ELSEVIER

Computers and Chemical Engineering 00 (2002) 1–16

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

Modeling of discrete/continuous optimization problems: characterization and formulation of disjunctions and their relaxations

Aldo Vecchietti [a], Sangbum Lee [b], Ignacio E. Grossmann [b,*]

[a] INGAR, Instituto de Desarrollo y Diseño, UTN, Facultad Regional Santa Fe, USA
[b] Department of Chemical Engineering, Carnegie Mellon University, 1107 Doherty Hall, Pittsburgh, PA 15213-3890, USA

Abstract

This paper addresses the relaxations in alternative models for disjunctions, big-M and convex hull model, in order to develop guidelines and insights when formulating Mixed-Integer Non-Linear Programming (MINLP), Generalized Disjunctive Programming (GDP), or hybrid models. Characterization and properties are presented for various types of disjunctions. An interesting result is presented for improper disjunctions where results in the continuous space differ from the ones in the mixed-integer space. A cutting plane method is also proposed that avoids the explicit generation of equations and variables of the convex hull. Several examples are presented throughout the paper, as well as a small process synthesis problem, which is solved with the proposed cutting plane method.

Keywords: Discrete–continuous optimization; Mixed-integer nonlinear programming; Generalized disjunctive programming; Big-M relaxation; Convex hull relaxation

1. Introduction

Developing optimization models with discrete and continuous variables is not a trivial task. The modeler has often several alternative formulations for the same problem, and each of them can have a very different performance in the efficiency on the problem solution. In the area of Process System Engineering models commonly involve linear and nonlinear constraints and discrete choices. The traditional model that has been used in the past corresponds to a mixed-integer optimization program whose representation can be expressed in the following equation form (Grossmann & Kravanja, 1997):

$$\min Z = f(x) + d^T y$$

$$\text{s.t.} \quad g(x) \le 0 \qquad \text{(PA)}$$

$$r(x) + Ly \le 0$$

$$Ay \ge a$$

$$x \in R^n, \quad y \in \{0,1\}^q$$

where $f(x)$, $g(x)$ and $r(x)$ are linear and/or nonlinear functions. In the model (PA) the discrete choices are represented with the binary variables y involving linear terms.

More recently, generalized disjunctive programming (Raman & Grossmann, 1994; Türkay & Grossmann, 1996) has been proposed as an alternative to the model (PA). A generalized disjunctive program can be formulated as follows:

$$\min Z = \sum_{k \in K} c_k + f(x)$$

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

$$\bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ h_{ik}(x) \le 0 \\ c_k = \gamma_{ik} \end{bmatrix} \quad k \in K \qquad \text{(GDP)}$$

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

$$x \in R^n, \quad Y_{ik} \in \{\text{True}, \ \text{False}\}^m, \quad c_k \ge 0$$

where the discrete choices are expressed with the Boolean variables Y_{ik} in terms of disjunctions, and logic propositions $\Omega(Y)$. The attractive feature of Generalized Disjunctive Programming (GDP) is that it allows a

* Corresponding author. Tel.: +1-412-268-2230; fax: +1-412-268-7139

E-mail addresses: aldovec@alpha.arcride.edu.ar (A. Vecchietti), grossmann@cmu.edu (I.E. Grossmann).

0098-1354/02/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0098-1354(02)00220-X

symbolic/quantitative representation of discrete and continuous optimization problems. Modeling language for GDP problem has been discussed by Vecchietti and Grossmann (2000).

An approach that combines the previous two models is a hybrid model proposed by Vecchietti and Grossmann (1999) where the discrete choices can be modeled as mixed-integer constraints and/or disjunctions. In this way we can potentially exploit the advantages of the two previous formulations by expressing part of it only in algebraic form, and the other in a symbolic/quantitative form. The hybrid formulation is as follows:

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

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

$$r(x)+Ly\leq 0$$

$$Ay\geq a \qquad \text{(PH)}$$

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

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

$$x\in R^n,\quad y\in\{0,1\}^q,\quad Y_{ik}\in\{\text{True},\ \text{False}\}^m,\quad c_k\geq 0$$

where $r(x)+Ly\leq 0$ is general mixed-integer constraints that can be linear/nonlinear equations/inequalities. These terms can be seen as disjunctions transformed into mixed-integer form. $Ay\geq a$ represents general integer equalities/inequalities transformed from former logic propositions.

An issue that is unclear is how the modeler should express the discrete choices, either as a symbolic disjunction, or in a mixed-integer form (Bockmayr & Kasper, 1998). One possible guideline for this decision is the gap between the optimal value of the continuous relaxation and the optimal integer value. Since several algorithms involve the solution of the relaxed problem, we will investigate in this paper the tightness of different relaxations for a disjunctive set: the big-M formulation (Nemhauser & Wolsey, 1988), the Beaumont surrogate (Beaumont, 1990) and the convex hull relaxation (Balas, 1979; Lee & Grossmann, 2000). The big-M formulation and the Beaumont surrogate can be regarded as 'obvious' constraints. However, the convex hull relaxation of a disjunction is tighter, and can be transformed into a set of mixed-integer constraints. The advantage of the convex hull relaxation is that the tight lower bound helps to reduce the search effort in the branch and bound procedure, in both nonlinear and linear problems (for examples of significant node reductions see Lee & Grossmann, 2000; Jackson & Grossmann, 2002). But the drawback with the convex hull formulation is that it increases the number of continuous variables and constraints of the original problem. This can potentially make a problem more expensive to solve, especially in large problems. The big-M relaxation is more convenient to use when the problem size does not increase substantially when compared with the convex hull relaxation (see Yeomans & Grossmann, 1999, who found the big-M to be more effective). But generally the lower bound by big-M relaxation is weaker, which may require longer CPU time than the convex hull relaxation. Therefore, depending on the case, there is a trade-off between the best possible relaxation and the problem size. In order to exploit the tightness of the convex hull relaxation, but without the substantial increase of the constraints, it will be shown that cutting planes can be used that correspond to a facet of the convex hull.

In this paper we first introduce the definition and properties of a disjunctive set. We then present the different relaxations and their properties. Finally, a cutting plane method is discussed, and illustrated with several small example problems. The goal of this paper is not to perform a detailed computational study, but rather to provide insights into the modeling and solution of disjunctive problems.

2. Definitions and properties of a disjunctive set

A disjunctive set F can be expressed as a set of constraints separated by the or ($\vee$) operator:

$$F=\bigvee_{i\in D}[h_i(x)\leq 0]\quad x\in R^n \tag{1}$$

It is assumed that $h_i(x)$ is a continuous convex function. F can be considered as a logical expression, which enforces only one set of inequalities. The feasible region of each disjunctive term can be expressed as the set of points that satisfy the inequality.

$$R_i=\{x|h_i(x)\leq 0\} \tag{2}$$

A disjunctive set can be expressed in other forms that are logically equivalent. F can also be expressed as the union of the feasible regions of the disjunctive terms, which is called Disjunctive Normal Form (DNF):

$$F=\bigcup_{i\in D}[h_i(x)\leq 0]\quad x\in R^n \tag{3}$$

$$F=\bigcup_{i\in D} R_i \tag{4}$$

If the union of the feasible regions of the disjunctive terms is equal to one of its terms, R_j, which is the largest feasible region, then the disjunctive set is called *improper*. Otherwise the disjunctive set is called *proper* (Balas, 1985). The *improper* disjunctive set can be written as follows:

$$F=\bigcup_{i\in D} R_i = R_j \tag{5}$$

The *improper* disjunctive set has also the following property:

$$R_i \subseteq R_j \quad \forall i \neq j \tag{6}$$

which means that the feasible regions i ($i \neq j$) in the disjunctive set F are included in the jth feasible region. Since F is expressed as the union of the different terms, an *improper* disjunctive set can be reduced to:

$$F = \{x | h_j(x) \leq 0\} \tag{7}$$

On the other hand, a *proper* disjunctive set is the one in which either the intersection of the feasible regions is empty, or else it is non-empty, but Eq. (5) does not apply. Therefore, for a *proper* disjunctive set, either there is no intersection among the feasible regions:

$$\bigcap_{i \in D} R_i = \emptyset \tag{8}$$

or else, there is some intersection, but no set R_j contains all of them:

$$\bigcap_{i \in D} R_i \neq \emptyset, \quad \bigcup_{i \in D} R_i \neq R_j \tag{9}$$

3. Relaxations of a disjunctive set

Given a disjunctive set as condition Eq. (1) there are a number of relaxations that can be derived, the big-M, the Beaumont surrogate and the convex hull relaxations. We consider below the case of convex nonlinear constraints, which easily simplifies to the linear case.

3.1. Big-M relaxation

Consider the following nonlinear disjunction:

$$F = \bigvee_{i \in D} [h_i(x) \leq 0] \quad x \in R^n \tag{10}$$

where $h_i(x)$ is a nonlinear convex function. For simplicity, and without loss of generality, it is assumed that each term in the disjunction Eq. (10) has only one inequality constraint. The big-M relaxation of Eq. (10) is given by:

$$h_i(x) \leq M_i(1 - y_i) \quad i \in D$$

$$\sum_{i \in D} y_i = 1$$

$$0 \leq y_i \leq 1, \quad i \in D \tag{11}$$

Again the tightest value for M_i can be calculated from:

$$M_i = \max\{h_i(x) | x^L \leq x \leq x^U\} \tag{12}$$

3.2. Beaumont relaxation

Beaumont (1990) proposed a valid inequality for the disjunctive set Eq. (10). A valid M_i value must be calculated as in Eq. (12). By dividing each constraint $i \in D$ in Eq. (11) by M_i and summing over $i \in D$, the Beaumont surrogate, which interestingly does not involve binary variables is given as follows:

$$\sum_{i \in D} \frac{h_i(x)}{M_i} \leq N - 1 \tag{13}$$

where $N = |D|$ in Eq. (10). Beaumont showed that Eq. (13) yields an equivalent relaxation as the big-M relaxation Eq. (11) projected onto the continuous x space when the constraints in Eq. (10) are linear.

