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 tight linear relaxations can be generated, and extends its application to nonlinear relaxations. This is particularly important for those cases in which the convex envelopes of the nonconvex functions arising in the formulations are nonlinear (e.g. linear fractional terms). This extension is now possible by using the latest developments in disjunctive convex 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 and 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) [20]. 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 that are suboptimal. In the last decade many global optimization algorithms for nonconvex problems have been proposed [15][14]. 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.

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 tighter 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 [10]. The main idea behind this methodology to find tight 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 [12] are used. We showed that a tighter 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 C). This leads to a nonlinear convex GDP whose relaxation can be strengthened by using recent results from the work of Ruiz and Grossmann [11]. Typical examples of nonlinear convex relaxations can be found in Appendix D.

This paper is organized as follows. In section 2 we define the general nonconvex GDP problem that we aim at solving and we review the traditional hull relaxation method to find relaxations. In section 3 we show how we can strengthen the relaxation of the traditional approach by presenting a

novel systematic procedure to generate a hierarchy of relaxations based on the recent developments in disjunctive convex programming. In section 4 we outline a set of rules that lead to a more efficient implementation of the method. Finally, in sections 5,6 and 7 we test the performance of the method in process networks, reactor networks and heat exchanger network problems.

2. Nonconvex Generalized Disjunctive Programs

The general structure of a nonconvex GDP, which we denote as (GDP_{NC}), is as follows,

$$
\begin{aligned}
&\min_{x,c_k,Y_{ik}} \quad Z = f(x) + \sum_{k \in K} c_k \\
&s.t. \qquad g^l(x) \leq 0 \qquad\qquad l \in L \\
&\qquad\quad \bigvee_{i \in D_k}
\begin{bmatrix}
Y_{ik} \\
r_{ik}^j(x) \leq 0 \quad j \in J_{ik} \\
c_k = \gamma_{ik}
\end{bmatrix}
\qquad k \in K \qquad (GDP_{NC}) \\
&\qquad\quad \bigvee_{i \in D_k} Y_{ik} \qquad\qquad k \in K \\
&\qquad\quad \Omega(Y) = True \\
&\qquad\quad x^{lo} \leq x \leq x^{up} \\
&\qquad\quad x \in R^n \,,\, c_k \in R^1,\, Y_{ik} \in \{True, False\}, i \in D_k, k \in K
\end{aligned}
$$

where $f : R^n \to R^1$ is a function of the continuous variables x in the objective function, $g^l : R^n \to R^1, l \in L$ belongs to the set of global constraints, the disjunctions $k \in K$, are composed of a number of terms $i \in D_k$, that are connected by the OR operator. In each term there is a Boolean variable Y_{ik}, a set of inequalities $r_{ik}^j(x) \leq 0, r_{ik}^j : R^n \to R^m$, and a cost variable c_k. If Y_{ik} is True, then $r_{ik}^j(x) \leq 0$ 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 $\Omega(Y) = \bigwedge_{t=1,2,..T} \left[\bigvee_{(i,k) \in R_t} (Y_{ik}) \bigvee_{(i,k) \in Q_t} (\neg Y_{ik}) \right]$ 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 constraints $\bigvee_{i \in D_k} Y_{ik}$ ensure that only one Boolean variable is True in each disjunction.

Note that as opposed to the convex case, a nonconvex GDP is defined by functions in the constraints that may be nonconvex. The approach followed

3

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

$$
\begin{aligned}
\min_{x,c_k,Y_{ik}} \quad & Z = \hat{f}(x) + \sum_{k \in K} c_k \\
s.t. \quad & \hat{g}^l(x) \leq 0 \quad l \in L \\
& \bigvee_{i \in D_k}
\begin{bmatrix}
Y_{ik} \\
\hat{r}_{ik}^j(x) \leq 0 \quad j \in J_{ik} \\
c_k = \gamma_{ik}
\end{bmatrix}
\quad k \in K \qquad (GDP_{CR}) \\
& \bigvee_{i \in D_k} Y_{ik} \\
& \Omega(Y) = True \\
& x^{lo} \leq x \leq x^{up} \\
& x \in R^n, \, c_k \in R^1, \, Y_{ik} \in \{True, False\}
\end{aligned}
$$

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 strong lower bounds to the global optimum. In the first step of this approach, we replace each nonconvex function with a valid convex under- and overestimator 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_{CR}. To illustrate this concept, let us consider the case where the feasible region is defined by a two term disjunction, namely, $[g_1(x) \leq 0] \vee [g_2(x) \leq 0]$, where g_1 and g_2 are nonconvex functions as shown in Figure 1. Note that we assume that g_1 and g_2 implicitly define the bounds on x_1 and x_2. Clearly, by replacing the nonconvex functions with suitable over- and underestimators, i.e. $[\hat{g}_1(x) \leq 0] \vee [\hat{g}_2(x) \leq 0]$, the result is a disjunctive convex set whose feasible region contains the feasible region described by the original nonconvex disjunctive set.

In order to predict strong lower bounds for the global optimum of (GDP_{NC}), we consider the hierarchy of relaxations for (GDP_{CR}) from the work of our
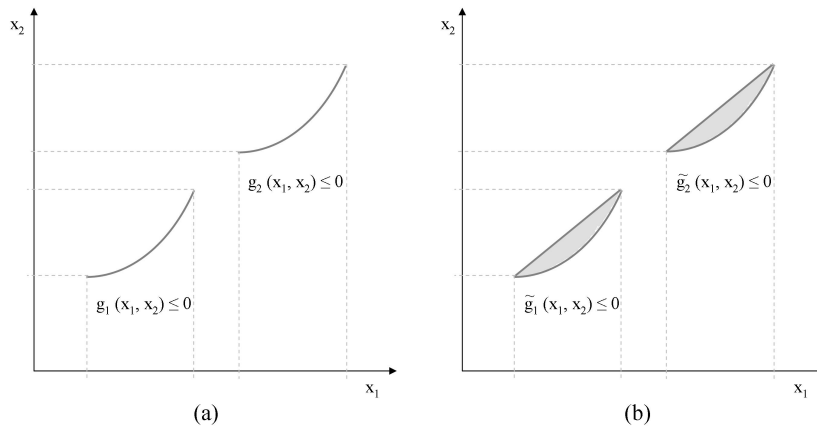
Figure 1: (a) GDP_{NC} (b) GDP_{CR}

previous paper [11]. In that paper we proved that any nonlinear Convex Generalized Disjunctive Program (CGDP) that involves Boolean and continuous variables can be equivalently formulated as a Disjunctive Convex Program (DCP) that only involves continuous variables (see Appendix A). This means that we are able to exploit the wealth of theory behind disjunctive convex programming in order to solve convex GDP.

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 disjunctive convex programming theory, as presented in a previous paper [11], is that we can systematically generate a set of equivalent disjunctive convex programs going from the CNF to the DNF by performing an operation called basic step that preserves regularity. A basic step is defined in Appendix B. Although the formulations obtained after the application of basic steps on the disjunctive sets are equivalent, their continuous relaxations are not. We denote the continuous relaxation of a disjunctive set $F = \bigcap_{j \in T} S_j$ in regular form, where each S_j is a union of convex sets, as the *hull-relaxation* of F (or $h - rel\, F$). Here $h - rel\, F := \bigcap_{j \in T} clconv\, S_j$ and $clconv\, S_j$ denotes the closure of the convex hull of S_j. That is, if $S_j = \bigcup_{i \in Q_j} P_i$, $P_i = \{x \in R^n, g_i(x) \leq 0\}$,

5

then the $clconvS_j$ is given by:

$$x = \sum_{i \in Q_j} \nu^i$$
$$\lambda_i g_i(\nu^i/\lambda_i) \leq 0, \qquad i \in Q_j$$
$$\sum_{i \in Q_j} \lambda_i = 1, \ \lambda_i \geq 0, \quad i \in Q_j \qquad \qquad (\text{DISJ}_{rel})$$
$$|\nu^i| \leq L\lambda_i \qquad \qquad i \in Q_j$$

where ν^i are disaggregated variables, λ_i are continuous variables between 0 and 1 and $\lambda_i g_i(\nu^i/\lambda_i)$ is the perspective function that is convex in ν and λ if the function $g(x)$ is also convex [19].

