Reformulations, Relaxations and Cutting Planes for Linear Generalized Disjunctive Programming

Nicolas Sawaya[1], Ignacio Grossmann

Department of Chemical Engineering, Carnegie Mellon University
Pittsburgh, PA, U.S.A, 15201.

Abstract

Generalized disjunctive programming (GDP), originally developed by Raman and Grossmann [1994], is an extension of the well-known disjunctive programming paradigm developed by Balas in the mid seventies in his seminal technical report [Balas, 1974], [see also Balas, 1998]. This mathematical representation of discrete-continuous optimization problems, which represents an alternative to the mixed integer program (MIP), led to the development of customized algorithms that successfully exploited the underlying logical structure of the problem, in both the linear [Raman & Grossmann, 1994] and nonlinear cases [Turkay & Grossmann, 1996; Lee & Grossmann, 2000]. The underlying theory of these methods, however, borrowed only in a limited way from the theories of disjunctive programming, and the unique insights from Balas' work have not been fully exploited.

In this paper, we establish new connections between the fields of disjunctive programming and generalized disjunctive programming for the linear case, which lead to new theoretical insights for linear GDP. We propose a novel family of MILP reformulations corresponding to the original linear GDP model that result in a hierarchy of tighter relaxations compared to those of Lee & Grossmann [2000] (for the linear case) and stronger cutting planes than the ones proposed by Sawaya & Grossmann [2005]. Furthermore, we integrate the latter works within the more general framework developed here for linear GDP.

[1] Current address: ExxonMobil Corporate Research Center, 1545 Route 22 East Annandale, NJ 08801

Section 1. Introduction

Generalized disjunctive programming, originally developed by Raman and Grossmann [1994], is an extension of the well-known disjunctive programming paradigm developed by Balas in the mid seventies in his seminal technical report [Balas, 1974], published 24 years later as an invited paper [Balas, 1998]. For additional work on disjunctive programming in the 1970s and 1980s, see [Balas, 1979, 1985; Balas, Tama & Tind, 1989; Blair, 1980; Jeroslow, 1977, 1987, 1989; Jeroslow & Lowe, 1984; Sherali & Shetty, 1980]. For extensions of disjunctive programming to the nonlinear case, see [Ceria & Soares, 1999; Stubbs & Mehrotra, 1999]. While disjunctive programming, for the most part, was originally developed by Balas as a unifying framework for the generation of polyhedral facets to be used in the solution of mixed-integer programs, the development of GDP in the chemical engineering community was spawned from an interest in developing an alternative modeling framework to the mixed integer program that was more adept at translating physical intuition into rigorous mathematical formalism. Indeed, while the mixed-integer programming (MIP) model is based entirely on algebraic equations and inequalities, the GDP model allows for a combination of algebraic and logical equations through disjunctions and logic propositions, which facilitates the representation of discrete decisions. This alternative mathematical representation of discrete-continuous optimization problems led to the development of customized algorithms that successfully exploited the underlying logical structure of the problem, in both the linear [Raman & Grossmann, 1994] and nonlinear cases [Turkay & Grossmann, 1996; Lee & Grossmann, 2000; Grossmann, 2002]. In particular, Raman and Grossmann [1994] developed a hybrid Branch and Bound (B&B) algorithm that can explicitly handle problems involving linear inequalities, disjunctions and symbolic logic relations for 0-1 variables. Turkay and Grossmann [1996] extended the Outer-Approximation (OA) method for solving mixed-integer nonlinear (MINLP) problems into a logical-equivalent algorithm. Lee and Grossmann [2000] (see also Grossmann & Lee [2003]) developed a disjunctive B&B method that relies on converting the nonlinear

GDP model into an equivalent MINLP model that obtains from the intersection of the convex hulls of every disjunction. Finally, Sawaya & Grossmann proposed a cutting plane method that relies on converting the GDP problem into an equivalent big-M reformulation that is successively strengthened by cuts generated from an LP in the linear case [2005].

Although good results were obtained using all these aforementioned methods for the solution of different problems in various areas of chemical engineering, including synthesis of process networks [Turkay & Grossmann, 1996; Lee & Grossmann, 2000], retrofit planning [Jackson & Grossmann, 2002; Sawaya & Grossmann, 2005], design of distillation columns [Yeomans & Grossmann, 2000; Jackson & Grossmann, 2001], and design of multi-product batch plants [Lee & Grossmann, 2000; Vecchietti & Grossmann, 2003], the underlying theory of these methods borrowed only in a limited way from the deep theoretical well of disjunctive programming, and the profound insights securely ensconced in Balas' work were never fully exploited.

The overall goal of this paper, then, is to remedy this situation. In this work, we establish new connections between the fields of disjunctive programming and generalized disjunctive programming for the linear case, which lead to new theoretical insights that allow us to exploit the rich theory developed for the former, in service of the latter. We propose a novel family of MILP reformulations corresponding to the original linear GDP model that result in a hierarchy of tighter relaxations compared to those of Lee & Grossmann [2000] (for the linear case) and stronger cutting planes than the ones proposed by Sawaya & Grossmann [2005]. Furthermore, we integrate the latter works within the more general framework developed here for linear GDP.

In section 1.1, we introduce the field of disjunctive programming as developed by Balas. We then present the mathematical formulation for the linear GDP model in section 1.2.

In section 2, we briefly review some of the underlying theory of disjunctive sets and their equivalent forms, and through newly established connections between disjunctive programming and linear GDP, extend that theory to the latter.

In section 3, we examine various MIP representations of the linear GDP model, and in particular, derive the traditional big-M reformulation and the formulation developed by Lee and Grossmann [2000] (for the linear case) through a unified disjunctive programming framework that leads to the development of a family of MIP reformulations for our linear GDP model.

In section 4, we develop a hierarchy of relaxations for linear GDP that mirror those developed by Balas for disjunctive programs [Balas, 1985]. We show that a subset of these relaxations yield tighter relaxations than the traditional big-M and Lee & Grossmann reformulations presented in section 3, and briefly review the inherent trade-offs between them.

In section 5, we generate valid cutting planes for linear GDP. We begin by describing the family of inequalities implied by the constraint set of GDP, before identifying the strongest ones amongst them (i.e. facets of the constraint set).

Finally, in section 6, we conclude with a critical review of our contributions and promising avenues for future research.

1.1. Disjunctive programming

Our presentation here is taken from that of Balas & Perregaard [2002]. Disjunctive programming is optimization over unions of polyhedra. The name reflects the fact that the objects investigated by this theory can be viewed as the solution sets of systems of linear inequalities joined by the logical operations of conjunction, negation (taking of complement) and disjunction.