3.3. Convex hull relaxation

The convex hull relaxation for the disjunctive set Eq. (10) can be written as follows (Lee & Grossmann, 2000):

$$x - \sum_{i \in D} v_i = 0 \quad x, v_i \in R^n$$

$$y_i h_i\left(\frac{v_i}{y_i}\right) \leq 0, \quad i \in D$$

$$\sum_{i \in D} y_i = 1$$

$$0 \leq y_i \leq 1 \quad i \in D$$

$$0 \leq v_i \leq v_i^U y_i, \quad i \in D \tag{14}$$

where v_i^U is a valid upper bound for the disaggregated variables v_i, usually chosen as x^U. The Eq. (14) define a convex set in the (x, v, y) space provided the inequalities $h_i(x) \leq 0$, $i \in D$ are convex and bounded. The convex hull in Eq. (14) can be proved to be tighter or at least as tight as the big-M relaxation (see Appendix A). Also, for case of linear disjunctions, $F = \vee_{i \in D}[a_i^T x \leq b_i] \quad x \in R^n$, Eq. (14) reduces to the equations by Balas (1979, 1988):

$$x - \sum_{i \in D} v_i = 0 \quad x, v_i \in R^n$$

$$a_i^T v_i - b_i y_i \leq 0, \quad i \in D$$

$$\sum_{i \in D} y_i = 1$$

$$0 \leq y_i \leq 1, \quad i \in D$$

$$0 \leq v_i \leq y_i v_i^{up}, \quad i \in D \tag{15}$$

3.4. Example 1

Consider the following nonlinear disjunction:

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

where $0\le x_1\le 5$ and $0\le x_2\le 5$. The feasible region is shown in Fig. 1. Figs. 2 and 3 show the feasible region of the big-M and the convex hull relaxations, respectively.

The big-M relaxation is given by:

$$(x_1-1)^2+(x_2-1)^2\le 1+31(1-y_1)$$
$$(x_1-4)^2+(x_2-2)^2\le 1+24(1-y_2)$$
$$(x_1-2)^2+(x_2-4)^2\le 1+24(1-y_3)$$
$$y_1+y_2+y_3=1$$
$$0\le x_1,\ x_2\le 5,\quad 0\le y_i\le 1,\quad i=1,\ 2,\ 3 \qquad (16)$$

where the big-M parameters are calculated by Eq. (12). The convex hull of Fig. 3 is given by the equations:

$$x_1=v_{11}+v_{12}+v_{13}$$
$$x_2=v_{21}+v_{22}+v_{23}$$
$$(y_1+\varepsilon)\left[\left(\frac{v_{11}}{y_1+\varepsilon}-1\right)^2+\left(\frac{v_{21}}{y_1+\varepsilon}-1\right)^2-1\right]\le 0$$
$$(y_2+\varepsilon)\left[\left(\frac{v_{12}}{y_2+\varepsilon}-4\right)^2+\left(\frac{v_{22}}{y_2+\varepsilon}-2\right)^2-1\right]\le 0$$
$$(y_3+\varepsilon)\left[\left(\frac{v_{13}}{y_3+\varepsilon}-2\right)^2+\left(\frac{v_{23}}{y_3+\varepsilon}-4\right)^2-1\right]\le 0$$
$$y_1+y_2+y_3=1$$

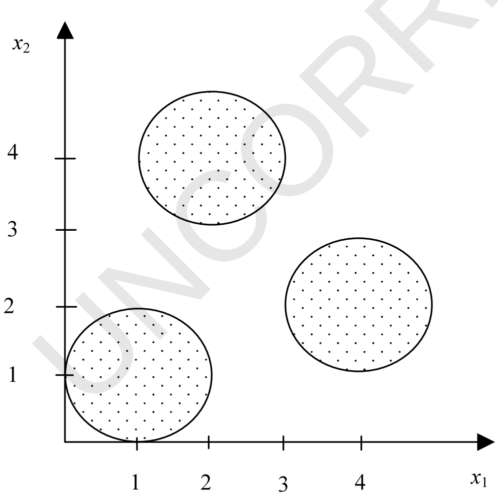

Fig. 1. Feasible region of example 1.

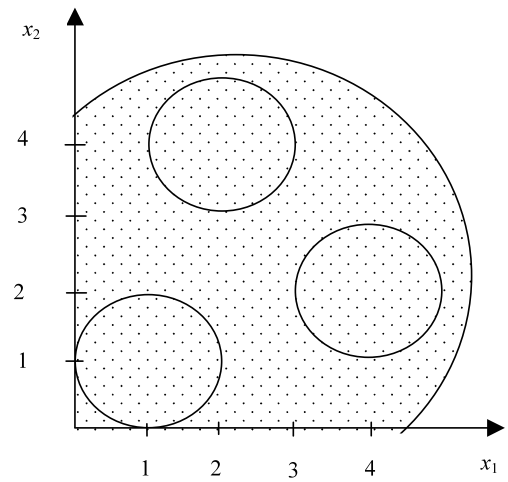

Fig. 2. Big-M relaxation of example 1.

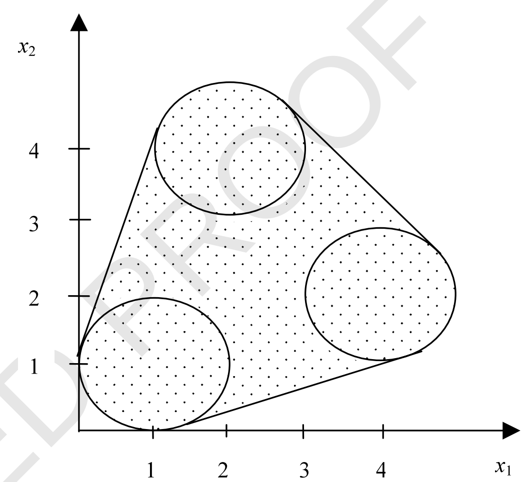

Fig. 3. Convex hull relaxation of example 1.

$$0\le y_i\le 1,\quad i=1,\ 2,\ 3$$
$$0\le v_{ji}\le 5y_i\quad \forall i,\forall j \qquad (17)$$

Note that to avoid division by zero ε is introduced in the nonlinear inequalities as a small tolerance (Lee & Grossmann, 2000). Typical values for ε are 0.001–0.0001. From Figs. 2 and 3 it is clear that the convex hull relaxation of the disjunctive set is tighter than the big-M relaxation for this example.

4. Impact of nature of disjunctions on relaxations in *x* space

Our aim in this section is to analyze different types of disjunctions for which it may be convenient or not to transform them into the convex hull formulation or a big-M formulation or Beaumont surrogate. Since the big-M formulation is as tight as the Beaumont surrogate, and it is more frequently used, we will compare the convex hull with only the big-M formulation. We will analyze the following cases: (a) *improper* disjunction; (b) *proper* disjunction. Within this last case we will analyze when the intersection of the feasible regions is empty and when it is non-empty.

If we denote the feasible region of the convex hull relaxation in the continuous x space as R_{CH}, the feasible region of the big-M relaxation as R_{BM}, and the feasible region of the Beaumont surrogate as R_B, then according to the properties shown in the previous section, the following can be established:

$$R_{CH} \subseteq R_{BM} \tag{18}$$

Beaumont (1990) has shown for the linear case that $R_{BM} = R_B$ where R_B is defined by constraint Eq. (13). In the Appendix A we show that $R_{BM} \subseteq R_B$ for nonlinear case. Therefore, the following property holds:

$$R_{BM} \subseteq R_B \tag{19}$$

It should be noted that properties Eqs. (19) and (20) apply in the space of the continuous variables x.

4.1. Improper disjunction

When the disjunctive set is *improper*, the property in Eq. (6) holds. Since the feasible region of one term contains the feasible regions of the other terms, the relaxations of the convex hull and of the big-M can be selected to be identical. The reason is that the redundant terms can be dropped and the disjunctive set can be represented by the term with the largest feasible region R_j. For example, suppose we have the following problem:

$$\min Z = (x_1 - 3.5)^2 + (x_2 - 4.5)^2$$

$$\text{s.t.} \quad \begin{bmatrix} Y_1 \\ 1 \le x_1 \le 3 \\ 2 \le x_2 \le 4 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ 2 \le x_1 \le 3 \\ 3 \le x_2 \le 4 \end{bmatrix} \tag{20}$$

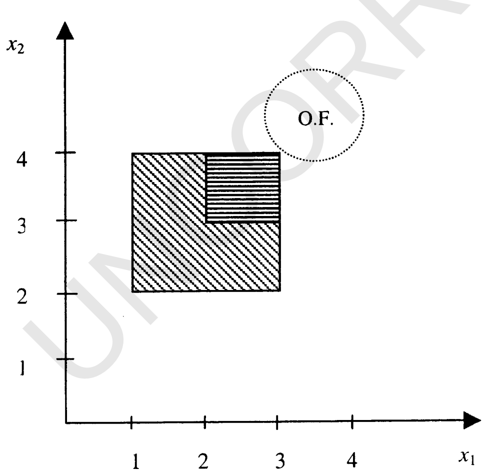

Fig. 4. Feasible region of disjunctive set Eq. (20).

The feasible region is shown in Fig. 4. Choosing the term with the largest feasible region, which is the first one, and solving the problem as an NLP we obtain the optimal solution $x = (3,4)$ and $Z = 0.5$. If we are not aware that the feasible regions are overlapped we can generate the big-M relaxation for this problem. If we use $M_i = 0.5$, $i = 1$, 2, and solve the relaxed Mixed-Integer Non-Linear Programming (MINLP) problem, the solution is $x = (3.25,4.25)$, $Z = 0.125$, $y = (0.5,0.5)$. If we choose $M_i = 1$ and solve the relaxed MINLP then the solution is $x = (3.5,4.5)$, $Z = 0$, $y = (0.5,0.5)$. Therefore, it is clear that arbitrary choice of M_i can yield a relaxation whose feasible region is larger than the disjunctive term with the largest feasible region. For the convex hull formulation it is clear that the resulting relaxation coincides with the region of the largest term in the x space, but at the expense of expressing it through disaggregated variables and additional constraints.

