Using Convex Nonlinear Relaxations in the Global Optimization of Nonconvex Generalized Disjunctive Programs

Juan P. Ruiz, Ignacio E. Grossmann

Carnegie Mellon University - Department of Chemical Engineering

Pittsburgh, PA 15213

Abstract

In this paper we present a framework to generate tight convex relaxations for nonconvex generalized disjunctive programs. The proposed methodology builds on our recent work on bilinear and concave generalized disjunctive programs for which strong linear relaxations can be generated and extends its application by allowing nonlinear relaxations. This is particular important for those cases in which the convex envelopes of the nonconvex functions arising in the formulations are nonlinear (e.g. fractional terms). This extension is now possible by using the latest developments in convex disjunctive programming. We test the performance of the method in three typical process systems engineering problems, namely, the optimization of process networks, reactor networks and heat exchanger networks.

Keywords:
Generalized Disjunctive Programs, Nonlinear Relaxations, Basic Steps

1. Introduction

Many problems in engineering can be reduced to finding a set of conditions that will lead to the best design or operation of a system. Often these problems are represented by a set of algebraic expressions using continuous and discrete variables, leading to a Mixed-integer Nonlinear Program (MINLP). In order to represent accurately the behavior of physical, chemical, biological, financial or social systems, nonlinear expressions are often used. In general, this leads to an MINLP where the solution space is nonconvex and hence difficult to solve since this may give rise to local solutions. In the last decade, many global optimization algorithms for nonconvex problems have been proposed [14][10]. However, most of them can be regarded as some particular implementation of the spatial branch and bound framework [2]. The efficiency of these methods heavily relies on having tight relaxations and that is why many of the contributions in this area have been related to this subject. However, in general, finding the global optimum of large-scale nonconvex MINLP models in reasonable computational time, remains a largely unsolved problem.

In 1994, Raman and Grossmann [6] presented an alternative way to represent discrete/continuous optimization problems by using not only algebraic expressions but also disjunctions and logic propositions giving birth to Generalized Disjunctive Programming (GDP), which can be regarded as an extension of Disjunctive Programming [1]. This higher abstraction level representation not only facilitates the modeling, but also keeps the underlying logical structure of the problem, which can be exploited to find stronger relaxations and hence, develop more efficient solution methods. With this in mind, we presented a method to solve nonconvex Generalized Disjunctive Programs that involve bilinear and concave functions [7]. The main idea behind this methodology to find strong relaxations relies on a two step procedure. In the first step the nonconvex functions are replaced with polyhedral relaxations leading to a Linear GDP. In the second step the results of Sawaya and Grossmann [9] are used. These authors showed that a stronger relaxation of Linear GDPs can be obtained by the application of basic steps, a process that consists of intersecting disjunctions to obtain equivalent disjunctive sets whose hull relaxation is tighter. Even though the method we presented showed significant improvements when compared to traditional approaches, the efficiency of the method depends on the strength of the polyhedral relaxations of the nonconvex functions. In this paper we aim at generalizing this approach by allowing the use of nonlinear convex relaxations, which in some cases have shown to be order of magnitude tighter than linear relaxations (see Appendix B). This leads to a nonlinear convex GDP which relaxation can be strengthen by using recently results from the work of Ruiz and Grossmann [8]. Typical examples of nonlinear convex relaxations can be found in Appendix C.

This paper is organized as follows. In section 2 we define the general nonconvex GDP problem that we aim at solving and the traditional method to find relaxations. In section 3 we show how we can strengthen the relaxation of the traditional approach by presenting a systematic procedure to generate a hierarchy of relaxations based on the recent developments in convex disjunctive programming. In section 4 we outline a set of rules that will lead to a more efficient implementation of the method. Finally in section 5,6 and 7 we test the performance of the method in process network, reactor network and heat exchanger network problems.

2. Nonconvex Generalized Disjunctive Programs

The general structure of a GDP can be represented as follows [6]:

$$\begin{array}{lll}
\min & Z = f(x) + \sum_{k\in K} c_k & \\
s.t. & g^l(x) \le 0 \quad l \in L & \\
& \underset{i\in D_k}{\vee} \left[\begin{array}{c} Y_{ik} \\ r^j_{ik}(x) \le 0 \;\; j \in J_{ik} \\ c_k = \gamma_{ik} \end{array} \right] & k \in K \qquad (GDP_{NC}) \\
& \underset{i\in D_k}{\underline{\vee}} \; Y_{ik} & k \in K \\
& \Omega(Y) = True & \\
& x^{lo} \le x \le x^{up} & \\
& x \in R^n \,,\; c_k \in R^1,\; Y_{ik} \in \{True, False\} &
\end{array}$$

where $f : R^n \to R^1$ is a function of the continuous variables x in the objective function, $g^l : R^n \to R^l$ 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 (exclusive). In each term there is a Boolean variable Y_{ik}, a set of inequalities $r^j_{ik}(x) \le 0$, $r^j_{ik} : R^n \to R^d$, and a cost variable c_k. If Y_{ik} is true, then $r^j_{ik}(x) \le 0$ and $c_k = \gamma_{ik}$ are enforced; otherwise they are ignored. Also, $\Omega(Y) = True$ are logic propositions for the Boolean variables expressed in the conjunctive normal form $\Omega(Y) = \underset{t=1,2,..T}{\wedge} \left[\underset{(i,k)\in R_t}{\vee} (Y_{ik}) \underset{(i,k)\in Q_t}{\vee} (\neg Y_{ik}) \right]$ where for each clause t=1,2 ... T, R_t is the subset of indices of Boolean variables that are non-negated, and Q_t is the subset of indices of Boolean variables that are negated. The logic expression $\underset{i\in D_k}{\underline{\vee}} Y_{ik}$ ensures that only one term in each disjunction is True.

In this work we are interested in those problems where the functions f,r and

g might be nonconvex. Traditionally, the approach followed to find relaxations for GDP_{NC} consists in replacing the nonconvex functions r^j_{ik},g^l and f with with suitable convex underestimators $\hat{r}^j_{ik}$, $\hat{r}^l$ and $\hat{f}$ leading to a convex GDP, i.e. GDP_{CR} [4].

$$
\begin{array}{lll}
\min & Z = \hat{f}(x) + \sum_{k\in K} c_k & \\
s.t. & \hat{g}^l(x) \leq 0 \quad l \in L & \\
 & \bigvee\limits_{i\in D_k} \left[\begin{array}{c} Y_{ik} \\ \hat{r}^j_{ik}(x) \leq 0 \quad j \in J_{ik} \\ c_k = \gamma_{ik} \end{array} \right] \qquad\qquad\qquad k \in K & \qquad (GDP_{CR}) \\
 & \underline{\vee}_{i\in D_k} \; Y_{ik} \qquad\qquad\qquad\qquad\qquad k \in K & \\
 & \Omega(Y) = True & \\
 & x^{lo} \leq x \leq x^{up} & \\
 & x \in R^n \, , \, c_k \in R^1, \, Y_{ik} \in \{True, False\} &
\end{array}
$$

In this work we will show that by the application of a systematic procedure we can improve the strength of the continuous relaxation of (GDP_{CR}), leading to stronger lower bound predictions for (GDP_{NC}).

