^{UP}
\end{aligned}$$

Intersecting the improper disjunctions that are given by the bounds on the demand and the relationship between production and demand, with the only proper disjunction present (i.e. by applying seven basic steps), we obtain the following GDP formulation,

$$\begin{aligned}
& \text{Min } Z = Cp - p_1Fc_1 - p_2Fc_2 \\
& \text{s.t.} \tag{GDP2_{RLP1}}
\end{aligned}$$

$$\left[\begin{array}{c} Y_1 \\ Fb = \alpha_1 Fa + \beta_1 \\ Cp = \gamma_1 Fb^* + \delta_1 \\ Fb^* \geq \frac{(Fb^{up})^{0.7} - (Fb^{lo})^{0.7}}{Fb^{up} - Fb^{lo}} (Fb - Fb^{lo}) + (Fb^{lo})^{0.7} \\ Fb^{lo1} \leq Fb \leq Fb^{up1} \\ Fb = Fc_1 + Fc_2 \\ Dc_1^{LO} \leq Fc_1 \leq Dc_1^{UP} \\ Dc_2^{LO} \leq Fc_2 \leq Dc_2^{UP} \end{array} \right] \vee \left[\begin{array}{c} Y_2 \\ Fb = \alpha_2 Fa + \beta_2 \\ Cp = \gamma_2 Fb^* + \delta_2 \\ Fb^* \geq \frac{(Fb^{up})^{0.7} - (Fb^{lo})^{0.7}}{Fb^{up} - Fb^{lo}} (Fb - Fb^{lo}) + (Fb^{lo})^{0.7} \\ Fb^{lo1} \leq Fb \leq Fb^{up1} \\ Fb = Fc_1 + Fc_2 \\ Dc_1^{LO} \leq Fc_1 \leq Dc_1^{UP} \\ Dc_2^{LO} \leq Fc_2 \leq Dc_2^{UP} \end{array} \right]$$

$$Y_1 \vee Y_2$$

Fig. 7 shows the actual feasible region of $(GDP2_{NC})$ and the projection onto the Cp-Fb space of the hull relaxation of $(GDP2_{RLP0})$ and $(GDP2_{RLP1})$. As can be seen the

relaxation of ($GDP2_{RLP1}$) predicts a stronger lower bound at 5.33 versus the weaker lower bound of ($GDP2_{RLP0}$), 4.90.

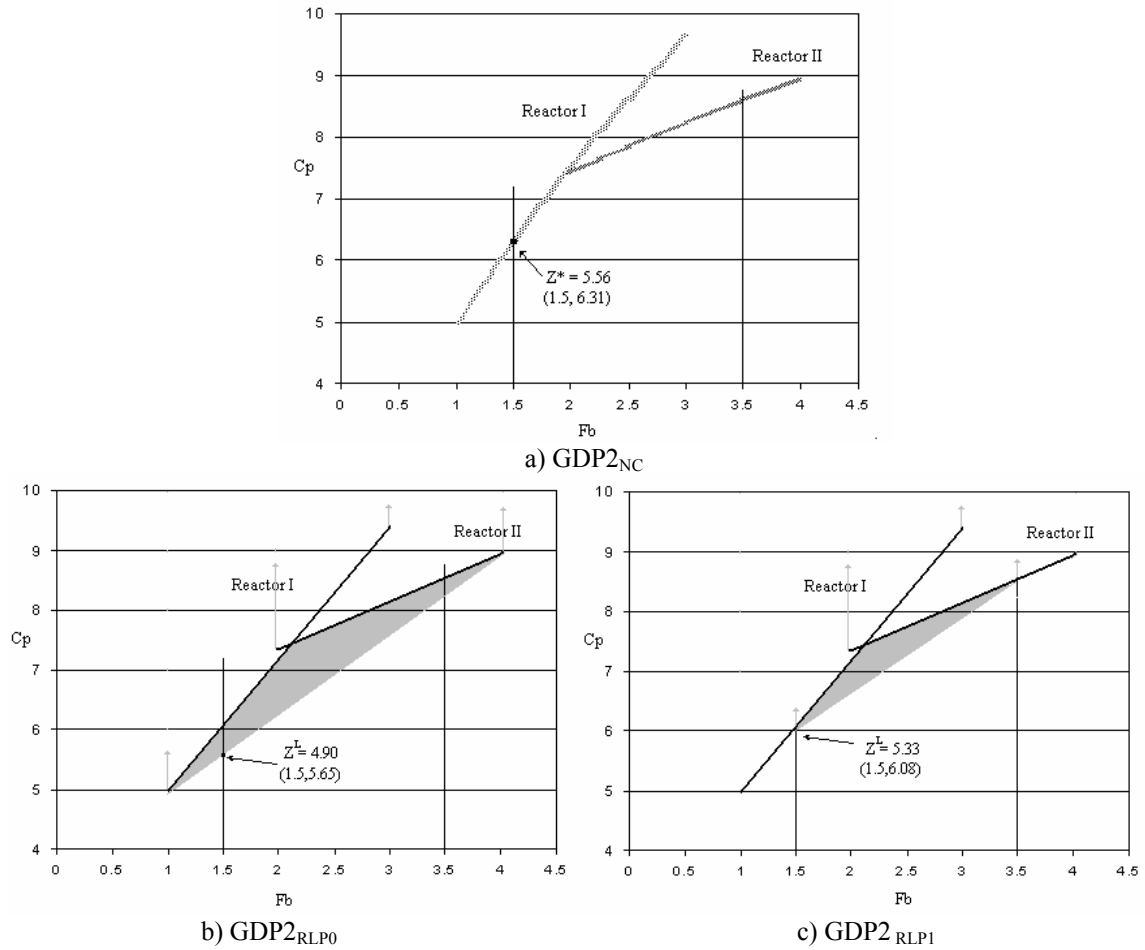


Fig 7. a) Projected feasible region of ($GDP2_{NC}$), b) Projected feasible region of relaxed ($GDP2_{RLP0}$)
c) Projected feasible region of relaxed ($GDP2_{RLP1}$)

6- Global Optimization algorithm with improved relaxations

The global optimization methodology of the GDP that we propose follows the well known spatial branch and bound method (Horst,1996) is obtained in the next section.