The constraint set of a disjunctive program, called a disjunctive set, can be expressed in many different forms, of which the following two extreme ones have special significance. Let

$$P_i := \{x \in \mathbf{R}^n : A^i x \geq a^i\},\ i \in Q$$

be convex polyhedra, with Q a finite index set and (A^i, a^i) an $m_i \times (n+1)$ matrix, $i \in Q$, and let $P := \{x \in \mathbf{R}^n : \widehat{A}x \geq \widehat{a}\}$ be the polyhedron defined by those inequalities (if any) common to all P_i, $i \in Q$. Then the disjunctive set $F = \bigcup_{i \in Q} P_i$ over which we wish to optimize some linear function can be expressed as

$$F = \left\{ x \in \mathbf{R}^n : \underset{i \in Q}{\vee} (A^i x \geq a^i) \right\}, \tag{1.1}$$

which is its *disjunctive normal form* (DNF) (i.e. a disjunction whose terms do not contain further disjunctions). The same disjunctive set can also be expressed as

$$F = \left\{ x \in \mathbf{R}^n : \widehat{A} x \geq \widehat{a}, \ \underset{h \in Q_j}{\vee} (d^h x \geq d_0^{\ h}), \ j = 1, \ldots, t \right\}, \tag{1.2}$$

which is its *conjunctive normal form* (CNF) (i.e. a conjunction whose terms do not contain further conjunctions). Here $(d^h, d_0^{\ h})$ is a $(n+1)$ vector for $h \in Q_j$, all j. The connection between (1.1) and (1.2) is that each term $A^i x \geq a^i$ of the disjunctive normal form (1.1) contains $\widehat{A} x \geq \widehat{a}$ and exactly one inequality $d^h x \geq d_0^{\ h}$ of each disjunction of (1.2) indexed by Q_j for $j = 1, \ldots, t$, and that all distinct systems $A^i x \geq a^i$ with this property are present among the terms of (1.1).

1.2. Linear generalized disjunctive programming

Consider the linear generalized disjunctive programming problem in (1.3), which is based on the work of Raman & Grossmann [1994] and is an extension of the work of Balas on disjunctive programming [Balas, 1974]:

$$\begin{aligned}
& Min \ Z = \sum_{k \in K} c_k + d^T x \\
& s.t. \qquad Bx \geq b \\
& \bigvee_{j \in J_k} \begin{bmatrix} Y_{jk} \\ A^{jk} x \geq a^{jk} \\ c_k = \gamma_{jk} \end{bmatrix} \qquad k \in K \\
& \underset{j \in J_k}{\underline{\vee}} \ Y_{jk} \qquad k \in K \\
& \Omega(Y) = True \\
& x^L \leq x \leq x^U \\
& Y_{jk} \in \{True, False\} \qquad j \in J_k, k \in K \\
& c_k \in \mathbf{R}^1 \qquad k \in K
\end{aligned} \tag{1.3}$$

Here, $x \in \mathbf{R}^n$ is a vector of continuous variables; $Y_{jk} \in \{True, False\}$ are Boolean variables; $c_k \in \mathbf{R}^1$ are continuous variables that represent the cost associated with each disjunction; γ_{jk} are fixed charges; x^U and x^L are parameters that corresponds to valid upper and lower bounds for *x*, respectively, where x^L can be negative; $Bx \le b$ are common constraints that must hold regardless of the discrete decisions that are selected, where (B,b) is an $m \times (n+1)$ matrix. A disjunction $k \in K$ is composed of several terms *j* $\in$ J_k, each containing a set of linear equations and/or inequalities $A^{jk} x \ge a^{jk}$, where (A^{jk}, a^{jk}) is an $m_{jk} \times (n+1)$ matrix, $j \in J_k$, $k \in K$. These inequalities represent the constraints of the problem, and are connected together in the continuous space by the logical OR operator ($\vee$) and in the Boolean space by the logic propositions $\underline{\vee}_{j \in J_k} Y_{jk}$, which ensure that exactly one term per disjunction is enforced. Thus, only the constraints inside disjunct $j \in J_k$ where Y_{jk} is true are enforced; otherwise, the corresponding constraints are not enforced, although they may still hold because of the non-exclusive nature of the logical OR operator ($\vee$) connecting the continuous space. Finally, $\Omega(Y) = True$ corresponds to logical propositions in terms of the Boolean variables that are expressed in Conjunctive Normal Form (CNF)

$$\Omega(Y) = \underset{l=1,2...L}{\wedge} \left[\underset{Y_{jk} \in R_l}{\vee} (Y_{jk}) \underset{Y_{jk} \in Q_l}{\vee} (\neg Y_{jk}) \right],$$

where for each clause $l, l = 1, 2, ..., L$, R_l is the subset of Boolean variables Y_{jk} that are non-negated, and Q_l is the subset of Boolean variables Y_{jk} that are negated.

Remark 1.1 We note that in previous works on generalized disjunctive programming (see [Raman & Grossmann, 1994; Lee & Grossmann, 2000; Sawaya & Grossmann, 2005]), the collection of logic propositions $\Omega(Y) = True$ implicitly included the propositions $\underline{\vee}_{j \in J_k} Y_{jk}$ that we have explicitly stated here.

1.2.1. Illustrative example: synthesis of process network with fixed charges

Consider the optimization of the process network shown in Fig. 1, where variables x represent material flow. The problem is to determine the selection of processes that maximizes the profit or minimizes the cost given an upper bound on the demand of product C. Linear mass balances are used at each node, while fixed cost charges c_i are assumed for every process $i \in \{1,2,3\}$. The GDP formulation for this problem is as follows:

$$Min\ Z = c_1 + c_2 + c_3 + d^T x$$

$$s.t.$$

$$x_1 = x_2 + x_4 \tag{1.4}$$

$$x_6 = x_3 + x_5 \tag{1.5}$$

$$\begin{bmatrix} Y_{11} \\ x_3 = p_1 x_2 \\ c_1 = \gamma_1 \end{bmatrix} \vee \begin{bmatrix} Y_{21} \\ x_3 = x_2 = 0 \\ c_1 = 0 \end{bmatrix} \tag{1.6}$$

$$\begin{bmatrix} Y_{12} \\ x_5 = p_2 x_4 \\ c_2 = \gamma_2 \end{bmatrix} \vee \begin{bmatrix} Y_{22} \\ x_5 = x_4 = 0 \\ c_2 = 0 \end{bmatrix} \tag{1.7}$$

$$\begin{bmatrix} Y_{13} \\ x_7 = p_3 x_6 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} Y_{23} \\ x_7 = x_6 = 0 \\ c_3 = 0 \end{bmatrix} \tag{1.8}$$