3. A hierarchy of convex relaxations for nonconvex GDP

In this section we present a general framework to obtain a hierarchy of convex nonlinear relaxations for the nonconvex GDP problem (GDP_{NC}) that can serve as a basis to predict tight lower bounds to the global optimum. In the first step of this approach, we replace each nonconvex function with a valid convex under- and over-estimator leading to a formulation of the form GDP_{CR}. Clearly, the feasible region defined by GDP_{NC} is contained in the feasible region defined by GDP_{RC}.

In order to predict strong lower bounds for the global optimum of (GDP_{NC}), we consider the hierarchy of relaxations for (GDP_{RC}) from the work of Ruiz and Grossmann [8]. These authors proved that any nonlinear Convex Generalized Disjunctive Program (CGDP) that involves Boolean and continuous variables can be equivalently formulated as a Convex Disjunctive Program (CDP) that only involves continuous variables (see Appendix A). This means that we are able to exploit the wealth of theory behind convex Disjunctive

Programming in order to solve convex GDP. Convex Disjunctive Programming can be regarded as the optimization over disjunctive sets. A disjunctive set is in regular form when it is described by the intersection of the union of convex sets. 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, and the Disjunctive Normal Form (DNF), which is expressed as the union of convex sets. One important result in Convex Disjunctive Programming theory, as presented in the work of Ruiz and Grossmann [8], is that we can systematically generate a set of equivalent convex Disjunctive Programs 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 3.1. *Let F be a disjunctive set in regular form. Then F can be brought to DNF by $|K|-1$ recursive applications of the following basic step which preserves regularity:*
For some $r, s \in K$, bring $S_r \cap S_s$ to DNF by replacing it with:

$$S_{rs} = \bigcup_{i \in D_r, j \in D_s} (P_i \cap P_j)$$

Although the formulations obtained after the application of basic steps on the disjunctive sets are equivalents, their continuous relaxations are not. We denote the continuous relaxation of a disjunctive set $F = \cap S_j$ in regular form where each S_j is a union of convex sets, as the hull-relaxations of F (or h-rel F). Where h-rel $F := \cap$ clconv S_j and clconv S_j denotes the closure of the convex hull of S_j. That is, if $S_j = \cup P_i$, $P_i = \{x \in R^n, g^i(x) \leq 0\}$, then the clconv S_j is given by :

$$\begin{aligned}
&x = \sum_{i \in D} x^i \\
&\lambda_i g_i(x^i/\lambda_i) \leq 0, \qquad i \in D \\
&\sum_{i \in D} \lambda_i = 1, \;\; \lambda_i \geq 0, \quad i \in D \qquad\qquad (\text{DISJ}_{rel}) \\
&|x^i| \leq L\lambda_i \qquad\qquad i \in D
\end{aligned}$$

As shown in the following Theorem from Ruiz and Grossmann [8], the application of a basic step leads to a new disjunctive set whose hull relaxation is at least as tight, if not tighter, than the original one.

Theorem 3.2. *For $i = 1, 2....k$ let $F_i = \cap_{k \in K} S_k$ be a sequence of regular forms of a disjunctive set such that F_i is obtained from F_{i-1} by the application of a basic step, then:*

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

Now we are ready to present one of the main results in this paper. Namely, a hierarchy of relaxations for GDP_{NC}. Let us suppose GDP_{RC0} is obtained by replacing the nonconvex functions with suitable relaxations as presented in section 2. Also, let us assume that GDP_{RCi} is the convex generalized disjunctive program whose defining disjunctive set is obtained after applying i basic steps on the disjunctive set of GDP_{RC0} and t the number of basic steps required to achieve the DNF. Then, the following relationship can be established,

$h\text{-}rel(F_0^{GDP_{RC}}) \supseteq h\text{-}rel(F_1^{GDP_{RC}})... \supseteq h\text{-}rel(F_i^{GDP_{RC}})\ ... \supseteq h\text{-}rel(F_t^{GDP_{RC}}) \supseteq F_t^{GDP_{RC}} \sim F_0^{GDP_{RC}} \supseteq F^{GDP_{NC}}$,

where $F_i^{GDP_{RC}}$ denotes the defining disjunctive set of GDP_{RCi} and $F^{GDP_{NC}}$ the defining disjunctive set of GDP_{NC}. Also, the symbol $\sim$ denotes equivalency.

It is important to note that solving the reformulated NLP using $(DISJ_{rel})$ would lead to numerical difficulties since the function $\lambda_i g^i(x^i/\lambda_i)$ is not differentiable at $\lambda_i = 0$. In order to overcome the latter issue we propose to use the approximation given by Sawaya (2006) et al. of the function :

$$\lambda_i g^i(x^i/\lambda_i) = ((1-\epsilon)\lambda_i + \epsilon)g^i(x^i/(1-\epsilon)\lambda_i + \epsilon) - \epsilon g^i(0)(1-\lambda_i) \le 0$$

which is exact for $\lambda_i = 0$ and $\lambda_i = 1$, independently of the values of ϵ. This parameter only has impact on the accuracy of the relaxation.

The approximation above fails if g_i is not defined at the zero point. However, we can overcome this issue by making a suitable variable transformation such that. In general:

$$\lambda_i g^i(x^i/\lambda_i) = ((1-\epsilon)\lambda_i + \epsilon)g'^i(y^i/(1-\epsilon)\lambda_i + \epsilon) - \epsilon g'^i(0)(1-\lambda_i) \le 0$$

where $g'^i(y) = g^i(y + \beta)$ and $y + \beta = x$. Note that convexity is preserved in this transformation provided the original function g^i is convex.

4. Rules to implement basic steps on convex Disjunctive Programs

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. As described in the work of Ruiz and Grossmann [7] the following propositions give sufficient conditions for a particular disjunctive set for which a basic step will not lead to a tighter relaxation.

Proposition 4.1. *Let $S = \bigcap_i P_i$, where $P_i, i = \{1, 2\}$ are convex sets defined in the x space, H is a half space defined by $ax + b \leq 0$ and H^* is a facet of H. If P_1 is a point such that $P_1 \subseteq H^*$ then h-rel$(S \cap H)$ = h-rel$(S) \cap$ h-rel(H)*

Proposition 4.2. *Let S_1 and S_2 be two disjunctive sets defined in 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 h-rel$(S_1 \cap S_2)$ = h-rel$(S_1) \cap$ h-rel(S_2) .*

A new rule developed in the work of Ruiz and Grossmann [8] consists in the inclusion of the convex objective function in the disjunctive set previous the application of basic steps. This has shown to be useful to strengthen the final relaxation of the disunctive set. However, one question that arises is whether this relaxation is still valid for the nonconvex case. The following proposition that we state without proof aims at tackling this question.