4.2. Proper disjunction

4.2.1. Non-empty intersecting feasible regions

When the feasible regions of the disjunctive terms have an intersection, it is not clear whether or not the convex hull and the big-M formulation could yield the same relaxation. Suppose we have disjunctions whose feasible regions are shown in Figs. 5 and 6. In Fig. 5 it is clear that the big-M relaxation, with a good selection of the M_i values can yield the same relaxation as the convex hull. For the case of Fig. 6 the convex hull will yield a tighter relaxation.

4.2.2. Disjoint disjunction

If the feasible region defined by each term in the disjunction has no intersection with others, then the disjunction is disjoint and *proper*. Fig. 7 shows an example of disjoint disjunction. In this case, it is clear that the convex hull relaxation should generally be

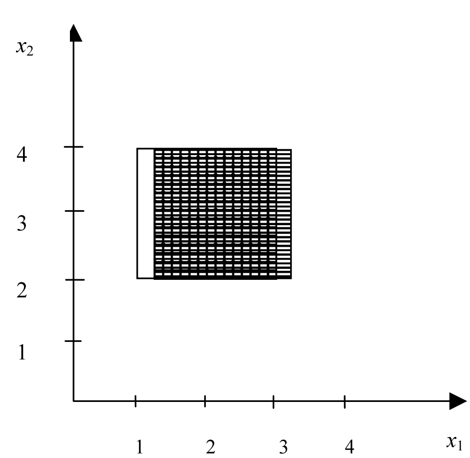

Fig. 5. Intersecting disjunction.

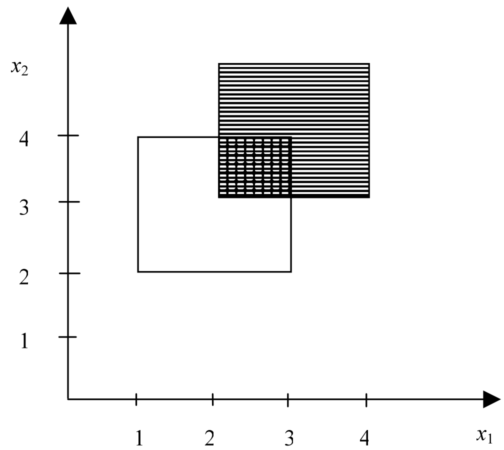

Fig. 6. Intersecting disjunction.

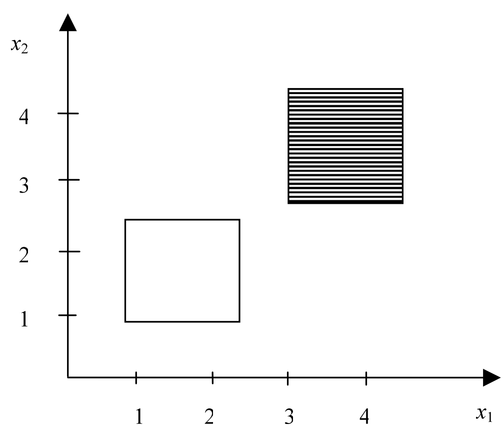

Fig. 7. Disjoint disjunction (general case).

tighter than the big-M relaxation (an exception is the particular case shown in Fig. 8). Also, in the special case shown in Fig. 9, where a disjunction has two terms with linear constraints and one of them yields zero point as a feasible region, the convex hull yields a cone with the zero point as the vertex. In this case, the convex hull relaxation can be simplified by not requiring disaggregated variables as given by the following:

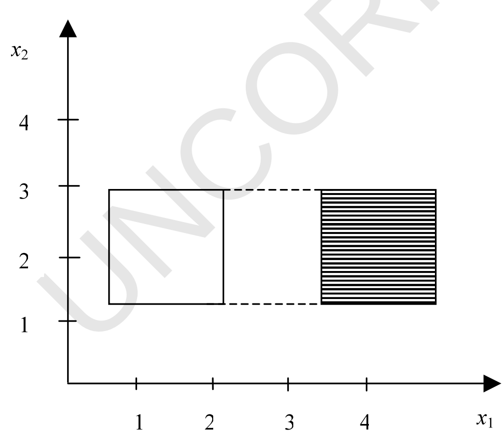

Fig. 8. Disjoint disjunction (particular case).

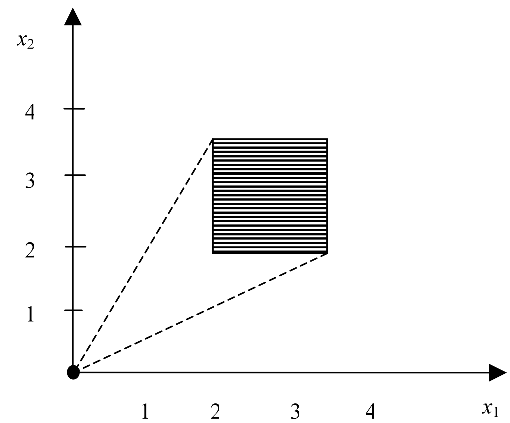

Fig. 9. Disjoint disjunction with zero point.

$$y_1 h_1\left(\frac{x}{y_1}\right) \le 0$$

$$0 \le x \le x^U y_1$$

$$0 \le y_1 \le 1 \qquad (21)$$

which includes the zero point as a feasible point. The above also applies to linear case.

5. Relaxation in $x-y$ space

The previous section analyzed the relation of relaxations for different types of disjunctions in the x space. When applying the big-M constraints Eq. (11) or the convex hull Eq. (14) these are written in the $x-y$ space. Therefore, an interesting question is whether or not the properties we noted in the previous section still apply in the $x-y$ space. Let us consider the following example, which has an *improper* disjunction.

5.1. Example 2

$$\min Z = (x_1 - 1.1)^2 + (x_2 - 1.1)^2 + c_1$$

$$\text{s.t.} \quad \begin{bmatrix} Y_1 \\ x_1^2 + x_2^2 \le 1 \\ c_1 = 1 \end{bmatrix} \vee \begin{bmatrix} \neg Y_1 \\ x_1 = x_2 = 0 \\ c_1 = 0 \end{bmatrix}$$

$$0 \le x_1,\ x_2 \le 1; \quad 0 \le c_1$$

$$Y_1 \in \{\text{true},\ \text{false}\} \qquad (22)$$

The optimal solution is $x = (0.707, 0.707)$, $Y_1 = \text{true}$ and $Z = 1.309$. The feasible region is shown in Fig. 10 and the feasible region of the second term, which is (0,0), is included in the feasible region of the first term. According to the previous section since this is an *improper* disjunction in the x space, it ought to be sufficient to use the first term only. However, when

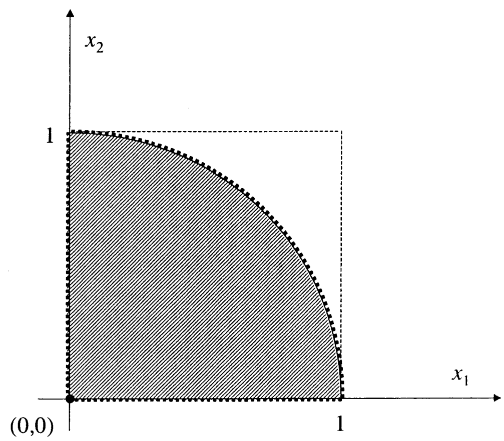

Fig. 10. Feasible region of example 2 in the x space.

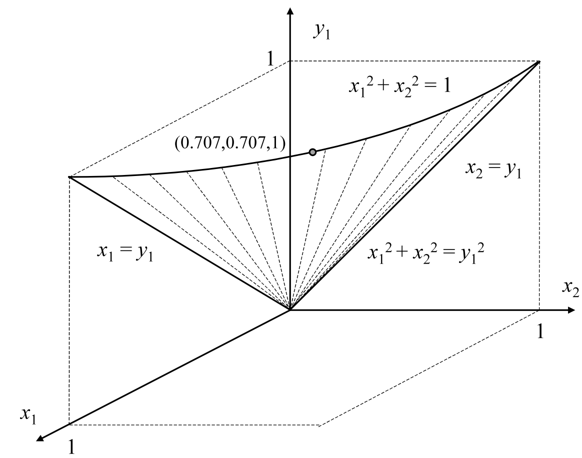

Fig. 11. Convex hull relaxation of example 2 in the $x-y$ space.

expressed algebraically, the big-M relaxation and the convex hull relaxation of the disjunction in Eq. (23) involve the additional variable y_1 as a continuous variable. In the case of the convex hull, we apply Eq. (22) to the first term. Rearranging the inequality $y_1[(x_1/y_1)^2+(x_2/y_1)^2-1]\le 0$ yields:

$$\min Z=(x_1-1.1)^2+(x_2-1.1)^2+y_1$$

$$\text{s.t.}\quad x_1^2+x_2^2\le y_1^2$$

$$0\le x_1\le y_1$$

$$0\le x_2\le y_1$$

$$0\le y_1\le 1 \qquad (23)$$

The big-M relaxation of Eq. (23) for the first term is given by:

$$\min Z=(x_1-1.1)^2+(x_2-1.1)^2+y_1$$

$$\text{s.t.}\quad x_1^2+x_2^2\le y_1$$

$$0\le x_1,\ x_2\le 1;\quad 0\le y_1\le 1 \qquad (24)$$

Figs. 11 and 12 show the convex hull relaxation and the big-M relaxation of Eq. (23) in the $x-y$ space, respectively. It is clear that Eqs. (24) and (25) are not identical due to the difference in the right hand side of the nonlinear inequality. In fact, the solution of Eq. (24) is $(x, y)=(0.707, 0.707, 1)$ and $Z=1.309$. Since the relaxed value of y_1 is 1, this solution is the optimal solution of Eq. (33), which is also shown in Fig. 11. On the other hand, the solution of Eq. (25) is $(x, y)=(0.55, 0.55, 0.605)$ and $Z=1.21$ which is weaker than the convex hull relaxation. This result can be seen by comparing Figs. 11 and 12. There is no difference between the feasible set of Eq. (24) and the feasible set of Eq. (25) projected in the x space as shown in Fig. 10. The difference, however, takes place in the $x-y$ space. Note that the nonlinear constraint in Eq. (25), $x_1^2+x_2^2\le y_1$, which is shown in Fig. 12, is weaker than $x_1^2+x_2^2\le y_1^2$ in Eq. (24) for $0\le y_1\le 1$. Therefore, even though the disjunction in Eq. (23) is *improper* in x space, the convex hull yields tighter relaxation than big-M relaxation in the $x-y$ space. Thus, this example demonstrates that for the case of *improper* nonlinear disjunctions, the convex hull may be tighter than the big-M constraint in the $x-y$ space even if they are identical in the projected x space.