I. GDP REFORMULATION: The first step in the procedure consists of making use of the framework proposed in section 3 and 4 to obtain a tight GDP formulation. In summary:

- Relax the nonconvex terms using suitable linear convex under/over-estimators. This will lead to the Linear GDP (GDP_{RLP0}).
- Apply basic steps according to the rules described in section 3 and 4 (i.e. Parallel basic steps between improper disjunctions with proper disjunctions following Proposition 3.4).

II. UPPER BOUND AND BOUND TIGHTENING: After a reformulation is obtained, the procedure continues by finding an optimal or suboptimal solution of the problem to obtain an upper bound. This is accomplished by solving the nonconvex GDP reformulated as a MINLP (either as big-M or convex hull formulation) with a local optimizer such as DICOPT++/GAMS (Viswanathan & Grossmann,1990). By using the result obtained in the previous step, a bound contraction of each continuous variable is performed (Zamora & Grossmann, 1999). This is done by solving min/max subproblems in which the objective function is the value of the continuous variable to be contracted subject to the condition that the objective of the original problem is less than the upper bound.

III. SPATIAL BRANCH AND BOUND: After the relaxed feasible region is contracted, a spatial branch and bound search procedure is performed. This technique consists of splitting the feasible region recursively into subproblems that are eliminated when it is established that their descendants cannot contain a better solution than the one that has been obtained so far. The splitting is based on a “branching rule”, and the decision about when to eliminate the subproblems is performed by comparing the lower bound LB (i.e. the solution of the subproblem) with the upper bound UB (i.e. the feasible solution with the lowest objective function value obtained so far). The latter can be obtained by solving an NLP with all the discrete variables fixed in the corresponding subproblem); if $UB-LB < \varepsilon$, where ε is a given tolerance, then the node (i.e. subproblem) is eliminated.

From the above outline of the algorithm, there are two features that characterize the particular branch and bound technique: the branching rule and the way to choose the next subproblem to split. In the implementation of this work we have chosen to first branch on the discrete variable which most violates the integrality condition in the relaxed LP (i.e. choosing the discrete variable closest to 1/2), and then on the continuous variables by choosing the one that most violates the feasible region in the original problem (i.e. the violation to the feasible region is computed by taking the difference between the nonconvex term and the associated relaxed variable). To generate the subproblems when branching on the continuous variables, we split their domain by using the bisection method. To choose the node to branch next, we followed the “Best First”

heuristic that consists in taking the subproblem with lowest LB. The search ends when no more nodes remain in the queue.

REMARK 1: The framework presented does not restrict the use of different set of basic steps than the one proposed in this algorithm (e.g. between proper disjunctions)

REMARK 2: There is an inherent gap between the best relaxation attainable by the application of basic steps to DNF form and the global optimal solution. The objective value of this relaxation can be obtained from the MIP solution of the Linear GDP (GDP_{RLP}) (See Corollary 2.1.2 in Balas (1998)).

REMARK 3: An alternative approach to the above algorithm is to consider only partial use of the proposed relaxation. This considers the fact that the tighter relaxations of the disjunctive sets obtained through the proposed framework are often followed by an increase in the size of the reformulation. This leads to a potentially higher computational effort that might not be compensated by the strength in the relaxation. In future work we propose to use the proposed relaxation to calculate new bounds on the variables and then feed these new bounds to the Lee & Grossmann relaxation, which is of lower dimension. In Fig. 8 we show a schematic of the two original approaches and the alternative proposed.

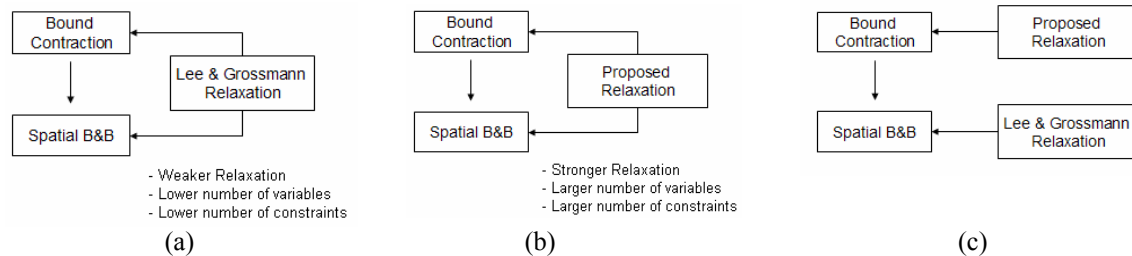


Fig. 8 a) Lee & Grossmann Method , b) Original Proposed Approach , c) Alternative Approach

7- Numerical performance

In this section we analyze the performance of the proposed algorithm through two sets of numerical examples. The first set consists of 4 instances of an analytical Bilinear GDP problem (i.e. Example 0) which aims at showing two of the main strengths of the proposed approach, namely, its capability to predict stronger lower bounds of the global optimum at the root node and its effect on the bound contraction procedure to produce

tighter lower and upper bounds for the continuous variables (See Appendix 3). Table 3 summarizes the characteristics of the instances.

Table 3 Characteristics and size of example problems

	Bilinear Terms	Discrete Variables	Continuous Variables
<i>Instance 1</i>	3	3	6
<i>Instance 2</i>	25	25	50
<i>Instance 3</i>	50	50	100
<i>Instance 4</i>	100	100	200

Table 4 Lower bounds of proposed framework

	Global Optimum	Lower Bound (Lee & Grossmann Relaxation)	Lower Bound (Proposed Relaxation)	Best Lower Bound
<i>Instance 1</i>	-1.5	-2.62	-2.25	-2.25
<i>Instance 2</i>	-12.5	-21.87	-18.75	-18.75
<i>Instance 3</i>	-25	-43.75	-37.5	-37.5
<i>Instance 4</i>	-50	-87.5	-75	-75

As it is shown in Table 4 a stronger lower bound is predicted by our approach. Moreover, as it can be seen in Table 5, the bound contraction achieved (i.e. 100%) in the continuous variables leads the global optimization technique to find the solutions at the root node. Note that for the larger instances, this leads to reasonable computational times.

Table 5 Performance of proposed methodology with spatial branch and bound.