Proposition 4.3. *Given the nonconvex GDP (GDP_{NC}) defined as $Z = \{\min_{x \in F} f(x)\}$ where F is a disjunctive set and $f(x)$ a nonconvex function, and the nonconvex GDP, $GDP_{NC'}$ defined as $Z = \{\min_{(x,\alpha) \in (F \cap D)} \alpha\}$ such that $D := \{(x, \alpha) \in R^{n+1} | \alpha \geq f(x)\}$. Then GDP_{NC} is equivalent to $GDP_{NC'}$ in the sense that every point $(x, \alpha) \in (F \cap D)$ has a corresponding $x \in F$ and every point $x \in F$ has a corresponding $(x, \alpha) \in (F \cap D)$ and the optimal solution of the problems are the same.*

Since $GDP_{NC'}$ is equivalent to GDP_{NC} the hierarchy of relaxations of $GDP_{NC'}$ is valid for GDP_{NC}.

Another important aspect to consider is the effect that a particular basic step has in the increase 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 convex sets in the disjunctive set, keeping the size of the formulation smaller.

5. Examples

In this section we present three main case studies, namely, the global optimization of a process network, a heat exchanger network and a reactor network. These problems are suitable for testing the performance of the method since they can be easily represented by GDPs and the nonconvex constraints have nonlinear convex envelopes or strong convex nonlinear relaxations. Note that this is different from the case of bilinear or concave constraints for which linear convex envelopes can be obtained.

5.1. Strengthening the lower bounds in process network models

Consider the optimization of a process network with fixed charges in Figure 1 [3]. The raw material A can be processed in either processes 2 or 3 to produce B which is required for the production of C. Alternatively, B can be purchased form the market eliminitaing processes 2 and 3 if it is not profitable. If any unit is selected, a fixed cost has to be paid for the selection. The objective function includes the operating costs and the revenue from the sales of product C. The superstructure involves three units and the Boolean variables, Y_2 and Y_3 represent the existence of units in the flowsheet. Design specifications of the problem requires that processes 2 and 3 cannot appear together in a feasible flowsheet.

Process 2 and 3 can be described by the following function $e^{x_{out}/\alpha} + 1 = x_{in}$, where x_{out} and x_{in} are the outlet and inlet flows respectively and α a parameter that defines the process. Note that process 2 and 3 work in different range of admissible values for x_{in}. Also there is a minimum value for the inlet stream that we need to process.

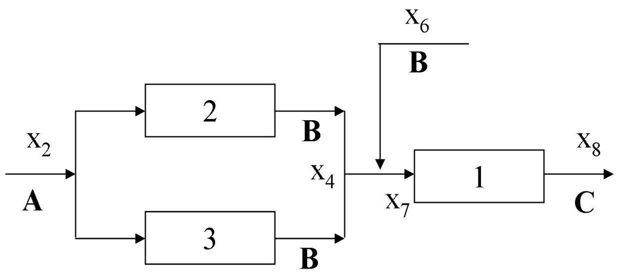

Figure 1: Example of a Process Network

The following is the Generalized Disjunctive Program that describes the above problem:

Objective Function:
$min\, Z = \theta_2 x_2 + \theta_7 x_7 - \theta_8 x_8 - \theta_6 x_6 + c_1 + c$

Global Constraints:
$x_4 + x_6 = x_7$
$x_2 \geq \beta$

Process Unit 1: $(PROCNET)$
$c_1 = \delta_7 x_7 + \delta_8 x_8$
$\xi x_7 = x_8$

Process Unit 2 and 3:

$$\begin{bmatrix} Y_2 \\ e^{x_4/\alpha_{42}} - 1 = x_2 \\ x_2^{lo2} \leq x_2 \leq x_2^{up2} \\ c = \delta_{42} x_4 + \delta_{22} x_2 + \gamma_2 \end{bmatrix} \vee \begin{bmatrix} Y_3 \\ e^{x_4/\alpha_{43}} - 1 = x_2 \\ x_2^{lo3} \leq x_2 \leq x_2^{up3} \\ c = \delta_{43} x_4 + \delta_{23} x_2 + \gamma_3 \end{bmatrix}$$

Logic Constraints:
$Y_2 \vee Y_3 = True$

Bounds:
$x_2, x_4, x_6 \geq 0$

Clearly, the problem is nonconvex where the nonconvexities arise from the nonlinear inequalities defining the process and from the disjunctive nature of the problem. In order to find a relaxation we propose first to find a relaxation for the nonlinear equalities. As a result a convex generalized disjunctive program is obtained. As a second step we will make use of the last developments in convex disjunctive programming theory to find strong relaxations for the convex disjunctive set.

5.1.1. STEP 1: Finding a convex GDP relaxation

Any nonlinear equation of the form $g(x) = 0$ where $g(x)$ is convex can be relaxed as $g(x) \leq 0$. The proof is trivial considering that $g(x) \leq 0$ is nothing but the epigraph of $g(x)$ and hence convex. Since $e^{x_{out}/\alpha} - 1 - x_{in}$ is a convex function (note that is the sum of convex funtions) then $e^{x_{out}/\alpha} - 1 - x_{in} \leq 0$ is a valid relaxation. In order to strengthen this relaxation we will add a cut that is obtained from the secant of $e^{x_{out}/\alpha} - 1 = x_{in}$ with extreme points $[x_{out}^{lo}, x_{out}^{up}]$, namely $\dfrac{e^{x_{out}^{up}/\alpha} - e^{x_{out}^{lo}/\alpha}}{x_{out}^{up} - x_{in}^{up}}(x_{out} - x_{out}^{lo}) + e^{x_{out}^{lo}/\alpha} - 1 \geq x_{in}$

Hence, the following convex disjunctive program is obtained:

$$\min \; Z = \Theta x_4 + \Gamma x_6 + c + \theta_2 x_2 \qquad\qquad (PROCNET_{step1})$$

$$x_2 \geq \beta$$

$$\begin{bmatrix} Y_2 \\ \Xi_2(x_4 - x_4^{lo}) + e^{x_4^{lo}/\alpha} - 1 \geq x_2 \\ e^{x_4/\alpha_{42}} - 1 \leq x_2 \\ x_2^{lo2} \leq x_2 \leq x_2^{up2} \\ c = \delta_{42}x_4 + \delta_{22}x_2 + \gamma_2 \end{bmatrix} \vee \begin{bmatrix} Y_3 \\ \Xi_3(x_4 - x_4^{lo}) + e^{x_4^{lo}/\alpha} - 1 \geq x_2 \\ e^{x_4/\alpha_{43}} - 1 \leq x_2 \\ x_2^{lo3} \leq x_2 \leq x_2^{up3} \\ c = \delta_{43}x_4 + \delta_{23}x_2 + \gamma_3 \end{bmatrix}$$

$$Y_2 \vee Y_3 = True$$

$$x_2, x_4, x_6 \geq 0$$

Note that for ease of notation we have reduced the space of the formulation by appropriate variable substitutions, where $\Theta = \delta_7 + \theta_7 - \theta_8\xi + \delta_8\xi$,

$$\Gamma = \delta_7+\theta_7-\theta_8\xi+\delta_8\xi-\theta_6, \Xi_2 = \frac{e^{x_4^{up}/\alpha_4 2} - e^{x_4^{lo}/\alpha_4 2}}{x_4^{up} - x_4^{lo}} \text{ and } \Xi_3 = \frac{e^{x_4^{up}/\alpha_4 3} - e^{x_4^{lo}/\alpha_4 3}}{x_4^{up} - x_4^{lo}}$$