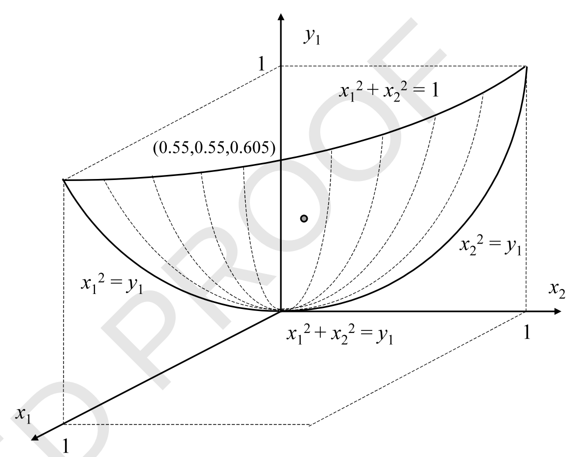

Fig. 12. Big-M relaxation of example 2 in the $x-y$ space.

For the linear case, we change the nonlinear constraint in the first term of the disjunction Eq. (23) by the following linear constraint:

$$\min Z=(x_1-1.1)^2+(x_2-1.1)^2+c_1$$

$$\text{s.t.}\quad \begin{bmatrix} Y_1 \\ x_1+x_2\le 1 \\ c_1=1 \end{bmatrix} \vee \begin{bmatrix} \neg\, Y_1 \\ x_1=x_2=0 \\ c_1=0 \end{bmatrix}$$

$$0\le x_1,\ x_2\le 1;\quad 0\le c_1$$

$$Y_1\in\{\text{true},\ \text{false}\} \qquad (25)$$

where the disjunction is *improper* in the x space. The optimal solution is $x=(0.5,0.5)$, $Y_1=\text{true}$ and $Z=1.72$.

The convex hull of the disjunction Eq. (26) yields a linear constraint:

$$x_1 + x_2 \le y_1 \tag{26}$$

After replacing the disjunction Eq. (26) with convex hull relaxation Eq. (27), the solution is $x = (0.5, 0.5)$, $y_1 = 1$ and $Z = 1.72$, which is exactly the optimal solution of Eq. (26). Since the disjunction Eq. (26) is *improper* in x space, only the first term is sufficient for the relaxation. The big-M relaxation of Eq. (26) is given by:

$$x_1 + x_2 - 1 \le M_1(1 - y_1) \tag{27}$$

This relaxation clearly depends on M_1 value. For example, if $M_1 = 1$ is used, then the relaxation yields $x = (0.67, 0.67)$, $y_1 = 0.67$ and $Z = 1.042$, which is weaker than the convex hull relaxation. The best M_1 value in this case is -1, which yields exactly the same solution as the convex hull relaxation. As shown with this example, even for the linear *improper* disjunction the big-M relaxation may have weaker relaxation than the convex hull depending on the big-M parameter value.

6. Cutting plane method

The two previous sections have analyzed the issue of determining in what cases it is worth to formulate disjunctions with the convex hull relaxation in order to obtain tighter relaxations when compared with the big-M relaxation. In this section, we present a numerical procedure for generating cutting planes, which potentially has the advantage of requiring much fewer variables and constraints than the convex hull relaxation. Cutting planes, which correspond to facets of the convex hull, can improve the tightness of the big-M relaxation. The proposed cutting planes can be used within a branch and cut enumeration procedure (Stubbs & Mehrotra, 1999), or as a way to strengthen an algebraic MINLP model before solving it with one of the standard methods.

Using as a basis the GDP model, the general form of the strengthened MINLP model (PC_n) at any iteration n will be as follows:

$$\min Z = \sum_{k \in K} \sum_{i \in D_k} \gamma_{ik} y_{ik} + f(x)$$

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

$$h_{ik}(x) \le M_{ik}(1 - y_{ik}), \quad i \in D_k, \quad k \in K \qquad (\mathrm{PC}_n)$$

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

$$Ay \le a$$

$$\beta_n^T x \le b_n, \quad n = 1, 2, \ldots, N$$

$$x \in R^n, \quad y_{ik} \in \{0, 1\}$$

where $\beta_n^T x \le b_n$ is the cutting plane at the iteration n. Let us denote the solution of the continuous relaxation of (PC_n) as $x_{\mathrm{R}}^{\mathrm{BM},n}$. In order to generate the cutting plane we consider the following separation problem, which has as an objective to find the point within the convex hull that is closest to the point $x_{\mathrm{R}}^{\mathrm{BM},n}$. This separation problem is given by the NLP:

$$\min \phi(x) = (x - x_{\mathrm{R}}^{\mathrm{BM},n})^T (x - x_{\mathrm{R}}^{\mathrm{BM},n})$$

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

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

$$y_{ik} h_{ik}\left(\frac{v_{ik}}{y_{ik}}\right) \le 0, \quad i \in D_k, k \in K \qquad (\mathrm{SP}_n)$$

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

$$Ay \le a$$

$$\beta_n^T x \le b_n, \quad n = 1, 2, \ldots, N$$

$$x, v_{ik} \in R^n, \quad 0 \le y_{ik} \le 1$$

Let the solution of the separation problem (SP_n) be $x^{\mathrm{S},n}$. A cutting plane $\beta_n^T x \le b_n$ can then be obtained from:

$$(x^{\mathrm{S},n} - x_{\mathrm{R}}^{\mathrm{BM},n})^T (x - x^{\mathrm{S},n}) \ge 0 \tag{28}$$

where the coefficient of x is a subgradient of the objective function of (SP_n) at $x^{\mathrm{S},n}$ (for derivation, see Stubbs & Mehrotra, 1999). Fig. 13 shows an example of a cutting plane generated with the points $x^{\mathrm{S},n}$ and $x_{\mathrm{R}}^{\mathrm{BM},n}$.

The cutting plane method can then be stated as follows:

1) Solve continuous relaxation of (PC_n).
2) Solve separation problem (SP_n).
   a) If $||x^{\mathrm{S},n} - x_{\mathrm{R}}^{\mathrm{BM},n}|| \le \varepsilon$, stop.

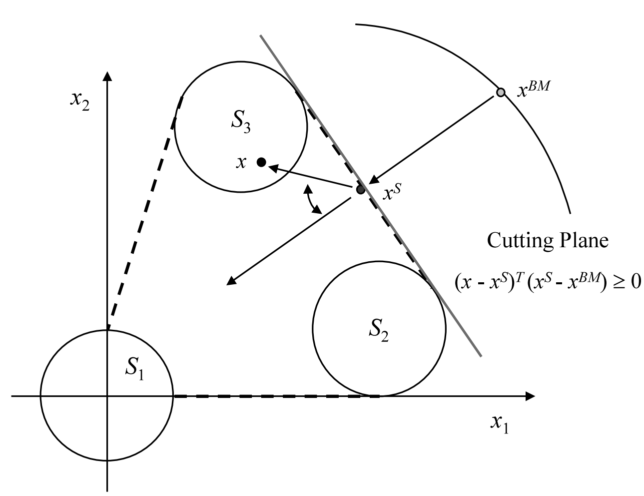

Fig. 13. Cutting plane generated by separation problem.

b) Else set $\beta_{n+1} = -(x^{S,n} - x_R^{BM,n})$ and $b_{n+1} = -(x^{S,n} - x_R^{BM,n})x^{S,n}$. Set $n = n+1$, return to Step 1.

This procedure can be used either in Branch and Cut enumeration method where a special case is to solve the separation problem only at the root node, or else it can be used to strengthen the MINLP model before applying methods such as Outer-Approximation (OA), Generalized Benders Decomposition (GBD), and Extended Cutting Plane (ECP). It is also interesting to note that cutting planes can be derived in the $x-y$ space. In example 2, when we consider the cutting plane in the x space, the big-M relaxation solution, $x = (0.55, 0.55)$ cannot be separated from the convex hull since it is feasible to the convex hull onto the x space. But when we consider the cutting plane in the $x-y$ space, then the big-M relaxation solution, $(x, y) = (0.55, 0.55, 0.605)$ can be separated from the convex hull since this point is infeasible to the convex hull relaxation Eq. (24). This suggests that the application of cutting planes in the $x-y$ space may be more effective than in the x space only for cutting off the big-M relaxation point from the convex hull.

Another application of the separation problem is for deciding whether it is advantageous or not to use the convex hull formulation. If the value of $||x^{S,n} - x_R^{BM,n}||$ is large, then it is an indication that this is the case. A small difference between $x^{S,n}$ and $x_R^{BM,n}$ would indicate that it might be better to use the big-M relaxation.

It should be also noted that the proposed cutting plane method can be extended to nonconvex disjunctive constraints using the global optimization procedure by Lee and Grossmann (2001). In this method the nonconvex constraints are replaced by convex under/overestimators, with which the convex hull relaxation or big-M relaxation can be used. Therefore, one can use the cutting plane method to tighten the relaxation of the bounding convex constraints.

7. Disjunctive programming examples

In this section we present a number of examples to illustrate the application of the main concepts in this paper.