	Global Optimum	Global Optimization Technique using Lee & Grossmann Relaxation			Global Optimization Technique using Proposed Relaxation		
		Nodes	Bound contract. (% Avg)	CPU Time (sec)	Nodes	Bound contract. (% Avg)	CPU Time (sec)
<i>Instance 1</i>	-1.5	94	48	41	1	100	11
<i>Instance 2</i>	-12.5	>5000	48	>3600	1	100	40
<i>Instance 3</i>	-25	>5000	48	>3600	1	100	99
<i>Instance 4</i>	-50	>5000	48	>3600	1	100	219

The second set of numerical examples consists of 6 problems that frequently arise in Process Systems, which include the illustrative problems in section 5 (i.e. Example 1 and Example 2) and four more (See Appendix 4). Example 3 and Example 6 deal with the optimization of a Heat Exchanger Network with discontinuous investment costs for the exchangers and can be represented by a nonconvex GDP with bilinear and concave constraints (Turkay & Grossmann, 1996). Example 4 deals with the optimization of a Wastewater Treatment Network whose associated nonconvex GDP formulation is a

bilinear GDP (Galan & Grossmann, 1998). Finally, Example 5 is a Pooling Design problem that can be also represented as a bilinear GDP (Lee & Grossmann, 2003).

Table 6 summarizes the characteristics and size of the examples, and Table 7 shows the computational performance of the Lee and Grossmann (2003) relaxation and the one proposed in this work.

Table 6 Characteristics and size of example problems

	Bilinear Terms	Concave Functions	Discrete Variables	Continuous Variables
<i>Example 1</i>	1	0	2	3
<i>Example 2</i>	0	2	2	5
<i>Example 3</i>	4	9	9	8
<i>Example 4</i>	36	0	9	114
<i>Example 5</i>	24	0	9	76
<i>Example 6</i>	11	24	24	24

Table 7 Lower bounds of proposed framework

	Global Optimum	Lower Bound (Lee & Grossmann Relaxation)	Lower Bound (Proposed Relaxation)	Best Lower Bound
<i>Example 1</i>	-1.01	-1.28	-1.10	-1.10
<i>Example 2</i>	5.56	4.90	5.33	5.33
<i>Example 3</i>	114384.78	91671.18	94925.77	97858.86
<i>Example 4</i>	1214.87	400.66	431.9	431.9
<i>Example 5</i>	-4640	-5515	-5468	-5241
<i>Example 6</i>	322122.09	260235.11	265361.46	281191.44

All the examples solved show an improvement in the lower bound. For instance, in Example 3 it increased from 91671 to 94925 which is a direct indication of the reduction of the relaxed feasible region. The column “Best Lower Bound”, as described previously, can be used as an indicator of the performance of our set of rules to apply basic steps. Note that in the Examples 1, 2 and 4, the lower bound proposed in our methodology is the best lower bound that can be obtained by solving the relaxed DNF, which is quite remarkable. A further indication of tightening is shown in Table 8 where numerical results of the branch and bound algorithm proposed in section 6 are presented. As it can be seen, the number of nodes that the spatial branch and bound algorithm requires for finding the global solution is significantly reduced in Examples 1, 3, 4 and 6. except in Example 5. This may be attributed to the location in which the bilinear terms arise in the formulation (i.e. bilinear terms inside the disjunctions) (See Appendix 5).

Table 8 Performance of proposed methodology with spatial branch and bound.

	Global Optimum	Global Optimization Technique using Lee & Grossmann Relaxation			Global Optimization Technique using Proposed Relaxation		
		Nodes	Bound contract. (% Avg)	CPU Time (sec)	Nodes	Bound contract. (% Avg)	CPU Time (sec)
<i>Example 1</i>	-1.01	5	35	2.1	1	38	1.4
<i>Example 2</i>	5.56	1	33	1.0	1	33	1.0
<i>Example 3</i>	114384.78	10	85	9.0	1	99	5.0
<i>Example 4</i>	1214.87	408	8	176	130	16	115
<i>Example 5</i>	-4640	162	1	89	140	1	93
<i>Example 6</i>	322122.09	18	98	24	5	99	18

Table 9 shows the size of the LP relaxation problem obtained with each methodology. Note that even when the proposed methodology leads to a significant increase in the size of the formulation, this is not translated proportionally to the solution time of the resulting LP. This is largely due to the preprocessing step which effectively reduces the size of the LP problem to be solved.

Table 9 Size of the LP relaxation for example problems

	Size of the LP Relaxation			
	Lee & Grossmann		Proposed	
	Constraints	Variables	Constraints	Variables
<i>Example 1</i>	23	15	28	15
<i>Example 2</i>	24	14	31	18
<i>Example 3</i>	87	52	206	106
<i>Example 4</i>	544	346	3424	1210
<i>Example 5</i>	3336	1777	4237	1777
<i>Example 6</i>	233	143	1137	693

8 – Conclusions

In this paper we have proposed a novel approach for obtaining stronger relaxations, and hence stronger lower bounds for the global optimization of bilinear and concave Generalized Disjunctive Programs. With this aim we proposed a general framework that combines the theory presented in the literature to obtain linear relaxations for nonconvex MINLPs (McCormick, 1976, Al-Khayyal & Falk, 1983, Tawarmalani & Sahinidis, 2002, Floudas, 2000) with the results of the work of Sawaya & Grossmann (2008) to obtain tighter relaxations for the case of Linear GDPs. We further exploited the theory behind Disjunctive Programming to guide the derivation of relaxations more efficiently by

considering the particular structure of the problems. The performance of this procedure was shown through a set of examples, six of which correspond to the Process Systems field. All of the examples showed improvements in the lower bounds at the root node, leading to a significant decrease in the enumerated nodes by the spatial branch and bound method. This is a direct indication of tightening that was achieved. Moreover, the fact that the lower bounds obtained by the proposed approach are close to the best achievable (i.e. bound obtained as a result of solving the DNF) is a further indication of the good performance of the method.

A major question that remains to be answered is how to implement the strong relaxations proposed in this work within a spatial branch and bound framework efficiently when dealing with large scale systems. As a future work we aim at tackling this problem by considering a hybrid approach that combines constraint programming techniques and cut generation strategies (Hooker, 2007).