5.1.2. STEP 2: Application of basic steps

By introducing the global constraints inside the disjunctions we obtain a new disjunctive set which hull relaxation is tighter, leading to a tighter relaxation for the nonconvex problem.

$$\min \ Z = \Theta x_4 + \Gamma x_6 + c + \theta_2 x_2 \qquad (PROCNET_{step2})$$

$$\left[\begin{array}{c} Y_2 \\ x_2 \geq \beta \\ \Xi_2(x_4 - x_4^{lo}) + e^{x_4^{lo}/\alpha} - 1 \geq x_2 \\ e^{x_4/\alpha_{42}} - 1 \leq x_2 \\ x_2^{lo2} \leq x_2 \leq x_2^{up2} \\ c = \delta_{42}x_4 + \delta_{22}x_2 + \gamma_2 \end{array}\right] \vee \left[\begin{array}{c} Y_3 \\ x_2 \geq \beta \\ \Xi_3(x_4 - x_4^{lo}) + e^{x_4^{lo}/\alpha} - 1 \geq x_2 \\ e^{x_4/\alpha_{43}} - 1 \leq x_2 \\ x_2^{lo3} \leq x_2 \leq x_2^{up3} \\ c = \delta_{43}x_4 + \delta_{23}x_2 + \gamma_3 \end{array}\right]$$

$Y_2 \vee Y_3 = True$
$x_2, x_4, x_6 \geq 0$

The final NLP relaxation used reads,

$$\begin{aligned}
\min Z \ & = \Theta x_4 + \Gamma x_6 + c^2 + c^3 + \theta_2 x_2 \\
s.t. \ & \\
& x_2 = x_2^2 + x_2^3 \\
& x_4 = x_4^2 + x_4^3 \\
& x_2^2 - \beta\lambda_2 \geq 0 \\
& x_3^2 - \beta\lambda_3 \geq 0 \qquad\qquad (PROCNET_{nlp}) \\
& \Xi_2(x_4^2 - x_4^{lo}\lambda_2) + (e^{x_4^2/\alpha_{42}} - 1)\lambda_2 \geq x_4^2 \\
& e^{(x_4^2/(\lambda_2\alpha_{42})} - \lambda_2 \leq x_2^2 \\
& x_2^{lo2}\lambda_2 \leq x_2^2 \leq x_2^{up2}\lambda_2 \\
& c^2 = \delta_{42}x_4^2 + \delta_{22}x_2^2 + \gamma_{22}\lambda_2 \\
& \Xi_3(x_4^3 - x_4^{lo}\lambda_3) + (e^{x_4^3/\alpha_{43}} - 1)\lambda_3 \geq x_4^3 \\
& e^{(x_4^3/(\lambda_3\alpha_{43})} - \lambda_3 \leq x_2^3 \\
& x_3^{lo2}\lambda_3 \leq x_2^3 \leq x_2^{up2}\lambda_2 \\
& c^3 = \delta_{43}x_4^3 + \delta_{23}x_2^3 + \gamma_{23}\lambda_3
\end{aligned}$$

$$\lambda_2 + \lambda_3 = 1$$
$$0 \le \lambda_2 \le 1, \le \lambda_3 \le 1$$
$$\lambda_2 x_2^{lo} \le x_2^2 \le \lambda_2 x_2^{lo}, \lambda_3 x_2^{lo} \le x_2^3 \le \lambda_3 x_2^{lo}$$
$$\lambda_2 x_4^{lo} \le x_4^2 \le \lambda_2 x_4^{lo}, \lambda_3 x_4^{lo} \le x_4^3 \le \lambda_3 x_4^{lo}$$

Note that to keep the notation clear we did not add the approximation of $\lambda_i g^i(x^i/\lambda_i)$.

5.2. Strengthening the lower bounds in reactor networks with non-elementary kinetics

Consider the reactor network illustrated in Figure 2. This simple network has two reactors where the reaction between components A and B to form C takes place. This reaction is non-elementary and the production rate of component C can be represented through the following posynomial function $r_C = kx_A^a x_B^b$. It is important to note that the reaction conditions in the two reactors are different which lead to different reaction constants k,a and b. Also, for safety reasons, the compositions of A and B should lie within a given safety curve. The problem consists in selecting the reactor that will maximize the profit in the production of C considering fixed costs for the reactors.

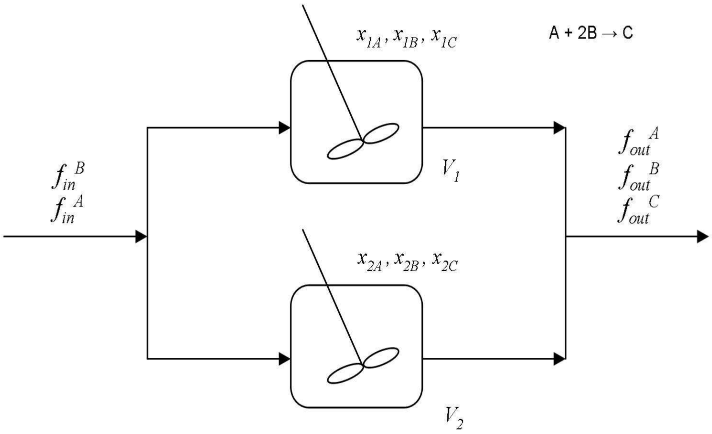

Figure 2: Example of a Reactor Network

The following is the Generalized Disjunctive Program that describes the above problem:

Objective Function:
$\min -C_C f_{out}^C + c$

Mass Balance:
$gen = f_{out}^C$

Safety zone constraint: (RXN)
$x_A + x_B \geq \beta$

Reactors:

$$\begin{bmatrix} Y_1 \\ gen = V_1 k_1 x_A^{a1} x_B^{b1} \\ x_A^{lo1} \leq x_A \leq x_A^{up1} \\ x_B^{lo1} \leq x_B \leq x_B^{up1} \\ c = \gamma_1 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ gen = V_2 k_2 x_A^{a2} x_B^{b2} \\ x_A^{lo2} \leq x_A \leq x_A^{up2} \\ x_B^{lo2} \leq x_B \leq x_B^{up2} \\ c = \gamma_2 \end{bmatrix}$$

Logic Constraints:

$Y_1 \vee Y_2 = True$

Bounds on variables:
$x_A^{LO} \leq x_A \leq x_A^{UP}$
$x_B^{LO} \leq x_B \leq x_B^{UP}$

Clearly, the problem is nonconvex where the nonconvexities arise from the posynoimial terms $x_A^a x_B^b$ and from the disjunctive nature of the problem. In order to find a relaxation we propose first to find a relaxation for the posynomial terms. As a result a convex generalized disjunctive program is obtained. As a second step we will make use of the last developments in convex disjunctive programming theory to find strong relaxations for the convex disjunctive set.