$$Y_{11} \veebar Y_{21} \tag{1.9}$$

$$Y_{12} \veebar Y_{22} \tag{1.10}$$

$$Y_{13} \veebar Y_{23} \tag{1.11}$$

$$Y_{11} \vee Y_{12} \Rightarrow Y_{13} \tag{1.12}$$

$$Y_{13} \Rightarrow Y_{11} \vee Y_{12} \tag{1.13}$$

$$Y_{21} \vee Y_{22} \tag{1.14}$$

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

$$Y_{11}, Y_{21}, Y_{12}, Y_{22}, Y_{13}, Y_{23} \in \{True, False\}$$

$$c_1, c_2, c_3 \in \mathbf{R}^1$$

Equations (1.4) and (1.5) represent linear mass balances around nodes N1 and N2. Disjunctions (1.6), (1.7) and (1.8) embody the discrete dichotomy of process selection, where a unit, along with its in-and-out flows (represented by linear equalities inside the disjunctions' terms) and fixed charge, is selected for inclusion in the final network only if its corresponding Boolean variable $Y_{1k'} = True, \text{ for some } k' \in \{1,2,3\}$; otherwise, as dictated by logic equations (1.9), (1.10) and (1.11), the unit is not selected $(Y_{2k'} = True, \text{ for some } k' \in \{1,2,3\})$, and its flows and fixed charge are set to 0. Finally, logic equations (1.12), (1.13) and (1.14) connect the three disjunctions together, expressing, in respective order, that the selection of process 1 or 2 for inclusion in the network must lead to the selection of process 3; that the selection of process 3 must lead to the selection of process 1 or 2; and that process 1 and 2 cannot both be selected.

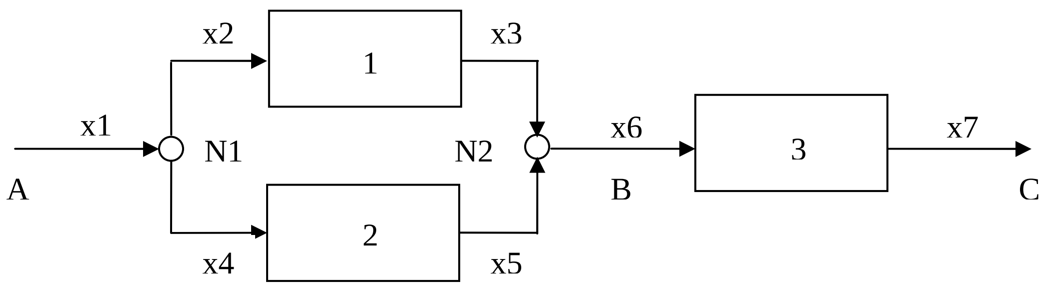

Figure 1. Superstructure for selection of processes

Section 2. Connections between disjunctive programming and linear GDP

In order to establish connections between disjunctive programming and linear generalized disjunctive programming, it is essential to convert the latter in terms of the former. This section is concerned with this objective and reviews some of the basic concepts on disjunctive programming proposed by Balas.

From Section 1, it is clear from the linear GDP model introduced in section 1.2 that it is intimately related to the disjunctive programming forms (CNF and DNF) presented in section 1.1. However, the difference between the two frameworks lies in the existence of a Boolean space in the case of GDP that allows disjunctions to be connected to one another through logical equations $\underset{j \in J_k}{\vee} Y_{jk}$ and $\Omega(Y) = True$. It becomes necessary, then, to convert a linear GDP model into an equivalent disjunctive programming model in order to exploit the theory developed by Balas. We show how to accomplish this here after briefly reviewing Balas' theory, and demonstrate how the latter theory thus becomes relevant to linear GDP.

2.1. Disjunctive programming and equivalent forms

Our presentation here is taken from Balas [1985]. The CNF and DNF forms presented in section 1.1, though two extremes of the spectrum of equivalent forms of a disjunctive set *F*, share a property *not* common to all forms: each of them is an *intersection of unions of polyhedra*. We will say that a disjunctive set that has this property is in *regular form* (RF). Thus the RF is

$$F = \underset{t \in T}{\cap} S_t \tag{2.1}$$

where for $t \in T$,

$$S_t = \underset{i \in Q_t}{\cup} P_i, \quad P_i \text{ a polyhedron, } i \in Q_t. \tag{2.2}$$

A disjunctive set S_t as in (2.2) will be called *improper* if $S_t = P_i$ for some $i \in Q_t$; *proper* otherwise. Any disjunctive set S_t such that $|Q_t| = 1$ is improper. If S_t is improper then it is convex (and polyhedral). The DNF form as in (1.1) is the RF in which $|T| = 1$. The CNF form as in (1.2), on the other hand, is the RF in which every S_t is *elementary*, i.e. every polyhedron P_i is a half-space $H^+ = \left\{ x \in \mathbf{R}^n \mid ax \geq a_0 \right\}$.We note that in (1.2), the polyhedron $\hat{A}x \geq \hat{a}$ corresponds to an improper disjunctive set that is composed of an intersection of half-spaces.

Next, we define an operation which, when applied to a disjunctive set in RF, results in another RF with one less conjuncts, i.e., an operation which brings the disjunctive set closer to the DNF. There are several advantages to having a disjunctive set in DNF, i.e., expressed as a union of polyhedra. Beyond this, the motivation for the basic step introduced here will become clearer when we discuss relaxations of disjunctive sets.

Theorem 2.1 (Theorem 2.1 in [Balas, 1985]). Let F be the disjunctive set in RF given by (2.1), (2.2). Then F can be brought to DNF by $|T|-1$ recursive applications of the following basic steps, which preserve regularity:
For some $r, s \in T, r \neq s,$ bring $S_r \cap S_s$ to DNF, by replacing it with:

$$S_{rs} = \bigcup_{\substack{i \in Q_r \\ t \in Q_s}} (P_i \cap P_t). \tag{2.3}$$

