A cutting plane method for solving linear generalized disjunctive programming problems

Nicolas W. Sawaya, Ignacio E. Grossmann *

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

Abstract

Raman and Grossmann [Raman, R., & Grossmann, I.E. (1994). Modeling and computational techniques for logic based integer programming. *Computers and Chemical Engineering*, *18*(7), 563–578] and Lee and Grossmann [Lee, S., & Grossmann, I.E. (2000). New algorithms for nonlinear generalized disjunctive programming. *Computers and Chemical Engineering*, *24*, 2125–2141] have developed a reformulation of Generalized Disjunctive Programming (GDP) problems that is based on determining the convex hull of each disjunction. Although the relaxation of the reformulated problem using this method will often produce a significantly tighter lower bound when compared with the traditional big-M reformulation, the limitation of this method is that the representation of the convex hull of every disjunction requires the introduction of new disaggregated variables and additional constraints that can greatly increase the size of the problem. In order to circumvent this issue, a cutting plane method that can be applied to linear GDP problems is proposed in this paper. The method relies on converting the GDP problem into an equivalent big-M reformulation that is successively strengthened by cuts generated from an LP or QP separation problem. The sequence of problems is repeatedly solved, either until the optimal integer solution is found, or else until there is no improvement within a specified tolerance, in which case one switches to a branch and bound method. The strip-packing, retrofit planning and zero-wait job-shop scheduling problems are presented as examples to illustrate the performance of the proposed cutting plane method.

1. Introduction

The most commonly used model in discrete/continuous optimization corresponds to a Mixed Integer Non Linear Program (MINLP). More recently, however, Generalized Disjunctive Programming (GDP), which is a generalization of disjunctive programming (Balas, 1998), has been proposed by Raman and Grossmann (1994) as an alternative model to the MINLP problem (Grossmann, 2002; Tawarmalani & Sahinidis, 2002). While the MINLP model is based entirely on algebraic equations and inequalities,

the GDP model allows a combination of algebraic and logical equations through disjunctions and logic propositions, which facilitates the representation of discrete decisions. Furthermore, Lee and Grossmann (2000) have shown that any GDP model can be converted into an equivalent MINLP reformulation. Currently, there are several algorithms and different approaches in the literature to tackle GDP problems (Grossmann, 2002), though some perform inefficiently under certain conditions and for certain classes of problems. We are thus interested in developing novel algorithms and solution methods aimed at solving both linear and non-linear GDP problems more efficiently, although we restrict ourselves exclusively to the linear case in this paper.

* Corresponding author. Tel.: +1 41 22683642; fax: +1 41 22687139.
E-mail address: grossmann@cmu.edu (I.E. Grossmann).

Lee and Grossmann (2000) have proposed a reformulation to solve convex non-linear GDP problems with multiple disjunctions based on determining the convex hull of each disjunction. Although the feasible region of the reformulated problem does not correspond to the true convex hull of the problem, we nonetheless have termed that reformulation in this paper as the "convex hull" reformulation for the sake of convenience. There exists other GDP to MINLP reformulations of which the traditional big-M reformulation is the most common. Grossmann and Lee (2003) have shown that the feasible region of the relaxation resulting from the convex hull reformulation projected onto the space of the big-M reformulation is always as tight as, or tighter than that of the big-M reformulation. The tightness of the relaxed feasible region, which is usually reflected in the lower bound of the problem (for minimization), is an important criterion when solving the original Mixed Integer Program, as tighter relaxed feasible regions reduce the search space of the solution algorithm. However, the representation of the convex hull requires the introduction of new disaggregated variables and additional constraints that can greatly increase the size of the problem, thus limiting the effectiveness of the method.

In order to circumvent the aforementioned problem, we present in this paper a cutting plane method that exploits the potentially tighter convex hull relaxed feasible region without the additional constraints and variables. This method can be applied to linear GDP problems that correspond to MIP problems, or else to master problems that are used in the solution of non-linear GDP problems (Turkay & Grossmann, 1996). We present the strip-packing, retrofit planning and zero-wait job-shop scheduling problems to illustrate the computational performance of the proposed method in solving these problems and compare all results obtained to those using the convex hull and big-M reformulations.

2. Background

Consider the linear generalized disjunctive programming problem (LGDP), which is based on the work of Raman and Grossmann (1994) and is an extension of the work of Balas (1998):

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

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

$$\bigvee_{\forall j \in J_k} \begin{bmatrix} Y_{jk} \\ A_{jk} x \leq a_{jk} \\ c_k = \gamma_{jk} \end{bmatrix} \qquad \forall k \in K$$

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

$$0 \leq x \leq U, c_k \in \mathbb{R}_+^1, Y_{jk} \in \{\text{True, False}\} \quad \forall j \in J_k, \forall k \in K$$
$$\text{(LGDP)}$$

Here, x is a vector of continuous variables bounded by a vector of upper bounds U, Y_{jk} are Boolean variables, $c_k \in \mathbb{R}_+^1$ are continuous variables that represent the cost associated with each disjunction and γ_{jk} are fixed charges. A disjunction $k \in K$ is composed of several disjuncts $j \in J_k$, each containing a set of linear equations and/or inequalities ($A_{jk} x \leq a_{jk}$) representing the constraints of the problem, connected together by the logical OR operator ($\vee$) that enforces the contents of only one disjunct. Discrete decisions are represented by the Boolean variables Y_{jk} in terms of disjunctions $k \in K$ and logic propositions $\Omega(Y)$ that are assumed to be expressed in Conjunctive Normal Form (CNF). Thus, only the constraints inside disjunct $j \in J_k$, where Y_{jk} is true, are enforced; otherwise, the corresponding constraints are not enforced. Finally, $Bx \leq b$ are common constraints that must hold regardless of the discrete decisions that are selected.

The linear GDP problem (LGDP) can be reformulated as a Mixed Integer Program (MIP) in different ways, including the two most common alternatives termed big-M (BM) and convex hull reformulations (CH). In order to obtain the big-M reformulation, problem LGDP is transformed into an MIP by replacing the Boolean variables Y_{jk} by binary variables y_{jk} and using big-M constraints. The logic constraints $\Omega(Y)$ are converted into linear inequalities (Williams, 1985), which leads to the following reformulation (Raman & Grossmann, 1994):

$$\text{Min } Z = \sum_{\forall k \in K} \sum_{\forall j \in J_k} \gamma_{jk} y_{jk} + d^{\mathrm{T}} x$$

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

$$A_{jk} x - a_{jk} \leq M_{jk}(1 - y_{jk}) \qquad \forall j \in J_k, \forall k \in K$$

$$\sum_{\forall j \in J_k} y_{jk} = 1 \qquad \forall k \in K$$

$$Dy \leq d$$

$$x \in \mathbb{R}_+^n, y_{jk} \in \{0, 1\} \qquad \forall j \in J_k, \forall k \in K$$
$$\text{(BM)}$$

Here, M_{jk} are the "big-M" parameters that render the jth system of inequalities in the kth disjunction redundant when $y_{jk} = 0$ (i.e. $Y_{jk} = \text{False}$). The inequalities $Dy \leq d$ can be systematically derived from their logical CNF form $\Omega(Y)$ as discussed by Williams (1985), Raman and Grossmann (1994), and Hooker (2000).

In order to obtain the convex hull reformulation (CH), problem LGDP is transformed into an MIP by replacing the Boolean variables Y_{jk} by binary variables y_{jk} and disaggregating the continuous variables $x \in \mathbb{R}_+^n$ into new variables $\nu \in \mathbb{R}_+^n$. Using the convex hull constraints for each disjunction (Balas, 1998; Raman & Grossmann, 1994), this leads to the following reformulation (Raman &
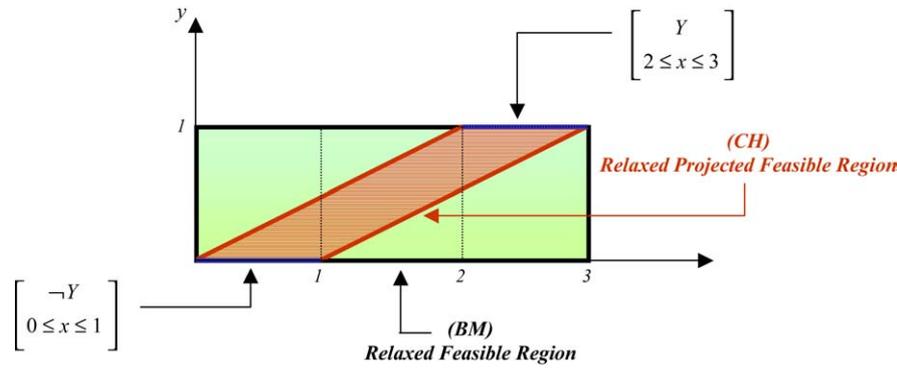
Fig. 1. Comparison between (BM) and (CH) relaxed feasible regions.

Grossmann, 1994):

$$\text{Min } Z = \sum_{\forall k \in K} \sum_{\forall j \in J_k} \gamma_{jk} y_{jk} + d^T x$$

s.t. $Bx \le b$

$$A_{jk} v_{jk} \le a_{jk} y_{jk} \qquad \forall j \in J_k, \forall k \in K$$

$$x = \sum_{\forall j \in J_k} v_{jk} \qquad \forall k \in K$$

$$v_{jk} \le y_{jk} U_{jk} \qquad \forall j \in J_k, \forall k \in K$$

$$\sum_{\forall j \in J_k} y_{jk} = 1 \qquad \forall k \in K$$

$$Dy \le d$$

$$x, v \in \mathbb{R}_+^n, y_{jk} \in \{0, 1\} \qquad \forall j \in J_k, \forall k \in K \tag{CH}$$

The new variables $v \in \mathbb{R}_+^n$ in (CH) are the disaggregated variables, while the parameters U_{jk} serve as their upper bounds. The latter are usually chosen so as to match the upper bounds on the continuous variables $x \in \mathbb{R}_+^n$. Note that $(y_{jk} = 0) \Rightarrow (v_{jk} = 0)$, and thus the jth system of inequalities in the kth disjunction is redundant.

In comparing the reformulations in (BM) and (CH), the following trade-offs can be observed. On the one hand, the relaxed feasible region of the (CH) reformulation is at least as tight, if not tighter, than that of the (BM) reformulation (see Fig. 1). This is reflected in the lower bounds of the aforementioned reformulations, where the lower bound of the relaxation of (CH) is equal to or greater than the lower bound of the relaxation of problem (BM), as has been proven by Grossmann and Lee (2003). The tightness of the feasible region, and by extension, the quality of the lower bound, affects the number of nodes being examined within the framework of a B&B algorithm. Thus a tighter feasible region, and by extension, a tighter lower bound, leads to a reduction in the search space of a particular problem, which usually translates into faster solution times. The example in Fig. 1 illustrates in the (x, y) space the convex hull and big-M relaxations of the

single disjunction $[0 \le x \le 1] \vee [2 \le x \le 3]$ that is expressed in terms of the Boolean variable Y to indicate whether the first term ($Y = $ False) or second term ($Y = $ True) applies. In this case, it is clear that the convex hull relaxation is tighter than the big-M relaxation, despite the optimal value of M having been chosen ($M = 3$).

On the other hand, the size of the (CH) reformulation is considerably larger than the size of the (BM) reformulation, which leads to an increase in solution time required per iteration at every node, and furthermore, to an increase in the number of total iterations per node. Hence, in general, it is very difficult to determine a priori when a given reformulation will be more effective than the other one in solving the problem (Vecchietti, Lee, & Grossmann, 2003). Therefore, it would appear that a desirable objective is to develop a method for generating cutting planes from the (CH) relaxation in order to strengthen the looser but smaller (BM) reformulation. In this fashion, one takes advantage of the tighter (CH) reformulation without incurring an increase in the number of variables and a significant increase in the number of constraints in the problem. Such an idea is proposed in the following section.

3. Cutting plane method

The basic idea of the proposed cutting plane method consists in solving a sequence of relaxed big-M MILPs with cutting planes that are successively generated from the convex hull relaxation projected onto the (x, y) space. More specifically, the cutting planes are determined by solving an LP (or QP) separation problem, whose feasible region corresponds to that of the convex hull relaxed reformulation. The separation problem has, as an objective, to find a point within the convex hull relaxed feasible region "closest" to the optimal solution point yielded by the relaxed big-M MILP. In essence, the objective of the separation problem consists of finding a cutting plane that corresponds to the most violated constraint of the convex hull that is projected onto the space of the original variables of the big-M MILP. A repeating sequence of relaxed big-M MILPs (with cuts up to that point) and LP (or QP) separation problems yielding new cuts is iteratively solved until the optimal solution for the original MILP

is found or until there is no improvement within a specified tolerance ε, in which case one switches to a B&B method for solving the resulting big-M MILP with all the cutting planes that have been generated.

It is interesting to note that the proposed cuts, though related to a certain extent to some of the work done on lift-and-project cutting planes originally developed by Balas, Ceria, and Cornuejols (1993), are different in important ways from the latter. The major distinction lies in the derivation and generation of our cuts, which are crucially based on a GDP formulation, as opposed to a 0–1 MIP formulation, as in the case of lift-and-project cuts. Furthermore, if we were to contextualize their work within a GDP framework, lift-and-project cuts would be generated through a sequential convexification procedure that would obtain by taking the convex hull of one disjunction at a time, as opposed to our proposed method which generates cuts based on a formulation that considers the intersection of the convex hulls of every disjunction. Thus, the elementary closure resulting from our cuts is not equivalent to that of lift-and-project cuts, since the latter's closure corresponds to the true convex hull of the original problem, as opposed to our case, where the elementary closure corresponds to the feasible region arising from the aforementioned strengthened formulation that considers all disjunctions simultaneously. We plan on comparing our cuts, both theoretically and computationally, to other cuts present in the literature, including lift-and-project cuts, mixed-integer Gomory cuts and mixed-integer rounding cuts, amongst others, in a subsequent paper.