5.2.1. STEP 1: Finding a convex GDP relaxation

Strong relaxations for posynomial terms have been studied by Bjoerk et. al [11] and Han-Li et al. [12] (see Appendix C). By following this technique a convex relaxation for $r_C = x_A^a x_B^b$ can be obtained as follows:

$$r_C \geq y_A^{-a} y_B^{-b}$$
$$1 \geq \frac{x_A}{x_A^{lo}} + x_A^{lo} y_A - \frac{x_A^{lo}}{x_A^{lo}}$$
$$1 \geq \frac{x_B}{x_B^{lo}} + x_B^{lo} y_B - \frac{x_B^{lo}}{x_B^{up}}$$

Clearly, for a and b positive $\leq x_A^{up\,a} x_B^{up\,b}$ is an upper bound for r_C.

By replacing the nonconvex terms with the relaxation proposed we obtain the following convex GDP:

$$\min -C_C f_{out}^C + c \qquad (RXN_{step1})$$

$$gen = f_{out}^C$$

$$x_A + x_b \geq \beta$$

$$\begin{bmatrix} Y_1 \\ gen \geq V_1 k_1 y_A^{-a1} y_B^{-b1} \\ gen \leq V_1 k_1 x_A^{up\,a1} x_B^{up\,b1} \\ x_A^{lo1} \leq x_A \leq x_A^{up1} \\ x_B^{lo1} \leq x_B \leq x_B^{up1} \\ 1 \geq \dfrac{x_A}{x_A^{lo}} + x_A^{lo} y_A - \dfrac{x_A^{lo}}{x_A^{up}} \\ 1 \geq \dfrac{x_B}{x_B^{lo}} + x_B^{lo} y_B - \dfrac{x_B^{lo}}{x_B^{up}} \\ c = \gamma_1 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ gen \geq V_2 k_2 y_A^{-a2} y_B^{-b2} \\ gen \leq V_1 k_2 x_A^{up\,a2} x_B^{up\,b2} \\ x_A^{lo2} \leq x_A \leq x_A^{up2} \\ x_B^{lo2} \leq x_B \leq x_B^{up2} \\ 1 \geq \dfrac{x_A}{x_A^{lo}} + x_A^{lo} y_A - \dfrac{x_A^{lo}}{x_A^{UP}} \\ 1 \geq \dfrac{x_B}{x_B^{lo}} + x_B^{lo} y_B - \dfrac{x_B^{lo}}{x_B^{up}} \\ c = \gamma_2 \end{bmatrix}$$

$$Y_1 \vee Y_2 = True$$

$$x_A^{lo} \leq x_A \leq x_A^{up}$$
$$x_B^{lo} \leq x_B \leq x_B^{up}$$

5.2.2. STEP 2: Application of basic steps

By introducing the global constraints inside the disjunctions we obtain a new disjunctive set which hull relaxation is tighter, leading to a tighter relaxation for the nonconvex problem. For space reasons we skip the explicit representation of the NLP relaxation.

5.3. Strengthening the lower bounds in heat exchanger network models

Consider the HEN illustrated in Figure 3 [5]. This network has four heat exchangers and consists of one cold stream, C1 which is split into exchangers 1 and 2. The two hot streams, H1 and H2, exchange heat in series with cold streams C2 and C1 and C3 and C1, respectively. The inlet and outlet temperatures and the heat capacity flow rates are given in Table D.5. Note that the outlet temperatures of C2 and C3 are not specified since these are assumed to correspond to streams that are in the last stage of the process and do not require any specific temperature. The objective of this problem is the minimization of the cost of the network given by the heat exchanger areas and some fixed costs for the exchangers.

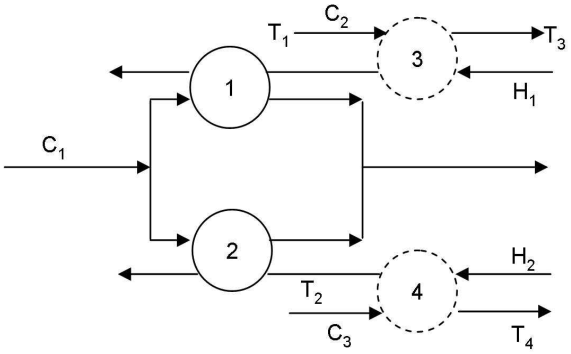

Figure 3: Example of a Heat Exchanger Network

The generalized disjunctive program that can be used to represent the problem is given by:

$$\min C = c_1A_1 + c_2A_2 + c_3A_3 + c_4A_4 + C_3 + C_4$$

$$\text{s.t. } A_1 = \frac{Q_1}{U_1 \Delta T_1} \, , A_2 = \frac{Q_2}{U_2 \Delta T_2} \qquad\qquad (HEN)$$

$$\begin{bmatrix} Y_1 \\ A_3 = \dfrac{Q_3}{U_3 \Delta T_3} \\ C_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ A_3 = 0 \\ C_3 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_4 \\ A_4 = \dfrac{Q_4}{U_4 \Delta T_4} \\ C_4 = \gamma_4 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ A_4 = 0 \\ C_4 = 0 \end{bmatrix}$$

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

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

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

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

$$Q_i \geq 0, \, \Delta T_i \geq EMAT \, , i = 1, ..., 4$$

Clearly, the problem is nonconvex where the nonconvexities arise from the fractional terms $Q_i/\Delta T_i$ and from the disjunctive nature of the problem. In order to find a relaxation we propose first to find a relaxation for the fractional terms. As a result a convex generalized disjunctive program is obtained. As a second step we will make use of the last developments in convex disjunctive programming theory to find strong relaxations for the convex disjunctive set.