As shown in Theorem 3.1 from paper [11], 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 3.1. *For $i = 1, 2....k$ let $F_i = \bigcap_{k \in K} S_k$ be a sequence of regular forms of a disjunctive set such that F_i is obtained from F_{i-1} by the application of a basic step, then:*

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

Figure 2 illustrates Theorem 3.1.

Now we are ready to present one of the main results in this section. Namely, a hierarchy of relaxations for GDP_{NC}. Let us suppose GDP_{CR0} is obtained by replacing the nonconvex functions with suitable relaxations as presented in section 2. Also, let us assume that GDP_{CRi} is the convex generalized disjunctive program whose defining disjunctive set is obtained after applying i basic steps on the disjunctive set of GDP_{CR0} and t the number of basic steps required to achieve the DNF. Note that $i \leq t$. Then, the following relationship can be established,

$$h\text{-}rel(F_0^{GDP_{CR}}) \supseteq h\text{-}rel(F_1^{GDP_{CR}})... \supseteq h\text{-}rel(F_i^{GDP_{CR}}) \ ... \supseteq$$

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

where $F_i^{GDP_{CR}}$ denotes the defining disjunctive set of GDP_{CRi} and $F^{GDP_{NC}}$ the defining disjunctive set of GDP_{NC}. Also, the symbol $\sim$ denotes equivalence. Figure 3 summarizes the proposed framework,
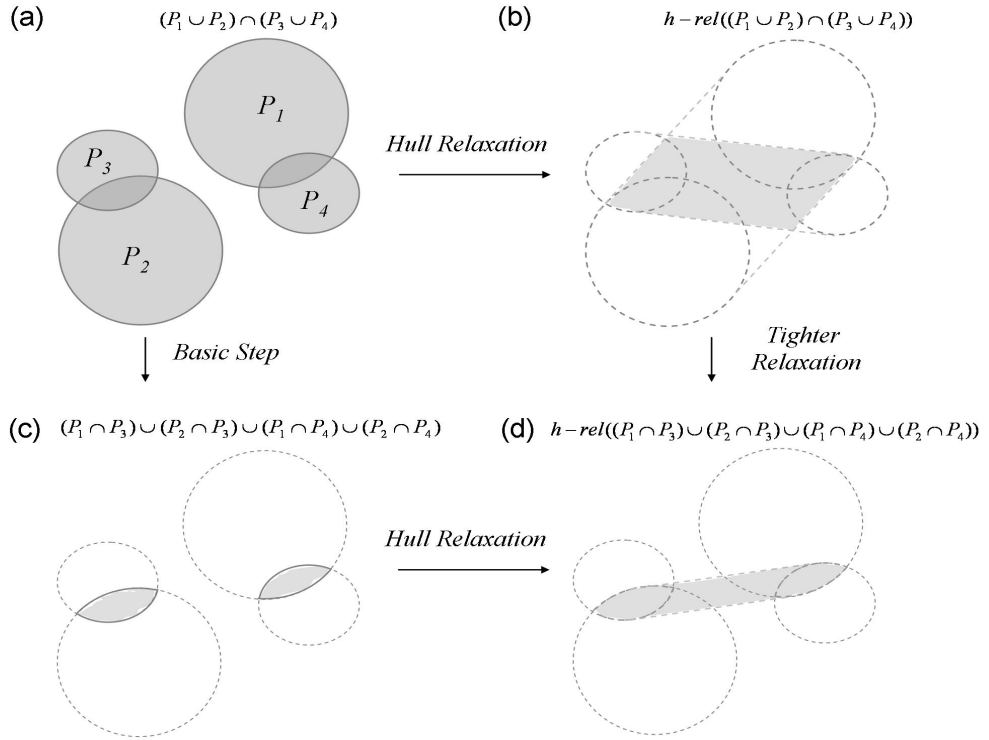
(a) $(P_1 \cup P_2) \cap (P_3 \cup P_4)$
(b) $h-rel((P_1 \cup P_2) \cap (P_3 \cup P_4))$

P_1
P_3
P_4
P_2

Hull Relaxation

Basic Step

Tighter Relaxation

(c) $(P_1 \cap P_3) \cup (P_2 \cap P_3) \cup (P_1 \cap P_4) \cup (P_2 \cap P_4)$
(d) $h-rel((P_1 \cap P_3) \cup (P_2 \cap P_3) \cup (P_1 \cap P_4) \cup (P_2 \cap P_4))$

Hull Relaxation

Figure 2: Example to illustrate Theorem 3.1

GDP_{NC} GDP_{CR0} GDP_{CRi} $h\text{-}rel\ GDP_{CRi}$

Step 1: Replace nonconvex functions with over and underestimators

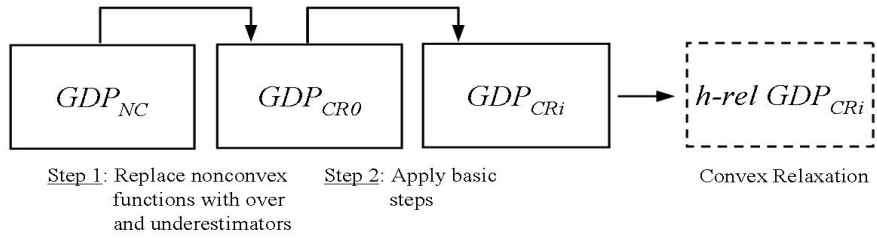Step 2: Apply basic steps

Convex Relaxation

Figure 3: Framework to obtain relaxations for GDP_{NC}

Note that solving the reformulated NLP using $(DISJ_{rel})$ would lead to numerical difficulties since the perspective function $\lambda_i g^i(\nu^i/\lambda_i)$ is not differ-

entiable at $\lambda_i = 0$. In order to overcome the latter issue we propose to use the approximation given by Sawaya [12]. This reformulation yields an exact approximation at $\lambda_i = 0$ and $\lambda_i = 1$ for *any* value of ε in the interval (0,1), and the feasibility and convexity of the approximating problem are maintained:

$$\lambda_i g_i(\nu^i/\lambda_i) \approx ((1 - \varepsilon)\lambda_i + \varepsilon)g_i(\nu^i/((1 - \varepsilon)\lambda_i + \varepsilon)) - \varepsilon g_i(0)(1 - \lambda_i) \qquad (3.1)$$

Note that this approximation assumes that $g_i(x)$ is defined at $x = 0$ and that the inequality $\lambda_i x^{lo} \leq \nu^i \leq \lambda_i x^{up}$ is enforced.

4. Rules to implement basic steps on disjunctive convex programs

In order to make good use of the hierarchy of relaxations described in the previous section, one important aspect is to understand what 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 [10], 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_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 $clconv(S_1 \cap S_2)$ $= clconv(S_1) \cap clconv(S_2)$.*

Proposition 4.2. *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 $clconv(S \cap H) = clconv(S) \cap clconv(H)$*

A new rule developed in [11] consists in the inclusion of the 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 disjunctive set. Note that this rule, different from the previous ones, has a particular effect when the disjunctive set is convex but nonlinear. However, one question that arises is whether this relaxation is still valid for the nonconvex case.

The following proposition, which we state without proof, aims at tackling this question.