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

Appendix 1. Lee & Grossmann GDP reformulation (Lee & Grossmann, 2000)

Consider the following GDP formulation:

$$\text{Min } Z = f(x) + \sum_k c_k$$

$$\text{s.t. } g(x) \leq 0$$

$$\bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ r_{ik}(x) \leq 0 \\ c_k = \gamma_{ik} \end{bmatrix} \quad k \in K$$

$$\Omega(Y) = \text{True}$$

$$x^{lo} \leq x \leq x^{up}$$

$$x \in R^n, c_k \in R^1, Y_{ik} \in \{\text{True}, \text{False}\}, k \in K, i \in D_k$$

The Big-M (BM) reformulation (see Nemhauser & Wolsey, 1988) is as follows:

$$\text{Min } Z = f(x) + \sum_{k \in K} \sum_{i \in D_k} \gamma_{ik} \lambda_{ik}$$

$$\text{s.t. } g(x) \leq 0$$

$$r_{ik}(x) \leq M(1 - \lambda_{ik}) \quad k \in K, i \in D_k \quad (BM)$$

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

$$A\lambda \geq a$$

$$x \in R^n, \lambda_{ik} \in \{0,1\}, k \in K, i \in D_k$$

where the variable λ_{ik} has a one-to-one correspondence with the Boolean variable Y_{ik} . Note that when $\lambda_{ik} = 0$ and the parameter M is large enough, the associated constraint becomes redundant; otherwise, it is enforced. Also, $A\lambda \geq a$ is the reformulation of the logic constraints in the discrete space, which can be easily accomplished as discussed in Raman and Grossmann (1994).

The convex hull reformulation proposed by Lee & Grossmann (2000) yields,

$$\text{Min } Z = f(x) + \sum_{k \in K} \sum_{i \in D_k} \gamma_{ik} \lambda_{ik}$$

$$\text{s.t. } g(x) \leq 0$$

$$x = \sum_{i \in D_k} v_{ik} \quad k \in K$$

$$\lambda_{ik} r_{ik}(v_{ik} / \lambda_{ik}) \leq 0 \quad k \in K, i \in D_k \quad (CH)$$

$$0 \leq v_{ik} \leq \lambda_{ik} U_v \quad k \in K, i \in D_k$$

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

$$A\lambda \geq a$$

$$x \in R^n, v_{ik} \in R^1, \lambda_{ik} \in \{0,1\}, k \in K, i \in D_k$$

where the function $\lambda_{ik} r_{ik}(v_{ik}/\lambda_{ik})$ is convex if $r_{ik}(\cdot)$ is convex. Furthermore, special treatment is required to implement this function as described in Sawaya and Grossmann (2006)

Appendix 2. Equivalence between Linear Generalized Disjunctive Programs and Disjunctive Programs (Sawaya & Grossmann, 2008).

Consider the following linear generalized disjunctive programming problem, (Raman & Grossmann, 1994)

$$\text{Min } Z = d^T x + \sum_k c_k$$

$$\text{s.t. } Bx \leq b$$

$$\bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ A_{ik} x \leq a_{ik} \\ c_k = \gamma_{ik} \end{bmatrix} \quad k \in K \quad (LGDP)$$

$$\Omega(Y) = \text{True}$$

$$x^{lo} \leq x \leq x^{up}$$

$$x \in R^n, c_k \in R^1, Y_{ik} \in \{\text{True}, \text{False}\}, k \in K, i \in D_k$$

Based on Proposition 2.2. below (for proof see Sawaya & Grossmann, 2008) (LGDP) can be stated as an equivalent Disjunctive Program (Balas, 1974) by replacing

Boolean variables $Y_{jk}, j \in J_k, k \in K$ inside the disjunctions by equalities $\lambda_{ik} = 1, i \in D_k, k \in K$, where λ is a vector of continuous variables whose domain is $[0, 1]$. Furthermore, the logical relations $\bigvee_{i \in D_k} Y_{ik}, k \in K$ and $\Omega(Y) = True$ are converted into algebraic equations, $\sum_{i \in D_k} \lambda_{ik} = 1, k \in K$, and $H\lambda \geq h$, respectively.

This yields the following model:

$$\text{Min } Z = d^T x + \sum_k c_k$$

$$\text{s.t. } Bx \leq b$$

$$\bigvee_{i \in D_k} \begin{bmatrix} \lambda_{ik} = 1 \\ A_{ik} x \leq a_{ik} \\ c_k = \gamma_{ik} \end{bmatrix} \quad k \in K \quad (DP)$$

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

$$H\lambda \geq h$$

$$x \in R^n, c \in R^k, \lambda_{ik} \in \{0,1\}, k \in K, i \in D_k$$

PROPOSITION 2.2 (Sawaya & Grossmann, 2008) The linear GDP model in (LGDP) is equivalent to the disjunctive program in (DP), in the sense that there exists a one-to-one correspondence between a feasible solution $(x, c, Y) \in R^{n+|K|} \times \{True, False\}^{\sum_{k \in K} |D_k|}$ to (LGDP) and a feasible solution $(x, c, \lambda) \in R^{n+|K| + \sum_{k \in K} |D_k|}$ to (DP).

Appendix 3

Since the strength of the relaxation of the nonconvex terms relies heavily on the bounds of each individual variable, it is of main importance to implement an efficient bound tightening procedure as part as the global solution method (Ryoo & Sahinidis, 1995, Tawarmalani & Sahinidis, 2002, Zamora & Grossmann, 1999). In this section we show how the relaxation procedure we proposed leads to a significant improvement in the

bound contraction performance when it is used in combination with the methodology proposed by Zamora & Grossmann (1999)

PROPOSITION 5: Let us consider the hyperrectangle $\Omega_0 = \{x : x_{lo} \leq x \leq x_{up}\}$ where x is a vector representing each complicating variable (i.e. variables that appear in at least one nonconvex term) and x_{lo} and x_{up} its lower and upper bounds explicitly defined in the formulation. Also, let us define $\Omega_{L\&G} = \{x : x_{lo}^{L\&G} \leq x \leq x_{up}^{L\&G}\}$ and $\Omega_{prop} = \{x : x_{lo}^{prop} \leq x \leq x_{up}^{prop}\}$ as the hyperrectangles obtained after applying the bound contraction procedure proposed by Zamora & Grossmann in combination with L&G relaxation and the proposed relaxation respectively. Then $\Omega_{L\&G} \supseteq \Omega_{prop}$

Proof

Since the L&G relaxation is contained in the proposed relaxation then the proof follows trivially.