For example, let $F = S_1 \cap S_2 \cap S_3$, where $S_1 = (P_{11} \cup P_{21})$, $S_2 = (P_{12} \cup P_{22})$ and $S_3 = (P_{13} \cup P_{23})$. Then F can be brought to DNF by two recursive applications of basic step (2.3). For instance, we first apply (2.3) to $S_1 \cap S_2 = (P_{11} \cup P_{21}) \cap (P_{12} \cup P_{22})$, thus replacing it with $S_{12} = (P_{11} \cap P_{12}) \cup (P_{11} \cap P_{22}) \cup (P_{21} \cap P_{12}) \cup (P_{21} \cap P_{22})$. We can then rewrite $F = S_1 \cap S_2 \cap S_3$ as $F = S_{12} \cap S_3$. Next, we apply (2.3) to $S_{12} \cap S_3 = ((P_{11} \cap P_{12}) \cup (P_{11} \cap P_{22}) \cup (P_{21} \cap P_{12}) \cup (P_{21} \cap P_{22})) \cap (P_{13} \cup P_{23})$, thus replacing it with $S_{123} = \begin{pmatrix} (P_{11} \cap P_{12} \cap P_{13}) \cup (P_{11} \cap P_{22} \cap P_{13}) \cup (P_{21} \cap P_{12} \cap P_{13}) \cup (P_{21} \cap P_{22} \cap P_{13}) \\ \cup (P_{11} \cap P_{12} \cap P_{23}) \cup (P_{11} \cap P_{22} \cap P_{23}) \cup (P_{21} \cap P_{12} \cap P_{23}) \cup (P_{21} \cap P_{22} \cap P_{23}) \end{pmatrix}$.
We can then rewrite $F = S_{12} \cap S_3$ as $F = S_{123}$, which is its equivalent DNF. We note that the sequence of basic steps to bring F to DNF is not unique, and so (2.3) could have been first applied to S_1 and S_3 (resulting in S_{13}) followed by S_2 (resulting in S_{123}), or first to S_2 and S_3 (resulting in S_{23}) followed by S_1 (resulting in S_{123}).

Every basic step reduces by one the number of conjuncts S_t in the RF to which it is applied. On the other hand, more often than not, a basic step applied to a pair of proper disjunctive sets results in an increase in the number of polyhedra whose union is taken. This was clearly observed in the preceding example. However, when one of the disjunctive sets, say S_r, is improper, then S_{rs} is the union of at most as many polyhedra as S_s. Indeed, if $S_r = P_{i_0}$ for some $i_0 \in Q_r$ (i.e. S_r improper), then

$$S_{rs} = \begin{cases} P_{i_0} & \text{if } i_0 \in Q_s, \\ \bigcup_{t \in Q_s} (P_{i_0} \cap P_t) & \text{otherwise.} \end{cases} \tag{2.4}$$

For example, let $F = S_1 \cap S_2$, where $S_1 = P_1$ and $S_2 = (P_{12} \cup P_{22})$. Then $S_{12} = (P_1 \cap P_{12}) \cup (P_1 \cap P_{22})$.

Because of the above property, it is often useful to carry out a parallel basic step, defined as follows:

For F given by (2.1), (2.2), and $S_r = P_{i_0}$ for some $i_0 \in Q_r$ (i.e. S_r improper), replace $\bigcap_{t \in T} S_t$ by $\bigcap_{t \in T \setminus \{r\}} S_{rt}$, where each S_{rt} is defined in (2.4).

For example, let $F = S_1 \cap S_2 \cap S_3$, where $S_1 = P_1$, $S_2 = (P_{12} \cup P_{22})$ and $S_3 = (P_{13} \cup P_{23})$. Then a parallel basic step would result in $F = S_{12} \cap S_{13}$, where $S_{12} = (P_1 \cap P_{12}) \cup (P_1 \cap P_{22})$ and $S_{13} = (P_1 \cap P_{13}) \cup (P_1 \cap P_{23})$. We note that if some of the basic steps of Theorem 2.1 are replaced by parallel basic steps, the total number of steps required to bring F to DNF remains the same. Indeed, in the previous example, the total number of steps to bring F to DNF is still two, as $F = S_{123}$ can be obtained from $F = S_{12} \cap S_{13}$ by applying one additional basic step as in (2.3). The motivation for performing parallel basic steps will become clearer when we examine relaxations of disjunctive sets in section 4.

2.2. Converting a linear GDP model into a disjunctive programming model

Having examined Balas' theory of equivalent forms for disjunctive sets, we aim to extend it to generalized disjunctive programming. In order to do so, we first need to convert the linear GDP model in (1.3) into an equivalent disjunctive programming model. This is accomplished by replacing Boolean variables Y_{jk}, $j \in J_k, k \in K$ inside the disjunctions by equalities $\lambda_{jk} = 1, j \in J_k, k \in K$, where λ is a vector of continuous variables whose domain is [0, 1], and converting logical relations $\underline{\vee}_{j \in J_k} Y_{jk}$, $k \in K$ and

$\Omega(Y) = True$ into algebraic equations $\sum_{j \in J_k} \lambda_{jk} = 1, \ k \in K$ and $H\lambda \geq h$, respectively. This yields the following model:

$$
\begin{aligned}
& Min\ Z = \sum_{k \in K} c_k + d^T x \\
& s.t. \qquad Bx \geq b \\
& \bigvee_{j \in J_k} \begin{bmatrix} \lambda_{jk} = 1 \\ A^{jk} x \geq a^{jk} \\ c_k = \gamma_{jk} \end{bmatrix} \qquad && k \in K \\
& \sum_{j \in J_k} \lambda_{jk} = 1 && k \in K \\
& H\lambda \geq h \\
& x^L \leq x \leq x^U \\
& 0 \leq \lambda_{jk} \leq 1 && j \in J_k, k \in K \\
& c_k \in \mathbf{R}^1 && k \in K
\end{aligned}
\tag{2.5}
$$

In the following proposition, we prove that the linear GDP model in (1.3) is equivalent to the disjunctive program in (2.5).

Proposition 2.2. The linear GDP model in (1.3) is equivalent to the disjunctive program in (2.5), in the sense that there exists a one-to-one correspondence between a feasible solution $(x, c, Y) \in \mathbf{R}^{n+|K|} \times \{True, False\}^{\sum_{k \in K} |J_k|}$ to (1.3) and a feasible solution $(x, c, \lambda) \in \mathbf{R}^{n+|K|+\sum_{k \in K} |J_k|}$ to (2.5).