Proposition 4.3. *Given GDP_{NC} defined as $Z = \{\min\limits_{x \in F} f(x)\}$ where F is a disjunctive set and $f(x)$ a nonconvex function, $GDP_{NC'}$ defined as $Z = \{\min\limits_{(x,\alpha) \in D} \alpha\}$ such that $D := \{(x, \alpha) \in R^{n+1} | \alpha \geq f(x), x \in F\}$. Then GDP_{NC} is equivalent to $GDP_{NC'}$ in the sense that the feasible regions of both problems projected onto the x space and the optimal value for the objective function are the same.*

Since $GDP_{NC'}$ is equivalent to GDP_{NC} the hierarchy of relaxations of $GDP_{NC'}$ is valid for GDP_{NC}. Figure 4 illustrates Proposition 4.3.
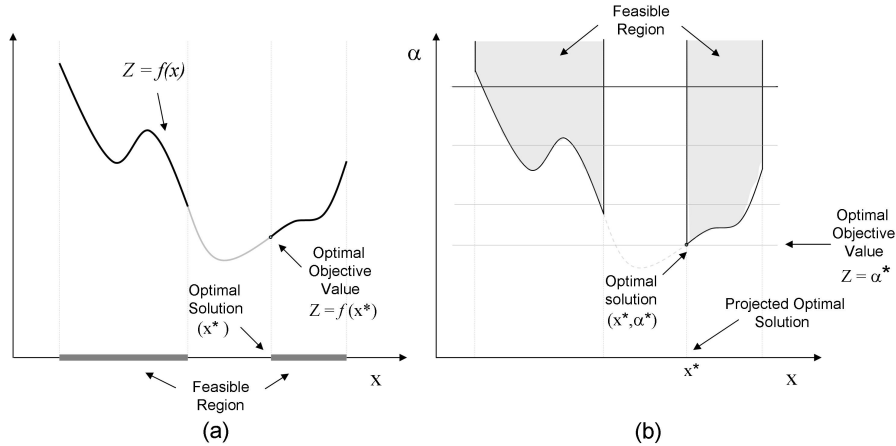

Figure 4: Equivalence between (a) GDP_{NC} and (b) $GDP_{NC'}$

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. As discussed in [10], 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. To illustrate this, let us consider the disjunctive set given by $F_1 = S_1 \cap S_2$

where $S_1 = P_1 \cup P_2$ and $S_2 = P_3$ are a proper and improper disjunction respectively. Clearly, the application of a basic step between S_1 and S_2 leads to the following disjunctive set $F_2 = S_{12} = (P_1 \cap P_3) \cup (P_2 \cap P_3)$ with no more convex sets than in F_1.

4.1. Illustrative Example

This example aims at illustrating the impact on the relaxation and, hence, in the lower bound prediction, when some of the proposed rules are used. Let us consider the following nonconvex GDP,

$$\min \quad Z = (x - 0.5)^2 + (y - 0.5)^2 - e^{0.8(x-0.7)} - e^{0.8(y-0.7)}$$

$$s.t.$$

$$\begin{bmatrix} Y \\ 0 \leq x \leq 0.4 \\ 0 \leq y \leq 0.4 \end{bmatrix} \vee \begin{bmatrix} \neg Y \\ 2.5 \leq x \leq 3 \\ 2.5 \leq y \leq 3 \end{bmatrix} \qquad (MGDP_{NC})$$

$$x \in R^1, y \in R^1, Y \in \{True, False\}$$

As it is shown in Figure 5, the objective function is clearly nonconvex, while the feasible region corresponds to a proper disjunction. Figure 6 displays the feasible region of the problem, which is described as two disjoint two-dimensional boxes. Note that the solution is obtained at the vertex of one of these boxes, i.e. (x=0.4, y=0.4), with an objective value of -1.55.

It is easy to show that the function $\hat{Z} = 0.4878(x^2 + y^2) - 1.3701(x + y) - 0.66$ is convex and always understimates the objective function of the $MGDP_{NC}$ (see Appendix E). Hence, the hull relaxation of the following convex GDP, as presented in the work of Lee and Grossmann [4], is a valid relaxation of $MGDP_{NC}$.

$$\min \quad Z = 0.4878(x^2 + y^2) - 1.3701(x + y) - 0.66$$

$$s.t.$$

$$\begin{bmatrix} Y \\ 0 \leq x \leq 0.4 \\ 0 \leq y \leq 0.4 \end{bmatrix} \vee \begin{bmatrix} \neg Y \\ 2.5 \leq x \leq 3 \\ 2.5 \leq y \leq 3 \end{bmatrix} \qquad (MGDP_{LG})$$

$$x \in R^1, y \in R^1, Y \in \{True, False\}$$
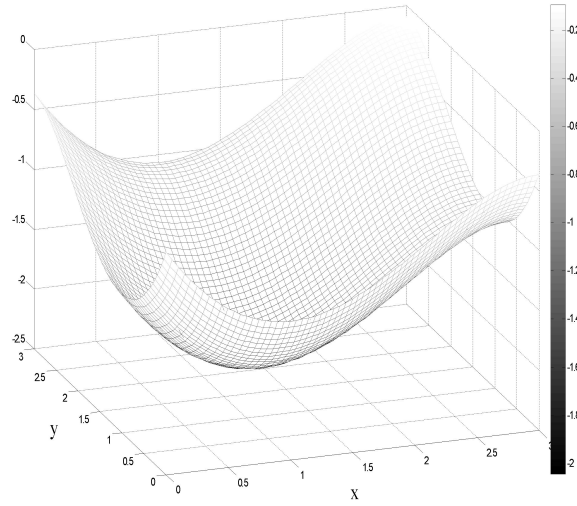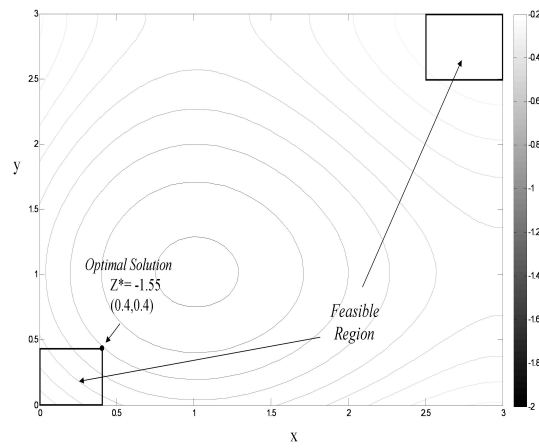
10

Figure 5: Objective function of problem $MGDP_{NC}$

Figure 6: Feasible region and optimal solution of problem $MGDP_{NC}$

11

The hull relaxation of $MGDP_{LG}$ reads,

$$\min \quad Z = 0.4878(x^2 + y^2) - 1.3701(x + y) - 0.66$$

$$s.t.$$

$$
\begin{aligned}
& x_1 + x_2 = x \\
& y_1 + y_2 = y \\
& 0 \le x_1 \le 0.4\lambda \qquad\qquad (MGDP_{LGRel}) \\
& 0 \le y_1 \le 0.4\lambda \\
& 2.5(1 - \lambda) \le x_2 \le 3(1 - \lambda) \\
& 2.5(1 - \lambda) \le y_2 \le 3(1 - \lambda)
\end{aligned}
$$

$$x_1, x_2, y_1, y_2, x, y \in R^1, \lambda \in [0, 1]$$

The solution of the above problem, i.e. (x=1.4,y=1.4) with an objective value of -2.58, is presented in Figure 7 and represents a lower bound on the global optimal solution.
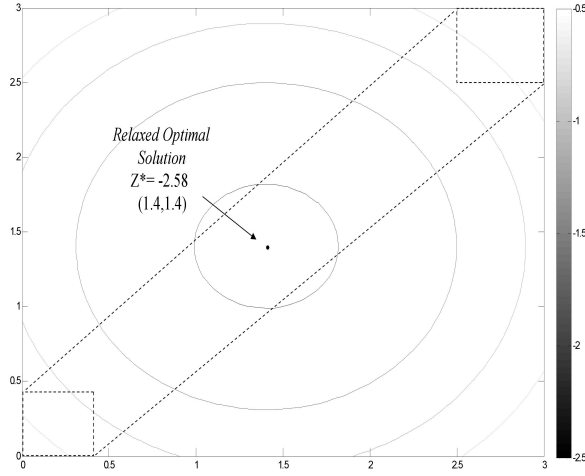

Figure 7: Optimal Solution of ($MGDP_{LGRel}$)

In the remaining of this section we show how by making use of the proposed framework we can significantly improve the predicted lower bound.