5.3.1. STEP 1: Finding a convex GDP relaxation

Finding strong relaxations for fractional terms has been a challenge for a few years until a thorough analysis was given by Tawarmalani and Sahinidis [10] (see Appendix). In this work they present the convex and concave envelopes for fractional terms. By following this technique, the convex and concave envelopes for $z_i = Q_i/\Delta T_i$ are as follows:

$$z_{pi} \geq \frac{Q_i^L(Q_i^L \Delta T_{pi} - Q_i(\Delta T_i^L + \Delta T_i^U) + Q_i^U(\Delta T_i^L + \Delta T_i^U - \Delta T_{pi}))}{(Q_i^U - Q_i^L)(\Delta T_i^L \Delta T_i^U)}$$

$$(z_i - z_{pi})(Q_i^U - Q_i^L)(\Delta T_i^U - \Delta T_i^L)^2 \geq Q_i^U (Q_i - Q_i^L)^2$$

$$\Delta T_i^L (Q_i^U - Q_i) \leq \Delta T_{pi}(Q_i^U - Q_i^L) \leq \Delta T_{pi}^U (Q_i^U - Q_i)$$

$$\Delta T_i^L (Q_i - Q_i^L) \leq (\Delta T_i - \Delta T_{pi})(Q_i^U - Q_i^L) \leq \Delta T_{pi}^U (Q_i - Q_i^L)$$

$$z_i - z_{pi}, z_{pi} \geq 0$$

By replacing $Q_i/\Delta T_i$ with z_i and introducing the equations presented above we obtain a convex disjunctive program whose hull relaxation is a relaxation for the original nonconvex program.

$$\min C = c_1A_1 + c_2A_2 + c_3A_3 + c_4A_4 + C_3 + C_4 \qquad (HEN_{step1})$$

s.t.

$$\begin{aligned}
&z_{pi} \geq \frac{Q_i^L(Q_i^L \Delta T_{pi} - Q_i(\Delta T_i^L + \Delta T_i^U) + Q_i^U(\Delta T_i^L + \Delta T_i^U - \Delta T_{pi}))}{(Q_i^U - Q_i^L)(\Delta T_i^L \Delta T_i^U)} \\
&(z_i - z_{pi})(Q_i^U - Q_i^L)(\Delta T_i^U - \Delta T_i^L)^2 \geq Q_i^U (Q_i - Q_i^L)^2 \\
&\Delta T_i^L (Q_i^U - Q_i) \leq \Delta T_{pi}(Q_i^U - Q_i^L) \leq \Delta T_{pi}^U (Q_i^U - Q_i) \qquad \text{i = 1,2} \\
&\Delta T_i^L (Q_i - Q_i^L) \leq (\Delta T_i - \Delta T_{pi})(Q_i^U - Q_i^L) \leq \Delta T_{pi}^U (Q_i - Q_i^L) \\
&z_i - z_{pi}, z_{pi} \geq 0
\end{aligned}$$

$$\begin{bmatrix} Y_1 \\ z_{p3} \geq \dfrac{Q_3^L(Q_3^L \Delta T_{p3} - Q_3(\Delta T_3^L + \Delta T_3^U) + Q_3^U(\Delta T_3^L + \Delta T_3^U - \Delta T_{p3}))}{(Q_3^U - Q_3^L)(\Delta T_3^L \Delta T_3^U)} \\ (z_3 - z_{p3})(Q_3^U - Q_3^L)(\Delta T_3^U - \Delta T_3^L)^2 \geq Q_3^U (Q_3 - Q_3^L)^2 \\ \Delta T_3^L (Q_3^U - Q_3) \leq \Delta T_{p3}(Q_3^U - Q_3^L) \leq \Delta T_{p3}^U (Q_3^U - Q_3) \\ \Delta T_3^L (Q_3 - Q_3^L) \leq (\Delta T_3 - \Delta T_{p3})(Q_3^U - Q_3^L) \leq \Delta T_{p3}^U (Q_3 - Q_3^L) \\ z_3 - z_{p3}, z_{p3} \geq 0 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ A_3 = 0 \\ c_3 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_1 \\ z_{p4} \geq \dfrac{Q_4^L(Q_4^L \Delta T_{p4} - Q_4(\Delta T_4^L + \Delta T_4^U) + Q_4^U(\Delta T_4^L + \Delta T_4^U - \Delta T_{p4}))}{(Q_4^U - Q_4^L)(\Delta T_4^L \Delta T_4^U)} \\ (z_4 - z_{p4})(Q_4^U - Q_4^L)(\Delta T_4^U - \Delta T_4^L)^2 \geq Q_4^U(Q_4 - Q_4^L)^2 \\ \Delta T_4^L(Q_4^U - Q_4) \leq \Delta T_{p4}(Q_4^U - Q_4^L) \leq \Delta T_{p4}^U(Q_4^U - Q_4) \\ \Delta T_4^L(Q_4 - Q_4^L) \leq (\Delta T_4 - \Delta T_{p4})(Q_4^U - Q_4^L) \leq \Delta T_{p4}^U(Q_4 - Q_4^L) \\ z_4 - z_{p4}, z_{p4} \geq 0 \\ c_4 = \gamma_4 \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ A_4 = 0 \\ c_4 = 0 \end{bmatrix}$$

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

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

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

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

$$Q_i \geq 0,\; \Delta T_i \geq EMAT\,,\; i = 1, ..., 4$$

5.3.2. STEP 2: Application of basic steps

By introducing the global constraints inside the disjunctions we obtain a new disjunctive set which hull relaxation is tighter, leading to a tighter relaxation for the nonconvex problem. As in the case of reactor networks, in order to keep the presentation clear, we will skip the explicit representation of the GDP after the application of basic steps and the final NLP relaxation.

6. Global Optimization algorithm with improved relaxations

In this section we describe the global optimization framework from the work of Ruiz and Grossmann [7] that we will use to test the new relaxations proposed. The global optimization methodology of the GDP follows the well

known spatial branch and bound method [15, 2] and is presented in this 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: a) Relax the nonconvex terms using suitable convex under/over-estimators. This will lead to the convex GDP (GDP_{RLP0}). b) Apply basic steps according to the rules described in section 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 [16]. By using the result obtained in the previous step, a bound contraction of each continuous variable is performed [17]. 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 descendents 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 $UB - LB < tol$, where tol 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 NLP (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.

7. Numerical Results

In this section we compare the performance of the method proposed to find tighter relaxations for nonconvex GDPs with the traditional approach, i.e. Lee and Grossmann relaxation [4]. In order to do this, we first consider the lower bound obtained at the root node as one of the main indicators of the strength of the relaxation produced. Also, to test the performance of the set of rules presented in section 4 to guide the generation of basic steps we compare the lower bound at the root node with the lower bound that we would obtain if we solve the DNF form of the convex disjunctive program.