3.1. Separation problem

The general form of the separation problem (SEP) is as follows (see Stubbs & Mehrotra, 1999; Vecchietti et al., 2003):

$$\text{Min } \phi(z) = ||z - z^{\text{bm}}||$$

s.t. $Bx \leq b$

$$A_{jk}v_{jk} \leq a_{jk}y_{jk} \qquad \forall j \in J_k, \forall k \in K$$

$$x = \sum_{\forall j \in J_k} v_{jk} \qquad \forall k \in K$$

$$v_{jk} \leq y_{jk}U_{jk} \qquad \forall j \in J_k, \forall k \in K \qquad \text{(SEP)}$$

$$\sum_{\forall j \in J_k} y_{jk} = 1 \qquad \forall k \in K$$

$$Dy \leq d$$

$$x, v \in \mathbb{R}_+^n, z \equiv [x, y] \in \mathbb{R}_+^n$$
$$\times \mathbb{R}_+^{\sum_{\forall k \in K} |J_k|}, 0 \leq y_{jk} \leq 1 \quad \forall j \in J_k, \forall k \in K$$

The objective function $\phi(z)$ corresponds to determining the point $z \in \mathbb{R}_+^n \times \mathbb{R}_+^{\sum_{\forall k \in K} |J_k|}$ within the convex hull relaxation "closest" to the point z^{bm}, which corresponds to the optimal solution of the relaxed big-M MILP. In order to represent distance in the function $\phi(z)$,

the 1-norm, $\phi(z) = ||z - z^{\text{bm}}||_1 \equiv \sum_i |z_i - z_i^{\text{bm}}|$,

the Euclidean norm, $\phi(z) = ||z - z^{\text{bm}}||_2 \equiv [(z - z^{\text{bm}})^{\text{T}} (z - z^{\text{bm}})]^{1/2}$,
or the infinity norm, $\phi(z) = ||z - z^{\text{bm}}||_\infty \equiv \max_i |z_i - z_i^{\text{bm}}|$ can be used. If either the 1-norm or the ∞-norm is used, then the separation problem is an LP; otherwise, using the Euclidean norm yields a QP.

3.2. Derivation of cutting planes

In this section, we present the derivation of the proposed cutting planes that are obtained from the separation problem (SEP). The proofs of the propositions presented can be found in Appendix A.

The first proposition formalizes the observation in Fig. 1 that the feasible region of the separation problem (SEP), which corresponds to that of the convex hull reformulation, is contained within the feasible region of problem (BM).

Proposition 1. *Let (FR-SEP) be the feasible region of the separation problem (SEP) in the (z, v) space, and let (FRP-SEP) represent the projection of (FR-SEP) onto the z-space. Then, (FRP-SEP)$\subseteq$(FR-BM), where (FR-BM) represents the feasible region of (BM) in the z-space. Furthermore, (FRP-SEP) is a convex set.*

The second proposition provides the general form of the valid inequality that corresponds to the cutting plane that is determined from the separation problem (SEP).

Proposition 2. *Let z^{bm} be the optimal solution of (BM) and z^{sep} be an optimal solution to (SEP). If $z^{\text{bm}} \notin$ (FRP-SEP), then $\exists \xi$ such that $\xi^{\text{T}}(z - z^{\text{sep}}) \geq 0$ is a valid linear inequality in z that cuts away z^{bm}, and such that ξ is a subgradient of $\phi(z)$ at z^{sep}, where $\phi(z)$ corresponds to the objective function of (SEP).*

Using the example previously shown in Fig. 1, Fig. 2 demonstrates how the proposed cut $\xi^{\text{T}}(z - z^{\text{sep}}) \geq 0$ cuts away z^{bm} and slices off part of the feasible region of problem (BM), thus strengthening the formulation. Also, note that the cut generated in this case corresponds to a facet of the feasible region of problem (SEP).

The third proposition shows that the subgradient of a differentiable function at a specific point corresponds to the gradient of the function at that same point.

Proposition 3. *Let (FRP-SEP)$\subset$S, where S is a convex set. If $\phi : S \to \mathbb{R}$ is differentiable over its entire domain, then the collection of subgradients of ϕ at z^{sep} is the singleton set $\partial^{\text{sep}}\phi \equiv \{\xi^{\text{sep}}|\xi^{\text{sep}} = \nabla\phi(z^{\text{sep}})\}$, which corresponds to the gradient of ϕ at z^{sep}.*

The last two propositions provide the specific expressions for the subgradient ξ in the inequality $\xi^{\text{T}}(z - z^{\text{sep}}) \geq 0$ for the Euclidean and ∞-norms, respectively. It should be noted that
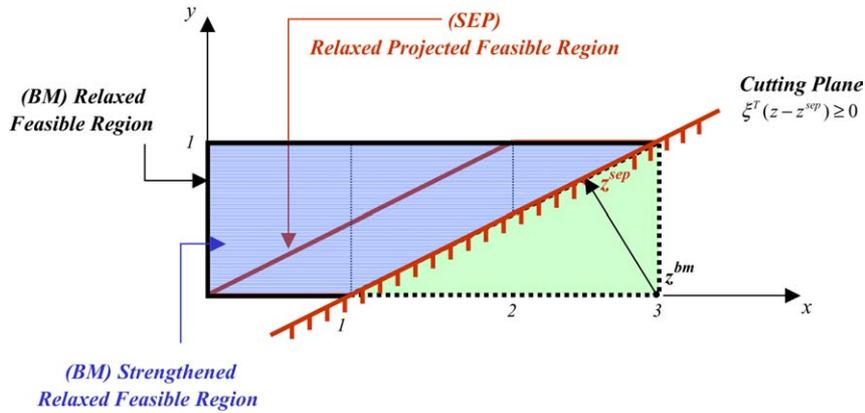
Fig. 2. Graphical representation of Proposition 2.

for the case of the 1-norm, the treatment is entirely similar as the ∞-norm.

Proposition 4. *Let (FRP-SEP)$\subset S$, where S is a convex set. If $\phi : S \rightarrow \mathbb{R}$ is defined as $\phi(z) = ||z - z^{bm}||_2^2$, then the collection of subgradients of ϕ at z^{sep} is the singleton set $\partial^{sep}\phi \equiv \{\xi^{sep}|\xi^{sep} = 2(z^{sep} - z^{bm})\}$.*

Proposition 5. *Let (FRP-SEP)$\subset S$, where S is a convex set. If $\phi : S \rightarrow \mathbb{R}$ is defined as $\phi(z) = ||z - z^{bm}||_\infty$, then the collection of subgradients of ϕ at z^{sep} is the set:*

$$\partial^{sep}\phi \equiv \{\xi^{sep}|\xi^{sep} = [\mu_+^{sep} - \mu_-^{sep}]\}$$

where μ_+^{sep} and μ_-^{sep} correspond to the optimal Lagrange multipliers of constraints (1) and (2) respectively, in the following problem (SEP2):

Min u

s.t. $u \geq z_i - z_i^{bm}$ $\forall i \in M$ (1)

$u \geq z_i^{bm} - z_i$ $\forall i \in M$ (2)

$R^1 z + R^2 v \leq r$

The cutting planes generated by the proposed method and based on Propositions 2, 4 and 5 can be used at the root node of the branch and bound tree in order to strengthen the corresponding relaxation of problem (BM). It is, of course, generally not obvious which of the different norms provides the deepest cut, as this is usually problem dependent. Furthermore, the depth of the cut will be affected, particularly in the cases of the 1-norm and the ∞-norm, by the selection of a specific set of (non-unique) optimal Lagrange multipliers, which is usually solver dependent. The following example provides a geometrical interpretation of the three different cuts when applied to the example in Fig. 1.

$$\text{Max } Z = x - (c_1 + c_2)$$

$$\text{s.t.} \begin{bmatrix} Y \\ 2 \leq x \leq 3 \\ c_1 = 1 \end{bmatrix} \lor \begin{bmatrix} \neg Y \\ 0 \leq x \leq 1 \\ c_2 = 0 \end{bmatrix}$$

In Fig. 3, the cut generated when the Euclidean norm or infinity norm is used corresponds to a facet of the convex
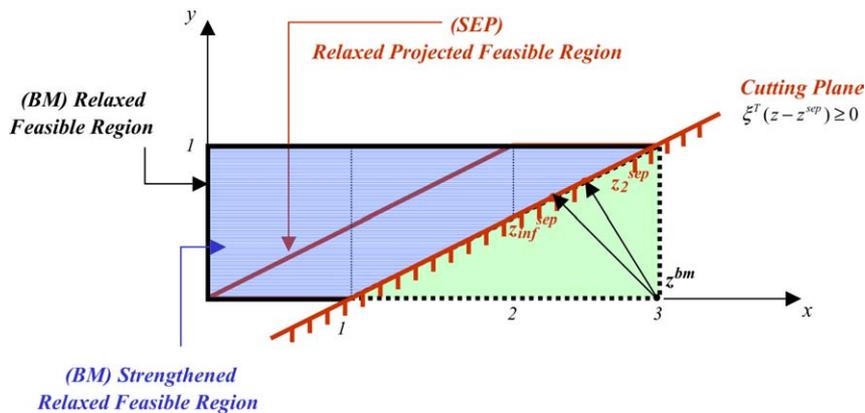
Fig. 3. Cutting plane generated when Euclidean or infinity norm are used in (SEP).
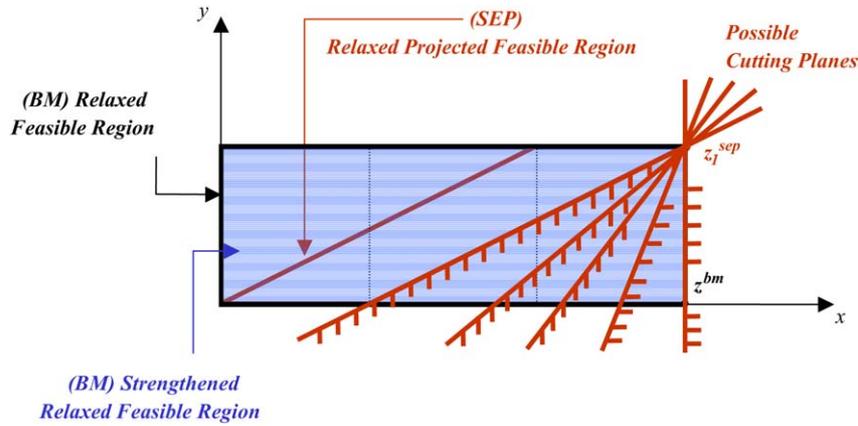
Fig. 4. Cutting plane(s) generated when 1-norm is used in (SEP).

hull of the disjunction. However, in Fig. 4, where the 1-norm is used, there is an infinite set of cutting planes that can be generated since the set of Lagrange multipliers corresponding to the optimal solution z^{sep} is not unique. Clearly then, the "quality" and depth of the cut generated from our procedure depends on the norm used in the separation problem. The difference in quality and depth of the cuts resulting from the use of different norms in the separation problem will be examined in more detail and rigor in a subsequent paper.

3.3. Cutting plane algorithm

Given the propositions of the previous section, the steps of the proposed algorithm are as follows:

(0) Specify a tolerance ε for the norm of the distance in the separation problem (SEP). Set $n = 0$, where n represents the iteration index.

(1) Solve the continuous relaxation of $(\text{BM})^n$ termed $(\text{RBM})^n$, which corresponds to the following problem. This yields the point $z^{\text{bm},n} \equiv [x, y]^{\text{bm},n}$.

$$\text{Min } Z = \sum_{\forall k \in K} \sum_{\forall j \in J_k} \gamma_{jk} y_{jk} + d^{\text{T}} x$$

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

$$A_{jk} x - a_{jk} \leq M_{jk}(1 - y_{jk}) \qquad \forall j \in J_k, \forall k \in K$$

$$\sum_{\forall j \in J_k} y_{jk} = 1 \qquad \forall k \in K$$

$$Dy \leq d$$

$$\xi^{l\text{T}}(z - z^{\text{sep},l}) \geq 0 \qquad l = 1, 2, \ldots n - 1$$

$$x \in \mathbb{R}^n_+, 0 \leq y_{jk} \leq 1 \qquad \forall j \in J_k, \forall k \in K$$

$$z \equiv [x, y] \in \mathbb{R}^n_+ \times \mathbb{R}^{\sum_{\forall k \in K} |J_k|}_+$$

$$(\text{RBM})^n$$

(2) Solve the separation problem $(\text{SEP})^n$, which corresponds to the following problem. This yields the point

$$z^{\text{sep},l} \equiv [x, y]^{\text{sep},l}.$$

$$\text{Min } \phi(z) = ||z - z^{\text{bm},n}||$$

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

$$A_{jk} v_{jk} \leq a_{jk} y_{jk} \qquad \forall j \in J_k, \forall k \in K$$

$$x = \sum_{\forall j \in J_k} v_{jk} \qquad \forall k \in K$$

$$v_{jk} \leq y_{jk} U_{jk} \qquad \forall j \in J_k, \forall k \in K$$

$$\sum_{\forall j \in J_k} y_{jk} = 1 \qquad \forall k \in K$$

$$Dy \leq d$$

$$x, v \in \mathbb{R}^n_+, z \equiv [x, y] \in \mathbb{R}^n_+$$
$$\times \mathbb{R}^{\sum_{\forall k \in K} |J_k|}_+, 0 \leq y_{jk} \leq 1 \quad \forall j \in J_k, \forall k \in K$$

$$(\text{SEP})^n$$

(a) If $\phi(z) = ||z - z^{\text{bm},n}|| \leq \varepsilon$, then stop, and proceed to the LP-based branch-and-bound solution of problem $(\text{BM})^n$ with all the added cutting planes.