7.1. Example 3

$$\min Z = (x_1 - 6)^2 + (x_2 - 4)^2$$

s.t.

$$\begin{bmatrix} Y_1 \\ (x_1-4)^2 + (x_2-2)^2 \le 0.5 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ (x_1-3)^2 + (x_2-4)^2 \le 1 \end{bmatrix} \vee \begin{bmatrix} Y_3 \\ (x_1-1)^2 + (x_2-1)^2 \le 1.5 \end{bmatrix}$$

$$0 \le x_1,\ x_2 \le 5 \qquad (29)$$

The feasible region is shown in Fig. 14. Note that the point (6,4), which is the minimizer of the objective function, lies outside the convex hull of the disjunction. The optimal solution is $x = (4,4)$, $Z = 4.0$, $Y = (\text{false, true, false})$.

To illustrate the cutting plane procedure, first we solve the big-M relaxation of Eq. (30) with $M = (19.5, 24, 30.5)$ from Eq. (12). The solution is $x^{BM} = (5, 4)$, $Z^{BM} = 1.0$, $y^{BM} = (0.209, 0.561, 0.230)$. Then we solve the separation problem (SP_n) with the relaxation point $x^{BM} = (5, 4)$:

$$\min Z = (x_1 - 5)^2 + (x_2 - 4)^2$$

$$\text{s.t.} \quad x_1 = v_{11} + v_{12} + v_{13}$$

$$x_2 = v_{21} + v_{22} + v_{23}$$

$$(y_1+\varepsilon)\left[\left(\frac{v_{11}}{y_1+\varepsilon} - 4\right)^2 + \left(\frac{v_{21}}{y_1+\varepsilon} - 2\right)^2 - 0.5\right] \le 0$$

$$(y_2+\varepsilon)\left[\left(\frac{v_{12}}{y_2+\varepsilon} - 3\right)^2 + \left(\frac{v_{22}}{y_2+\varepsilon} - 4\right)^2 - 1\right] \le 0$$

$$(y_3+\varepsilon)\left[\left(\frac{v_{13}}{y_3+\varepsilon} - 1\right)^2 + \left(\frac{v_{23}}{y_3+\varepsilon} - 1\right)^2 - 1.5\right] \le 0$$

$$y_1 + y_2 + y_3 = 1$$

$$0 \le y_i \le 1, \quad i = 1,\ 2,\ 3$$

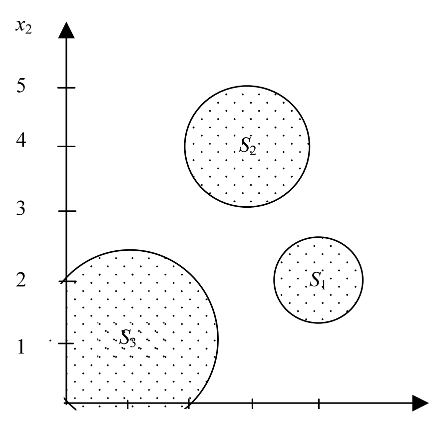

Fig. 14. Feasible region of example 3.

$0 \le v_{ji} \le 5y_i \quad \forall i, \forall j$

$$0 \le x_1,\ x_2 \le 5 \tag{30}$$

The solution of problem Eq. (31) is $x^S = (4.16,\ 3.70)$ with the objective value of 0.791. Therefore, the cutting plane is given as follows:

$$\begin{bmatrix} 4.16-5.0 \\ 3.70-4.0 \end{bmatrix}^T \begin{bmatrix} x_1-4.16 \\ x_2-3.70 \end{bmatrix} \ge 0 \tag{31}$$

which can be simplified as $-0.84(x_1-4.16)-0.3(x_2-3.70) \ge 0$. We add Eq. (32) to the big-M relaxation and solve it again. The solution of this augmented big-M relaxation is $x^{CP} = (4.27,\ 3.4)$, $Z^{CP} = 3.37$, $y^{CP} = (0.294,\ 0.676,\ 0.029)$. For comparison, we solve the convex hull relaxation, obtaining $x^{CH} = (4.27,\ 3.4)$, $Z^{CH} = 3.37$, $y^{CH} = (0.442,\ 0.558,\ 0)$. Note that the solution x^{CP} and the objective value Z^{CP} are identical to x^{CH} and Z^{CH}. The difference in (x^{BM}, Z^{BM}) and (x^{CH}, Z^{CH}) is a clear indication that the convex hull is significantly tighter than big-M relaxation. For this example, only one cutting plane yields the same tightness of the relaxation as the convex hull. The numerical results are shown in Table 1. Note that the big-M relaxation yields the lowest objective value to the optimal solution, 4.0. Fig. 15 shows the convex hull and cutting plane. As shown in Fig. 15, the cutting plane is a facet of the convex hull. From Table 1 it can be seen that the big-M relaxation with a cutting plane yields a competitive relaxation compared with the convex hull.

7.2. Cutting planes in $x-y$ space: example 2

Let us revisit example 2. If we apply the separation problem (SP_n) to the big-M relaxation solution $x_R^{BM} = (0.55, 0.55)$, the objective value of the separation problem is zero since x_R^{BM} is feasible to the convex hull relaxation of Eq. (23) in the x space. However, if we treat the binary variable y as continuous variable and then extend the dimension of the solution to the $x-y$ space, we have the following separation problem with $(x, y)_R^{BM} = (0.55,\ 0.55,\ 0.605)$:

$$\min Z = [(x_1-0.55)^2 + (x_2-0.55)^2 + (y_1-0.605)^2]$$

s.t.

$$x_1^2 + x_2^2 \le y_1^2 \tag{SP1}$$

$$0 \le x_1 \le y_1$$

$$0 \le x_2 \le y_1$$

$$0 \le y_1 \le 1$$

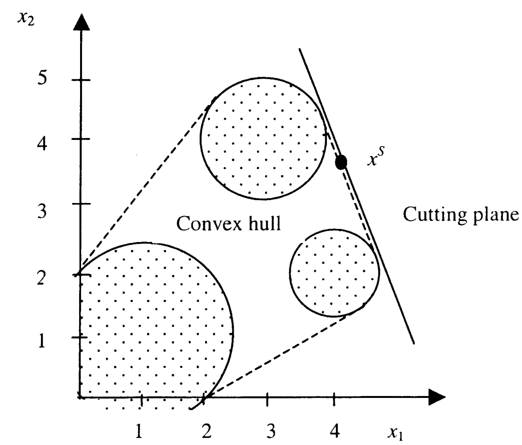

Fig. 15. Convex hull and cutting plane for example 3.

The solution is $Z = 0.015$ and $(x, y)^S = (0.489,\ 0.489,\ 0.691)$, which means that $(x, y)_R^{BM}$ is infeasible in the convex hull relaxation Eq. (24) in the $x-y$ space. The cutting plane is now given by $(0.489-0.55)(x_1-0.489)+(0.489-0.55)(x_2-0.489)+(0.691-0.605)(y_1-0.691) \ge 0$.When this cutting plane is added to the big-M relaxation Eq. (25), the optimal solution is $(x, y) = (0.707,\ 0.707,\ 1)$ and $Z = 1.309$, which is identical to the solution of the convex hull relaxation Eq. (24) and is also the optimal solution of Eq. (23). This shows that the cutting plane method applied to the $x-y$ space can yield tighter relaxations than the cutting plane in the x space only.

7.3. Example 4

Consider the synthesis of a process network (Türkay & Grossmann, 1996) where the following disjunctive set is used to model the problem:

$$\begin{bmatrix} Y_k \\ h_{ik}(x) = 0 \\ c_k = \gamma_k \end{bmatrix} \vee \begin{bmatrix} \neg Y_k \\ B_{ik}x = 0 \\ c_k = 0 \end{bmatrix} \quad i \in D_k,\ k \in K \tag{32}$$

It means that if the kth unit is selected ($Y_k = \text{true}$) then the first term of the disjunction applies, if it is not ($\neg Y_k$) then a subset of the x variables is set to zero.

Table 1
Comparisons of the relaxations for example 3

Relaxation	M	x_1	x_2	y_1	y_2	y_3	Z
Big-M	(19.5, 24, 30.5)	5.0	4.0	0.209	0.561	0.023	1.0
Convex hull	–	4.27	3.40	0.442	0.558	0.0	3.37
Cutting plane	–	4.27	3.40	0.294	0.676	0.029	3.37
Optimal solution	–	4.0	4.0	0	1	0	4.0

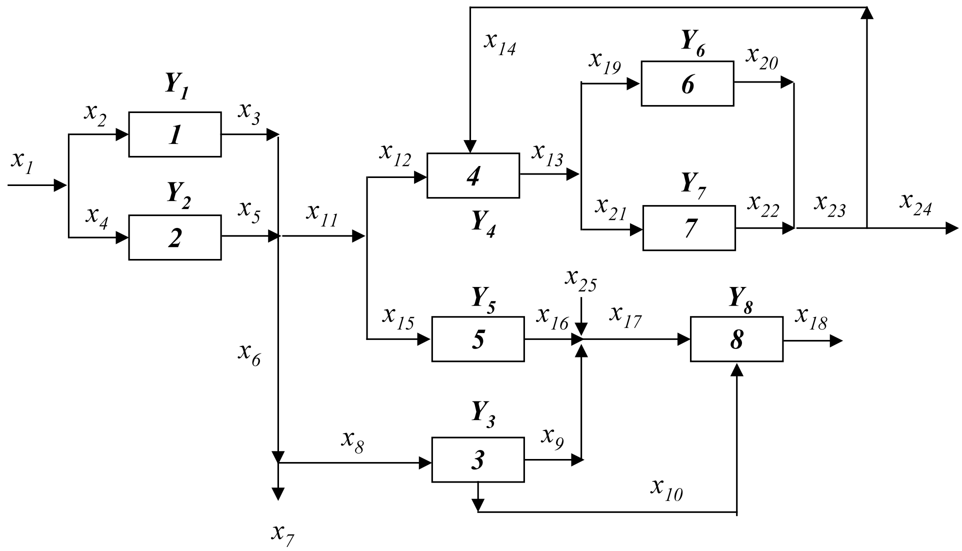

Fig. 16. Process superstructure of example 4.