Let us consider the following example (Example 0):

$$\text{Min } -\sum_i y_i$$

$$\text{s.t. } x_i y_i = 0.25$$

$$\begin{bmatrix} Y_{1i} \\ 0 \leq x_i \leq 0.5 \\ 0 \leq y_i \leq 0.5 \end{bmatrix} \vee \begin{bmatrix} Y_{2i} \\ 0.5 \leq x_i \leq 1 \\ 0.5 \leq y_i \leq 1 \end{bmatrix} \quad i=1,2,\dots,I$$

$$Y_{1i} \vee Y_{2i} = \text{true}$$

$$0 \leq x_i \leq 1$$

$$0 \leq y_i \leq 1$$

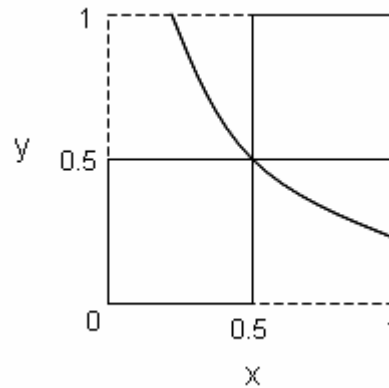


Fig 13 Feasible region $x=0.5$ $y=0.5$

Clearly the feasible region is defined by the point $x_i = 0.5$, $y_i = 0.5$ for $i = 1, 2, \dots, I$

Below we show the relaxed feasible region obtained after the bound contraction procedure when the Lee and Grossmann relaxation and the Proposed Relaxation are used

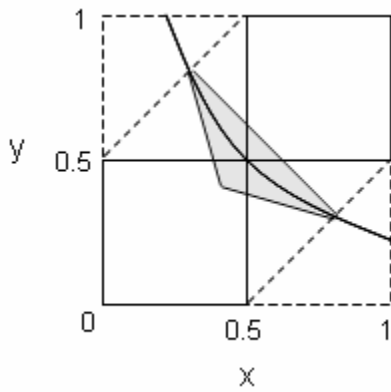


Fig 14: Relaxed Feasible Region using Lee & Grossmann relaxation for bounding. ($0.31 \leq x, y \leq 0.81$)

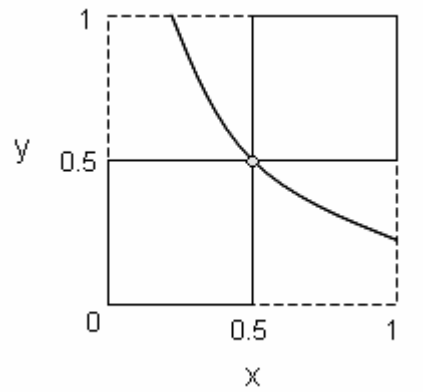


Fig 15: Relaxed Feasible Region using Proposed Relaxation for bounding. ($0.5 \leq x, y \leq 0.5$)

REMARK:

Although simple in its representation, the above example is inherently difficult to solve. For $I > 25$, the solver GAMS/BARON (Sahinidis, 1996) fails to close more than 10% GAP after 10000 sec.

Appendix 4. Case studies

EXAMPLE 3: Process Systems with discontinuous Investment Cost-Multiple Size Regions (Turkay & Grossmann, 1996).

This problem consists of finding the heat loads of utilities, the intermediate temperatures and the area of each exchanger that minimizes the investment and operation cost. Note that the structure of the network is fixed and that we use the arithmetic mean temperature as the heat driving force.

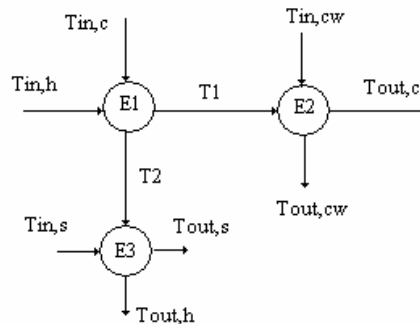


Fig 9 HEN structure of Example 3

Table 9 Data for Example 3

Heat Exchanger	Area (m ²)	Investment Cost (\$/yr)
1,2 and 3	0 ≤ A ≤ 10	2750 A ^{0.6} + 3000
	10 ≤ A ≤ 25	1500 A ^{0.6} + 15000
	25 ≤ A ≤ 50	600 A ^{0.6} + 46500

Heat Exchanger	Overall Heat Transfer
1	1.5
2	0.5
3	1

Stream	FCP(kW/K)	T _{in} (K)	T _{out} (K)	Cost(\$/kW yr)
Hot	10.0	500	340	
Cold	7.5	350	560	
Cooling Water		300	320	20
Steam		600	600	80

The problem can be formulated as follows:

$$\text{Min } Z = \sum_i CP_i + FCP_h(T_1 - T_{out,h})C_{cu} + FCP_c(T_{out,c} - T_2)C_{hu}$$

s. t.

$$FCP_h(T_{in,h} - T_1) = A_1 U_1 \frac{(T_{in,h} - T_2) + (T_1 - T_{in,c})}{2}$$

$$FCP_h(T_{in,h} - T_1) = A_2 U_2 \frac{(-T_{in,cw} + T_{out,h}) + (T_1 - T_{out,cw})}{2}$$

$$FCP_c(T_{out,c} - T_2) = A_3 U_3 \frac{(T_s - T_2) + (T_s - T_{out,c})}{2}$$

$$FCP_h(T_{in,h} - T_1) = FCP_c(T_2 - T_{in,c})$$

$$\left[\begin{array}{c} Y_{1i} \\ CP_i = 2750A_i^{0.6} + 3000 \\ 0 \leq A_i \leq 10 \end{array} \right] \vee \left[\begin{array}{c} Y_{2i} \\ CP_i = 1500A_i^{0.6} + 15000 \\ 10 \leq A_i \leq 25 \end{array} \right] \vee \left[\begin{array}{c} Y_{3i} \\ CP_i = 600A_i^{0.6} + 46500 \\ 25 \leq A_i \leq 50 \end{array} \right] \quad i=1,2,3$$

$$Y_{1i} \vee Y_{2i} \vee Y_{3i} = True \quad i = 1,2,3$$

$$T_1^{lo} \leq T_1 \leq T_1^{up}$$

$$T_2^{lo} \leq T_2 \leq T_2^{up}$$

$$Y_{ij} \in \{True, False\} \quad i = 1,2,3 \quad j = 1,2,3$$

$$T_i \in R^1 \quad i = 1,2$$

$$A_i \in R^1 \quad i = 1,2,3$$

Notation:

Parameters

FCP_i : Heat Capacity of stream i , where $i \in \{h,c\}$

$T_{in,i}$: Inlet temperature conditions of the stream i , where $i \in \{h,c\}$

$T_{out,i}$: Outlet temperature conditions of the stream i , where $i \in \{h,c\}$

U_i : Overall Heat Transfer Coefficient for exchanger E_i , where $i \in \{1,2,3\}$

C_{hu} : Cost of hot utility

C_{cw} : Cost of cooling water

Variables

$T_{1,2}$: Intermediate temperatures

A_i : Area of heat exchanger E_i , where $i \in \{1,2,3\}$

This problem is a nonconvex GDP where the nonconvexities arise from the bilinear terms inside the global constraints and concave functions inside the disjunctions.