(b) Else, generate cutting plane $\xi^{n\text{T}}(z - z^{\text{sep},n}) \geq 0$, and add to $(\text{RBM})^n$. Set $n = n + 1$, go to 1.

The effectiveness of the above algorithm is dependent on the trade-off between the amount of time spent on the cut generation procedure versus the amount of time "saved" in the B&B tree relative to either the big-M formulation without cuts, or to the convex hull formulation. In the former case, these savings usually result from a tightening of the relatively loose big-M feasible region, while in the latter case, these savings result from the smaller number of variables and constraints in the big-M plus cuts formulation. The cut generation procedure involves solving a sequence of separation problems $(\text{SEP})^n$ that are LPs or QPs, so it can be expected that it will take a reasonable amount of time to solve these problems in order to generate the proposed cuts. It is important to note that the feasible region of the strengthened big-M problem will be, at the very best, as tight as that of the convex hull formulation, and thus, intuitively at

least, lead us to believe that the cut generation procedure is finite. In other words, no further cuts will be generated either when the optimal point of (RBM)^n lies within the feasible region of (SEP)^n, or when the feasible region of the strengthened big-M problem is exactly equivalent to that of (SEP)^n. A rigorous proof of the previous claim will be performed in a subsequent paper. The question that remains, however, is whether these cuts are effective enough in slicing off parts of the big-M feasible region that are superfluous. This is examined in detail in the next section, and we apply the above algorithm to three different problems that highlight some of the major strengths and weaknesses of the proposed method.

4. Numerical results

In this section, we present the results of the proposed cutting plane algorithm on the strip-packing, retrofit planning and zero-wait job-shop scheduling problems. The strip-packing problem is an example within a class of problems suitably solved by the proposed method, while the last two problems serve to highlight an important characteristic regarding the usefulness of the method, notably the degree of tightness exhibited by the convex hull relaxation.

We present results for the strip-packing problem using the proposed method with all norms, although the discussion is mostly focused on results obtained using the *infinity norm* as the latter turned out to be the most efficient norm. This observation also holds true for the retrofit planning and zero-wait job-shop scheduling problems, thus, we only present and discuss results obtained using the *infinity norm*. All results obtained using the proposed cutting plane method are discussed and compared with those obtained using the aforementioned convex hull and big-M reformulations, where optimal values of the big-M parameters were used (i.e. equal to $\max_x(A_{jk}x - a_{jk})$).

All example problems were solved with GAMS (Brooke, Kendrick, Meeraus, & Raman, 1997) on a 2.8 GHz Pentium IV PC (512 MB of RAM). The CPLEX solver (v. 8.1) was used for the infinity norm for all three problems and for all comparisons between reformulations with all MIP options turned off and with default options turned on, while the CPLEX solver (v. 9.0) was used for the 1-norm and Euclidean norm. Note that the LP pre-solver was turned off during the cut-generation procedure for reasons of computational efficiency. Finally, the cuts generated are added only at the root node of the B&B tree.

4.1. Strip-packing problem

Cutting and packing problems belong to a well-known family of combinatorial NP-hard optimization problems that arise in numerous applications of computer science, industrial engineering, and operations management (Hifi, 1998). One important problem in this family is the strip-packing

problem, where a given set of small rectangles is packed into a strip of fixed width W but unknown length L. The aim is to minimize the length of the strip while fitting all rectangles without any overlap and without rotation. We propose the following general linear GDP model for problem (SP-GDP):

$$\text{Min} \quad lt \tag{1}$$

$$\text{s.t.} \quad lt \geq x_i + L_i \quad \forall i \in N \tag{2}$$

$$
\begin{bmatrix} Y_{ij}^1 \\ x_i + L_i \leq x_j \end{bmatrix} \vee
\begin{bmatrix} Y_{ij}^2 \\ x_j + L_j \leq x_i \end{bmatrix} \vee
\begin{bmatrix} Y_{ij}^3 \\ y_i - H_i \geq y_j \end{bmatrix}
$$
$$
\vee \begin{bmatrix} Y_{ij}^4 \\ y_j - H_j \geq y_i \end{bmatrix} \quad \forall i, j \in N, i < j \tag{3}
$$

$$x_i \leq UB_i - L_i \quad \forall i \in N \tag{4}$$

$$H_i \leq y_i \leq W \quad \forall i \in N \tag{5}$$

$lt, x_i, y_i \in \mathbb{R}_+^1, Y_{ij}^1, Y_{ij}^2, Y_{ij}^3, Y_{ij}^4 \in \{\text{True, False}\}$

$\forall i, j \in N, i < j$

The objective in this problem consists of minimizing the length of the strip lt (1) and (2) by representing every rectangle by its coordinates in the (x, y) space such that no overlap occurs between rectangles. Thus, every rectangle $i \in N$ has length L_i, height H_i, and coordinates (x_i, y_i), where the point of reference corresponds to the upper left corner of every rectangle. By constraining every pair of rectangles (i, j) where $(i, j \in N, i < j)$ such that no overlap occurs, we obtain a series of disjunctions with four disjuncts each, where each disjunct represents the position of rectangle i in relation to rectangle j (3). Note that the y-coordinate of every rectangle is bounded from above by the fixed width of the strip W (5), and that the upper bound UB_i, which in a best case scenario would correspond to the optimal value of lt, is obtained using a bottom-left rectangle-placing heuristic and serves as an upper bound for the x-coordinate of every rectangle (4).

We consider first a 12-rectangle instance of the strip-packing problem (SP-GDP) with the following ordered lengths and heights for every rectangle: (1, 10), (2, 9), (3, 8), (4, 4), (5, 5), (9, 6), (7, 7), (6, 3), (5, 2), (12, 1), (3, 1), (2, 3). The problem was transformed into an MIP model by using both the big-M and convex hull reformulations that are given in Appendix B. The problem sizes for both reformulations are listed in Table 1, and the graphical solution to this problem is presented in Fig. 5.

Table 1
Problem sizes for 12-rectangle strip-packing problem

	Total number of constraints	Total number of variables	Number of discrete variables
Convex hull	1663	1346	264
Big-M	343	290	264
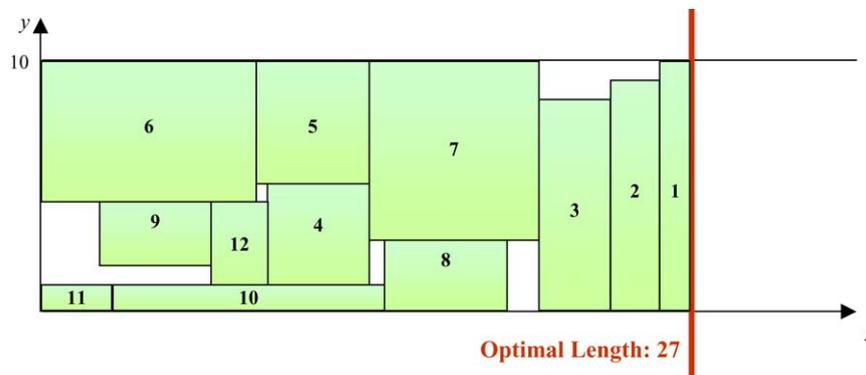
Fig. 5. Graphical solution for 12-rectangle strip-packing problem.

Table 2
Results for 12-rectangle strip-packing problem (∞-norm, MIP options off)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	12	27	55.55	682464	0	1286.39	530.52
Big-M	12	–	–	54244296	0	>10800	5022.36
Big-M + 40 cuts	12	–	–	41831856	2.44	>10800	3873.32
Big-M + 50 cuts	12	27	55.55	10289250	3.05	2986.33	3448.97
Big-M + 60 cuts	12	27	55.55	694596	3.66	191.95	3688.96
Big-M + 70 cuts	12	27	55.55	320535	4.27	97.61	3434.05
Big-M + 80 cuts	12	27	55.55	502727	4.88	154.23	3366.09
Big-M + 87 cuts	12	27	55.55	72677	5.31	27.51	3273.73

[a] Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP.

We also solved the problem using the cutting plane method with the infinity norm, and compared the resulting solutions to those from the convex hull and big-M reformulations in Tables 2 and 3. We first examine the results with all MIP algorithmic options turned off (see Table 2). This is done in order to better gauge the effect of the proposed cuts on solution time and number of nodes examined during the B&B procedure.

The optimal solution of the problem is 27. The lower bound obtained from the relaxation is equal to 12 for both (BM) and (CH) reformulations, but the problem was solved in 682 464 nodes using the (CH) reformulation, as opposed to the big-M reformulation, which failed to solve the prob-

lem after 54 244 296 nodes. This is due to the tighter relaxed feasible region of (CH) when compared to that of (BM), which results in substantial savings in computational time (1286.39 s versus >10 800 s). Note however that the LP at every node of the (CH) B&B tree is about 10 times more expensive to solve than that of the (BM) reformulation as seen by the amount of nodes computed per second for both reformulations (530.52 versus 5022.36). This is due to the larger number of variables and constraints present in the (CH) reformulation. After the addition of 50 cutting planes to the (BM) reformulation, we are able to solve the problem in less than the self-imposed limit of 3 h (2986.33 s) while examining 10 289 250 nodes in the B&B tree. Upon the successive ad-

Table 3
Results for 12-rectangle strip-packing problem (∞-norm, default options on)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	12	–	–	2887380	0	>10800	267.35
Big-M	12	27	55.55	73225	0	59.00	1241.10
Big-M + 40 cuts	12	27	55.55	13361	2.44	12.21	1368.25
Big-M + 50 cuts	12	27	55.55	9008	3.05	11.01	1131.65
Big-M + 60 cuts	12	27	55.55	20247	3.66	20.61	1194.51
Big-M + 70 cuts	12	27	55.55	11405	4.27	14.05	1166.15
Big-M + 80 cuts	12	27	55.55	10225	4.88	14.41	1072.92
Big-M + 87 cuts	12	27	55.55	6397	5.31	11.12	1101.03

[a] Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP.

Table 4
Results for 12-rectangle strip-packing problem (1-norm, MIP options off)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	12	27	55.55	682464	0	1286.39	530.52
Big-M	12	–	–	54244296	0	>10800	5022.36
Big-M + 50 cuts	12	–	–	17145216	6.0	>10800	1587.52
Big-M + 100 cuts	12	–	–	7373700	12.0	>10800	682.75
Big-M + 200 cuts	12	–	–	253800	22.0	>10800	23.5

[a] Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP.

dition of more cuts, the number of nodes examined is further reduced, which results in a further decrease in total solution time. Finally, with 87 cuts, the proposed cutting plane algorithm solves this problem in 72 677 nodes. Although the resulting strengthened MIP is still not as tight as the (CH) MIP, it has much fewer variables. This compromise is key to the success of the proposed algorithm and results in improved total computational times (27.51 s versus 1286.39 s). Note that the time required to generate the 87 cutting planes was only 5.31 s.

We now examine the results with default options turned on. This is done in order to demonstrate the effectiveness of the proposed cuts in aiding the branch-and-cut routine of a powerful MIP solver like CPLEX (see Table 3).

We see a noticeable improvement in the number of nodes examined and solution times upon the addition of the cuts. After 87 cuts, the problem was solved in 6397 nodes and 11.12 s compared to 73 225 nodes and 59.00 s for the big-M. However, CPLEX failed to solve the CH reformulation in less than 3 h, which is odd considering that one would expect an improvement in nodes examined and solution times when default options are turned on. This phenomenon could have been caused by many factors, although we believe that poor CPLEX-generated cuts are the most likely culprits. As more CPLEX cuts are generated, they tend to become shallower and to flatten out, and upon their addition to the matrix of the problem, create dependent rows and affect the conditioning number of the matrix thus resulting in numerical difficulties. This hypothesis will be investigated in future work. Nonetheless, the results demonstrate the effectiveness of the cutting plane algorithm when it is considered that CPLEX (with options turned on) may have difficulties in solving this problem when posed as a convex hull reformulated MIP.

We now briefly present and discuss the results when the 1-norm (see Table 4) and the Euclidean norm (see Table 5) are used. Note that while the use of the 1-norm in problem (SEP) still results in an LP, the use of the Euclidean norm results in a QP. Also, only results obtained with all MIP options turned off are presented.

Using the 1-norm or the Euclidean norm for the objective function in the separation problem does not yield good results as CPLEX failed to solve the problem to optimality. Furthermore, and in both cases, the cut generation routine was terminated after the self-imposed limit of 200 cuts without having reduced the objective function value to zero in the separation problem. The problem in both cases is that the cuts generated are weak and do not tighten the feasible region enough. Moreover, upon the addition of more cuts in the hope of strengthening the formulation and improving solution times, we observe the same phenomenon that occurred when we attempted to solve the convex hull formulation with default MIP options on. In other words, the addition of more of these "poor" cuts negatively affects the computational performance of the algorithm because of numerical difficulties. In light of these observations in this case and other cases, we will only report results using the infinity norm for the remainder of this paper.

Let us now consider a 21-rectangle instance of the strip-packing problem (SP-GDP) with the following ordered lengths and heights for every rectangle: (1, 5), (2, 2), (3, 2), (2, 7), (5, 1), (6, 6), (5, 10), (4, 3), (3, 2), (9, 5), (4, 2), (1, 1), (2, 3), (3, 1), (2, 6), (2, 2), (1, 2), (2, 1), (2, 1), (1, 1), (1, 1). This was the largest instance of the strip-packing problem that was solvable in less than 3 h. The problem sizes for the (BM) and (CH) reformulations are listed in Table 6, and we present the graphical solution to this problem in Fig. 6.

The results using the proposed cutting plane method are presented only with default MIP options turned on (see Table 7) as CPLEX failed to solve this problem when options were turned off.