Fig. 16 shows the superstructure of example 4, which has eight units. The corresponding GDP model is as follows:

$$\min Z = \sum_{k=1}^{8} c_k + a^T x + 122$$

s.t. Mass balances:

$$x_1 = x_2 + x_4, \quad x_6 = x_7 + x_8$$

$$x_3 + x_5 = x_6 + x_{11}$$

$$x_{11} = x_{12} + x_{15}, \quad x_{13} = x_{19} + x_{21}$$

$$x_9 + x_{16} + x_{25} = x_{17}$$

$$x_{20} + x_{22} = x_{23}, \quad x_{23} = x_{14} + x_{24}$$

Specifications:

$$x_{10} - 0.8x_{17} \le 0, \quad x_{10} - 0.4x_{17} \ge 0$$

$$x_{12} - 5x_{14} \le 0, \quad x_{12} - 2x_{14} \ge 0$$

Disjunctons:

$$\begin{bmatrix} Y_1 \\ \exp(x_3) - 1 - x_2 \le 0 \\ c_1 = 5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_1 \\ x_3 = x_2 = 0 \\ c_1 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_2 \\ \exp\left(\frac{x_5}{1.2}\right) - 1 - x_4 \le 0 \\ c_2 = 5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_2 \\ x_4 = x_5 = 0 \\ c_2 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_3 \\ 1.5x_9 + x_{10} - x_8 = 0 \\ c_3 = 6 \end{bmatrix} \vee \begin{bmatrix} \neg Y_3 \\ x_9 = 0, \quad x_8 = x_{10} \\ c_3 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_4 \\ 1.25(x_{12} + x_{14}) - x_{13} = 0 \\ c_4 = 10 \end{bmatrix} \vee \begin{bmatrix} \neg Y_4 \\ x_{12} = x_{13} = x_{14} = 0 \\ c_4 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_5 \\ x_{15} - 2x_{16} = 0 \\ c_5 = 6 \end{bmatrix} \vee \begin{bmatrix} \neg Y_5 \\ x_{15} = x_{16} = 0 \\ c_5 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_6 \\ \exp\left(\frac{x_{20}}{1.5}\right) - 1 - x_{19} \le 0 \\ c_6 = 7 \end{bmatrix} \vee \begin{bmatrix} \neg Y_6 \\ x_{19} = x_{20} = 0 \\ c_6 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_7 \\ \exp(x_{22}) - 1 - x_{21} \le 0 \\ c_7 = 4 \end{bmatrix} \vee \begin{bmatrix} \neg Y_7 \\ x_{21} = x_{22} = 0 \\ c_7 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_8 \\ \exp(x_{18}) - 1 - x_{10} - x_{17} \le 0 \\ c_8 = 5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_8 \\ x_{10} = x_{17} = x_{18} = 0 \\ c_8 = 0 \end{bmatrix} \tag{33}$$

Logic propositions:

$Y_1 \Rightarrow Y_3 \vee Y_4 \vee Y_5 \qquad Y_5 \Rightarrow Y_8$
$Y_2 \Rightarrow Y_3 \vee Y_4 \vee Y_5 \qquad Y_6 \Rightarrow Y_4$
$Y_3 \Rightarrow Y_1 \vee Y_2 \qquad Y_7 \Rightarrow Y_4$
$Y_3 \Rightarrow Y_8 \qquad Y_8 \Rightarrow Y_3 \vee Y_5 \vee (\neg Y_3 \wedge \neg Y_5)$
$Y_4 \Rightarrow Y_1 \vee Y_2 \qquad Y_1 \veebar Y_2$
$Y_4 \Rightarrow Y_6 \vee Y_7 \qquad Y_4 \veebar Y_5$
$Y_5 \Rightarrow Y_1 \vee Y_2 \qquad Y_6 \veebar Y_7$

Problem data:

$a^T = (a_1 = 0,\ a_2 = 10,\ a_3 = 1,\ a_4 = 1,\ a_5 = -15,\ a_6 = 0,\ a_7 = 0,\ a_8 = 0,\ a_9 = -40,\ a_{10} = 15,\ a_{11} = 0,\ a_{12} = 0,\ a_{13} = 0,\ a_{14} = 15,\ a_{15} = 0,\ a_{16} = 0,\ a_{17} = 80,\ a_{18} = -65,\ a_{19} =$

25, $a_{20} = -60$, $a_{21} = 35$, $a_{22} = -80$, $a_{23} = 0$, $a_{24} = 0$, $a_{25} = -35$); $x_j^{\text{lo}} = 0,\ \forall j$.

Before introducing the big-M relaxation, it should be noted that in the disjunctions we have the following properties:

i) The disjunctions are *improper* since the feasible region of the second term belongs to the feasible region of the first term in x space (except the cost term).
ii) In the second term of the disjunctions a subset of the continuous variables x are zero.
iii) No continuous variable x is repeated in the second term ($\neg Y_k$) of the disjunctions.

Because of these properties, it is possible to rewrite the disjunctions as follows:

$$\exp(x_3) - 1 - x_2 \le 0$$

$$\exp\left(\frac{x_5}{1.2}\right) - 1 - x_4 \le 0$$

$$1.5x_9 + x_{10} - x_8 = 0$$

$$1.25(x_{12} + x_{14}) - x_{13} = 0$$

$$x_{15} - 2x_{16} = 0$$

$$\exp\left(\frac{x_{20}}{1.5}\right) - 1 - x_{19} \le 0$$

$$\exp(x_{22}) - 1 - x_{21} \le 0$$

$$\exp\left(\frac{x_{18}}{1.5}\right) - 1 - x_{10} - x_{17} \le 0$$

Disjunctions:

$$\begin{bmatrix} Y_1 \\ 0 \le x_2 \le x_2^{\text{up}} \\ 0 \le x_3 \le x_3^{\text{up}} \\ c_1 = 5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_1 \\ x_3 = x_2 = 0 \\ c_1 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_2 \\ 0 \le x_4 \le x_4^{\text{up}} \\ 0 \le x_5 \le x_5^{\text{up}} \\ c_2 = 5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_2 \\ x_4 = x_5 = 0 \\ c_2 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_3 \\ 0 \le x_9 \le x_9^{\text{up}} \\ c_3 = 6 \end{bmatrix} \vee \begin{bmatrix} \neg Y_3 \\ x_9 = 0 \\ c_3 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_4 \\ 0 \le x_{12} \le x_{12}^{\text{up}} \\ 0 \le x_{13} \le x_{13}^{\text{up}} \\ 0 \le x_{14} \le x_{14}^{\text{up}} \\ c_4 = 10 \end{bmatrix} \vee \begin{bmatrix} \neg Y_4 \\ x_{12} = x_{13} = x_{14} = 0 \\ c_4 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_5 \\ 0 \le x_{15} \le x_{15}^{\text{up}} \\ 0 \le x_{16} \le x_{16}^{\text{up}} \\ c_5 = 6 \end{bmatrix} \vee \begin{bmatrix} \neg Y_5 \\ x_{15} = x_{15} = 0 \\ c_5 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_6 \\ 0 \le x_{19} \le x_{19}^{\text{up}} \\ 0 \le x_{20} \le x_{20}^{\text{up}} \\ c_6 = 7 \end{bmatrix} \vee \begin{bmatrix} \neg Y_6 \\ x_{19} = x_{20} = 0 \\ c_6 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_7 \\ 0 \le x_{21} \le x_{21}^{\text{up}} \\ 0 \le x_{22} \le x_{22}^{\text{up}} \\ c_7 = 4 \end{bmatrix} \vee \begin{bmatrix} \neg Y_7 \\ x_{21} = x_{22} = 0 \\ c_7 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_8 \\ 0 \le x_{10} \le x_{12}^{\text{up}} \\ 0 \le x_{17} \le x_{13}^{\text{up}} \\ 0 \le x_{18} \le x_{14}^{\text{up}} \\ c_8 = 5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_8 \\ x_{10} = x_{17} = x_{18} = 0 \\ c_8 = 0 \end{bmatrix} \quad (34)$$

It should be noted that constraints Eq. (35) consist of global constraints (nonlinear) and disjunctions (linear). The convex hull of the above disjunctions can be reduced to linear constraints for the big-M relaxation which are given by:

$$0 \le x_j \le x_j^{\text{up}} y_k, \quad j \in J, \quad k \in K \quad (35)$$

$$c_k = \gamma_k y_k, \quad k \in K$$

$$0 \le y_k \le 1, \quad k \in K$$

which means that if the first term of the disjunction is true ($y_k = 1$) then the continuous variables x_j can have a value between its bounds and the fixed cost is activated, else if the second term is true ($y_k = 0$) then the continuous variables become zero that still satisfies the global constraints (condition i).

The GDP problem Eq. (34) is solved with the convex hull relaxation. The upper bounds used are $x_3^{\text{up}} = 2$, $x_5^{\text{up}} = 2$, $x_9^{\text{up}} = 2$, $x_{10}^{\text{up}} = 1$, $x_{14}^{\text{up}} = 1$, $x_{17}^{\text{up}} = 2$, $x_{19}^{\text{up}} = 2$, $x_{21}^{\text{up}} = 2$, $x_{25}^{\text{up}} = 3$, and for the rest of the variables, $x_j^{\text{up}} = 6.5$. The objective function value $Z = 64.8$ was obtained from the convex hull relaxation, and the corresponding NLP requires 0.07 CPU s with CONOPT/GAMS. Applying the big-M relaxation to the modified GDP formulation Eq. (35) and the same bounds, we obtained $Z = 49.9$ as the solution value. Therefore, the convex hull relaxation of the original GDP model yields a much tighter lower bound. The difference between these two relaxation values comes from the fact that the feasible region by the convex hull relaxation of nonlinear disjunctions Eq. (34) in the $x-y$ space is tighter than the feasible region by big-M relaxation of Eq. (35). However, it should be noted that their projections onto the x space are identical since the disjunctions are *improper*. If the disjunctions are linear, then both relaxations can be identical in the $x-y$ space if appropriate big-M parameters are used.