Proof: Let $(x', c', Y') \in \mathbf{R}^{n+|K|} \times \{True, False\}^{\sum_{k \in K} |J_k|}$ be a feasible solution to the constraint set of (1.3). Then, from the set of constraints

$$\begin{aligned} &\bigvee_{j\in J_k}\begin{bmatrix} Y_{jk} \\ A^{jk}x \geq a^{jk} \\ c_k = \gamma_{jk} \end{bmatrix} \qquad k\in K \\ &\underset{j\in J_k}{\underline{\vee}}\, Y_{jk} \qquad\qquad\qquad k\in K \end{aligned},$$

we can deduce that for any $k' \in K$, exactly one $\lambda_{jk'} = 1, j \in J_{k'}$. Without loss of generality, let that j be *j'*. Thus, $Y'_{j'k'} = True$ for $j' \in J_{k'}, k' \in K$, which implies that the constraints of term $j' \in J_{k'}$ for disjunction $k' \in K$ are enforced while the constraints of terms $j \setminus j' \in J_{k'}$ are not. Now let $(x^*, c^*, \lambda^*) \in \mathbf{R}^{n+|K|+\sum_{k\in K}|J_k|}$ be a feasible solution to the constraint set of (2.5). Then from the constraints

$$\begin{aligned} &\bigvee_{j\in J_k}\begin{bmatrix} \lambda_{jk} = 1 \\ A^{jk}x \geq a^{jk} \\ c_k = \gamma_{jk} \end{bmatrix} \qquad k\in K \\ &\sum_{j\in J_k}\lambda_{jk} = 1 \qquad\qquad k\in K \end{aligned}$$

we can deduce that for any $k^* \in K$, exactly one $\lambda_{jk^*} = 1, j \in J_{k^*}$. Without loss of generality, let that j be *j**. If we set $k^* = k' \in K$ and $j^* = j' \in J$, then $\lambda^*_{j'k'} = 1$, and the constraints of term $j' \in J_{k'}$ for disjunction $k' \in K$ are enforced while the constraints of terms $j \setminus j' \in J_{k'}$ are not because of the algebraic relation $\sum_{j\in J_{k'}}\lambda_{jk'} = 1,\ k' \in K$, just as the previous case. This implies that for every $Y'_{j'k'} = True$ for $j' \in J_{k'}, k' \in K$, there exists some equivalent $\lambda_{jk^*} = 1, j \in J_{k^*}, k^* \in K$, where $k^* = k' \in K$ and $j^* = j' \in J$, and such that $(x', c') \in \mathbf{R}^{n+|K|}$ is equal to $(x^*, c^*) \in \mathbf{R}^{n+|K|}$. Finally, because the above two sets of constraints force any feasible λ_{jk} to be equal to 0 or 1 exclusively (despite the fact that these variables are continuous), the inequalities $H\lambda \geq h$ can be systematically derived from their logical CNF form $\Omega(Y) = True$, (as discussed by Williams [1985], Raman and Grossmann [1994], and Hooker [2000]), and are thus equivalent.

■

It is less straightforward to see that the model in (2.5) corresponds to a disjunctive program in an intermediary form between its disjunctive and conjunctive normal forms described in (1.1) and (1.2), respectively. This, however, becomes more evident if we rewrite the feasible region of (2.5) as:

$$F = \left\{ z := (x, \lambda, c) \in \mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|} : \bigcap_{i\in\overline{T}} \overline{b}^i z \geq \overline{b}_0^{\,i} \bigcap_{k\in K} \bigcup_{j\in J_k} (\overline{A}^{jk} z \geq \overline{a}^{\,jk}) \right\}, \tag{2.6}$$

where $(\overline{b}^i, \overline{b}_0^{\,i})$ is a $1 \times (n + \sum_{k\in K} |J_k| + |K| + 1)$ vector, $i \in \overline{T}$, that corresponds to a single row of the following matrix $(\overline{B}, \overline{b})$, and where

$$\overline{B} := \begin{pmatrix} \bigcap_{i\in I_B} \big((b^i)^T & 0 & 0 \big) \\ \bigcap_{k\in K_S} \big(0 & (e^{\forall j\in J_k})^T & 0 \big) \\ \bigcap_{k\in K_S} \big(0 & (-e^{\forall j\in J_k})^T & 0 \big) \\ \bigcap_{i\in I_H} \big(0 & (h^i)^T & 0 \big) \\ \bigcap_{i\in I_X} \big((\hat{e}^i)^T & 0 & 0 \big) \\ \bigcap_{i\in I_X} \big((-\hat{e}^i)^T & 0 & 0 \big) \\ \bigcap_{(j,k)\in L} \big(0 & (e^{jk})^T & 0 \big) \\ \bigcap_{(j,k)\in L} \big(0 & (-e^{jk})^T & 0 \big) \end{pmatrix}, \quad \overline{b} := \begin{pmatrix} \bigcap_{i\in I_B} b_0^{\,i} \\ 1_{|K_S|} \\ -1_{|K_S|} \\ \bigcap_{i\in I_H} h_0^{\,i} \\ \bigcap_{i\in I_X} x_i^L \\ \bigcap_{i\in I_X} -x_i^U \\ 0_{|L|} \\ -1_{|L|} \end{pmatrix},$$

such that the vectors $(b^i)^T$, $i \in I_B$ correspond to the rows of the $|I_B| \times n$ matrix *B*, where $|I_B| = m$; the vectors $(e^{\forall j\in J_k})^T$, $k \in K_S$ correspond to the vectors $e^T := (1...1)^{\sum_{k\in K}|J_k|}$ with 1s only at *every* position $j \in J_k$, for some $k \in K_S$, where $|K_S| = |K|$; the vectors $(h^i)^T$, $i \in I_H$ correspond to the rows of the $|I_H| \times \sum_{k\in K} |J_k|$ matrix *H*; the vectors $(\hat{e}^i)^T$, $i \in I_X$ correspond to the rows of the $|I_X| \times |I_X|$ unit matrix $\hat{I}$, where $|I_X| = n$; the vectors $(e^{jk})^T$, $(j,k) \in L$ correspond to the rows of the $|L| \times |L|$ unit matrix I, where L is the set of all pairs $j \in J_k, k \in K$ with $|L| = (\sum_{k\in K} |J_k|)$; the scalars $b_0^{\,i}$, $i \in I_B$ correspond to the rows of the $|I_B| \times 1$ vector *b*; the vector $1_{|K_S|}$ is the vector $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}^{|K_S|}$; the scalars $h_0^i,\ i \in I_H$ correspond to the rows of the $|I_H| \times 1$ vector *h*; the scalars $x_i^L,\ i \in I_X$ correspond to the rows of the $|I_X| \times 1$ vector x^L; the scalars $x_i^U,\ i \in I_X$ correspond to the rows of the $|I_X| \times 1$ vector x^U; the vector $0_{|L|}$ is the vector $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}^{|L|}$; the vector $1_{|L|}$ is the vector $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}^{|L|}$; and $\overline{T} = I_B \cup K_S \cup I_H \cup I_X \cup L$ corresponds to the index set of the rows of matrix $(\overline{B}, \overline{b})$, where $|\overline{T}| = (|I_B| + 2|K_S| + |I_H| + 2|I_X| + 2|L|)$. Furthermore, $(\overline{A}^{jk}, \overline{a}^{jk})$ is an $(m_{jk} + 4) \times (n + \sum_{k \in K} |J_k| + |K| + 1)$ matrix, $j \in J_k,\ k \in K$, where

$$\overline{A}^{jk} := \begin{pmatrix} 0 & (e^{jk})^T & 0 \\ 0 & (-e^{jk})^T & 0 \\ A^{jk} & 0 & 0 \\ 0 & 0 & (\hat{e}^k)^T \\ 0 & 0 & (-\hat{e}^k)^T \end{pmatrix}, \quad \overline{a}^{jk} := \begin{pmatrix} 1 \\ -1 \\ a^{jk} \\ \gamma_{jk} \\ -\gamma_{jk} \end{pmatrix}, \qquad j \in J_k, k \in K,$$

such that the vector $(\hat{e}^k)^T$ corresponds to the vector $\hat{e}^T$ with a 1 at position $k \in K$.