Table 5
Results for 12-rectangle strip-packing problem (2-norm, MIP options off)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	12	27	55.55	682464	0	1286.39	530.52
Big-M	12	–	–	54244296	0	>10800	5022.36
Big-M + 50 cuts	12	–	–	15932676	5.4	>10800	1475.03
Big-M + 100 cuts	12	–	–	7604280	12.0	>10800	704.1
Big-M + 200 cuts	12	–	–	3482008	28.0	>10800	322.41

[a] Total solution time includes times for relaxed MIP(s) + QP(s) from separation problem + MIP.
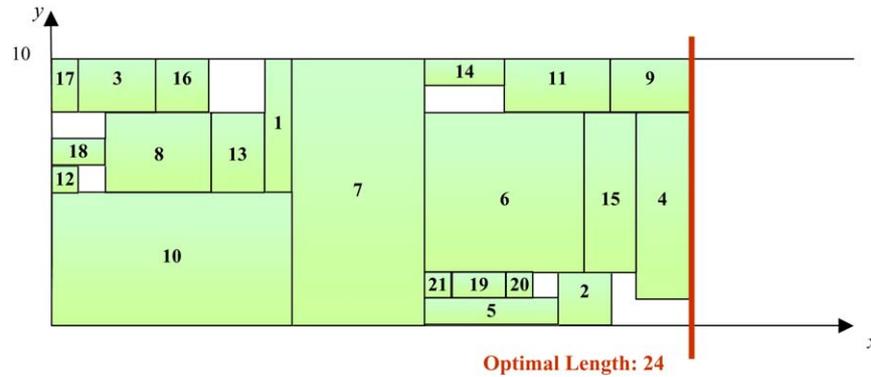
Fig. 6. Graphical solution for 21-rectangle strip-packing problem.

Table 6
Problem sizes for 21-rectangle strip-packing problem

	Total number of constraints	Total number of variables	Number of discrete variables
Convex hull	5272	4244	840
Big-M	1072	884	840

The optimal solution of the problem is 24. Although the lower bound obtained from the relaxation is equal to 9 for the (BM) and to 9.1786 for the (CH) reformulations, CPLEX failed to solve the latter after 968 652 nodes (same reasoning as previously) while solving the former in 1 416 137 nodes. Upon the addition of the cuts, we obtain noticeable improvements in the number of nodes examined and total solution time. The problem is solved in 32 185 nodes and 91.4 s upon the addition of 62 cutting planes compared to 1 416 137 nodes and 4093.39 s when no cuts are added. This again demonstrates the efficiency of the proposed cutting plane algorithm in solving different instances of the strip-packing problem.

4.2. Retrofit planning problem

The retrofit planning problem essentially consists in the redesign of existing plants (Jackson & Grossmann, 2002). Processes can be retrofitted to achieve goals such as increasing throughput, reducing energy consumption, improving yields and reducing waste generation. Work in retrofit design has been limited because of the difficulties in dealing with the many constraints of a pre-existing operation, such as layout,

available space, piping and operating conditions, and also because of the many modification possibilities which causes the problem to greatly grow in size. For a general review of retrofit issues, see work by Grossmann, Westerberg, and Biegler (1987). In this paper, we assume that an existing process network is given where each process can possibly be retrofitted for improvements such as higher yield, increased capacity, and reduced energy consumption. Given limited capital investments to make process improvements and cost estimations over a given time horizon, the problem consists of identifying those modifications that yield the highest economic improvement in terms of economic potential, which is defined as the income from product sales minus the cost of raw materials, energy and process modifications. We propose the following linear model for this problem (RP-GDP), which is a modification of work done by Jackson and Grossmann (2002):

$$\text{Min} \sum_{\forall t \in T} \sum_{\forall s \in S_{\text{prod}}} \text{PR}_s^t mf_s^t - \sum_{\forall t \in T} \sum_{\forall s \in S_{\text{raw}}} \text{PR}_s^t mf_s^t$$

$$- \sum_{\forall t \in T} \text{PRSTqst}^t - \sum_{\forall t \in T} \text{PRWTqwt}^t$$

$$- \sum_{\forall t \in T} \sum_{\forall p \in P} fc_P^t - \sum_{\forall t \in T} \text{ec}^t \qquad (6)$$

s.t. $\quad mf_s^t = f_s^t \text{MW}_s \quad \forall s \in S, \forall t \in T \qquad (7)$

$mf_s^t \geq \text{DEM}_s^t \quad \forall s \in S_{\text{prod}}, \forall t \in T \qquad (8)$

Table 7
Results for 21-rectangle strip-packing problem (default options on)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	9.1786	–	–	968652	0	>10800	89.69
Big-M	9	24	62.5	1416137	0	4093.39	345.95
Big-M + 20 cuts	9.1786	24	61.75	306029	3.74	917.79	334.80
Big-M + 40 cuts	9.1786	24	61.75	547828	7.48	1063.51	518.76
Big-M + 60 cuts	9.1786	24	61.75	28611	11.22	79.44	419.32
Big-M + 62 cuts	9.1786	24	61.75	32185	11.59	91.4	403.27

[a] Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP.

$$mf_s^t \leq \text{SUP}_s^t \quad \forall s \in S_{\text{raw}}, \forall t \in T \tag{9}$$

$$\sum_{\forall s \in S_{n_{\text{in}}}} mf_s^t = \sum_{\forall s \in S_{n_{\text{out}}}} mf_s^t \quad \forall n \in N, \forall t \in T \tag{10}$$

$$\sum_{\forall s \in S_{p_{\text{in}}}} mf_s^t = \sum_{\forall s \in S_{p_{\text{out}}}} mf_s^t + \text{unrct}_p^t \quad \forall p \in P, \forall t \in T \tag{11}$$

$$\bigvee_{\forall m \in M} \left[\begin{array}{c} Y_{pm}^t \\ f_s^t = f_{p_{lmt}}^t \left(\dfrac{\text{GMA}_s^t}{\text{GMA}_{p_{lmt}}} \right) \text{ETA}_{pm}^t \\ \sum_{\forall s \in S_{p_{\text{in}}}} mf_s^t \leq \text{CAP}_{pm}^t \end{array} \right]$$
$$\forall s \in S_{p_{\text{out}}}, \forall p \in P, \forall t \in T \tag{12}$$

$$\bigvee_{\forall m \in M} \left[\begin{array}{c} W_{pm}^t \\ fc_p^t = \text{FC}_{pm}^t \end{array} \right] \quad \forall p \in P, \forall t \in T \tag{13}$$

$$q_{sk}^t = mf_s^t \text{CP}_s (T_{s_{\text{out}_k}}^t - T_{s_{\text{in}_k}}^t) \quad \forall s \in S_{\text{cold}}, \forall k \in K, \forall t \in T \tag{14}$$

$$q_{sk}^t = mf_s^t \text{CP}_s (T_{s_{\text{in}_k}}^t - T_{s_{\text{out}_k}}^t) \quad \forall s \in S_{\text{hot}}, \forall k \in K, \forall t \in T \tag{15}$$

$$\left[\begin{array}{c} X_1^t \\ \text{qst}^t = \sum_{\forall k \in K} \sum_{\forall s \in S_{\text{cold}}} q_{sk}^t \\ \text{qwt}^t = \sum_{\forall k \in K} \sum_{\forall s \in S_{\text{hot}}} q_{sk}^t \end{array} \right]$$
$$\vee \left[\begin{array}{c} X_2^t \\ r_k^t - r_{k-1}^t - \text{qst}_k^t + \text{qwt}_k^t = \sum_{\forall s \in S_{\text{hot}}} q_{sk}^t - \sum_{\forall s \in S_{\text{cold}}} q_{sk}^t \\ \text{qst}^t = \sum_{\forall k \in K} \text{qst}_k^t \\ \text{qwt}^t = r_{|K|}^t + \sum_{\forall k \in K} \text{qwt}_k^t \end{array} \right]$$
$$\forall k \in K, \forall t \in T \tag{16}$$

$$\left[\begin{array}{c} V_1^t \\ \text{ec}^t = \text{EFC}_1^t \end{array} \right] \vee \left[\begin{array}{c} V_2^t \\ \text{ec}^t = \text{EFC}_2^t \end{array} \right] \quad \forall t \in T \tag{17}$$

$$\sum_{\forall p \in P} fc_p^t + \text{ec}^t + \sum_{\forall s \in S_{\text{raw}}} \text{PR}_s^t mf_s^t + \text{PRST} \text{qst}^t$$
$$+ \text{PRWT} \text{qwt}^t \leq \text{INV}^t \quad \forall t \in T \tag{18}$$

$$Y_{pm}^t \rightarrow \bigwedge_{\tau > t} Y_{pm}^\tau \quad \forall p \in P, \forall t \in T, \forall m \in M \backslash m_1 \tag{19}$$

$$W_{pm}^t \rightarrow \bigwedge_{\tau \neq t} W_{p1}^t \quad \forall p \in P, \forall t \in T, \forall m \in M \backslash m_1 \tag{20}$$

$$Y_{p1}^t \rightarrow W_{p1}^t \quad \forall p \in P, \forall t \in T \tag{21}$$

$$Y_{pm}^t \bigwedge_{\tau < t} \neg Y_{pm}^\tau \rightarrow W_{pm}^t \quad \forall p \in P, \forall t \in T, \forall m \in M \backslash m_1 \tag{22}$$

$$X_j^t \rightarrow \bigwedge_{\tau > t} X_j^\tau \quad \forall t \in T, \forall j \in J \backslash j_1 \tag{23}$$

$$V_j^t \rightarrow \bigwedge_{\tau \neq t} V_1^t \quad \forall t \in T, \forall j \in J \backslash j_1 \tag{24}$$

$$X_1^t \rightarrow V_1^t \quad \forall t \in T \tag{25}$$

$$X_j^t \bigwedge_{\tau < t} \neg X_j^\tau \rightarrow V_j^t \quad \forall t \in T, \forall j \in J \backslash j_1 \tag{26}$$

$$mf_s^t, f_s^t \in \mathbb{R}_+^1 \quad \forall s \in S, \forall t \in T$$

$$f_{lmt,p}^t, \text{unrct}_p^t, fc_p^t \in \mathbb{R}_+^1 \quad \forall p \in P, \forall t \in T$$

$$q_{sk}^t \in \mathbb{R}_+^1 \quad \forall s \in S, \forall k \in K, \forall t \in T$$

$$\text{qst}^t, \text{qwt}^t, \text{ec}^t \in \mathbb{R}_+^1 \quad \forall t \in T$$

$$\text{qst}_k^t, \text{qwt}_k^t, r_k^t \in \mathbb{R}_+^1 \quad \forall k \in K, \forall t \in T$$

$$Y_{pm}^t, W_{pm}^t \in \{\text{True}, \text{False}\} \quad \forall p \in P, \forall t \in T, \forall m \in M$$

$$X_j^t, V_j^t \in \{\text{True}, \text{False}\} \quad \forall j \in J, \forall t \in T$$

The objective function (6) includes revenues from sales, costs of raw material, utility costs, as well as capital costs fc_p^t and energy costs ec^t over time periods $t \in T$. Equation (7) represents an equivalence relation between mass and molar flow rates, equations (8) and (9) ensure that mass flow rates for products and raw materials are respectively bounded by demand and supply parameters, and equations (10) and (11) serve as mass balances around nodes $n \in N$ and processes $p \in P$, respectively. The first set of disjunctions (12) selects one of the operating modes for the retrofit project $m \in M$, for every process $p \in P$, in every time period $t \in T$, where projects m include modifying either nothing at all ($m_1 \in M$), process conversion ($m_2 \in M$), capacity ($m_3 \in M$) or both ($m_4 \in M$). The second set of disjunctions (13) enforces the cost of the aforementioned modifications, where capital costs are set to zero ($fc_p^t = 0$) if nothing is modified. Equations (14) and (15) serve as equivalence relations between energy and mass flow rate variables, while disjunction(s) (16) select the appropriate operating mode $X_j^t \, \forall j \in J$ so that X_1^t corresponds to no energy integration and X_2^t enforces the transshipment equations (Biegler, Grossmann, & Westerberg, 1997). Through Boolean variables V_j^t, the set of disjunctions (17) enforce the cost associated with energy reduction, where these costs are set to zero ($\text{ec}^t = 0$) if nothing is modified ($V_1^t = \text{True}$). Equation (18) limits the expenses for the retrofit project. Equations

(21) and (22), (25) and (26) are logical conditions that connect, respectively, disjunctions (12) to (13) and disjunctions (16) to (17) with each other, and equations (19) and (20), (23) and (24) impose logical conditions between disjuncts in every set of corresponding disjunctions. Essentially, these logical equations constrain the problem such that costs associated with conversion and/or capacity are enforced exactly once for every process $p \in P$ in every time period $t \in T$, and such that costs associated with energy reduction are enforced exactly once per time period $t \in T$.

We consider in Fig. 7 a 10 process instance of the retrofit planning problem (RP-GDP) that involves the production of products (G, H, I, J, K, L, M) from raw materials (A, B, C, D, E).

We use a 1-year planning horizon of four time periods each consisting of 3 months. Modifications for increased conversion and capacity only are considered, and black-box (input/output) models are used for each process. We do not include explicit data for this problem because of its size. The problem was transformed into an MIP model by using both the big-M and convex hull reformulations given in Appendix C. Problem sizes for both reformulations are listed in Table 8, while the graphical solution to this problem is presented in Fig. 8.

We solved this problem using the proposed cutting plane algorithm, and compared the resulting solutions to those from

Table 8
Problem sizes for 10-process retrofit planning problem

	Total number of constraints	Total number of variables	Number of discrete variables
Convex hull	2505	1417	320
Big-M	1957	697	320

the convex hull and big-M reformulations in Tables 9 and 10. We first examine the results with all MIP algorithmic options turned off (see Table 9).