Since the convex hull relaxation yields a significant increase in the number of additional constraints and variables, we consider the generation of cutting planes to strengthen the big-M relaxation. As outlined in Section 6, a separation problem is solved. And the solution of the separation problem is used to build a cutting plane as in example 4. The big-M relaxation of Eq. (35) is then solved again with this cutting plane. Since the cutting plane is a facet of the convex hull, it will tighten the lower bound. Table 2 shows the increase of the lower bound as cutting planes are added to the big-M relaxation. The first column shows the number of cutting planes added. The second column shows the relaxation value. Note that the optimal solution of example 4 is 68.01. The third column shows the objective value of the separation problem. As more cutting planes are added, the objective value of the separation problem decreases, implying that the solution point of the augmented big-M relaxation gets closer to the convex hull. The fourth column shows the CPU time of separation problem. The fifth and sixth column show the MINLP solution results by DICOPT++ with the corresponding cuts. In all cases, the optimal solution is found in the second major iteration. Since this problem is a convex MINLP, the Outer-Approximation (OA) algorithm stops when the crossover occurs. The CPU time is less than 1 s on a Pentium III PC 600 MHz with 128 Mbytes RAM memory. After adding seven cutting planes, the lower bound improved significantly compared with the case when no cutting plane is used (62.5 vs. 49.9). The advantage of the cutting plane method is that only one linear constraint is added to the big-M relaxation at each step. However, there is a cost for building a cutting plane and that is to solve a separation problem, which is a convex NLP problem (SP_n).

7.4. Example 5

To illustrate the application of the cutting plane method with a branch and bound algorithm, we have constructed the following GDP problem with linear/nonlinear *proper* disjunctions.

$$\min Z = \sum_{k=1}^{9} c_k + a^T x$$

$$-0.6\log(x_{12}+1) + 0.8(x_{13}-8)^2 + 0.7\exp(-x_{14}+1) - 0.5\log(x_{15}+2)$$

s.t. Mass balances:

$$x_1 = x_5 + x_6, \quad x_4 = x_7 + x_8$$

$$x_{10} = x_{19} + x_{20}, \quad x_{11} = x_{17} + x_{18}$$

$$x_{14} = x_{21} + x_{22}, \quad x_9 = x_{23} + x_{24}$$

$$x_{12} = x_{25} + x_{26}$$

Specifications:

$$x_1 + x_2 + x_3 + x_4 \le 30$$

$$x_9 + x_{10} + x_{11} \le 25$$

$$x_{12} + x_{13} + x_{14} + x_{15} + x_{16} \le 20$$

$$\begin{bmatrix} Y_1 \\ x_9 \le 1.7\log(x_2 + x_5 + 1) \\ x_9 \ge 0.1 + 0.2x_5 \\ x_5 \ge 2x_2 \\ c_1 = 2 \end{bmatrix} \vee \begin{bmatrix} \neg Y_1 \\ x_2 = x_5 = x_9 = 0 \\ c_1 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_2 \\ x_{10} = 0.9x_3 + 0.8x_7 \\ 1 \le x_3 + x_7 \\ x_7 \ge x_3 \\ c_2 = 1 \end{bmatrix} \vee \begin{bmatrix} \neg Y_2 \\ x_3 = x_7 = x_{10} = 0 \\ c_2 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_3 \\ 1.5x_{11} = x_6 + x_8 \\ x_6 = x_8 \\ x_{11} \ge 1 \\ c_3 = 9 \end{bmatrix} \vee \begin{bmatrix} \neg Y_3 \\ x_6 = x_8 = x_{11} = 0 \\ c_3 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_4 \\ x_{25} \le \log(x_{23} + 1) + 0.1 \\ x_{25} \ge 1 \\ c_4 = 1.5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_4 \\ x_{23} = x_{25} = 0 \\ c_4 = 0 \end{bmatrix}$$

Table 2
Numerical results of cutting plane method for example 4

Number of cutting planes	Big-M relaxation	Separation problem solution	Separation CPU (s)	DICOPT++ major iterations	CPU (s)
0	49.9	0.545	0.043	2	0.139
1	51.7	0.701	0.078	2	0.129
2	52.2	0.576	0.078	2	0.121
3	53.2	0.163	0.027	2	0.139
4	61.2	0.010	0.039	2	0.248
5	61.9	0.004	0.051	2	0.151
6	62.4	0.005	0.051	2	0.143
7	62.5	0.002	0.051	2	0.157

$$\begin{bmatrix} Y_5 \\ x_{26} \le 1.5\,\log(x_{24}+1) \\ x_{26} \ge 1 \\ c_5 = 4 \end{bmatrix} \vee \begin{bmatrix} \neg\, Y_5 \\ x_{24} = x_{26} = 0 \\ c_5 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_6 \\ (x_{17}-4)^2 + (x_{21}-4)^2 \le 12 \\ x_{21} \ge 1 \\ c_6 = 3.7 \end{bmatrix} \vee \begin{bmatrix} \neg\, Y_6 \\ x_{17} = x_{21} = 0 \\ c_6 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_7 \\ x_{13} \le 7 - 1.2(x_{20}-3)^2 \\ x_{22} \le 8 - (x_{20}-3)^2 \\ x_{20} \ge 1 \\ c_7 = 7.4 \end{bmatrix} \vee \begin{bmatrix} \neg\, Y_7 \\ x_{13} = x_{20} = x_{22} = 0 \\ c_7 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_8 \\ x_{15} \le 1.2\,\log(x_{19}+2) \\ x_{15} \ge 1 + 0.2x_{19} \\ x_{19} \ge 1 \\ c_8 = 6.5 \end{bmatrix} \vee \begin{bmatrix} \neg\, Y_8 \\ x_{15} = x_{19} = 0 \\ c_8 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_9 \\ x_{16} + x_{18} \ge 5 \\ x_{16} \le 6 + 2\,\log(x_{18}+1) \\ x_{18} \ge 1 \\ c_9 = 5.2 \end{bmatrix} \vee \begin{bmatrix} \neg\, Y_9 \\ x_{16} = x_{18} = 0 \\ c_9 = 0 \end{bmatrix} \qquad (36)$$

Logic proposition:

$Y_1 \vee Y_2 \vee Y_3$

$\neg(Y_1 \wedge Y_2 \wedge Y_3)$

$\neg\, Y_4 \vee \neg\, Y_5$

$Y_1 \Rightarrow Y_4 \vee Y_5$

$Y_4 \Rightarrow Y_1$

$Y_5 \Rightarrow Y_1$

$Y_2 \Rightarrow Y_7 \vee Y_8$

$Y_3 \Rightarrow Y_6 \vee Y_9$

$Y_6 \Rightarrow Y_3$

$\neg\, Y_8 \vee \neg\, Y_9$

$Y_9 \Rightarrow Y_3$

$[Y_4 \vee Y_5] \Rightarrow [Y_7 \vee Y_8 \vee Y_9]$

$[\neg\, Y_4 \wedge \neg\, Y_5] \Rightarrow [Y_7 \wedge Y_8] \vee [Y_8 \wedge Y_9] \vee [Y_7 \wedge Y_9]$

$0 \le x_j \le 9 \quad j = 1, \ldots, 26; \quad 0 \le c_k, \quad Y_k \in \{\text{true}, \text{false}\}, \quad k = 1, \ldots, 9$

The optimal solution is $Z = -197.3$, $Y_2, Y_3, Y_6, Y_7, Y_9 = \text{true}$ and $x = (1.15, 0, 1.56, 2.72, 0, 1.15, 1.56, 1.15, 0, 2.67, 1.53, 0, 6.87, 9, 0, 7.38, 0.53, 1, 0, 2.67, 4.02, 4.98, 0, 0, 0, 0)$. The big-M relaxation of Eq. (36) yields a lower bound of −326.4. The convex hull relaxation of problem Eq. (36) yields a lower bound of −209. Table 3 shows the results of cutting plane method applied to big-M relaxation of Eq. (36). As more cutting planes are added, the lower bound of big-M relaxation increases and the objective value of the separation problem decreases. After adding ten cutting planes, the lower bound significantly improved (−219.7). Table 4 shows the branch and bound search results when cutting planes are added before starting the branch and bound search. First, the big-M MINLP problem is solved with branch and bound search. Nineteen nodes are searched and the optimal solution −197.3 is found. Secondly, four cutting planes are added to big-M MINLP problem at the root node of branch and bound tree. Note that the relaxation value, which is the objective value at the root node, is −239.5 and 13 nodes are searched to find the optimal solution. The decrease in the number of search nodes is due to the tighter relaxation value. When eight cutting planes are added, the relaxation value is −221.4 and only seven nodes are searched. For comparison, the convex hull relaxation of Eq. (36) is solved and the number of nodes is seven, which is same as in the case of eight cutting planes. The CPU time for each case is also shown in Table 4 and less CPU time is spent with fewer number of nodes. The CPU time for generating eight cutting planes is about 2 s. This example clearly shows that the cutting planes can tighten the relaxation and thus reduce the number of search nodes in branch and bound method. Although the example presented is rather small, the proposed cutting plane method should be promising for solving larger problems. This will be the subject of our future work.