12

By Proposition 4.3 the following GDP is equivalent to $MGDP_{NC}$.

$$\min \quad Z = \alpha$$
$$s.t. \quad (x - 0.5)^2 + (y - 0.5)^2 - e^{0.8(x-0.7)} - e^{0.8(y-0.7)} \leq \alpha$$

$$\begin{bmatrix} Y \\ 0 \leq x \leq 0.4 \\ 0 \leq y \leq 0.4 \end{bmatrix} \vee \begin{bmatrix} \neg Y \\ 2.5 \leq x \leq 3 \\ 2.5 \leq y \leq 3 \end{bmatrix} \qquad (MGDP_{NC'})$$

$$\alpha, x \in R^1, y \in R^1, Y \in \{True, False\}$$

Clearly, by replacing $(x-0.5)^2 + (y-0.5)^2 - e^{0.8(x-0.7)} - e^{0.8(y-0.7)}$ with the convex function $0.4878(x^2 + y^2) - 1.3701(x + y) - 0.66$ we obtain the convex GDP,

$$\min \quad Z = \alpha$$
$$s.t. \qquad\qquad\qquad\qquad\qquad\qquad\qquad (MGDP_{step1})$$

$$0.4878(x^2 + y^2) - 1.3701(x + y) - 0.66 \leq \alpha$$

$$\begin{bmatrix} Y \\ 0 \leq x \leq 0.4 \\ 0 \leq y \leq 0.4 \end{bmatrix} \vee \begin{bmatrix} \neg Y \\ 2.5 \leq x \leq 3 \\ 2.5 \leq y \leq 3 \end{bmatrix}$$
$$\alpha, x, y \in R^1, Y \in \{True, False\}$$

By applying a basic step between the proper disjunction and the improper disjunction, we obtain the following equivalent GDP whose hull relaxation can be used to obtain bounds for the global optimum.

$$\min \quad Z = \alpha$$
$$s.t. \qquad\qquad\qquad\qquad\qquad\qquad\qquad (MGDP_{step2})$$

$$\begin{bmatrix} Y \\ 0.4878(x^2 + y^2) - 1.3701(x + y) - 0.66 \leq \alpha \\ 0 \leq x \leq 0.4 \\ 0 \leq y \leq 0.4 \end{bmatrix} \vee$$

$$\vee \begin{bmatrix} \neg Y \\ 0.4878(x^2 + y^2) - 1.3701(x + y) - 0.66 \leq \alpha \\ 2.5 \leq x \leq 3 \\ 2.5 \leq y \leq 3 \end{bmatrix}$$

$$\alpha, x, y \in R^1, Y \in \{True, False\}$$

13

The hull relaxation reads,

$$\min \ Z = \alpha$$

$$s.t. \hspace{8cm} (MGDP_{NLP})$$

$$x_1 + x_2 = x$$
$$y_1 + y_2 = y$$
$$\alpha_1 + \alpha_2 = \alpha$$
$$0.4878\lambda((x_1/\lambda)^2 + (y_1/\lambda)^2) - 1.3701(x_1 + y_1) - 0.66\lambda \leq \alpha_1$$
$$0.4878(1-\lambda)((x_2/(1-\lambda))^2 + (y_2/(1-\lambda))^2) - 1.3701(x_2 + y_2) - 0.66(1-\lambda) \leq \alpha_2$$
$$0 \leq x_1 \leq 0.4\lambda$$
$$0 \leq y_1 \leq 0.4\lambda$$
$$2.5(1-\lambda) \leq x_2 \leq 3(1-\lambda)$$
$$2.5(1-\lambda) \leq y_2 \leq 3(1-\lambda)$$

$$\alpha_1, \alpha_2, \alpha, x_1, x_2, y_1, y_2, x, y \in R^1, \lambda \in [0,1]$$

By solving $(MGDP_{NLP})$, using (3.1) to approximate the corresponding perspective function, a lower bound on the objective fuction of -1.6 is obtained, which is a significant improvement over the lower bound of -2.58. Furthermore, this lower bound is very close to the global optimum. Figure 8 shows a qualitative schematic of the feasible region lifted in the α space and the projection of the solution onto the (x,y) space. Note that the projected solution lies on the vertex (x=0.4, y=0.4).

5. Examples

In this section we present three main case studies, namely, the global optimization of a process network, a reactor network and a heat exchanger 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 GDP for which tight linear relaxations can be obtained [10].

5.1. Process network models with exponential functions

Consider the optimization of a process network with fixed charges in Figure 9 [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 from the market eliminating processes 2 and 3 if they are not
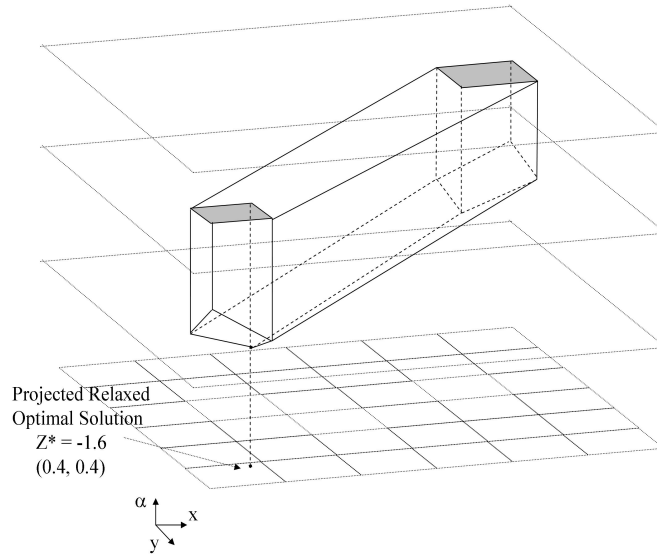
Figure 8: Optimal Solution of ($MGDP_{NLP}$)

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 require that processes 2 and 3 cannot appear together in a feasible flowsheet.
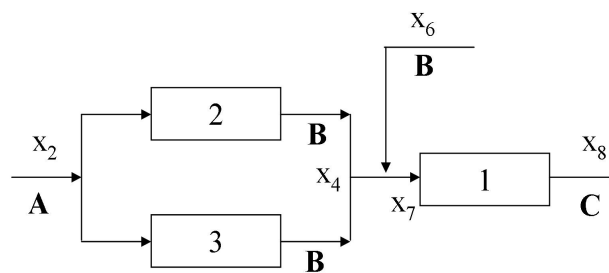

Figure 9: Example of a Process Network

15

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.

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:
$Y_2 \underline{\vee} Y_3$

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 recent developments in disjunctive convex programming theory to find strong relaxations for the convex disjunctive set. See Appendix G for the details in

16

the implementation.

5.2. Reactor networks with non-elementary kinetics described through posynomial functions

Consider the reactor network illustrated in Figure 10. 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.
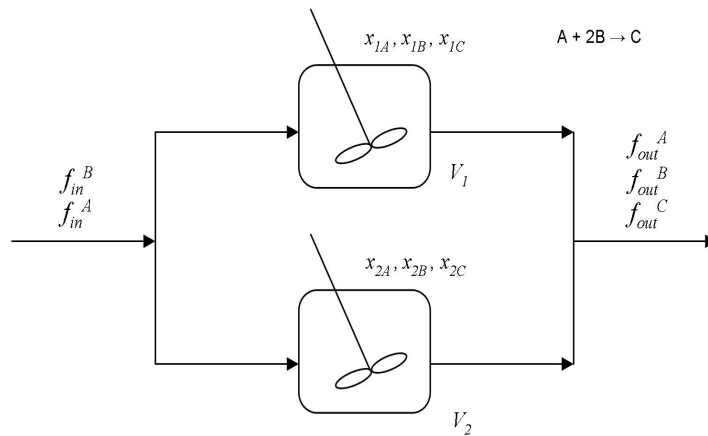

Figure 10: Example of a Reactor Network

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

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

17

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:
$$Y_1 \underline{\vee} Y_2$$

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 posynomial 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 recent developments in disjunctive convex programming theory to find strong relaxations for the convex disjunctive set. See Appendix G for the details in the implementation.

5.3. Heat exchanger network models with linear fractional terms

Consider the HEN illustrated in Figure 11 [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 1. 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

18

is the minimization of the cost of the network given by the heat exchanger areas and some fixed costs for the exchangers.