The optimal solution of the problem is US$ 7 868 786.32. The upper bound obtained from the relaxation is equal to US$ 11 743 915.93 for the (BM) reformulation and US$ 7 868 786.32 for the (CH) reformulation, and the problem was solved in 1 607 486 nodes using the (BM) reformulation, as opposed to the (CH) reformulation, which required only 2155 nodes. Clearly, the (CH) feasible region is tighter than that of the big-M, which results in large savings in computational time (5.8 s versus 1913.67 s). After the addition of 40 cutting planes to the (BM) reformulation, we are able to reduce the relaxation gap by nearly 40% and solve the problem in 656.15 s while examining 403 463 nodes in the B&B tree. Upon the successive addition of more cuts, the number of nodes examined is further reduced, which results in a further decrease in total solution time. Finally, 196 cuts were
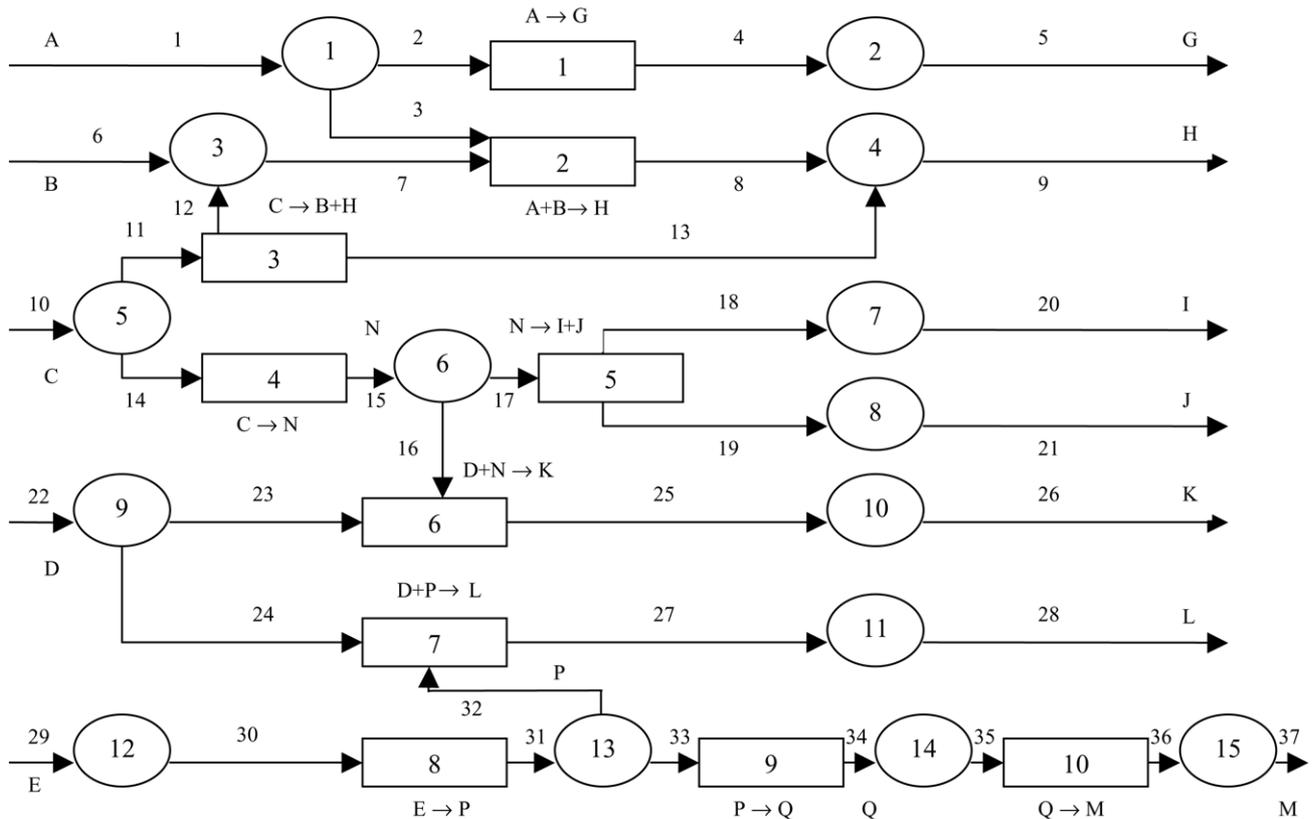
Fig. 7. Ten process retrofit planning problem flowsheet.
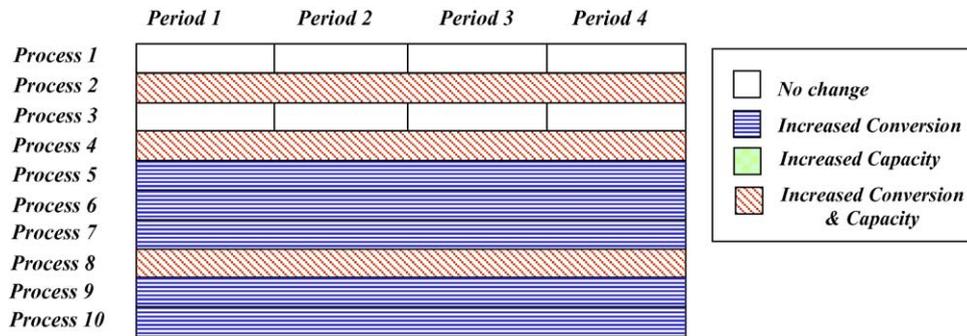
Fig. 8. Graphical solution for 10-process retrofit planning problem.

Table 9
Results for 10-process retrofit planning problem (MIP options off)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	7868786.32	7310873.99	7.63	2155	0	5.8	371.55
Big-M	11743915.93	7310873.99	60.64	1607486	0	1913.67	840.00
Big-M + 40 cuts	8975184.67	7310873.99	22.76	403463	5.6	656.15	620.18
Big-M + 80 cuts	8110107.97	7310873.99	10.93	59601	11.2	134.9	481.81
Big-M + 120 cuts	7930714.76	7310873.99	8.48	46249	16.8	107.96	507.33
Big-M + 160 cuts	7888443.72	7310873.99	7.90	15280	22.4	68.73	329.80
Big-M + 196 cuts	7868786.32	7310873.99	7.63	13669	27.44	60.3	415.97

[a] Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP.

generated from the proposed cutting plane algorithm (requiring 27.44 s) and the problem was solved to optimality in a total of 60.3 s while examining only 13 669 nodes. Furthermore, upon the addition of all cuts generated, the relaxation gap was reduced to 7.63%, identical to that of the (CH) relaxation. Although this demonstrates the efficiency of the cuts, and allows the problem to be solved in much less time than without cuts (60.3 s versus 1913.67 s), the solution time required by the (CH) reformulation is less still (5.8 s). This is due to the extremely tight region generated by the (CH) reformulation, which justifies the additional variables incurred by the reformulation and allows the problem to be solved in faster times than our method. This leads us to believe that classes of problems with extremely tight (CH) reformulations are solved more efficiently as (CH) MIPs through traditional B&B solvers without requiring the additional cut generation technique that we have developed. On the other hand, the example shows very good improvement of the big-M formulation with the addition of cutting planes. The results when default MIP options are turned on present similar trends as previously discussed and are shown in Table 10.

4.3. Zero-wait job-shop scheduling problem

Consider a job-shop scheduling problem where a set of jobs $i \in I$ must be processed sequentially on a set of consecutive stages $j \in J$, where all jobs can be sequenced on a subset of stages $j \in J(i)$. Furthermore, zero-wait transfer is assumed between stages, and the objective is to obtain a schedule that minimizes the makespan, ms. The following model (JS-GDP) from Raman and Grossmann (1994) is proposed:

$$\text{Min} \quad \text{ms} \tag{27}$$

$$\text{s.t. ms} \geq t_i + \sum_{\forall j \in J(i)} \text{TAU}_{ij} \quad \forall i \in I \tag{28}$$

Table 10
Results for 10-process retrofit planning problem (default options on)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	7868786.32	7310873.99	7.63	35	0	0.578	60.55
Big-M	11743915.93	7310873.99	60.64	400612	0	518.64	772.42
Big-M + 40 cuts	8975184.67	7310873.99	22.76	326864	5.6	591.4	557.97
Big-M + 80 cuts	8110107.97	7310873.99	10.93	37464	11.2	95.04	446.85
Big-M + 120 cuts	7930714.76	7310873.99	8.48	7695	16.8	38.41	356.08
Big-M + 160 cuts	7888443.72	7310873.99	7.9	3391	22.4	33.6	302.76
Big-M + 196 cuts	7868786.32	7310873.99	7.63	1857	27.44	35.89	219.76

[a] Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP.
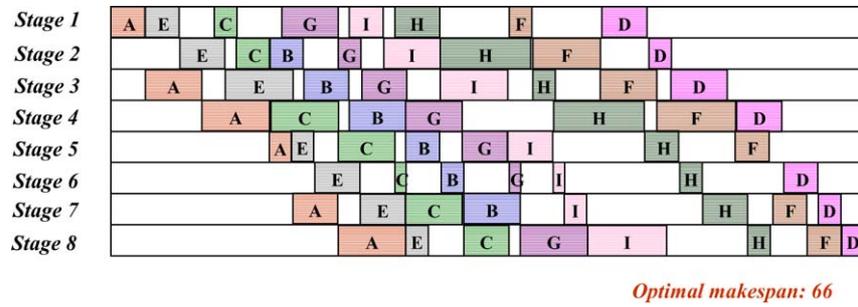
Fig. 9. Graphical solution for 9-job/8-stage zero-wait job-shop scheduling problem.

$$
\begin{bmatrix}
Y_{ik}^1 \\
t_i + \sum_{\substack{\forall m \in J(i) \\ m \le j}} \mathrm{TAU}_{im} \le t_k + \sum_{\substack{\forall m \in J(k) \\ m < j}} \mathrm{TAU}_{km}
\end{bmatrix}
$$

$$
\vee
\begin{bmatrix}
Y_{ik}^2 \\
t_k + \sum_{\substack{\forall m \in J(k) \\ m \le j}} \mathrm{TAU}_{km} \le t_i + \sum_{\substack{\forall m \in J(i) \\ m < j}} \mathrm{TAU}_{im}
\end{bmatrix}
$$

$$
\forall j \in C_{ik}, \ \forall i, k \in I, \ i < k \tag{29}
$$

$$
ms, t_i \in \mathbb{R}_+^1, \ Y_{ik}^1, Y_{ik}^2 \in \{\text{True, False}\} \quad \forall i, k \in I, \ i < k
$$

Equations (27) and (28) correspond to the objective function and aim to minimize the makespan ms, where t_i is the start time of job i and TAU_{ij} is the processing time of job i in stage j. Equation (29) ensures that no clash between jobs occurs at any stage at the same time, where for each pair of jobs i, k, the stages with potential clashes are $C_{ik} = \{J(i) \cap J(k)\}$.

Table 11
Problem sizes for 9-job/8-stage zero-wait job-shop scheduling problem

	Total number of constraints	Total number of variables	Number of discrete variables
Convex hull	681	226	72
Big-M	465	82	72

We consider an instance of the zero-wait job-shop scheduling problem with 9-jobs and 8-stages. We do not include explicit data for this problem because of its size. The problem was transformed into an MIP model by using both the big-M and convex hull reformulations given in Appendix D. The problem sizes for both reformulations are presented in Table 11, while the graphical solution to this problem is presented in Fig. 9.

The problem was solved with the proposed cutting plane algorithm and compared with the results from the convex hull and big-M reformulations in Tables 12 and 13. We first examine the results with all MIP algorithmic options turned off (see Table 12).

Table 12
Results for 10-job/8-stage job-shop scheduling problem (MIP options off)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	35.25	66	46.59	27402	0	12.03	2277.81
Big-M	33	66	50.0	37260	0	7.47	4987.95
Big-M + 10 cuts	35.25	66	46.59	45970	0.46	10.43	4610.83
Big-M + 20 cuts	35.25	66	46.59	38153	0.92	9.17	4624.61
Big-M + 30 cuts	35.25	66	46.59	25547	1.38	7.01	4619.7

[a] Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP.

Table 13
Results for 10-job/8-stage job-shop scheduling problem (default options on)

	Relaxation	Optimal solution	Gap (%)	Total nodes in MIP	Solution time for cut generation (s)	Total solution time[a] (s)	Number of nodes per second
Convex hull	35.25	66	46.59	58599	0	38.34	1528.4
Big-M	33	66	50.0	9757	0	1.75	5575.43
Big-M + 10 cuts	35.25	66	46.59	10900	0.46	2.55	5207.84
Big-M + 20 cuts	35.25	66	46.59	6040	0.92	2.05	5368.88
Big-M + 30 cuts	35.25	66	46.59	5562	1.38	2.44	5247.17

[a] Total solution time includes times for relaxed MIP(s) + LP(s) from separation problem + MIP.

The optimal solution of the problem is 66. The lower bound obtained from the relaxation is equal to 33 for the (BM) reformulation and 35.25 for the (CH) reformulation, and the problem was solved in 37 260 nodes using the (BM) reformulation, as opposed to the (CH) reformulation, which required 27 402 nodes. It is clear that the (CH) feasible region is not much tighter than that of the (BM) as seen from the poor relaxation value and the number of nodes examined in the B&B tree. This causes the solution time of (CH) to be larger than that of (BM) due to the greater number of variables and constraints in the formulation (12.03 s versus 7.47 s). Furthermore, one can conjecture that since the (CH) feasible region is not much tighter than that of the (BM), the effect that the cuts will have on overall solution times and number of nodes examined will be minimal. In fact, after the addition of 30 cutting planes to the (BM) reformulation, we are able to solve the problem in 7.01 s while examining 25 547 nodes in the B&B tree, only slight improvements on the results obtained using the (BM) reformulation. This leads us to believe that classes of problems with extremely loose (CH) reformulations are solved efficiently enough as (BM) MIPs through traditional B&B solvers. The proposed cutting plane algorithm does not improve solution times for this class of problems since the amount of time required to generate the cuts does not justify the (loose) tightening the cuts provide. The results when default MIP options are turned on present similar trends as previously discussed and are shown in Table 13 (note once again, as in the case of the strip-packing problem, the poor results obtained using the (CH) reformulation).