We observe, then, that disjunctive set (2.6) is in *regular form* since it corresponds to an intersection of unions of polyhedra, and furthermore, is in an intermediary form between the DNF and the CNF since it is corresponds to the intersection of multiple elementary disjunctive sets $\overline{b}^i z \geq \overline{b}_0^i,\ i \in \overline{T}$ with *multiple non-elementary* disjunctive sets $\overline{A}^{jk} z \geq \overline{a}^{jk},\ j \in J_k\ k \in K$ (in contrast, the DNF corresponds to a single non-elementary disjunctive set, while the CNF corresponds to the intersection of multiple elementary disjunctive sets).

2.3. Generalized disjunctive programming and equivalent forms

Having examined the procedure to convert the linear GDP model in (1.3) into an equivalent disjunctive programming model, we extend Balas' theory of equivalent forms to generalized disjunctive programming. The following corollary follows from Theorem 2.1:

Corollary 2.3. *Let F be the disjunctive set given by (2.6). Then F can be brought to DNF by* $|\overline{T}|+|K|-1$ *recursive applications of the basic step defined in (2.3).*

Proof: We have shown that the disjunctive set given in (2.6) is in regular form. Thus, Corollary 1 follows from Proposition 1 if we replace $|T|$ with $|\overline{T}|+|K|$.

■

The above corollary implies that it is possible to develop a *family* of disjunctive equivalent forms for linear GDP by recursively applying, one at a time, and up to $|\overline{T}|+|K|-1$ times, (parallel) basic step (2.3) to the disjunctive form in (2.6). In light of this, we present, next, a formulation that captures all possible equivalent forms of that in (2.6), such that a specific form obtains from the instantiation of certain index sets as a result of performing certain (parallel) basic steps as in (2.3). This most general form for linear GDP is then as follows:

$$F=\left\{z:=(x,\lambda,c)\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\bigcap_{i\in\widehat{T}}\overline{b}^{i}z\geq\overline{b}_0^{i}\bigcap_{n\in N}\bigcup_{m\in J_n}\left(\bigcap_{i\in\widehat{T}_n}\overline{b}^{i}z\geq\overline{b}_0^{i}\bigcap_{(j,k)\in M_{mn}}(\overline{A}^{jk}z\geq\overline{a}^{jk})\right)\right\}, \quad (2.7)$$

where $\widehat{T}\subseteq\overline{T}$ represents the index set of those rows $(\overline{b}^{i},\overline{b}_0^{i})$, $i\in\overline{T}$ that were *not* intersected with any disjunctive sets $(\overline{A}^{jk}z\geq\overline{a}^{jk})$, $j\in J_k, k\in K$, while $\widehat{T}_n\subseteq\overline{T}$, $n\in N$ represent the index sets of those rows $(\overline{b}^{i},\overline{b}_0^{i})$, $i\in\overline{T}$ that were intersected with some disjunctive set $(\overline{A}^{jk}z\geq\overline{a}^{jk})$, $j\in J_k, k\in K$ through the application of some (parallel) basic step(s) as in (2.3). Furthermore, N represents the index set of those disjunctions $n\in N$ that are either:

(a) identical to some old disjunctions $k\in K$ (i.e. those disjunctions to which no basic step was applied), or that obtain from the intersection of disjunctive sets $(\overline{A}^{jk}z \geq \overline{a}^{jk}),\ j \in J_k, k \in K$ either:

(b) with one another, or

(c) with those rows $(\overline{b}^i, \overline{b}_0^i),\ i \in \overline{T}_n$, or

(d) with both, through the application of some (parallel) basic step(s) as in (2.3).

Finally, the index sets $M_{mn}, m \in J_n, n \in N$ contain the collection of pairs (j,k) that correspond to the indices of those constraints $(\overline{A}^{jk}z \geq \overline{a}^{jk}),\ j \in J_k, k \in K$ that are present in terms $m \in J_n$ of disjunctions $n \in N$.

Remark 2.1. For disjunctions $n^{'} \in N$ in category

(a), $\widehat{T}_{n^{'}} = \varnothing$ and $|M_{mn^{'}}| = 1,\ m \in J_{n^{'}}$

(b), $\widehat{T}_{n^{'}} = \varnothing$ and $|M_{mn^{'}}| > 1,\ m \in J_{n^{'}}$

(c), $\widehat{T}_{n^{'}} \neq \varnothing$ and $|M_{mn^{'}}| = 1,\ m \in J_{n^{'}}$

(d), $\widehat{T}_{n^{'}} \neq \varnothing$ and $|M_{mn^{'}}| > 1,\ m \in J_{n^{'}}$.

Remark 2.2. If $\widehat{T} = \overline{T}$ (which implies that $\widehat{T}_n = \varnothing,\ n \in N$ since $\widehat{T} \underset{n \in N}{\cup} \widehat{T}_n = \overline{T}$), and $|M_{mn}| = 1,\ m \in J_n,\ n \in N$, then (2.7) is equivalent to (2.6). If $\widehat{T} = \varnothing$ and $|\widehat{K}| = 1$, then (2.7) is equivalent to the DNF of (2.6).