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

$$\min C = c_1 A_1 + c_2 A_2 + c_3 A_3 + c_4 A_4 + C_3 + C_4$$

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

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

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

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

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

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

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

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

Clearly, the problem is nonconvex where the nonconvexities arise from the linear 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 recent developments
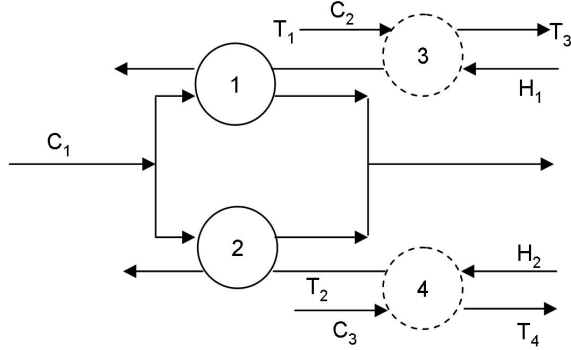
Figure 11: Example of a Heat Exchanger Network

Table 1: 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

in disjunctive convex programming theory to find strong relaxations for the convex disjunctive set. See Appendix G for the details in the implementation.

6. Global Optimization algorithm with improved relaxations

In this section we describe the global optimization framework from the work of Ruiz and Grossmann [10] that we will use to test the new proposed relaxations. The global optimization methodology of the GDP follows the well known spatial branch and bound method [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) If the objective function is nonlinear, introduce it in the set of constraints following Proposition 4.3. b) Relax the nonconvex terms using suitable convex under/over-estimators. This will lead to the convex GDP (GDP_{RLP0}). c) 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. Note that this technique converges in a finite number of iterations. See [14] where sufficient conditions for finite convergence of the spatial branch and bound are presented.

7. Numerical Results

In this section we compare the performance of the proposed method to find tighter relaxations for nonconvex GDPs with the traditional approach, i.e. the Lee and Grossmann relaxation [4], in two sets of problem instances. In the first set we consider a mathematical problem described in Appendix E with $|I| = 1,10$ and 100, which is in fact a generalization of the illustrative example in section 4.1. In the second set we consider instances of the process design problems described in section 5. In order to assess the perfomance of the method, we first consider the lower bound obtained at the root node as one of the main indicators of the strength of the relaxation that is 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 disjunctive convex program, which is the tightest bound attainable by the application of basic steps. All problems were solved using a Pentium(R) CPU 3.40GHz and 1GB of RAM.

In Table 2 we show the size and characteristics of the first set of instances. Clearly, Instance 1,2 and 3 are defined through 1, 10 and 100 disjunctions. In Table 3 we present the size and characteristics of the relaxations that arise by using the proposed approach with basic steps and the Lee and Grossmann approach that only uses the hull relaxation.

As is shown in Table 4, the lower bounds obtained using the proposed approach are tighter than the ones predicted using the Lee and Grossmann relaxation. For example, in Instance 1, the optimal value of the objective function is -1.55. By using the proposed relaxation we are able to obtain a bound of -1.6, whereas the Lee and Grossmann relaxation predicts a bound

Table 2: Size and characteristics of the mathematical examples formulated as GDP

Example	Cont. Vars.	Boolean Vars.	Logic Const.	Disj. Const.	Global. Const.
Instance 1	2	1	0	1	0
Instance 2	20	10	0	10	0
Instance 3	200	100	0	100	0

Table 3: Size of different reformulations for the mathematical example

Example	Lee and Grossmann			Proposed Approach		
	Bin	Con	Const	Bin	Con	Const
Instance 1	1	6	10	1	9	13
Instance 2	10	60	100	10	81	130
Instance 3	100	600	1000	100	801	1300

of -2.58. Moreover, the proposed bounds are identical to the best bounds we would obtain by solving the relaxation of the DNF form.

Table 4: Performance of the method to find tight relaxations for mathematical example

Example	Global Optimum	LB (Lee and Grossmann)	LB (Proposed Approach)	LB of DNF
Instance 1	-1.55	-2.58	-1.6	-1.6
Instance 2	-15.53	-25.84	-16.0	-16.0
Instance 3	-155.30	-258.40	-160.0	-160.0

A further analysis of the performance of the relaxations proposed was carried out considering their effect when used within a spatial branch and bound framework as described in section 6. The results for the first set of instances can be seen in Table 5. The number of nodes required to find the solution is significantly reduced when using the proposed relaxation, leading to a significant decrease in the computational times. This is due to the fact that a very strong lower bound can be predicted at the root node. Note that the solution of nonconvex problems will not be found in general at the root node. However, this extreme example is useful to illustrate some of the potential advantages of this method.

In Table 6 we show the size and characteristics of the second set of instances and in Table 7 we present the size and characteristics of the relaxations that arise by using the proposed approach and the Lee and Grossmann approach.

Table 5: Performance of the relaxation within a spatial branch and bound framework

		Lee and Grossmann		Proposed Approach	
Example	Opt.	Nodes	CPU-time(s)	Nodes	CPU-time(s)
Instance 1	-1.55	2	0.5	1	0.2
Instance 2	-15.53	286	20	1	0.3
Instance 3	-155.3	>12000	>1000	1	1.9

Table 6: Size and characteristics of the process design 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
HEN 2	18	2	2	2	21

Table 7: Size of different reformulations for the process design problems

	Lee and Grossmann			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
HEN 2	2	53	90	2	117	273

Clearly, from Table 8 we observe a significant improvement in the predicted lower bound in all instances. For instance, in Procnet 1 our approach predicts 16.01 as a lower bound whereas the approach based only on the hull relaxation is only able to obtain 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 DNF form. For example, HEN 1, reaches a lower bound of 48230, which is close to the maximum attainable 48531.

The results for the second set of instances when using a spatial branch and bound method as described in section 6 can be seen in Table 9. Note that the modest reduction in the number of nodes necessary to find to solution is due to the fact that the problems are small in size. Furthermore, the computational times are essentially the same due to the small size of the problems as well as the small size of the NLP subproblems as seen in Table 7. However, a clear indication of a tighter proposed relaxation is observed in the column "Bounding %" which refers to how much reduction in the upper

Table 8: 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
HEN 2	45460	35460	45281	45281

Table 9: Performance of the relaxation within a spatial branch and bound framework

Example	Opt.	Lee and Grossmann			Proposed Approach		
		Bounding %	Nds	T(s)	Bounding %	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
HEN 2	45460	7.5	3	14	97	1	23

and lower bounds of the variables can be predicted. More precisely, the column "Bounding %" refers to $100(1 - \sum_i \frac{x_i^{up*} - x_i^{lo*}}{|I|(x_i^{up} - x_i^{lo})})$ 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% with respect to the original bounds, whereas by using the Lee and Grossmann 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

In this paper we have proposed a framework to generate tight relaxations for the global optimization of nonconvex generalized disjunctive programs. We extended the method proposed in [10] for bilinear and concave disjunctive programming by allowing the use of nonlinear relaxations. Even though linear relaxations are in general desired due to the robustness and efficiency

of linear programming, in some cases using nonlinear relaxations can lead to significant improvements in the predicted lower bounds for the global optimum. Furthermore, for certain classes of functions (e.g. linear fractional), the convex envelopes are nonlinear. The use of nonlinear relaxations is now possible due to the latest developments in convex disjunctive programming presented in [11]. In that paper we presented a framework to generate a hierarchy of relaxations for convex generalized disjunctive programs. In this work we showed that this framework is also valid for nonconvex GDP if the nonconvex functions are first replaced by convex relaxations. One of the questions that arises is whether the nonconvexity of the objective function will have an effect in the validity of the relaxations. As presented in Proposition 4.3, the validity is guaranteed if the objective function is included in the set of constraints prior to the application of basic steps.

We have tested the method in several instances that showed significant improvements in the predicted lower bounds and bound contraction, which is a direct indication of tightening. Although we are confident that we have set the basis for an efficient algorithm, a major question that still remains to be answered is how to implement the strong proposed relaxations within a spatial branch and bound framework efficiently when dealing with large-scale GDP problems. Even though the very large increase in the size of the proposed extended formulation is avoided by using a set of rules when applying basic steps, a polynomial increase in the size of the reformulation cannot be prevented in general. This leads to a potentially higher computational effort that might not be compensated by the strength in the relaxation. Hence, for large-scale GDP problems other alternative mechanisms might be necessary. Since the strength of the relaxations heavily depends on the bounds of the variables, one approach we aim at exploring in the future is to consider the strong relaxations to calculate new bounds for the variables that belong to nonconvex terms. These stronger bounds should then be introduced in the Lee and Grossmann [4] relaxation, which is of lower dimension. Another approach that might find fruitful results is the consideration of the use of cutting planes inferred from the strong proposed relaxations. Even though an extensive analysis of different cut generation strategies should be performed, a good starting point would be to follow the idea of Lee and Grossmann [18] or Sawaya and Grossmann [12], where the cutting planes are generated by considering how far is the incumbent solution of the weak relaxation from the feasible region of the tighter relaxation.