EXAMPLE 4: Wastewater treatment network design

This example consists of the selection of the equipment necessary to process a given amount of contaminated water at the lowest cost to satisfy predetermined limits (Galan & Grossmann, 1998). See Figure 10 and Tables 10 and 11.

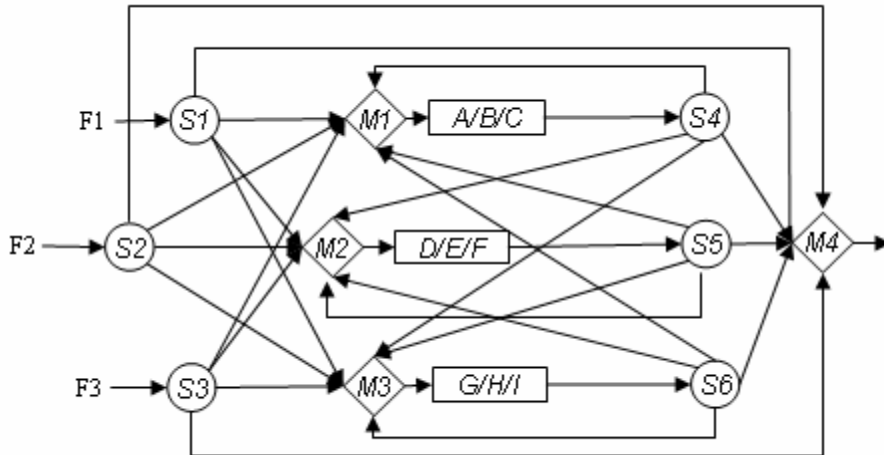


Fig. 10: System superstructure

Table 10: Stream quality

Stream	Flowrate(ton/h)	Pollutant	ppm
F1	2.0	A	1100
		B	300
F2	1.5	A	300
		B	700
F3	0.5	A	500
		B	1000

Table 11: Equipment data

Treatment Unit k	Equipment h	Removal Ratio (%)		Cost Function $\alpha (0.4922F + 0.5479) \pi + \gamma F$		
		A	B	α	γ	π
1	A	90	0	3480	0	0.1
	B	50	70	469	10	0.1
	C	0	80	26	1	0.1
2	D	0	90	726	0.0089	0.1
	E	0	99	1260	0.018	0.1
	F	50	99	5000	5.8	0.1
3	G	80	0	320	6	0.1
	H	50	50	58	15	0.1
	I	0	70	10	1	0.1

Notation:

Indices

i Streams
 k Units

j Components
 h Equipment Units

Sets

J Components j
 MU Mixers k
 PU Process Units k
 S_k Stream belong to k

SU Splitters k
 M_k Streams belong to mixer k
 IUP_k Stream input to process unit k
 OUP_k Stream output to PU k

Parameters

δ_{ik} Cost coefficient

Variables

f_i^j Individual component j of stream i
 f_k^j Individual component j of mixer k
 ζ_i^k Split fraction of splitter k into stream i

YP_k^h Selection of eq. h in k
 CP_k Cost of treatment for k

The problem can be formulated as a BGDGP as follows:

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

s.t.

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

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

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

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

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

$$0 \leq \zeta_i^k \leq 1 \quad \forall j, k$$

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

$$0 \leq CP_k \quad \forall k$$

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

This problem is a nonconvex GDP where the nonconvexities arise from the bilinear terms inside the global constraints.

EXAMPLE 5: Pooling problem

This example considers the optimal selection of a stream supply, pools and flows to minimize the total cost of a pooling network while satisfying the product requirements. The superstructure of the system as well as the data are shown below:

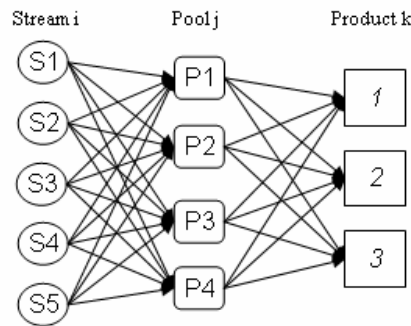


Fig. 11: System superstructure

Table 12 Supply stream quality and fix cost

w/i	S1	S2	S3	S4	S5
A	0.13	0.89	0.69	0.28	0.35
B	0.87	0.11	0.31	0.72	0.65
$\alpha_i(\$)$	260	70	150	190	110

Table 13 Cost of flows (\$/kg) and fix cost of pools

i/j	P1	P2	P3	P4
S1	6.2	9.4	7.6	10.2
S2	1.67	2.53	2.05	2.75
S3	3.58	5.42	4.39	5.89
S4	4.53	6.87	5.55	7.45
S5	2.62	3.98	3.22	4.32
$\gamma_j(\$)$	310	470	380	510

Table 14 Product quality , price and demand

w/k	1	2	3
A	0.56	0.3	0.41
B	0.44	0.7	0.59
$d_k(\$/kg)$	10.5	11.2	12.5
$S_k(kg)$	229	173	284

Notation:

Indices

i	Stream	j	Pools
k	Products	w	Components

Sets

I	Supply streams I	J	Pools j
K	Products k	W	Components w

Parameters

S_k	Demand of Product	α_i	Fixed cost for supply stream i
γ_j	Fixed cost for pool j	d_k	Price of product k
c_{ij}	Cost for flow from i to j	λ_{iw}	Quality of w from i
Z_{wk}	Quality requirement of w in k	$f^{up,lo}$	Upper/Lower bound of flows

Variables

f_{ijw}	Individual flow of w from i to j	f_{jkw}	Individual flow w from j to k
ζ_j^k	Split fraction from j to k	YST_i	Existence of supply stream i
YP_j	Existence of pool j	CST_i	Cost of supply stream i
CP_j	Cost of pool j		

The problem can be formulated as a BGDP as follows:

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

$$\text{s.t. } \sum_{i \in I} \sum_{w \in W} f_{ijw} = \sum_{k \in K} \sum_{w \in W} f_{jkw} \quad \forall j \in J$$