For the sake of presentational clarity in subsequent sections, we concatenate the contents of disjunctions $n \in N$ and rewrite (2.7) as

$$F = \left\{ z := (x, \lambda, c) \in \mathbf{R}^{n + \sum_{k \in K} |J_k| + |K|} : \bigcap_{i \in \widehat{T}} \overline{b}^i z \geq \overline{b}_0^i \bigcap_{n \in N} \bigcup_{m \in J_n} (\hat{A}^{mn} z \geq \hat{a}^{mn}) \right\}. \qquad (2.8)$$

Here, $(\overline{b}^i, \overline{b}_0^i)$ is a $1 \times (n + \sum_{k \in K} |J_k| + |K| + 1)$ vector, $i \in \widehat{T}$, that corresponds to a single row of the following matrix $(\widehat{B}, \widehat{b})$, where

$$\widehat{B} := \begin{pmatrix} \bigcap_{i\in I_{B_1}} \left((b^i)^T \quad 0 \quad 0 \right) \\ \bigcap_{k\in K_{S_1}} \left(0 \quad (e^{\forall j\in J_k})^T \quad 0 \right) \\ \bigcap_{k\in K_{S_1}} \left(0 \quad (-e^{\forall j\in J_k})^T \quad 0 \right) \\ \bigcap_{i\in I_{H_1}} \left(0 \quad (h^i)^T \quad 0 \right) \\ \bigcap_{i\in I_{X_1}} \left((\widehat{e}^i)^T \quad 0 \quad 0 \right) \\ \bigcap_{i\in I_{X_1}} \left((-\widehat{e}^i)^T \quad 0 \quad 0 \right) \\ \bigcap_{(j,k)\in L_1} \left(0 \quad (e^{jk})^T \quad 0 \right) \\ \bigcap_{(j,k)\in L_1} \left(0 \quad (-e^{jk})^T \quad 0 \right) \end{pmatrix}, \quad \widehat{b} := \begin{pmatrix} \bigcap_{i\in I_{B_1}} b_0{}^i \\ 1_{|K_{S_1}|} \\ -1_{|K_{S_1}|} \\ \bigcap_{i\in I_{H_1}} h_0{}^i \\ \bigcap_{i\in I_{X_1}} x_i{}^L \\ \bigcap_{i\in I_{X_1}} -x_i{}^U \\ 0_{|L_1|} \\ -1_{|L_1|} \end{pmatrix},$$

such that $I_{B_1} \subseteq I_B$; $K_{S_1} \subseteq K_S$; $I_{H_1} \subseteq I_H$; $I_{X_1} \subseteq I_X$; $L_1 \subseteq L$; and $\widehat{T} = I_{B_1} \cup K_{S_1} \cup I_{H_1} \cup I_{X_1} \cup L_1$, where $|\widehat{T}| = (|I_{B_1}| + 2|K_{S_1}| + |I_{H_1}| + 2|I_{X_1}| + 2|L_1|)$. Furthermore, $(\hat{A}^{mn}, \hat{a}^{mn})$ is an $(|M_{mn}|(m_{jk}+4) + |I_{B_{2_n}}| + 2|K_{S_{2_n}}| + |I_{H_{2_n}}| + 2|I_{X_{2_n}}| + 2|L_{2_n}|) \times (n + \sum_{k\in K} |J_k| + |K| + 1)$ matrix, $m \in J_n$, $n \in N$, where

$$\hat{A}^{mn} := \begin{pmatrix} \underset{(j,k)\in M_{mn}}{\cap} \overline{A}^{jk} \\ \underset{i\in I_{B_{2_n}}}{\cap} \begin{pmatrix} (b^i)^T & 0 & 0 \end{pmatrix} \\ \underset{k\in K_{S_{2_n}}}{\cap} \begin{pmatrix} 0 & (e^{\forall j\in J_k})^T & 0 \end{pmatrix} \\ \underset{k\in K_{S_{2_n}}}{\cap} \begin{pmatrix} 0 & (-e^{\forall j\in J_k})^T & 0 \end{pmatrix} \\ \underset{i\in I_{H_{2_n}}}{\cap} \begin{pmatrix} 0 & (h^i)^T & 0 \end{pmatrix} \\ \underset{i\in I_{X_{2_n}}}{\cap} \begin{pmatrix} (\widehat{e}^i)^T & 0 & 0 \end{pmatrix} \\ \underset{i\in I_{X_{2_n}}}{\cap} \begin{pmatrix} (-\widehat{e}^i)^T & 0 & 0 \end{pmatrix} \\ \underset{(j,k)\in L_{2_n}}{\cap} \begin{pmatrix} 0 & (e^{jk})^T & 0 \end{pmatrix} \\ \underset{(j,k)\in L_{2_n}}{\cap} \begin{pmatrix} 0 & (-e^{jk})^T & 0 \end{pmatrix} \end{pmatrix}, \quad \hat{a}^{mn} := \begin{pmatrix} \underset{(j,k)\in M_{mn}}{\cap} \overline{a}^{jk} \\ \underset{i\in I_{B_{2_n}}}{\cap} b_0^{\ i} \\ 1_{|K_{S_{2_n}}|} \\ -1_{|K_{S_{2_n}}|} \\ \underset{i\in I_{H_{2_n}}}{\cap} h_0^{\ i} \\ \underset{i\in I_{X_{2_n}}}{\cap} x_i^{\ L} \\ \underset{i\in I_{X_{2_n}}}{\cap} -x_i^{\ U} \\ 0_{|L_{2_n}|} \\ -1_{|L_{2_n}|} \end{pmatrix}, \qquad m\in J_n, n\in N,$$

such that $I_{B_{2_n}} \subseteq I_B$; $K_{S_{2_n}} \subseteq K_S$; $I_{H_{2_n}} \subseteq I_H$; $I_{X_{2_n}} \subseteq I_X$; $L_{2_n} \subseteq L$; and $\widehat{T}_n = I_{B_{2_n}} \cup K_{S_{2_n}} \cup I_{H_{2_n}} \cup I_{X_{2_n}} \cup L_{2_n}$, $n \in N$, where $|\widehat{T}_n| = (|I_{B_{3_n}}| + 2|K_{S_{2_n}}| + |I_{H_{2_n}}| + 2|I_{X_{2_n}}| + 2|L_{2_n}|)$, $n \in N$.

Remark 2.3. The set $I_B = I_{B_1} \underset{n\in N}{\cup} I_{B_{2_n}}$, while the set $I_{B_1} \cap \underset{n\in N}{\cup} I_{B_{2_n}} = \varnothing$. Furthermore, the set $\underset{n\in N}{\cap} I_{B_{2_n}}$ is not necessarily empty, and this occurs when a *parallel* basic step is applied to rows $(\overline{b}^i, \overline{b}_0^{\ i})$, $i \in \overline{T}$, intersecting the latter with constraints $\overline{A}^{mn} z \geq \overline{a}^{mn}$, $n \in N$. The same logic holds for index sets $K_S = K_{S_1} \underset{n\in N}{\cup} K_{S_{2_n}}$, $I_H = I_{H_1} \underset{n\in N}{\cup} I_{H_{2_n}}$, $I_X = I_{X_1} \underset{n\in N}{\cup} I_{X_{2_n}}$ and $L = L_1 \underset{n\in N}{\cup} L_{2_n}$.