Appendix A. Equivalence between disjunctive convex programs and convex generalized disjunctive programs [12] [11]

Any convex generalized disjunctive program can be represented as a disjunctive convex program. The transformation that allows this, which is equivalent to the one proposed by Sawaya and Grossmann [12] 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 $\underset{i \in D_k}{\vee} Y_{ik}, k \in K$ and $\Omega(Y) = True$ into algebraic equations $\underset{i \in D_k}{\sum} \lambda_{ik} = 1, k \in K$ and $H\lambda \geq h$, respectively. This yields the following equivalent disjunctive model:

$$
\begin{aligned}
\min \quad & Z = f(x) + \sum_{k \in K} c_k \\
s.t. \quad & g^l(x) \leq 0 \ l \in L \\
& \underset{i \in D_k}{\vee} \begin{bmatrix} \lambda_{ik} = 1 \\ r_{ik}^j(x) \leq 0 \ j \in J_{ik} \\ c_k = \gamma_{ik} \end{bmatrix} \qquad k \in K \qquad (CGDP) \\
& H\lambda \geq h \\
& x^{lo} \leq x \leq x^{up} \\
& x \in R^n,\ c_k \in R^1,\ \lambda_{ik} \in [0,1]
\end{aligned}
$$

Appendix B. Basic Steps

In this paper we define Regular Form (RF) as the form represented by the intersection of the union of convex sets. Hence, the regular form is:

$$F = \bigcap_{k \in K} S_k$$

where for $k \in K$, $S_k = \underset{i \in D_k}{\bigcup} P_i$ and P_i a convex set for $i \in D_k$.

The following theorem, as stated in [11], defines an operation that takes a disjunctive set to an equivalent disjunctive set with less number of conjuncts.

Theorem 2.1 [11] *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:

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

Appendix C. 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, nonlinear relaxations can have a great impact in reducing the time required 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. $y_i = x_i^2 \quad \forall i$

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

Solving this problem with linear relaxations using BARON [13], it takes more than 1000 seconds. On the other hand, by using the relaxation $y_i \geq x_i^2, \forall i$, which leads to a nonlinear convex inequality, the solution can be found at the root node in 1 second. This example then shows the importance of exploiting nonlinear convex relaxations.

Appendix D. Typical nonlinear relaxation of nonconvex constraints

Appendix D.1. Fractional Terms
The convex hull of a set $NC = \{(x, y, f) | f = x/y, x^{lo} \leq x \leq x^{up} \, y^{lo} \leq y \leq y^{up}\}$ is given in Tawarmalani and Sahinidis [14].

$f_p \geq \frac{x^{lo}(x^{lo}y_p - x(y^{lo} + y^{up}) + x^{up}(y^{lo} + y^{up} - y_p))}{(x^{up} - x^{lo})(y^{lo}y^{up})}$

$(f - f_p)(x^{up} - x^{lo})(y^{up} - y^{lo})^2 \geq x^{up}(x - x^{lo})^2$

$$y^{lo}(x^{up} - x) \leq y_p(x^{up} - x^{lo}) \leq y_p^{up}(x^{up} - x)$$

$$y^{lo}(x - x^{lo}) \leq (y - y_p)(x^{up} - x^{lo}) \leq y_p^{up}(x - x^{lo})$$

$$f - f_p \geq 0$$

Appendix D.2. Convex functions in equality constraints

The convex hull of a set $NC = \{x, g(x)|g(x) = 0, x^{lo} \leq x \leq x^{up}, g(x) \; convex\}$ 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 D.3. Posynomials

Several strategies have been proposed in literature to tackle the relaxation of posynomials [7] [8]. Here we present the method proposed by Han-Lin et al. (2008). 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 D.4. Trigonometric Functions

Convex underestimators for the function $f(x) = \alpha sin(x + s)$, $x^{lo} \leq x \leq x^{up}$ have been studied by Caratzoulas and Floudas (2005) [9], leading to nonlinear 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 E. Numerical Example

Appendix E.1. Convex Relaxation

In this section we show that the function $\hat{Z} = 0.4878(x^2 + y^2) - 1.3701(x + y) - 0.66$ is a valid convex relaxation of $Z = (x - 0.5)^2 + (y - 0.5)^2 - e^{0.8(x-0.7)} - e^{0.8(y-0.7)}$ in the interval $0 \leq x \leq 3$, $0 \leq y \leq 3$. Namely, $\hat{Z}$ is convex and it always understimates Z in the given interval.

Let us define the functions $Z_1 = (x - 0.5)^2 - e^{0.8(x-0.7)}$, $Z_2 = (y - 0.5)^2 - e^{0.8(y-0.7)}$, $\hat{Z}_1 = 0.4878x^2 - 1.3701x - 0.33$ and $\hat{Z}_2 = 0.4878y^2 - 1.3701y - 0.33$. From Figure E.12,

$$\hat{Z}_1 \leq Z_1$$
$$\hat{Z}_2 \leq Z_2$$

Then,

$$\hat{Z} = \hat{Z}_1 + \hat{Z}_2 \leq Z_1 + Z_2 = Z$$

Since the Hessian of $\hat{Z}$ $\left(i.e. \begin{bmatrix} 0.4878 & 0 \\ 0 & 0.4878 \end{bmatrix} \right)$ is clearly possitive definite, we can conclude that $\hat{Z}$ is convex.

Appendix E.2. Large-scale instances

The illustrative example in section 4 is used to generate instances of non-convex GDPs with larger number of variables and constraints. The main structure of these problems is presented below,
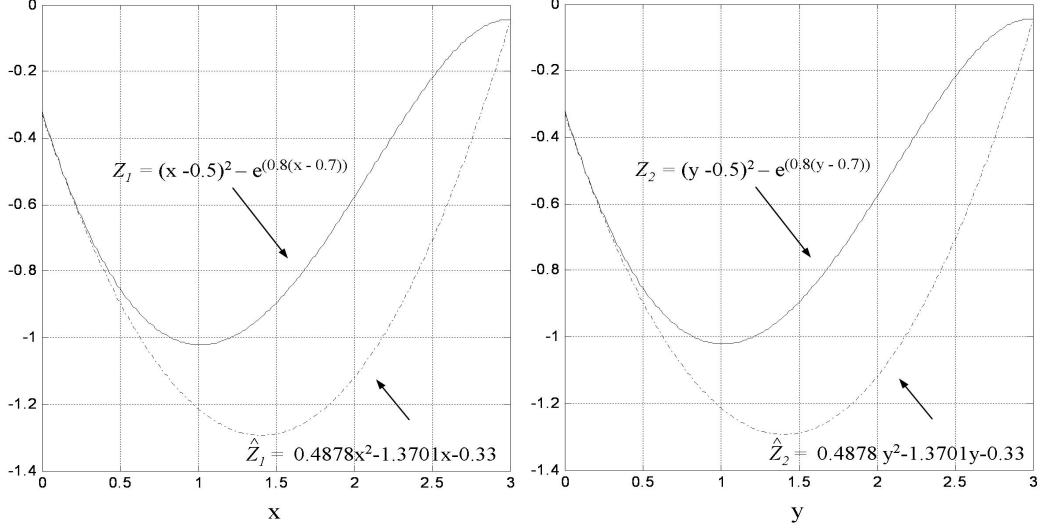
Figure E.12: Convex relaxations of Z_1 and Z_2

$$\min \quad Z = \sum_i (x_i - 0.5)^2 + (y_i - 0.5)^2 - e^{0.8(x_i - 0.7)} - e^{0.8(y_i - 0.7)}$$

$$s.t.$$

$$
\begin{bmatrix}
Y_i \\
0 \le x_i \le 0.4 \\
0 \le y_i \le 0.4
\end{bmatrix}
\vee
\begin{bmatrix}
\neg Y_i \\
2.5 \le x_i \le 3 \\
2.5 \le y_i \le 3
\end{bmatrix}
\qquad (MGDP_{NCI})
$$

$$x_i \in R^1, y_i \in R^1, Y_i \in \{True, False\}$$

where i = 1,2,...I.