5. Conclusion

We have presented in this paper a cutting plane method that adds cuts generated from a separation problem to a big-M reformulation of a linear GDP problem. We have rigorously derived the cuts, and applied the method to the strip-packing, retrofit planning and zero-wait job-shop scheduling problems. The results demonstrate the efficiency of the proposed method for a class of problems where the convex hull relaxation is tighter than that of the big-M, but not tight enough to justify the additional variables required by the (CH) reformulation. An example within that class is the strip-packing problem where excellent results were obtained. Furthermore, we have also highlighted some of the drawbacks of the method regarding other classes of problems which include the retrofit planning and zero-wait job-shop scheduling problems, where the convex hull relaxation was either too tight or too loose respectively.

We intend to examine in the future different methods that could improve the algorithm, specifically as pertaining to the judicious selection of those Lagrange multipliers (for the infinity norm) that generate the "best" cuts for our specific problem. We also intend to compare the proposed cuts, both theoretically and computationally, to those already present in the literature and to extend the work to solution methods for convex non-linear GDP problems.

Appendix A. Proofs of propositions

Proof of Proposition 1.

(1) (*FR-SEP*)⊆(*FR-BM*): See Grossmann and Lee (2003), Proposition 4.
(2) (*FRP-SEP*) *is convex*: From Ceria and Soares (1999) and Grossmann and Lee (2003), we know that (FR-SEP) is a convex set. Thus, since (FRP-SEP) is the projection of (FR-SEP) from the (z, v) space onto the z space, and projection preserves convexity, then (FRP-SEP) is also convex. $\square$

Proof of Proposition 2.

(1) Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be defined as $||z - z^{\mathrm{bm}}||$. Then ϕ is a convex function for the 1, 2 and ∞ norms over all its domain. Also, from Proposition 1, we know that (FRP-SEP) is a convex set. Furthermore, let $(z^{\mathrm{sep}}, v^{\mathrm{sep}})$ be the optimal solution of (SEP). Clearly, from the properties of projection, z^{sep} would be the optimal solution of (SEP1), where (SEP1) is as follows:

$$\text{Min} \quad \phi(z) = ||z - z^{\mathrm{bm}}||$$
$$\text{s.t.} \quad z \in (\text{FRP-SEP})$$

From Theorem 3.4.3 in Bazarra and Shetty (1979), if z^{sep} is an optimal solution to (SEP1), then ϕ has a subgradient ξ at z^{sep} such that $\xi^{\mathrm{T}}(z - z^{\mathrm{sep}}) \geq 0 \quad \forall z \in (\text{FRP-SEP})$.

(2) From Proposition 1, we know that (FR-SEP)⊆(FR-BM). In the case where $z^{\mathrm{bm}} \in$ (FRP-SEP), obviously no cut can be generated. Otherwise, $z^{\mathrm{bm}} \notin$ (FRP-SEP) and we show that the above inequality cuts off z^{bm}. From (1) we know that $\xi^{\mathrm{T}}(z - z^{\mathrm{sep}}) \geq 0 \quad \forall z \in$ (FRP-SEP). Furthermore, ξ is a subgradient of $\phi(z)$ at z^{sep}. By definition of subgradient (Nemhauser & Wolsey, 1999), for convex $\phi(z)$, we have:

$$\phi(z) - \phi(z^{\mathrm{sep}}) \geq \xi^{\mathrm{T}}(z - z^{\mathrm{sep}}) \quad \forall z \in (\text{FR-SEP})$$
$$\Leftrightarrow \quad \left\|z - z^{\mathrm{bm}}\right\| - \left\|z^{\mathrm{sep}} - z^{\mathrm{bm}}\right\| \geq \xi^{\mathrm{T}}(z - z^{\mathrm{sep}})$$
$$\forall z \in (\text{FR-SEP})$$

if $z \equiv z^{\mathrm{bm}}$, then

$$||z^{\mathrm{bm}} - z^{\mathrm{bm}}|| - ||z^{\mathrm{sep}} - z^{\mathrm{bm}}|| \geq \xi^{\mathrm{T}}(z^{\mathrm{bm}} - z^{\mathrm{sep}})$$
$$\Leftrightarrow \xi^{\mathrm{T}}(z^{\mathrm{bm}} - z^{\mathrm{sep}}) \leq -||z^{\mathrm{sep}} - z^{\mathrm{bm}}|| < 0$$

Thus, z^{bm} does not satisfy $\xi^T(z - z^{sep}) \geq 0$ and is therefore cut off by the inequality. $\square$

Proof of Proposition 3. If ϕ is differentiable over S, then ϕ is differentiable at z^{sep}. It follows from Lemma 3.3.2 in Bazarra and Shetty (1979) that the only element of the subdifferential of ϕ is $\{\nabla\phi(z^{sep})\}$. $\square$

Proof of Proposition 4. If ϕ is defined as $\phi(z) = ||z - z^{bm}||_2^2$, then ϕ is differentiable and from Proposition 3, the collection of subgradients of ϕ at z^{sep} is the singleton set $\{\nabla\phi(z^{sep})\}$. Thus,

$$\phi(z^{sep}) = (z^{sep} - z^{bm})^T(z^{sep} - z^{bm}) \text{ and}$$

$$\nabla\phi(z^{sep}) = 2(z^{sep} - z^{bm}) \quad \square$$

Proof of Proposition 5. Let $\phi : S \rightarrow \mathbb{R}$ be defined as $\phi(z) = ||z - z^{bm}||_\infty$ in (SEP). Then (SEP) can be rewritten as:

Min u

s.t. $u \geq z_i - z_i^{bm} \quad \forall i \in M$

$\quad\quad u \geq z_i^{bm} - z_i \quad \forall i \in M$ (A.1)

(FR-SEP)

From Proposition 1, we know that (FR-SEP) is convex. Furthermore, all the constraints in (FR-SEP) are linear. Thus, (FR-SEP) corresponds to a polytope in the (z, v) space, and $\exists$ matrices R^1, R^2 with dimensions $m \times (n \times \sum_{k \in K} |J_k|)$, $m \times n$, respectively, and vector $r \in \mathbb{R}^m$ such that (FR-SEP) $\equiv \{(z, v)|R^1z + R^2v \leq r\}$. Note that the non-negativity constraints for z and v are taken into account in the construction of R^1, R^2.

We can thus write (A.1) as:

Min u

s.t. $u \geq z_i - z_i^{bm} \quad \forall i \in M$

$\quad\quad u \geq z_i^{bm} - z_i \quad \forall i \in M$ (A.2)

$\quad\quad R^1z + R^2v \leq r$

The appropriate Lagrangean function of (A.2) is as follows:

$$L = u + \sum_{i \in M} \mu_{i+}(z_i - z_i^{bm} - u) + \sum_{i \in M} \mu_{i-}(z_i^{bm} - z_i - u)$$

$$+ \rho^T(R^1z + R^2v - r)$$

and it is implied at $(z^{sep}, v^{sep}, u^{sep})$ that multipliers μ_+^{sep}, μ_-^{sep}, and ρ^{sep} exist such that:

$$\frac{\partial L}{\partial u}(z^{sep}, v^{sep}, u^{sep}) = 0 \Rightarrow 1 - \sum_{i \in M^{sep}} \mu_{i+} - \sum_{i \in M^{sep}} \mu_{i-} = 0 \Rightarrow \sum_{i \in M^{sep}} (\mu_{i+} + \mu_{i-}) = 1$$

$$\nabla_z L(z^{sep}, v^{sep}, u^{sep}) = 0 \Rightarrow [\mu_+ - \mu_-] + R^{1T}\rho^{sep} = 0$$ (A.3)

$$\nabla_v L(z^{sep}, v^{sep}, u^{sep}) = 0 \Rightarrow R^{2T}\rho^{sep} = 0$$

$$\mu_i^+, \mu_i^-, \rho^{sep} \geq 0 \quad \forall i \in M^{sep}, \text{ where } M^{sep} \equiv \text{active } M \equiv \{i|(u = z_i - z_i^{bm}) \vee (u = z_i^{bm} - z_i) \forall i \in M\}$$

Now let us define the following matrix H with columns h_i as $H \equiv [I| - I]$ and the following vector μ as $\mu \equiv \begin{bmatrix} \mu_+ \\ \mu_- \end{bmatrix}$.

We claim that if $\xi = H\mu$, then the existence of a vector $\mu^{sep} \equiv \begin{bmatrix} \mu_+^{sep} \\ \mu_-^{sep} \end{bmatrix} \equiv \begin{bmatrix} \mu_{i+} \\ \mu_{i-} \end{bmatrix} \quad \forall i \in M^{sep}$ in (A.3) is equivalent to the existence of a vector ξ^{sep} in the set:

$$\partial^{sep}\phi \equiv \{\xi^{sep}|\xi^{sep} \in \operatorname*{conv}_{i \in N^{sep}} h_i\}$$ (A.4)

such that

$$\xi + R^{1T}\rho^{sep} = 0$$

$$R^{2T}\rho^{sep} = 0$$

$$\rho^{sep} \geq 0$$

where (A.4) is the subdifferential of ϕ, ξ^{sep} is a subgradient of $\phi(z^{sep})$ and $N^{sep} \equiv \{i : |z_i - z_i^{bm}| \text{ is maximized}\}$, according to Section 14.1 and Lemma 14.2.2 in Fletcher (1987).

In essence, we are claiming that in order to obtain a subgradient vector ξ^{sep} of $\phi(z^{sep})$ in (A.4), one needs only to obtain a set of Lagrange multipliers μ^{sep} from (A.3) (thus, the existence of one is equivalent to the existence of the other). We prove the claim as follows.

From (A.4), we have,

$$\partial^{sep}\phi \equiv \{\xi^{sep}|\xi^{sep} \in \operatorname*{conv}_{i \in N^{sep}} h_i\}$$

such that

$$\xi + R^{1T}\rho^{sep} = 0$$

$$R^{2T}\rho^{sep} = 0$$

$$\rho^{sep} \geq 0$$

$$\Leftrightarrow \xi^{sep} = \sum_{i \in N^{sep}} \alpha_i h_i \text{ with } \sum_{i \in N^{sep}} \alpha_i = 1,$$

$$\alpha_i \geq 0 \text{ from the convex hull definition}$$

such that

$$\xi + R^{1T}\rho^{sep} = 0$$

$$R^{2T}\rho^{sep} = 0$$

$$\rho^{\text{sep}} \geq 0$$

$$\Leftrightarrow \xi^{\text{sep}} = H^{\text{sep}}\alpha^{\text{sep}} = [I^{\text{sep}}| - I^{\text{sep}}] \begin{bmatrix} \alpha_+^{\text{sep}} \\ \alpha_-^{\text{sep}} \end{bmatrix}$$

$$= [\alpha_+^{\text{sep}} - \alpha_-^{\text{sep}}]$$

such that

$$\sum_{i \in N^{\text{sep}}} \alpha_{i+} + \alpha_{i-} = 1, \alpha^{\text{sep}} \geq 0$$

$$\xi + R^{1\text{T}}\rho^{\text{sep}} = 0 \tag{A.5}$$

$$R^{2\text{T}}\rho^{\text{sep}} = 0$$

$$\rho^{\text{sep}} \geq 0$$

If $\xi = H\mu \Leftrightarrow \xi = [I| - I] \begin{bmatrix} \mu_+ \\ \mu_- \end{bmatrix} = [\mu_+ - \mu_-]$

then

$$\xi^{\text{sep}} = [\mu_+^{\text{sep}} - \mu_-^{\text{sep}}] \quad \text{and} \quad \xi_i = [\mu_{i+} - \mu_{i-}] \quad \forall i \in N^{\text{sep}}$$

but from (A.5) we know that $\xi^{\text{sep}} = [\alpha_+^{\text{sep}} - \alpha_-^{\text{sep}}]$, so $\alpha = \mu$ and (A.5) becomes:

$$\xi^{\text{sep}} = [\mu_+^{\text{sep}} - \mu_-^{\text{sep}}]$$

s.t. $\sum_{i \in N^{\text{sep}}} \mu_{i+} + \mu_{i-} = 1, \mu_i \geq 0$

$$[\mu_+ - \mu_-] + R^{1\text{T}}\rho^{\text{sep}} = 0 \tag{A.6}$$

$$R^{2\text{T}}\rho^{\text{sep}} = 0$$

$$\rho^{\text{sep}} \geq 0$$

Clearly, $N^{\text{sep}} \equiv M^{\text{sep}}$ and we have thus recovered (A.3) and shown that the form of the subgradient is indeed $\xi^{\text{sep}} = [\mu_+^{\text{sep}} - \mu_-^{\text{sep}}]$.   □

Appendix B. Reformulations of strip-packing problem (SP-GDP)

B.1. Big-M reformulation of (SP-GDP)

Min lt

s.t. $lt \geq x_i + L_i$ $\forall i \in N$

$x_i + L_i \leq x_j + \text{BIGM}_{ij}^1(1 - w_{ij}^1)$ $\forall i, j \in N, i < j$

$x_j + L_j \leq x_i + \text{BIGM}_{ij}^2(1 - w_{ij}^2)$ $\forall i, j \in N, i < j$

$y_i - H_i \geq y_j - \text{BIGM}_{ij}^3(1 - w_{ij}^3)$ $\forall i, j \in N, i < j$

$y_j - H_j \geq y_i - \text{BIGM}_{ij}^4(1 - w_{ij}^4)$ $\forall i, j \in N, i < j$

$\sum_{d \in D} w_{ij}^d = 1$ $\forall i, j \in N, i < j$

$x_i \leq UB_i - L_i$ $\forall i \in N$

$H_i \leq y_i \leq W$ $\forall i \in N$

$lt, x_i, y_i \in \mathbb{R}_+^1, w_{ij}^d \in \{0, 1\}$ $\forall d \in D, \forall i, j \in N, i < j$

where $D = \{1, 2, 3, 4\}$.