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

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

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

$$\left[\begin{array}{c} YST_i \\ f^{lo} \leq \sum_{j \in J} \sum_{w \in W} f_{ijw} \\ CST_i = \alpha_i \end{array} \right] \vee \left[\begin{array}{c} \neg YST_i \\ f_{ijw} = 0 \\ CST_i = 0 \end{array} \right] \quad \forall i \in I$$

$$\left[\begin{array}{c} YP_j \\ f^{lo} \leq \sum_{i \in I} \sum_{w \in W} f_{ijw} \\ \sum_{k \in K} f_{jkw} = \sum_{i \in I} f_{ijw}, \forall w \in W \\ f_{jkw} = \zeta_j^k \sum_{i \in I} f_{ijw}, \forall w \in W, k \in K \\ \sum_{k \in K} \zeta_j^k = 1 \\ CP_j = \gamma_j \end{array} \right] \vee \left[\begin{array}{c} \neg YP_j \\ f_{ijw} = 0, \forall i \in I, w \in W \\ f_{jkw} = 0, \forall k \in K, w \in W \\ CP_j = \gamma_j \end{array} \right] \quad \forall j \in J$$

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

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

This problem is a nonconvex GDP where the nonconvexities arise from the bilinear terms inside the disjunctions.

EXAMPLE 6: Process Systems with discontinuous Investment Cost-Multiple Size Regions (Turkay & Grossmann, 1996)

As in the example 3, this problem consists of finding the heat loads of utilities, the intermediate temperatures and the area of each exchanger that minimizes the investment and operation cost. Note that the structure of the network is fixed (See Fig. 12), and that we use the arithmetic mean temperature as the heat driving force. The data for this problem is presented in Table 15

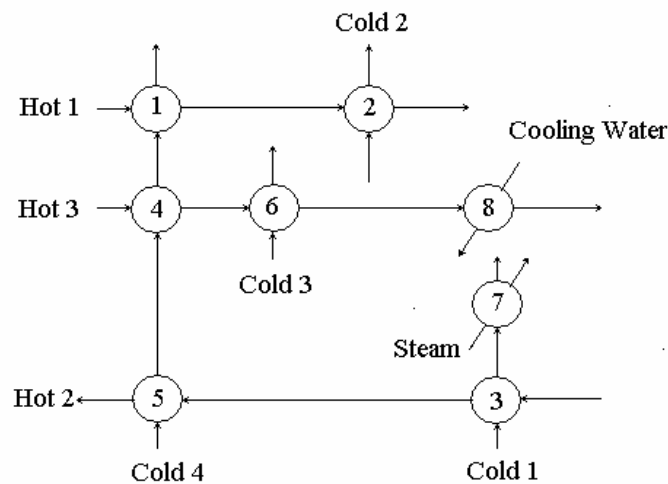


Fig 12

Table 15: Problem Data for Example 6

Heat Exchanger	Area (m ²)	Investment Cost (\$/yr)
1,2 and 3	$0 \leq A \leq 10$	$670 A^{0.83} + 2000$
	$10 \leq A \leq 25$	$640 A^{0.83} + 8000$
	$50 \leq A \leq 100$	$600 A^{0.83} + 16000$

Heat Exchanger	Overall Heat Transfer
1	1.5
2	0.2
3	0.06
4	1.6
5	0.04
6	0.3
7	0.6
8	1.7

Stream	FCP(kW/K)	T _{in} (K)	T _{out} (K)	Cost(\$/kW yr)
Hot 1	30.0	626	586	
Hot 2	13.5	620	519	
Hot 3	20.0	528	353	
Cold 1	31.0	525	613	
Cold 2	5.0	405	576	
Cold 3	28.0	353	386	
Cold 4	11.0	313	545	
Cooling Water		293	308	20
Steam		650	650	80

This problem is a nonconvex GDP where the nonconvexities come from bilinear terms representing energy balances and concave terms representing cost functions.

Appendix 5

PROPOSITION 6: When the bilinear terms are present outside the disjunctions, a tightening effect can be expected from the application of basic steps.

PROOF: The proof follows trivially from Theorem 4.3 (Balas, 1985). □

Illustrative example:

Given the following equivalent formulations (GDP_{B1}) and (GDP_{B2})

a) Bilinear term outside the disjunction

$$\text{Min } Z = d^T x$$

$$x_1 x_2 \leq 2$$

$$\left[\begin{array}{c} Y_1 \\ 1.3 \leq x_1 \leq 1.8 \\ 1.4 \leq x_2 \leq 1.9 \end{array} \right] \vee \left[\begin{array}{c} Y_2 \\ x_1 = 0 \\ x_2 = 0 \end{array} \right] \quad (GDP_{B1})$$

$$Y_1 \vee Y_2 = \text{True}$$

$$0 \leq x_1 \leq 2$$

$$0 \leq x_2 \leq 2$$

$$Y_{1,2} = \{\text{True}, \text{False}\}, x_{1,2} \in R^1$$

b) Bilinear term inside the disjunction

$$\text{Min } Z = d^T x$$

$$\left[\begin{array}{c} Y_1 \\ x_1 x_2 \leq 2 \\ 1.3 \leq x_1 \leq 1.8 \\ 1.4 \leq x_2 \leq 1.9 \end{array} \right] \vee \left[\begin{array}{c} Y_2 \\ x_1 = 0 \\ x_2 = 0 \end{array} \right] \quad (GDP_{B2})$$

$$Y_1 \vee Y_2 = \text{True}$$

$$0 \leq x_1 \leq 2$$

$$0 \leq x_2 \leq 2$$

$$Y_{1,2} = \{\text{True}, \text{False}\}, x_{1,2} \in R^1$$

It is clear from Fig.16 that the application of basic steps on (GDP_{B1}) leads to a tighter continuous relaxation. On the other hand, no improvement in the relaxation is observed when the application of basic steps is performed on (GDP_{B2})