Appendix F. Method Implementation

In this section we present a detailed description of the method applied to the Examples in section 5.

Appendix F.1. Process network models with exponential functions

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 $g(x) \le 0$. The proof is trivial considering that $g(x) \le 0$ is nothing

31

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 bound 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 disjunctive convex program is obtained:

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

$$x_2 \geq \beta$$

$$
\begin{bmatrix}
Y_2 \\
\Lambda_2(x_4 - x_4^{lo}) + e^{x_4^{lo}/\alpha} - 1 \geq x_2 \\
e^{x_4/\alpha_{42}} - 1 \leq x_2 \\
x_2^{lo2} \leq x_2 \leq x_2^{up2} \\
c = \delta_{42}x_4 + \delta_{22}x_2 + \gamma_2
\end{bmatrix}
\vee
\begin{bmatrix}
Y_3 \\
\Lambda_3(x_4 - x_4^{lo}) + e^{x_4^{lo}/\alpha} - 1 \geq x_2 \\
e^{x_4/\alpha_{43}} - 1 \leq x_2 \\
x_2^{lo3} \leq x_2 \leq x_2^{up3} \\
c = \delta_{43}x_4 + \delta_{23}x_2 + \gamma_3
\end{bmatrix}
$$

$$Y_2 \underline{\vee} Y_3$$

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

Note that for ease of notation we have expressed the formulation in reduced form by appropriate variable substitutions for x_7 and x_8, where
$\Theta = \delta_7 + \theta_7 - \theta_8\xi + \delta_8\xi$, $\Gamma = \delta_7 + \theta_7 - \theta_8\xi + \delta_8\xi - \theta_6, \Lambda_2 = \dfrac{e^{x_4^{up}/\alpha_{42}} - e^{x_4^{lo}/\alpha_{42}}}{x_4^{up} - x_4^{lo}}$
and $\Lambda_3 = \dfrac{e^{x_4^{up}/\alpha_{43}} - e^{x_4^{lo}/\alpha_{43}}}{x_4^{up} - x_4^{lo}}$
STEP 2: Application of basic steps

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

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

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

$Y_2 \underline{\vee} Y_3$

$x_2, x_4, x_6 \geq 0$

By applying the convex hull of the disjunction as in $(DISJ_{rel})$, the final NLP reads as follows,

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

Note that to keep the notation clear we did not add the approximation of $\lambda_i g^i(x^i/\lambda_i)$ in (3.1).

STEP 1: Finding a convex GDP relaxation

Tight relaxations for posynomial terms have been studied by Bjoerk et. al [7] and Han-Li et al. [8] (see Appendix D). By following this technique, a convex relaxation for $r_C = x_A^a x_B^b$ can be obtained as follows:

$$r_C \geq y_A^{-a} y_B^{-b}$$

$$1 \geq \frac{x_A}{x_A^{up}} + x_A^{lo} y_A - \frac{x_A^{lo}}{x_A^{up}}$$

$$1 \geq \frac{x_B}{x_B^{up}} + x_B^{lo} y_B - \frac{x_B^{lo}}{x_B^{up}}$$

Clearly, for a and b positive $x_A^{upa} x_B^{upb}$ is an upper bound for r_C.

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

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

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

$$x_A + x_b \geq \beta$$

$$
\begin{bmatrix}
Y_1 \\
gen \geq V_1 k_1 y_A^{-a1} y_B^{-b1} \\
gen \leq V_1 k_1 x_A^{upa1} x_B^{upb1} \\
x_A^{lo1} \leq x_A \leq x_A^{up1} \\
x_B^{lo1} \leq x_B \leq x_B^{up1} \\
1 \geq \frac{x_A}{x_A^{up}} + x_A^{lo} y_A - \frac{x_A^{lo}}{x_A^{up}} \\
1 \geq \frac{x_B}{x_B^{up}} + x_B^{lo} y_B - \frac{x_B^{lo}}{x_B^{up}} \\
c = \gamma_1
\end{bmatrix}
\vee
\begin{bmatrix}
Y_2 \\
gen \geq V_2 k_2 y_A^{-a2} y_B^{-b2} \\
gen \leq V_1 k_2 x_A^{upa2} x_B^{upb2} \\
x_A^{lo2} \leq x_A \leq x_A^{up2} \\
x_B^{lo2} \leq x_B \leq x_B^{up2} \\
1 \geq \frac{x_A}{x_A^{up}} + x_A^{lo} y_A - \frac{x_A^{lo}}{x_A^{up}} \\
1 \geq \frac{x_B}{x_B^{up}} + x_B^{lo} y_B - \frac{x_B^{lo}}{x_B^{up}} \\
c = \gamma_2
\end{bmatrix}
$$

$$Y_1 \underline{\vee} Y_2$$

$$x_A^{lo} \leq x_A \leq x_A^{up}, (x_A^{up})^{-1} \leq y_A \leq (x_A^{lo})^{-1}$$

$$x_B^{lo} \leq x_B \leq x_B^{up}, (x_B^{up})^{-1} \leq y_B \leq (x_B^{lo})^{-1}$$

STEP 2: Application of basic steps By introducing the global constraints inside the disjunctions, we obtain a new disjunctive set whose hull relaxation is tighter, leading to a tighter relaxation for the nonconvex problem.

$$\min Z = -C_C gen + c \qquad\qquad (RXN_{step2})$$

$$
\begin{bmatrix}
Y_1 \\
x_A + x_B \geq \beta \\
gen \geq V_1 k_1 y_A^{-a1} y_B^{-b1} \\
gen \leq V_1 k_1 x_A^{upa1} x_B^{upb1} \\
x_A^{lo1} \leq x_A \leq x_A^{up1} \\
x_B^{lo1} \leq x_B \leq x_B^{up1} \\
1 \geq \dfrac{x_A}{x_A^{up}} + x_A^{lo} y_A - \dfrac{x_A^{lo}}{x_A^{up}} \\
1 \geq \dfrac{x_B}{x_B^{up}} + x_B^{lo} y_B - \dfrac{x_B^{lo}}{x_B^{up}} \\
c = \gamma_1
\end{bmatrix}
\vee
\begin{bmatrix}
Y_2 \\
x_A + x_B \geq \beta \\
gen \geq V_2 k_2 y_A^{-a2} y_B^{-b2} \\
gen \leq V_1 k_2 x_A^{upa2} x_B^{upb2} \\
x_A^{lo2} \leq x_A \leq x_A^{up2} \\
x_B^{lo2} \leq x_B \leq x_B^{up2} \\
1 \geq \dfrac{x_A}{x_A^{up}} + x_A^{lo} y_A - \dfrac{x_A^{lo}}{x_A^{up}} \\
1 \geq \dfrac{x_B}{x_B^{up}} + x_B^{lo} y_B - \dfrac{x_B^{lo}}{x_B^{up}} \\
c = \gamma_2
\end{bmatrix}
$$

$$Y_1 \veebar Y_2$$
$$x_A^{lo} \leq x_A \leq x_A^{up}$$
$$x_B^{lo} \leq x_B \leq x_B^{up}$$