B.2. Convex hull reformulation of (SP-GDP)

Min lt

s.t. $lt \geq x_i + L_i$ $\forall i \in N$

$x_k = \sum_{d \in D} v_{kij}^d$ $\forall i, j, k \in N, i < j, k = i \vee j$

$y_k = \sum_{d \in D} \omega_{kij}^d$ $\forall i, j, k \in N, i < j, k = i \vee j$

$v_{iij}^1 - v_{jij}^1 \leq -L_i w_{ij}^1$ $\forall i, j \in N, i < j$

$v_{jij}^2 - v_{iij}^2 \leq -L_j w_{ij}^2$ $\forall i, j \in N, i < j$

$\omega_{iij}^3 - \omega_{jij}^3 \geq H_i w_{ij}^3$ $\forall i, j \in N, i < j$

$\omega_{jij}^4 - \omega_{iij}^4 \geq H_j w_{ij}^4$ $\forall i, j \in N, i < j$

$\sum_{d \in D} w_{ij}^d = 1$ $\forall i, j \in N, i < j$

$v_{kij}^d \leq \text{UB}_{kij}^d w_{ij}^d$ $\forall d \in D, \forall i, j, k \in N, i < j, k = i \vee j$

$\omega_{kij}^d \leq \text{UB}_{kij}^d w_{ij}^d$ $\forall d \in D, \forall i, j, k \in N, i < j, k = i \vee j$

$x_i \leq \text{UB}_i - L_i$ $\forall i \in N$

$H_i \leq y_i \leq W$ $\forall i \in N$

$lt, x_i, y_i, v_{kij}^d, \omega_{kij}^d$
    $\in \mathbb{R}_+^1, w_{ij}^d \in \{0, 1\}$ $\forall d \in D, \forall i, j, k \in N, i < j, k = i \vee j$

where $D = \{1, 2, 3, 4\}$.

Appendix C. Reformulations of retrofit planning problem (RP-GDP)

C.1. Big-M reformulation of (RP-GDP)

$$\text{Min } Z = \sum_{t \in T} \sum_{s \in S_{\text{prod}}} \text{PR}_s^t mf_s^t - \sum_{t \in T} \sum_{s \in S_{\text{raw}}} \text{PR}_s^t mf_s^t - \sum_{t \in T} \text{PRST}qst^t$$

$$- \sum_{t \in T} \text{PRWT}qwt^t - \sum_{t \in T} \sum_{m \in M} \sum_{p \in P} \text{FC}_{pm}^t w_{pm}^t - \sum_{t \in T} \sum_{j \in J} \text{EFC}_j^t$$

s.t. $\quad mf_s^t = f_s^t \text{MW}_s$ $\hspace{4cm} \forall s \in S, \forall t \in T$

$mf_s^t \geq \text{DEM}_s^t$ $\hspace{4cm} \forall s \in S_{\text{prod}}, \forall t \in T$

$mf_s^t \leq \text{SUP}_s^t$ $\hspace{4cm} \forall s \in S_{\text{raw}}, \forall t \in T$

$$\sum_{s \in S_{n_{\text{in}}}} mf_s^t = \sum_{s \in S_{n_{\text{out}}}} mf_s^t \hspace{2.5cm} \forall n \in N, \forall t \in T$$

$$\sum_{s \in S_{p_{\text{in}}}} mf_s^t = \sum_{s \in S_{p_{\text{out}}}} mf_s^t + \text{unrct}_p^t \hspace{1.5cm} \forall p \in P, \forall t \in T$$

$$f_s^t \leq f_{p_{lmt}}^t \left(\frac{\text{GMA}_s^t}{\text{GMA}_{p_{lmt}}} \right) \text{ETA}_{pm}^t + \text{BIGM}_{pm}^t (1 - y_{pm}^t) \hspace{1cm} \forall s \in S_{p_{\text{out}}}, \forall p \in P, \forall m \in M, \forall t \in T$$

$$f_s^t \geq f_{p_{lmt}}^t \left(\frac{\text{GMA}_s^t}{\text{GMA}_{p_{lmt}}} \right) \text{ETA}_{pm}^t - \text{BIGM}_{pm}^t (1 - y_{pm}^t) \hspace{1cm} \forall s \in S_{p_{\text{out}}}, \forall p \in P, \forall m \in M, \forall t \in T$$

$$\sum_{s \in S_{p_{\text{in}}}} mf_s^t \leq \text{CAP}_{pm}^t + \text{BIGM}_{pm}^t (1 - y_{pm}^t) \hspace{1.5cm} \forall p \in P, \forall m \in M, \forall t \in T$$

$fc_p^t \leq \text{FC}_{pm}^t + \text{BIGM}_{pm}^t (1 - w_{pm}^t)$ $\hspace{2cm} \forall p \in P, \forall m \in M, \forall t \in T$

$fc_p^t \geq \text{FC}_{pm}^t - \text{BIGM}_{pm}^t (1 - w_{pm}^t)$ $\hspace{2cm} \forall p \in P, \forall m \in M, \forall t \in T$

$q_{sk}^t = mf_s^t \text{CP}_s (T_{s_{\text{out}_k}}^t - T_{s_{\text{in}_k}}^t)$ $\hspace{2cm} \forall s \in S_{\text{cold}}, \forall k \in K, \forall t \in T$

$q_{sk}^t = mf_s^t \text{CP}_s (T_{s_{\text{in}_k}}^t - T_{s_{\text{out}_k}}^t)$ $\hspace{2cm} \forall s \in S_{\text{hot}}, \forall k \in K, \forall t \in T$

$$qst^t \leq \sum_{k \in K} \sum_{s \in S_{\text{cold}}} q_{sk}^t + \text{BIGM}_{1,1}^t (1 - x_1^t) \hspace{1cm} \forall t \in T$$

$$qst^t \geq \sum_{k \in K} \sum_{s \in S_{\text{cold}}} q_{sk}^t - \text{BIGM}_{1,1}^t (1 - x_1^t) \hspace{1cm} \forall t \in T$$

$$qwt^t \leq \sum_{k \in K} \sum_{s \in S_{\text{hot}}} q_{sk}^t + \text{BIGM}_{2,1}^t (1 - x_1^t) \hspace{1cm} \forall t \in T$$

$$qwt^t \geq \sum_{k \in K} \sum_{s \in S_{\text{hot}}} q_{sk}^t - \text{BIGM}_{2,1}^t (1 - x_1^t) \hspace{1cm} \forall t \in T$$

$$r_k^t - r_{k-1}^t - qst_k^t + qwt_k^t \leq \sum_{s \in S_{\text{hot}}} q_{sk}^t - \sum_{s \in S_{\text{cold}}} q_{sk}^t + \text{BIGM}_{1,2}^t (1 - x_2^t) \hspace{0.5cm} \forall k \in K, \forall t \in T$$

$$r_k^t - r_{k-1}^t - qst_k^t + qwt_k^t \geq \sum_{s \in S_{\text{hot}}} q_{sk}^t - \sum_{s \in S_{\text{cold}}} q_{sk}^t - \text{BIGM}_{1,2}^t (1 - x_2^t) \hspace{0.5cm} \forall k \in K, \forall t \in T$$

$$\text{qst}^t \le \sum_{k \in K} \text{qst}_k^t + \text{BIGM}_{2,2}^t (1 - x_2^t) \qquad \forall t \in T$$

$$\text{qst}^t \ge \sum_{k \in K} \text{qst}_k^t - \text{BIGM}_{2,2}^t (1 - x_2^t) \qquad \forall t \in T$$

$$\text{qwt}^t \le r_{|K|}^t + \sum_{k \in K} \text{qwt}_k^t + \text{BIGM}_{3,2}^t (1 - x_2^t) \qquad \forall t \in T$$

$$\text{qwt}^t \ge r_{|K|}^t + \sum_{k \in K} \text{qwt}_k^t - \text{BIGM}_{2,2}^t (1 - x_2^t) \qquad \forall t \in T$$

$$\text{ec}^t \le \text{EFC}_j^t + \text{BIGM}_j^t (1 - v_j^t) \qquad \forall j \in J, \forall t \in T$$

$$\text{ec}^t \ge \text{EFC}_j^t - \text{BIGM}_j^t (1 - v_j^t) \qquad \forall j \in J, \forall t \in T$$

$$\sum_{p \in P} fc_p^t + \text{ec}^t + \sum_{s \in S_{\text{raw}}} \text{PR}_s^t \, \text{mf}_s^t + \text{PRST} \, \text{qst}^t + \text{PRWT} \, \text{qwt}^t \le \text{INV}^t \qquad \forall t \in T$$

$$\sum_{m \in M} y_{pm}^t = 1 \qquad \forall p \in P, \forall t \in T$$

$$\sum_{m \in M} w_{pm}^t = 1 \qquad \forall p \in P, \forall t \in T$$

$$\sum_{j \in J} x_j^t = 1 \qquad \forall t \in T$$

$$\sum_{j \in J} v_j^t = 1 \qquad \forall t \in T$$

$$y_{pm}^t \le y_{pm}^\tau \qquad \forall p \in P, \forall t < \tau \in T, \forall m \in M \backslash m_1$$
$$w_{pm}^t \le w_{p1}^\tau \qquad \forall p \in P, \forall t \ne \tau \in T, \forall m \in M \backslash m_1$$
$$y_{p1}^t \le w_{p1}^t \qquad \forall p \in P, \forall t \in T$$

$$y_{pm}^t \le w_{pm}^t + \sum_{\tau=1}^{|T|-1} y_{pm}^{t-\tau} \qquad \forall p \in P, \forall t \in T, \forall m \in M \backslash m_1$$

$$x_2^t \le x_1^\tau \qquad \forall t < \tau \in T$$
$$v_2^t \le v_1^\tau \qquad \forall t \ne \tau \in T$$
$$x_1^t \le v_1^t \qquad \forall t \in T$$

$$x_2^t \le v_2^t + \sum_{\tau=1}^{|T|-1} x_2^{t-\tau} \qquad \forall t \in T$$

$$mf_s^t, f_s^t \in \mathbb{R}_+^1 \qquad \forall s \in S, \forall t \in T$$
$$f_{p_{\text{lmt}}}^t, \text{unrct}_p^t, fc_p^t \in \mathbb{R}_+^1 \qquad \forall p \in P, \forall t \in T$$
$$q_{sk}^t \in \mathbb{R}_+^1 \qquad \forall s \in S, \forall k \in K, \forall t \in T$$
$$\text{qst}^t, \text{qwt}^t, \text{ec}^t \in \mathbb{R}_+^1 \qquad \forall t \in T$$
$$\text{qst}_k^t, \text{qwt}_k^t, r_k^t \in \mathbb{R}_+^1 \qquad \forall k \in K, \forall t \in T$$
$$y_{pm}^t, w_{pm}^t \in \{0, 1\} \qquad \forall p \in P, \forall t \in T, \forall m \in M$$
$$x_j^t, v_j^t \in \{0, 1\} \qquad \forall j \in J, \forall t \in T$$

C.2. Convex hull reformulation of (RP-GDP)

$$\text{Min } Z = \sum_{t \in T} \sum_{s \in S_{\text{prod}}} \text{PR}_s^t mf_s^t - \sum_{t \in T} \sum_{s \in S_{\text{raw}}} \text{PR}_s^t mf_s^t - \sum_{t \in T} \text{PRST} qst^t - \sum_{t \in T} \text{PRWT} qwt^t$$
$$- \sum_{t \in T} \sum_{m \in M} \sum_{p \in P} \text{FC}_{pm}^t w_{pm}^t - \sum_{t \in T} \sum_{j \in J} \text{EFC}_j^t$$

s.t. $\quad mf_s^t = f_s^t \text{MW}_s \qquad\qquad\qquad\qquad\qquad\qquad \forall s \in S, \forall t \in T$

$mf_s^t \geq \text{DEM}_s^t \qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall s \in S_{\text{prod}}, \forall t \in T$

$mf_s^t \leq \text{SUP}_s^t \qquad\qquad\qquad\qquad\qquad\qquad\qquad\;\; \forall s \in S_{\text{raw}}, \forall t \in T$

$\displaystyle\sum_{s \in S_{P_{\text{in}}}} mf_s^t = \sum_{s \in S_{P_{\text{out}}}} mf_s^t \qquad\qquad\qquad\qquad\quad \forall n \in N, \forall t \in T$

$\displaystyle\sum_{s \in S_{P_{\text{in}}}} mf_s^t = \sum_{s \in S_{P_{\text{out}}}} mf_s^t + \text{unrct}_p^t \qquad\qquad\quad \forall p \in P, \forall t \in T$

$\displaystyle f_s^t = \sum_{m \in M} zf_{sm}^t \qquad\qquad\qquad\qquad\qquad \forall s \in S_{P_{\text{out}}}, \forall p \in P, \forall t \in T$

$\displaystyle f_{P_{lmt}}^t = \sum_{m \in M} zf_{P_{lmt}m}^t \qquad\qquad\qquad\qquad\qquad \forall p \in P, \forall t \in T$

$\displaystyle mf_s^t = \sum_{m \in M} zmf_{sm}^t \qquad\qquad\qquad\qquad \forall s \in S_{p_{\text{in}}}, \forall p \in P, \forall t \in T$

$\displaystyle fc_p^t = \sum_{m \in M} zfc_{pm}^t \qquad\qquad\qquad\qquad\qquad\quad \forall p \in P, \forall t \in T$

$\displaystyle qst^t = \sum_{j \in J} zqst_j^t \qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall t \in T$

$\displaystyle qwt^t = \sum_{j \in J} zqwt_j^t \qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall t \in T$