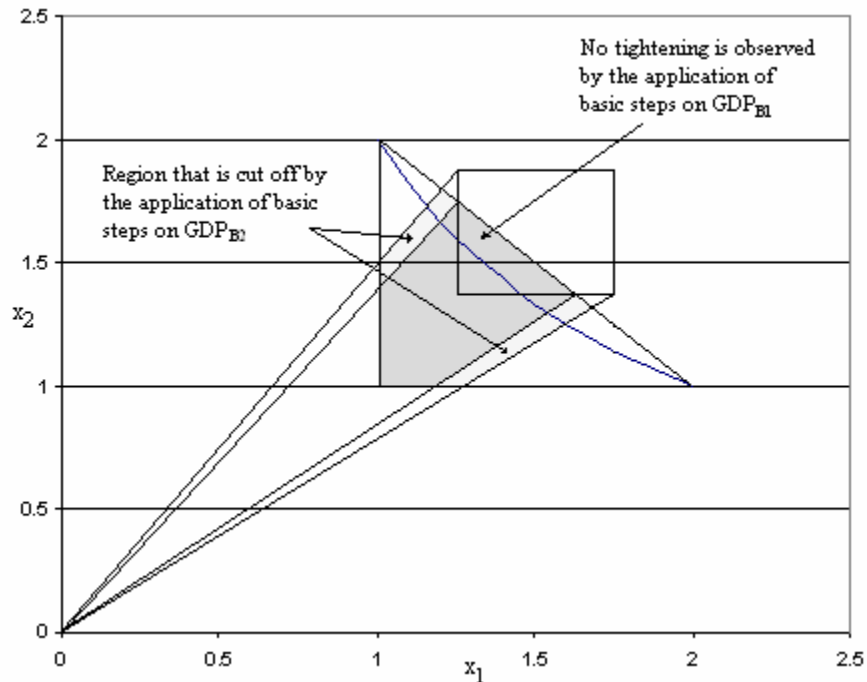


Fig.16

Balas, E. (1979) Disjunctive Programming . Annals of Discrete Mathematics 5 (1979) 3-51

Balas, E. (1985) Disjunctive Programming and a hierarchy of relaxations for discrete optimization problems. SIAM Journal on Algebraic and Discrete Methods, 6, 466-486.

Balas, E. (1998) Disjunctive Programming: Properties of the convex hull of feasible points. Discrete Applied Mathematics 89 (1998) 3-44

Benders J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. Numer. Math. 4, 238-252

Bonami, P., L.T. Biegler, A.R. Conn, G. Cornuejols, I.E. Grossmann, C.D. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, A. Wächter (2008) An algorithmic framework for convex mixed integer nonlinear programs, Discrete Optimization 5, 186-204.

Chen, 1987 J.J.J. Chen, Comments on improvements on a replacement for the logarithmic mean. Letters to the editor, *Chemical Engineering Science* 42 (1987) (10), pp. 2488–2489.

Duran, M. A.; Grossmann, I. E. (1986) An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming* 36, 307.

Floudas C.A. (2000), *Deterministic Global Optimization: Theory, Methods and Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Galan, B. & Grossmann, I. E. (1998). Optimal Design of Distributed Wastewater Treatment Networks. *Ind. Eng. Chem. Res.* 37, 4036-4048

Geoffrion A. M. (1972) Generalized Benders decomposition. *JOTA* 10, 237-260

- Grossmann, I. E. (2002). Review of nonlinear mixed-integer and disjunctive programming techniques. *Optimization and Engineering*, 3, 227-252.
- Gruber, P.M. & Kenderov (1982) Approximation of convex bodies by polytopes. *Rend Circ Mat Palermo* 31, 195-225, 1982
- Horst, R. & Tuy, H. (1996) *Global Optimization deterministic approaches* (3rd Ed) Berlin: Springer-Verlag
- Hooker, J (2007) *Integrated Methods for Optimization* : Springer Science
- Karuppiah, R. , Grossmann, I. E. (2006) Global optimization for the synthesis of integrated water systems in chemical processes. *Computers and Chemical Engineering* 20, 650-673
- Kesavan, P.; Allgor, R. J.; Gatzke, E. P.; Barton, P. I. (2004), Outer Approximation Algorithms for Separable Nonconvex Mixed-Integer Nonlinear Programs. *Mathematical Programming* 100, (3), 517.
- Leyffer, S. (2001) Integrating SQP and branch and bound for mixed integer nonlinear programming. *Computational Optimization and Applications* 18, 295.
- Lee, S. & Grossmann I. E. (2000) New algorithms for nonlinear generalized disjunctive programming. *Computers and Chemical Engineering* 24, 2125-2141
- Lee, S. & Grossmann I. E. (2003) Global optimization of nonlinear generalized disjunctive programming with bilinear inequality constraints: application to process networks. *Computers and Chemical Engineering* 27 , 1557-1575
- McCormick, G. P. (1976) Computability of global solutions to factorable nonconvex programs. Part I. Convex underestimating problems. *Mathematical Programming*, 10, 146-175.
- Meyer C. and C.A. Floudas (2006), "Global Optimization of a Combinatorially Complex Generalized Pooling Problem", *AIChE Journal*, 52, 1027-1037.
- Quesada, I. & Grossmann I. E. (1995b). Global optimization of bilinear process networks with multicomponent flows. *Computers and Chemical Engineering*, 19 (12), 1219-1242.
- Raman, R., & Grossmann, I. E. (1994). Modeling and Computational Techniques for logic based integer programming. *Computers and Chemical Engineering* 18 (7), 563-578.
- Sahinidis, N. (1996). BARON: A general purpose global optimization software package. *Journal of Global Optimization* , 8(2) 201-205.
- Sawaya & Grossmann (2008): "Reformulations, Relaxations and Cutting Planes for Linear Generalized Disjunctive Programming". Submitted for publication (2008)
- Tawarmalani, M., Sahinidis, N. (2002) *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming*. , Kluwer Academic Publishers
- Turkay M. & Grossmann I.E. (1996) Disjunctive Programming Techniques for the Optimization of Process Systems with Discontinuous Investment Costs-Multiple Size Regions. *Ind. Eng. Chem. Res.* 1996, 35, 2611-2623
- Vecchiotti, A., Lee, S. , Grossmann, I. E. (2003) Modeling of discrete/continuous optimization problems: characterization and formulation of disjunctions and their relaxations. *Computers and Chemical Engineering* 27, 433-448

Westerlund, T. and Pettersson, F. (1995) , "An extended cutting plane method for solving convex MINLP problems," Computers and Chemical Engineering, vol. 19(Suppl.), pp. S131-S136.

Viswanathan & Grossmann I.E. (1990) A combined penalty function and outer-approximation method for MINLP optimization Computers & Chemical Engineering, Volume 14, Issue 7, July 1990, Pages 769-782

Zamora, J. M. , & Grossmann, I. E. (1999). A branch and bound algorithm for problems with concave univariate , bilinear and linear fractional terms. Journal of Global Optimization 14 (3), 217-249