Table 1: Size and characteristics of the examples formulated as GDP

Example	Cont. Vars.	Boolean Vars.	Logic Const.	Disj. Const.	Global. Const.
Procnet 1	5	2	1	1	3
Procnet 2	5	2	1	1	3
RXN 1	4	2	1	1	6
RXN 2	4	2	1	1	6
HEN 1	18	2	2	2	21

Table 2: Size of different reformulations

	Traditional Approach			Proposed Approach		
Example	Bin	Con	Const	Bin	Con	Const
Procnet 1	2	12	19	2	12	20
Procnet 2	2	12	19	2	12	20
RXN 1	2	17	31	2	20	37
RXN 2	2	17	31	2	20	37
HEN 1	2	53	90	2	117	273

Clearly, from Table 3 we observe in all instances a significant improvement in the lower bound predicted. For instance in Procnet 1 our approach predicts 16.01 as a lower bound whereas the traditional approach is only able to guarantee a bound of 11.85. Moreover, the lower bounds obtained are close if not the same as the one we would obtain if we solved the relaxation of the

Table 3: Performance of the method to find tight relaxations

Example	Global Optimum	LB (Lee and Grossmann)	LB (Proposed Approach)	LB of DNF
Procnet 1	18.61	11.85	16.01	16.01
Procnet 2	19.48	12.38	17.07	17.07
RXN 1	42.89	-337.5	-320.0	-320.0
RXN 2	76.47	22.5	40.0	40.0
HEN 1	48531	38729.27	48230	48531

Table 4: Performance of the relaxation within a spatial branch and bound framework

		Lee and Grossmann			Proposed Approach		
Example	Opt.	ABR %	Nds	T(s)	ABR %	Nds	T(s)
Procnet 1	18.61	51.3	3	6	67.0	2	4
Procnet 2	19.48	40.5	2	4	47.2	2	4
RXN 1	42.89	51.0	2	7	66.0	2	7
RXN 2	76.46	51.0	2	6	66.0	2	6
HEN 1	48531	13.8	3	15	35.0	1	14

DNF form. For example, HEN 1, reaches a lower bound of 48230, which is close to the maximum attainable 48531.

A further analysis of the performance of the relaxations proposed was carried out considering their effect when used within a spatial branch and bound freamework as described in section 6. The results can be seen in Table 4. Note that the slight improvement in the number of nodes necessary to find to solution or the computational time required is due to the fact that the problems are small in size. However, a clear indication of a tighter relaxation proposed is observed in the column "ABR %" which refers to how much reduction in the upper and lower bounds of the variables can be predicted. More precisely, the column "ABR %" refers to $1 - \sum_i \frac{x_i^{up*} - x_i^{lo*}}{|I|(x_i^{up} - x_i^{lo})}$ where $x_i^{up/lo}$ and $x_i^{up*/lo*}$ refer to the upper/lower bound of the variable x_i before and after the bound contraction procedure respectively and $|I|$ the number of variables x_i considered in the contraction. For instance, in HEN 1, we are able to reduce the bounds of the variables 35% respect to the original bounds, whereas by using the traditional approach we can only contract the bounds 13.8%. Note that the strength of the relaxations of the nonconvex functions heavily depend on the bounds of the variables on which they are defined and that is why it is very important to count on an efficient procedure to find these bounds.

8. Conclusions

We have proposed a framework to generate tight relaxation for the global optimization of nonconvex generalized disjunctive programs. We extended the method proposed in a recent work on bilinear and concave disjunctive programming by allowing the use of nonlinear relaxations. This is now possible from the latest developments in convex disjunctive programming. We tested the method in several instances and show significant improvements in the lower bounds predicted, which is a direct indication of tightening.

Appendix A. Equivalency between convex disjunctive programs and convex generalized disjunctive programs [9] [8]

Any convex generalized disjunctive program can be represented as a convex disjunctive program. The transformation that allows this, which is equivalent to the one proposed by Sawaya and Grossmann [9] for linear GDP, consists in first replacing the boolean variables $Y_{ik}, i \in D_k, k \in K$ 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], and finally convert logical relations $\underset{i \in D_k}{\veebar} Y_{ik}, k \in K$ and $\Omega(Y) = True$ into algebraic equations $\sum\limits_{i \in D_k} \lambda_{ik} = 1, k \in K$ and $H\lambda \geq h$, respectively. This yields the following equivalent disjunctive model:

$$
\begin{aligned}
\min \quad & Z = f(x) + \textstyle\sum_{k \in K} c_k \\
s.t. \quad & g^l(x) \leq 0 \;\; l \in L \\
& \bigvee_{i \in D_k} \begin{bmatrix} \lambda_{ik} = 1 \\ r^j_{ik}(x) \leq 0 \;\; j \in J_{ik} \\ c_k = \gamma_{ik} \end{bmatrix} \qquad\qquad k \in K \qquad\quad (CGDP) \\
& \sum_{i \in D_k} \lambda_{ik} = 1, k \in K \\
& A\lambda \geq a \\
& x^{lo} \leq x \leq x^{up} \\
& x \in R^n \, , \, c_k \in R^1, \, \lambda_{ik} \in 0, 1
\end{aligned}
$$

Appendix B. Nonlinear relaxations vs polyhedral relaxations

The maturity of linear programming solvers have encouraged the use of linear relaxations for nonconvex functions. Even though this approach has shown to be useful in several cases, many times it has a great impact in the time consumed to find the solution. The following is a simple example where we aim at showing this effect:

$\min \sum_{i=1}^{100} 10(x_i - 5)^2 + y_i^2$

$s.t. \;\; y_i = x_i^2 \qquad \forall i$

$|y_i| \leq 100, |x_i| \leq 100$

Solving this problem using BARON with linear relaxations, it takes more than 1000 seconds. On the other hand, by using the relaxation $y_i \geq x_i^2, \forall i$ the solution can be found at the root node in less than a second.

Appendix C. Typical nonlinear relaxation of nonconvex constraints

Appendix C.1. Fractional Terms
The convex hull of a set $NC = \{(x, y, f) | f = x/y, x^L \geq x \geq x^U, y^L \geq y \geq y^U\}$ is given in Tawarmalani and Sahinidis [10].

$$f_p \geq \frac{x^L(x^L y_p - x(y^L + y^U) + x^U(y^L + y^U - y_p))}{(x^U - x^L)(y^L y^U)}$$

$$(f - f_p)(x^U - x^L)(y^U - y^L)^2 \geq x^U(x - x^L)^2$$

$$y^L(x^U - x) \leq y_p(x^U - x^L) \leq y_p^U(x^U - x)$$

$$y^L(x - x^L) \leq (y - y_p)(x^U - x^L) \leq y_p^U(x - x^L)$$

$$f - f_p \geq 0$$

Appendix C.2. Convex functions in equality constraints
The relaxation for a set $NC = \{(x, g(x)) | g(x) = 0, x^L \leq x \leq x^U, g(x) \; convex\}$ can be relaxed as:

$g(x) \leq 0$

The proof is trivial considering that $g(x) \leq 0$ is nothing but the epigraph of $g(x)$ and hence convex.