Table 3
Numerical results of cutting plane method for example 5

Number of cutting planes	Big-M relaxation solution	Separation problem solution
0	−326.4	91.2
1	−265.6	5.97
2	−255.8	9.76
3	−245.5	7.12
4	−239.5	4.34
5	−238.0	4.35
6	−224.4	2.39
7	−223.4	1.31
8	−221.4	0.91
9	−220.8	0.25
10	−219.7	0.19
Convex hull relaxation	−209.0	0

8. Conclusions

The purpose of this paper has been to analyze the different alternatives of modeling the discrete choices as disjunctions or as mixed-integer (0–1) inequalities, in order to provide guidelines on this decision. The resulting model can correspond to one of the three formulations: mixed-integer constraints (PA), disjunctive constraints (GDP) or hybrid (PH). For the analysis,

Table 4
Comparisons of branch and bound search results for example 5

Model	Big-M MINLP no cutting planes	Big-M MINLP +4 cutting planes	Big-M MINLP +8 cutting planes	Convex hull relaxation
Relaxation value	−326.4	−239.5	−221.4	−209.0
Optimal solution	−197.3	−197.3	−197.3	−197.3
Number of nodes	19	13	7	7
CPU s	3.39	2.53[a]	1.56[a]	1.62

[a] CPU time for generating cutting planes is not included.

we considered three different possible relaxations of a disjunctive set, the convex hull, the big-M relaxation and the Beaumont surrogate. The analysis was performed mainly on the first two since the big-M formulation is widely used.

Although it was proved that the convex hull relaxation yields a tighter relaxation than the traditional 0–1 big-M relaxation, there are several cases when the big-M relaxation can compete with the convex hull relaxation. As a general rule, the big-M model is competitive when good bounds can be provided for the variables, and for large problems where it is important to keep the number of equations and variables as small as possible. For convex *improper* disjunction both the convex hull and the big-M model give the same relaxation in the x space, but this may not be true in the x–y space as was demonstrated with examples. For *proper* disjunctions where the feasible regions have some intersection, the objective function plays an important role, if the minimizer of the objective function is inside the feasible region of the disjunctive set, both the big-M and the convex hull relaxation may yield the same relaxation value. Otherwise the convex hull should be generally better, but the big-M constraints with appropriate bounds can be competitive. For *proper* disjunctions with an empty intersection on the feasible regions (disjoint terms) the convex hull is generally better than the big-M relaxation. Although these conclusions are not general, we believe they help to provide some insight in the modeling of discrete/continuous optimization problems.

Finally, to address the problem of formulating tight models without generating the explicit equations of the convex hull, a cutting plane algorithm has been proposed. A number of examples have been presented to illustrate the various ideas in this paper as well as the cutting plane method.

Acknowledgements

The authors would like to acknowledge financial from the NSF Grant CTS-9710303.

Appendix A: Property of relaxations

Property 1. Let R_{BM} be the feasible set of big-M relaxation of a given disjunctive set projected onto the x space. Let R_{CH} be the feasible set of convex hull relaxation projected onto the x space. Let R_B be the feasible set of the Beaumont surrogate that is defined in the x space. Then $R_{CH} \subseteq R_{BM} \subseteq R_B$.

Proof. First consider $R_{BM} \subseteq R_B$. For the linear case, Beaumont (1990) proved that $R_{BM} = R_B$. Therefore, $R_{BM} \subseteq R_B$ holds. For the nonlinear case, we consider one disjunction for simplicity. Given a nonlinear disjunctive set:

$$F = \vee_{i \in D} [h_i(x) \le 0] \quad x \in R^n \tag{A1}$$

where $h_i(x)$ are assumed to be convex bounded functions. The big-M relaxation of Eq. (A1) is as follows:

$$h_i(x) \le M_i(1 - y_i), \quad i \in D \tag{A2}$$

$$\sum_{i \in D} y_i = 1 \tag{A3}$$

$$0 \le y_i \le 1, \quad i \in D \tag{A4}$$

where $M_i = \max\{h_i(x) | x^L \le x \le x^U\}$. Let $R^F_{BM}(x, y)$ be the feasible set defined by Eqs. (A2), (A3) and (A4). The Beaumont surrogate of Eq. (A1) is given by:

$$\sum_{i \in D} \frac{h_i(x)}{M_i} \le N - 1 \tag{A5}$$

where $N = |D|$ and M_i are assumed to be same as in Eq. (A2). Let $R^F_B(x, y)$ be the feasible set defined by Eqs. (A5) and (A4). Since Eq. (A5) is given by a linear combination of Eqs. (A2) and (A3), any feasible point $(x^*, y^*) \in R^F_{BM}(x, y)$ also satisfies Eqs. (A5) and (A4). Hence, $(x^*, y^*) \in R^F_B(x, y)$. Therefore, $R^F_{BM}(x, y) \subseteq R^F_B(x, y)$. Since R_{BM} and R_B are the projection of $R^F_{BM}(x, y)$ and $R^F_B(x, y)$ onto the x space, it follows that:

$$R_{BM} \subseteq R_B \tag{A6}$$

Secondly, we consider $R_{CH} \subseteq R_{BM}$ for linear and nonlinear case. The convex hull relaxation of Eq. (A1) is given by:

$$x - \sum_{i \in D} v_i = 0 \quad x, v_i \in R^n \tag{A7}$$

$$y_i h_i\left(\frac{v_i}{y_i}\right) \le 0, \quad i \in D \tag{A8}$$

$$\sum_{i\in D} y_i = 1 \tag{A9}$$

$$0 \le y_i \le 1, \quad i \in D \tag{A10}$$

$$0 \le v_i \le v_i^U y_i, \quad i \in D \tag{A11}$$

Let $R^{\mathrm{F}}_{\mathrm{CH}}(x, y, v)$ be the feasible set defined by Eqs. (A7), (A8), (A9), (A10) and (A11). Consider any feasible point $(x^*, y^*, v^*) \in R^{\mathrm{F}}_{\mathrm{CH}}(x, y, v)$. From Eq. (A7), there exist μ_i such that:

$$y_i \mu_i = v_i, \quad i \in D \tag{A12}$$

$$h_i(\mu_i) \le 0, \quad i \in D \tag{A13}$$

Since $h_i(x)$ are convex functions, for any $l \in D$:

$$h_l(x) = h_l\left(\sum_{i\in D} y_i \mu_i\right) \le \sum_{i\in D} y_i h_i(\mu_i) \tag{A14}$$

For $h_l(\mu_l) \le 0$ and $h_l(\mu_i)_{i\ne l} \le M_l$, it follows from Eqs. (A14), (A10) and (A11):

$$h_l(x) \le \sum_{i\in D, i\ne l} y_i M_l = M_l(1 - y_l) \tag{A15}$$

Eq. (A15) is identical to Eq. (A2) in the big-M relaxation for $l \in D$. Hence, any feasible point $(x^*, y^*, v^*) \in R^{\mathrm{F}}_{\mathrm{CH}}(x, y, v)$ has a corresponding feasible point (x^*, y^*) which satisfies Eqs. (A2), (A3) and (A4). Therefore, $(x^*, y^*) \in R^{\mathrm{F}}_{\mathrm{BM}}(x, y)$. Since R_{BM} and R_{CH} are the projection of $R^{\mathrm{F}}_{\mathrm{BM}}(x, y)$ and $R^{\mathrm{F}}_{\mathrm{CH}}(x, y, v)$ onto the x space, it follows that:

$$R_{\mathrm{CH}} \subseteq R_{\mathrm{BM}} \tag{A16}$$

From Eqs. (A6) and (A16), $R_{\mathrm{CH}} \subseteq R_{\mathrm{BM}} \subseteq R_{\mathrm{B}}$. This completes the proof.

References

Balas, E. (1979). *Disjunctive programming. Discrete optimizations II, annals of discrete mathematics*, vol. 5. Amsterdam: North Holland.

Balas, E. (1985). Disjunctive Programming and a hierarchy of relaxations for discrete optimization problems. *SIAM Journal of Algebric and Discrete Mathematics 6* (3), 466–485.

Balas, E. (1988). On the convex hull of the union of certain polyhedra. *Operations Research Letters 7*, 279–284.

Beaumont, N. (1990). An algorithm for disjunctive programs. *European Journal of Operational Research 48*, 362–371.

Bockmayr, A. & Kasper, T. (1998). Branch-and-infer: a unifying framework for integer and finite domain constraint programming. *INFORMS Journal on Computing 10* (3), 287–300.

Grossmann, I. E. & Kravanja, Z. (1997). Mixed-integer nonlinear programming: a survey of algorithms and applications. In L. T. Biegler, et al. (Ed.), *Large-scale optimization with applications, part II: optimal design and control*, pp. 73–100). Springer.

Jackson, J. R. & Grossmann, I. E. (2002). High-level optimization model for the retrofit planning of process networks. *Industrial Engineering and Chemical Research 41* (16), 3762–3770.

Lee, S. & Grossmann, I. E. (2000). New algorithms for nonlinear generalized disjunctive programming. *Computers and Chemical Engineering 24* (9–10), 2125–2141.

Lee, S. & Grossmann, I. E. (2001). A global optimization algorithm for nonconvex generalized disjunctive programming and applications to process systems. *Computers and Chemical Engineering 25* (11–12), 1675–1697.

Nemhauser, G. L. & Wolsey, L. A. (1988). *Integer and combinatorial optimization*. Wiley.

Raman, R. & Grossmann, I. E. (1994). Modeling and computational techniques for logic based integer programming. *Computers and Chemical Engineering 18* (7), 563–578.

Stubbs, R. & Mehrotra, S. (1999). A branch-and-cut method for 0–1 mixed convex programming. *Mathematical Programming 86* (3), 515–532.

Türkay, M. & Grossmann, I. E. (1996). Logic-based algorithms for the optimal synthesis of process networks. *Computers and Chemical Engineering 20* (8), 959–978.

Vecchietti, A. & Grossmann, I. E. (1999). LOGMIP: a disjunctive 0–1 nonlinear optimizer for process system models. *Computers and Chemical Engineering 23*, 555–565.

Vecchietti, A. & Grossmann, I. E. (2000). Modeling issues and implementation of language for disjunctive programming. *Computers and Chemical Engineering 24* (9–10), 2143–2155.

Yeomans, H. & Grossmann, I. E. (1999). Nonlinear disjunctive programming models for the synthesis of heat integrated distillation sequences. *Computers and Chemical Engineering 23* (9), 1135–1151.