And the NLP relaxation reads:

$$
\begin{aligned}
\min \quad & Z = -C_C gen + \gamma_1 \lambda_1 + \gamma_2 \lambda_2 \qquad\qquad (RXN_{NLP}) \\
s.t. \quad & x_A = x_{1A} + x_{2A} \\
& x_B = x_{1B} + x_{2B} \\
& y_A = y_{1A} + y_{2A} \\
& y_B = y_{1B} + y_{2B} \\
& gen = gen_1 + gen_2 \\
& x_{1A} + x_{1B} \geq \beta \lambda_1 \\
& \lambda_1 \left(\frac{gen_1}{\lambda_1} - V_1 k_1 \left(\frac{y_{1A}}{\lambda_1} \right)^{-a1} \left(\frac{y_{1B}}{\lambda_1} \right)^{-b1} \right) \geq 0 \\
& \lambda_1 \left(\frac{gen_1}{\lambda_1} - V_1 k_1 x_A^{upa1} x_B^{upb1} \right) \leq 0 \\
& x_A^{lo1} \lambda_1 \leq x_{2A} \leq x_{1A}^{up1} \lambda_1 \\
& x_B^{lo1} \lambda_1 \leq x_{2B} \leq x_{1B}^{up1} \lambda_1
\end{aligned}
$$

$$\lambda_1(1 - \frac{x_{1A}}{\lambda_1 x_A^{up}} - x_A^{lo}\frac{y_A}{\lambda_1} + \frac{x_A^{lo}}{x_A^{up}}) \geq 0$$

$$\lambda_1(1 - \frac{x_{1B}}{\lambda_1 x_B^{up}} - x_B^{lo}(\frac{y_{2B}}{\lambda_1}) + \frac{x_B^{lo}}{x_B^{up}}) \geq 0$$

$$x_{2A} + x_{2B} \geq \beta\lambda_2$$

$$\lambda_2(\frac{gen_2}{\lambda_2} - V_2 k_2(\frac{y_{2A}}{\lambda_2})^{-a2}(\frac{y_{2B}}{\lambda_2})^{-b2}) \geq 0$$

$$\lambda_2(\frac{gen_2}{\lambda_2} - V_2 k_2 x_A^{upa2} x_B^{upb2}) \leq 0$$

$$x_A^{lo2}\lambda_2 \leq x_{2A} \leq x_{2A}^{up2}\lambda_2$$

$$x_B^{lo2}\lambda_2 \leq x_{2B} \leq x_{2B}^{up2}\lambda_2$$

$$\lambda_2(1 - \frac{x_{2A}}{\lambda_2 x_A^{up}} - x_A^{lo}\frac{y_A}{\lambda_2} + \frac{x_A^{lo}}{x_A^{up}}) \geq 0$$

$$\lambda_2(1 - \frac{x_{2B}}{\lambda_2 x_B^{up}} - x_B^{lo}(\frac{y_{2B}}{\lambda_2}) + \frac{x_B^{lo}}{x_B^{up}}) \geq 0$$

$$\lambda_1 + \lambda_2 = 1$$

$$x_A^{lo}\lambda_1 \leq x_{1A} \leq x_A^{lo}\lambda_1$$

$$x_B^{lo}\lambda_1 \leq x_{1B} \leq x_B^{lo}\lambda_1$$

$$x_A^{lo}\lambda_2 \leq x_{2A} \leq x_A^{lo}\lambda_2$$

$$x_B^{lo}\lambda_2 \leq x_{2B} \leq x_B^{lo}\lambda_2$$

$$x_A^{-up}\lambda_1 \leq y_{1A} \leq x_A^{-lo}\lambda_1$$

$$x_B^{-up}\lambda_1 \leq y_{1B} \leq x_B^{-lo}\lambda_1$$

$$x_A^{-up}\lambda_2 \leq y_{2A} \leq x_A^{-lo}\lambda_2$$

$$x_B^{-up}\lambda_2 \leq y_{2B} \leq x_B^{-lo}\lambda_2$$

$$V_1 k_1 x_A^{loa1} x_B^{lob1}\lambda_1 \leq gen_1 \leq V_1 k_1 x_A^{upa1} x_B^{upb1}\lambda_1$$

$$V_2 k_2 x_A^{loa2} x_B^{lob2}\lambda_2 \leq gen_2 \leq V_2 k_2 x_A^{upa2} x_B^{upb2}\lambda_2$$

Appendix F.3. Heat exchanger network models with linear fractional terms

STEP 1: Finding a convex GDP relaxation

Finding tight relaxations for fractional terms has been a challenge for a few years until a thorough analysis was given by Tawarmalani and Sahinidis [14] (see Appendix D). In that 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:

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

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

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

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

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

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

$$\min C = c_1 A_1 + c_2 A_2 + c_3 A_3 + c_4 A_4 + C_3 + C_4 \qquad (HEN_{step1})$$
s.t.

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

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

$$
\begin{bmatrix}
Y_3 \\
z_{p3} \geq \frac{Q_3^L(Q_3^L \Delta T_{p3} - Q_3(\Delta T_3^L + \Delta T_3^U) + Q_3^U(\Delta T_3^L + \Delta T_3^U - \Delta T_{p3}))}{(Q_3^U - Q_3^L)(\Delta T_3^L \Delta T_3^U)} \\
(z_3 - z_{p3})(Q_3^U - Q_3^L)(\Delta T_3^U - \Delta T_3^L)^2 \geq Q_3^U(Q_3 - Q_3^L)^2 \\
\Delta T_3^L(Q_3^U - Q_3) \leq \Delta T_{p3}(Q_3^U - Q_3^L) \leq \Delta T_3^U(Q_3^U - Q_3) \\
\Delta T_3^L(Q_3 - Q_3^L) \leq (\Delta T_3 - \Delta T_{p3})(Q_3^U - Q_3^L) \leq \Delta T_3^U(Q_3 - Q_3^L) \\
z_3 - z_{p3}, z_{p3} \geq 0 \\
c_3 = \gamma_3
\end{bmatrix}
\vee
\begin{bmatrix}
\neg Y_3 \\
A_3 = 0 \\
c_3 = 0
\end{bmatrix}
$$

$$
\begin{bmatrix}
Y_4 \\
z_{p4} \geq \frac{Q_4^L(Q_4^L \Delta T_{p4} - Q_4(\Delta T_4^L + \Delta T_4^U) + Q_4^U(\Delta T_4^L + \Delta T_4^U - \Delta T_{p4}))}{(Q_4^U - Q_4^L)(\Delta T_4^L \Delta T_4^U)} \\
(z_4 - z_{p4})(Q_4^U - Q_4^L)(\Delta T_4^U - \Delta T_4^L)^2 \geq Q_4^U(Q_4 - Q_4^L)^2 \\
\Delta T_4^L(Q_4^U - Q_4) \leq \Delta T_{p4}(Q_4^U - Q_4^L) \leq \Delta T_4^U(Q_4^U - Q_4) \\
\Delta T_4^L(Q_4 - Q_4^L) \leq (\Delta T_4 - \Delta T_{p4})(Q_4^U - Q_4^L) \leq \Delta T_4^U(Q_4 - Q_4^L) \\
z_4 - z_{p4}, z_{p4} \geq 0 \\
c_4 = \gamma_4
\end{bmatrix}
\vee
\begin{bmatrix}
\neg Y_4 \\
A_4 = 0 \\
c_4 = 0
\end{bmatrix}
$$

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

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

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

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

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

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 the sake of simplicity in the presentation, we will skip the explicit representation of the GDP after the application of basic steps and the final NLP relaxation.

[1] BALAS E., *Disjunctive Programming*, Annals of Discrete Mathematics, **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*, Industrial and Engineering Chemistry Research, **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] Bjoerk K.J., Lindberg P.O.,Westerlund T., *Some convexifications in global optimization of problems containing signomial terms*, Computers and Chemical Engineering, **27**, 669-679, 2003.

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

[9] Caratzoulas S., Floudas C.A., *A trigonometric convex underestimator for the base functions in Fourier space*, J. Optim. Theory Appl., **124**, 339–362, 2005.

[10] 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.

[11] Ruiz J.P. and Grossmann I.E., *A hierarchy of relaxations for nonlinear convex generalized disjunctive programming*, European Journal of Operational Research, **218**:1, 38-47, 2012.

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

[13] Sahinidis N.V., *BARON: A General Purpose Global Optimization Software Package*, Journal of Global Optimization, **8**:2, 201-205, 1996.

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

[15] Floudas, C. A., Deterministic global optimization: Theory methods and applications., Dordrecht, The Netherlands: Kluwer Academic Publishers, 2000.

[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*, Journal of Global Optimization, **14**:3, 217-249, 1999.

[18] LEE S. AND GROSSMANN I.E., *New Algorithms for Nonlinear Generalized Disjunctive Programming*, Computers and Chemical Engineering, **24**, 2125-2141, 2000.

[19] STUBBS R. AND MEHROTRA,S. , *A Branch-and-Cut Method for 0-1 Mixed Convex Programming*, Math Programming, **86**:3, 515-532, 1999.

[20] GROSSMANN I.E., *Review of Non-Linear Mixed Integer and Disjunctive Programming Techiques for Process Systems Engineering*, Optimization and Engineering, **3**, 227-252, 2002.