Appendix C.3. Posynomials

Several strategies have been proposed in literature to tackle the relaxation of posynomials [11] [12]. Here we present the method proposed by Han-Lin et al. Given a function of the form $f(x) = dx_1^{\alpha_1}x_2^{\alpha_2}x_3^{\alpha_3}......x_n^{\alpha_n}$ where $0 < x_i^{lo} \leq x_i \leq x_i^{up}$ and $d > 0$ $\alpha_i < 0, i = 1, 2...m$ and $\alpha_i > 0, i = m+1, ...n$. A lower bound for $f(x)$ is obtained by:

$$f(x) \geq dx_1^{\alpha_1}x_2^{\alpha_2}x_m^{\alpha_m}y_{m+1}^{-\alpha_{m+1}}......y_n^{-\alpha_n}$$
$$1 \geq \frac{x_i}{x_i^{up}} + x_i^{lo}y_i - \frac{x_i^{lo}}{x_i^{up}} \quad i = m+1, m+2, ...n$$

Appendix C.4. Trigonometric Functions

Convex underestimators for the function $f(x) = \alpha sin(x+s)$, $x^L \geq x \geq x^U$ have been studied by Caratzoulas and Floudas (2005) [13], leading to non-linear relaxations. Here we present as an illustration the relaxation for:

$$f(x) = sin(x) \; 0 \leq x \leq \pi$$

Clearly the convex hull of the set $NC = \{x, f | f = sin(x),\ 0 \leq x \leq \pi\}$ is given by:

$$f \leq sin(x)$$
$$f \geq 0$$
$$0 \leq x \leq \pi$$

Appendix D. Data for numerical examples

Appendix D.1. Process network models

The parameters that are used in both instance 1 and 2 of the process network problem are: $\Theta = -3.05$, $\Gamma = 2.95$, $\theta_2 = 0.125$, $\delta_{22} = 5$, $\delta_{42} = 5$,

$\gamma_2 = 3$,$\delta_{23} = 3$, $\delta_{43} = 3$, $\gamma_3 = 4$, $x_2^{lo2} = 1$, $x_2^{up2} = 3$,$x_2^{lo3} = 5$, $x_2^{up3} = 8$. For the particular case of instance 1, $\alpha_{42} = 2.5$ and $\alpha_{43} = 1.8$, whereas for instance 2, $\alpha_{42} = 2$ and $\alpha_{43} = 5/3$

Appendix D.2. Reactor network problems

The parameters that are used in both instance 1 and 2 of the reactor network problem are: $\beta = 0.5$, $V_1k_1 = 2$, $V_2k_2 = 2$, $x_A^{lo1} = 0.4$,$x_A^{up1} = 0.6$,$x_B^{lo1} = 0.2$,$x_B^{up1} = 0.6$, $x_A^{lo2} = 0.1$,$x_A^{up2} = 0.2$,$x_B^{lo2} = 0.1$,$x_B^{up2} = 0.2$, $\gamma_1 = 80$, $\gamma_2 = 60$. For the particular case of instance 1, $a_1 = 1, a_2 = 1.2, b_1 = 0.5, b_2 = 0.3$ and $C_c = 40$ whereas for instance 2, $a_1 = 1, a_2 = 1.2, b_1 = 0.6, b_2 = 0.5$ and $C_c = 4$.

Appendix D.3. Heat Exchanger network problems

Table D.5: Data for HEN problem

stream	Fcp(kW/K)	Temp. inlet	Temp. outlet
C_1	10	300	400
C_2	4.545	365	
C_3	3.571	358	
H_1	3.555	575	395
H_2	2.125	718	398

The cost coefficients for the exchangers are c1=\$120/m2, c2=\$220/m2, c3 = \$440/m2 and c4 = 500/m2 and the overall heat-transfer coefficients are U1 = U2 = 0.1 kW/(K m2) and U3 = U4 = 1 kW / (K m2). The fixed costs for exchanger 3 is \$7000 whereas for exchanger 4 is \$1000. Also note that the EMAT is assumed to be 5K.

References

[1] Balas E., *Disjunctive Programming*, **5**, 3-51, 1979.

[2] Horst, R. and Tuy, H., *Global Optimization deterministic approaches (3rd Ed)*, Berlin: Springer-Verlag, 1996.

[3] Kocis G.R. and Grossmann I.E., *Relaxation Strategy for the structural optimization of Process Flow Sheets*, Industrial and Engineering Chemistry Research, **26**,1869, 1987.

[4] Lee S. and Grossmann I.E., *Global optimization of nonlinear generalized disjunctive programming with bilinear inequality constraints: application to process networks*, Computers and Chemical Engineering, **27**, 1557-1575, 2003.

[5] Quesada I. and Grossmann I.E., *Global optimization algorithm for heat exchanger networks*, Ind. Eng. Chem. Res., **32**, 487-499, 1993.

[6] Raman R. and Grossmann I.E., *Modelling and Computational Techniques for Logic-Based Integer Programming*, Computers and Chemical Engineering, **18**, 563, 1994.

[7] Ruiz J.P. and Grossmann I.E., *Strengthening the lower bounds for bilinear and concave GDP problems*, Computers and Chemical Engineering, **34**:3, 914-930, 2010.

[8] Ruiz J.P. and Grossmann I.E., *A hierarchy of relaxations for convex disjunctive programs*, Paper in preparation.

[9] Sawaya N., *Thesis: Reformulations, relaxations and cutting planes for generalized disjunctive programming*, Carnegie Mellon University, 2006.

[10] Tawarmalani, M. and Sahinidis, N., *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming.*, Kluwer Academic Publishers , 2002.

[11] Bjoerk, K.J., Lindberg, P.O.,Westerlund, T., *Some convexifications in global optimization of problems containing signomial terms.*, Comput. Chem. Eng. 27, 669679,2003.

[12] Han-Lin L., Jugn-Fa T., Floudas C. A., *Convex underestimation for posynomial functions of positive variables*, Optimization Letters, 2, 333-340, 2008.

[13] Caratzoulas, S., Floudas, C.A., *A trigonometric convex underestimator for the base functions in Fourier space.*, J. Optim. Theory Appl., 124(2), 339362, 2005.

[14] FLOUDAS, C. A., DETERMINISTIC GLOBAL OPTIMIZATION: THEORY METHODS AND APPLICATIONS., Dordrecht, The Netherlands: Kluwer Academic Publishers, 2000.

[15] AL-KHAYYAL, F.A. AND FALK, J.E.,*Jointly constrained biconvex programming.*, Mathematics of Operations Research, **8**(2):273-286, 1983.

[16] VISWANATHAN AND GROSSMANN I.E., *A combined penalty function and outer-approximation method for MINLP optimization*, Computers and Chemical Engineering, **14**, 769-782, 1990.

[17] ZAMORA J.M. AND GROSSMANN I.E., *A branch and bound algorithm for problems with concave univariate , bilinear and linear fractional terms*, **14**:3, 217-249, 1999.