$\displaystyle q_{sk}^t = \sum_{j \in J} zq_{skj}^t \qquad\qquad\qquad \forall s \in S_{\text{cold}} \cup S_{\text{hot}}, \forall k \in K, \forall t \in T$

$\displaystyle r_k^t = \sum_{j \in J} zr_{kj}^t \qquad\qquad\qquad\qquad\qquad\qquad \forall k \in K, \forall t \in T$

$\displaystyle qst_k^t = \sum_{j \in J} zqst_{kj}^t \qquad\qquad\qquad\qquad\qquad\quad \forall k \in K, \forall t \in T$

$\displaystyle qwt_k^t = \sum_{j \in J} zqwt_{kj}^t \qquad\qquad\qquad\qquad\qquad\;\; \forall k \in K, \forall t \in T$

$\displaystyle ec^t = \sum_{j \in J} zec_j^t \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \forall t \in T$

$\displaystyle zf_{sm}^t = zf_{P_{lmt}m}^t \left(\frac{\text{GMA}_s^t}{\text{GMA}_{P_{lmt}}}\right) \text{ETA}_{pm}^t \qquad \forall s \in S_{P_{\text{out}}}, \forall p \in P, \forall m \in M, \forall t \in T$

$\displaystyle\sum_{s \in S_{P_{\text{in}}}} zmf_{sm}^t \leq \text{CAP}_{pm}^t y_{pm}^t \qquad\qquad\qquad \forall p \in P, \forall m \in M, \forall t \in T$

$zfc_{pm}^t = \text{FC}_{pm}^t w_{pm}^t \qquad\qquad\qquad\qquad \forall p \in P, \forall m \in M, \forall t \in T$

$q_{sk}^t = mf_s^t \text{CP}_s (T_{S_{\text{out}_k}}^t - T_{S_{\text{in}_k}}^t) \qquad\qquad \forall s \in S_{\text{cold}}, \forall k \in K, \forall t \in T$

$q_{sk}^t = mf_s^t \text{CP}_s (T_{S_{\text{in}_k}}^t - T_{S_{\text{out}_k}}^t) \qquad\qquad \forall s \in S_{\text{hot}}, \forall k \in K, \forall t \in T$

$$zqst_1^t = \sum_{k \in K} \sum_{s \in S_{cold}} zq_{sk1}^t \qquad \forall t \in T$$

$$zqwt_1^t = \sum_{k \in K} \sum_{s \in S_{hot}} zq_{sk1}^t \qquad \forall t \in T$$

$$zr_{k2}^t - zr_{k-1,2}^t - zqst_{k2}^t + zqwt_{k2}^t = \sum_{s \in S_{hot}} zq_{sk2}^t - \sum_{s \in S_{cold}} zq_{sk2}^t \qquad \forall k \in K, \forall t \in T$$

$$zqst_2^t = \sum_{k \in K} zqst_{k2}^t \qquad \forall t \in T$$

$$zqwt_2^t = zr_{|K|2}^t + \sum_{k \in K} zqwt_{k2}^t \qquad \forall t \in T$$

$$zec_j^t = EFC_j^t v_j^t \qquad \forall j \in J, \forall t \in T$$

$$\sum_{p \in P} fc_p^t + ec^t + \sum_{s \in S_{raw}} PR_s^t mf_s^t + PRSTqst^t + PRWTqwt^t \le INV^t \qquad \forall t \in T$$

$$zf_{sm}^t \le BND_{sm}^t y_{pm}^t \qquad \forall s \in S_{p_{out}}, \forall p \in P, \forall m \in M, \forall t \in T$$

$$zf_{P_{lmt}m}^t \le BND_{P_{lmt}m}^t y_{pm}^t \qquad \forall p \in P, \forall m \in M, \forall t \in T$$

$$zmf_{sm}^t \le BND_{sm}^t y_{pm}^t \qquad \forall s \in S_{p_{in}}, \forall p \in P, \forall m \in M, \forall t \in T$$

$$zfc_{pm}^t \le BND_{pm}^t w_{pm}^t \qquad \forall p \in P, \forall m \in M, \forall t \in T$$

$$zqst_j^t \le BND_j^t x_j^t \qquad \forall j \in J, \forall t \in T$$

$$zqwt_j^t \le BND_j^t x_j^t \qquad \forall j \in J, \forall t \in T, \forall t \in T$$

$$zq_{skj}^t \le BND_{skj}^t x_j^t \qquad \forall s \in S_{cold} \cup S_{hot}, \forall k \in K, \forall j \in J, \forall t \in T$$

$$zr_{kj}^t \le BND_{kj}^t x_j^t \qquad \forall k \in K, \forall j \in J, \forall t \in T$$

$$zqst_{kj}^t \le BND_{kj}^t x_j^t \qquad \forall k \in K, \forall j \in J, \forall t \in T$$

$$zqwt_{kj}^t \le BND_{kj}^t x_j^t \qquad \forall k \in K, \forall j \in J, \forall t \in T$$

$$zec_j^t \le BND_j^t v_j^t \qquad \forall j \in J, \forall t \in T$$

$$\sum_{m \in M} y_{pm}^t = 1 \qquad \forall p \in P, \forall t \in T$$

$$\sum_{m \in M} w_{pm}^t = 1 \qquad \forall p \in P, \forall t \in T$$

$$\sum_{j \in J} x_j^t = 1 \qquad \forall t \in T$$

$$\sum_{j \in J} v_j^t = 1 \qquad \forall t \in T$$

$$y_{pm}^t \le y_{pm}^\tau \qquad \forall p \in P, \forall t < \tau \in T, \forall m \in M \setminus m_1$$

$$w_{pm}^t \le w_{p1}^\tau \qquad \forall p \in P, \forall t \ne \tau \in T, \forall m \in M \setminus m_1$$

$$y_{p1}^t \le w_{p1}^t \qquad\qquad \forall p \in P, \forall t \in T$$

$$y_{pm}^t \le w_{pm}^t + \sum_{\tau=1}^{|T|-1} y_{pm}^{t-\tau} \qquad\qquad \forall p \in P, \forall t \in T, \forall m \in M \backslash m_1$$

$$x_2^t \le x_1^\tau \qquad\qquad \forall t < \tau \in T$$
$$v_2^t \le v_1^\tau \qquad\qquad \forall t \ne \tau \in T$$

$$x_1^t \le v_1^t \qquad\qquad \forall t \in T$$

$$x_2^t \le v_2^t + \sum_{\tau=1}^{|T|-1} x_2^{t-\tau} \qquad\qquad \forall t \in T$$

$$mf_s^t, f_s^t \in \mathbb{R}_+^1 \qquad\qquad \forall s \in S, \forall t \in T$$

$$f_{P_{lmt}}^t, \mathrm{unrct}_p^t, fc_p^t \in \mathbb{R}_+^1 \qquad\qquad \forall p \in P, \forall t \in T$$

$$q_{sk}^t \in \mathbb{R}_+^1 \qquad\qquad \forall s \in S, \forall k \in K, \forall t \in T$$

$$\mathrm{qst}^t, \mathrm{qwt}^t, \mathrm{ec}^t \in \mathbb{R}_+^1 \qquad\qquad \forall t \in T$$

$$\mathrm{qst}_k^t, \mathrm{qwt}_k^t, r_k^t \in \mathbb{R}_+^1 \qquad\qquad \forall k \in K, \forall t \in T$$

$$zmf_{sm}^t, zf_{sm}^t \in \mathbb{R}_+^1 \qquad\qquad \forall s \in S, \forall m \in M, \forall t \in T$$

$$zf_{P_{lmt}m}^t, zfc_{pm}^t \in \mathbb{R}_+^1 \qquad\qquad \forall p \in P, \forall m \in M, \forall t \in T$$

$$zq_{skm}^t \in \mathbb{R}_+^1 \qquad\qquad \forall s \in S, \forall k \in K, \forall m \in M, \forall t \in T$$

$$zqst_m^t, zqwt_m^t, zec_m^t \in \mathbb{R}_+^1 \qquad\qquad \forall m \in M, \forall t \in T$$

$$zqst_{km}^t, zqwt_{km}^t, zr_{km}^t \in \mathbb{R}_+^1 \qquad\qquad \forall k \in K, \forall m \in M, \forall t \in T$$

$$y_{pm}^t, w_{pm}^t \in \{0, 1\} \qquad\qquad \forall p \in P, \forall t \in T, \forall m \in M$$

$$x_j^t, v_j^t \in \{0, 1\} \qquad\qquad \forall j \in J, \forall t \in T$$

Appendix D. Reformulations of zero-wait job-shop scheduling problem (JS-GDP)

D.1. Big-M reformulation of (JS-GDP)

Min   ms

s.t. $\mathrm{ms} \ge t_i + \sum_{j \in J(i)} \mathrm{TAU}_{ij} \qquad\qquad \forall i \in I$

$$t_i + \sum_{\substack{m \in J(i) \\ m \le j}} \mathrm{TAU}_{im} \le t_k + \sum_{\substack{m \in J(k) \\ m \le j}} \mathrm{TAU}_{km} + \mathrm{BIGM}(1 - y_{ik}^1) \quad \forall j \in C_{ik}, \forall i, k \in I, i < k$$

$$t_k + \sum_{\substack{m \in J(k) \\ m \le j}} \mathrm{TAU}_{km} \le t_i + \sum_{\substack{m \in J(i) \\ m \le j}} \mathrm{TAU}_{im} + \mathrm{BIGM}(1 - y_{ik}^2) \quad \forall j \in C_{ik}, \forall i, k \in I, i < k$$

$$\sum_{d \in D} y_{ik}^d = 1 \qquad\qquad \forall i, k \in I, i < k$$

$$\mathrm{ms}, t_i \in \mathbb{R}_+^1, y_{ik}^d \in \{0, 1\} \qquad\qquad \forall i, k \in I, i < k, \forall d \in D$$

D.2. Convex hull reformulation of (JS-GDP)

Min   ms

$$\text{s.t.} \quad \text{ms} \geq t_i + \sum_{j \in J(i)} \text{TAU}_{ij} \qquad \forall i \in I$$

$$t_n = \sum_{d \in D} v^d_{nik} \qquad \forall i, k \in I, i < k, \forall n \in I, n = i \vee k$$

$$v^1_{iik} - v^1_{kik} \leq \left(\sum_{\substack{m \in J(k) \\ m < j}} \text{TAU}_{km} - \sum_{\substack{m \in J(i) \\ m \leq j}} \text{TAU}_{im} \right) y^1_{ik} \quad \forall j \in C_{ik}, \forall i, k \in I, i < k$$

$$v^2_{kik} - v^2_{iik} \leq \left(\sum_{\substack{m \in J(k) \\ m < j}} \text{TAU}_{im} - \sum_{\substack{m \in J(i) \\ m \leq j}} \text{TAU}_{km} \right) y^2_{ik} \quad \forall j \in C_{ik}, \forall i, k \in I, i < k$$

$$v^d_{nik} \leq \text{UB}^d_{ik} y^d_{ik} \qquad \forall i, k \in I, i < k, \forall n \in I, n = i \vee k$$

$$\sum_{d \in D} y^d_{ik} = 1 \qquad \forall i, k \in I, i < k$$

$$\text{ms}, t_i \in \mathbb{R}^1_+, y^d_{ik} \in \{0, 1\} \qquad \forall i, k \in I, i < k, \forall d \in D$$

References

Balas, E. (1998). Disjunctive programming: Properties of the convex hull of feasible points. *Invited paper in Discrete Applied Mathematics*, *89*, 1–44.

Balas, E., Ceria, S., & Cornuejols, G. (1993). A lift-and-project cutting plane algorithm for mixed 0–1 programs. *Mathematical Programming*, *58*(3), 295–324.

Bazarra, M. S., & Shetty, C. M. (1979). *Non-linear programming: Theory and algorithms*. John Wiley & Sons.

Biegler, L., Grossmann, I. E., & Westerberg, A. W. (1997). *Systematic methods of chemical process design*. Prentice Hall.

Brooke A., Kendrick D., Meeruas A., & Raman R., GAMS language guide, Version 98, GAMS Development Corporation, Washington, D.C.

Ceria, S., & Soares, J. (1999). Convex programming for disjunctive optimization. *Mathematical Programming*, *86*(3), 595–614.

Fletcher, R. (1987). *Practical methods of optimization* (2nd ed.). John Wiley & Sons.

Grossmann, I. E., Westerberg, A. W., & Biegler, L. T. (1987). Retrofit design of chemical processes. In G. V. Reklaitis & H. D. Spriggs (Eds.), *Proceedings of foundations of computer aided process operations* (p. 403). Elsevier.

Grossmann, I. E. (2002). Review of non-linear mixed-integer and disjunctive programming techniques. *Optimization and Engineering*, *3*, 227–252.

Grossmann, I. E., & Lee, S. (2003). Generalized disjunctive programming: Nonlinear convex hull relaxation and algorithms. *Computational Optimization and Applications*, *26*, 83–100.

Hifi, M. (1998). Exact algorithms for the guillotine strip cutting/packing problem. *Computers and Operations Research*, *25*(11), 925–940.

Hooker, J. (2000). *Logic-based methods for optimization: Combining optimization and constraint satisfaction*. John Wiley & Sons.

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

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

Nemhauser, G. L., & Wolsey, L. A. (1999). *Integer and combinatorial optimization*. Wiley Interscience.

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

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

Tawarmalani, M., & Sahinidis, N. V. (2002). *Convexification and global optimization in continuous and mixed-integer nonlinear programming*. Dordrecht: Kluwer Academic Publishers.

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

Vecchietti, A., Lee, S., & Grossmann, I. E. (2003). Modeling of discrete/continuous optimization problems: Characterization and formulation of disjunctions and their relaxations. *Computers and Chemical Engineering*, *27*, 433–448.

Williams, H. P. (1985). *Model building in mathematical programming*. John Wiley & Sons.