Section 3. MIP reformulations for linear GDP

A solution approach to linear GDP problems is to reformulate them as mixed-integer linear programs, rather than developing specific solution methods. The reformulation, however, is not unique. Ideally one would like to obtain that reformulation that leads to the tightest LP relaxation, while keeping the problem at reasonable size. This, however, is non-trivial to achieve. This section provides a framework based on disjunctive programming to formulate mixed-integer linear programs.

Mixed-integer programming reformulations for the linear GDP problem in (1.3) from Section 1 can be obtained in various ways. As such, we use the disjunctive form in (2.6) of Section 2 to derive the traditional big-M reformulation as presented in Raman & Grossmann [1994]. Furthermore, we exploit the results of section 2.3 to reconstruct the reformulation presented in Lee & Grossmann (for the linear case) that obtains from the intersection of the convex hulls of every disjunction (see also Grossmann & Lee [2003], Sawaya & Grossmann for the linear case [2005]), and show that it belongs to a family of MIP reformulations that are equivalent to our original GDP model.

3.1. Big-M reformulation for GDP

Raman & Grossmann [1994] proposed the following mixed integer big-M reformulation for the linear generalized disjunctive program in (1.3):

$$
\begin{aligned}
& Min\ Z = \sum_{k \in K} \sum_{j \in J_k} \gamma_{jk} y_{jk} + d^T x \\
& s.t. \qquad Bx \geq b \\
& A^{jk} x \geq a^{jk} - M^{jk}(1 - y_{jk}) \qquad j \in J_k, k \in K \\
& \sum_{j \in J_k} y_{jk} = 1 \qquad k \in K \\
& Hy \geq h \\
& x^L \leq x \leq x^U \\
& y_{jk} \in \{0,1\} \qquad j \in J_k, k \in K
\end{aligned}
\tag{3.1}
$$

In the following proposition, we show that the above reformulation obtains from disjunctive form (2.6).

Proposition 3.1. The Raman & Grossmann mixed-integer big-M reformulation in (3.1) for the linear generalized disjunctive program in (1.3) obtains from disjunctive form (2.6).

Proof: Consider the disjunctive form in (2.6). The big-M reformulation for this form can be obtained by replacing the constraints of disjunctions $j \in J_k, k \in K$ with constraints that make use of "big-M" parameters M^{jk} and binary variables $y_{jk}, j \in J_k, k \in K$, such that the j^{th} system of inequalities in the k^{th} disjunction is enforced when $y_{jk} = 1$, or rendered redundant when $y_{jk} = 0$. This results in the following MIP model:

$$
\begin{array}{lll}
Min\ Z = \sum_{k \in K} c_k + d^T x & & \\
s.t. \qquad Bx \geq b & & \\
\lambda_{jk} \leq 1 + M_1^{jk}(1 - y_{jk}) & j \in J_k, k \in K & (3.2) \\
\lambda_{jk} \geq 1 - M_2^{jk}(1 - y_{jk}) & j \in J_k, k \in K & (3.3) \\
A_{jk} x \geq a_{jk} - M_3^{jk}(1 - y_{jk}) & j \in J_k, k \in K & (3.4) \\
c_k \leq \gamma_{jk} + M_4^{jk}(1 - y_{jk}) & j \in J_k, k \in K & (3.5) \\
c_k \geq \gamma_{jk} - M_5^{jk}(1 - y_{jk}) & j \in J_k, k \in K & (3.6) \\
\sum_{j \in J_k} \lambda_{jk} = 1 & k \in K & (3.7) \\
\sum_{j \in J_k} y_{jk} = 1 & k \in K & (3.8) \\
H\lambda \geq h & & \\
x^L \leq x \leq x^U & & \\
0 \leq \lambda_{jk} \leq 1 & j \in J_k, k \in K & (3.9) \\
c_k \in \mathbf{R}^1 & k \in K & \\
y_{jk} \in \{0,1\} & j \in J_k, k \in K &
\end{array}
$$

The above model can be simplified if we observe that $y_{jk} = \lambda_{jk}, \forall j \in J_k, \forall k \in K$. Indeed,

$$y_{j'k'} = 1, \text{ for some } j' \in J_{k'}, k' \in K$$

$$\Rightarrow \begin{cases} y_{jk'} = 0, j \setminus j' \in J_{k'}, k' \in K \text{ from (3.8)} \\ \lambda_{j'k'} = 1 \text{ from (3.2) \& (3.3)} \\ \qquad \Rightarrow \lambda_{jk'} = 0, j \setminus j' \in J_{k'}, k' \in K \text{ from (3.7)} \end{cases}$$

Similarly, the above can be shown for $y_{j'k'} = 0,$ for some $j' \in J_{k'}, k' \in K.$

Thus, we can replace all instances of λ with y, and remove redundant constraints (3.2), (3.3), (3.7) and (3.9). Furthermore, we note that (3.5) and (3.6) can be replaced by the equivalent expression

$$c_k = \sum_{j \in J_k} \gamma_{jk} y_{jk}, \; k \in K, \tag{3.10}$$

whose values at $y = 1$ and $y = 0$ coincide exactly with those of constraints (3.5) and (3.6). This is shown as follows:

For (3.5) and (3.6)

$$y_{j'k'} = 1, \text{ for some } j' \in J_{k'}, k' \in K$$

$$\Rightarrow \begin{cases} c_{k'} = \gamma_{j'k'} \; j' \in J_{k'}, k' \in K \text{ from (3.5) \& (3.6)} \\ y_{jk'} = 0, j \setminus j' \in J_{k'}, k' \in K \text{ from (3.8)} \\ \qquad \Rightarrow c_{k'} \in \mathbf{R}^1, j \setminus j' \in J_{k'}, k' \in K \text{ from (3.5) \& (3.6)} \end{cases}$$

$$\Rightarrow c_{k'} = \gamma_{j'k'}$$

For (3.10)

$$y_{j'k'} = 1, \text{ for that same } j' \in J_{k'}, k' \in K$$

$$\Rightarrow y_{jk'} = 0, j \setminus j' \in J_{k'}, k' \in K \text{ from (3.8)}$$

$$\Rightarrow c_{k'} = \gamma_{j'k'} \text{ from (3.10)}$$

Similarly, the above can be shown for $y_{j'k'} = 0,$ for some $j' \in J_{k'}, k' \in K$

The equivalence between (3.10) and (3.5) & (3.6) allows for the elimination of the latter constraints and for the substitution of the former into the objective function, thus removing the variable c from the program. These modifications lead exactly to reformulation in (3.1).

■

3.2. Lee & Grossmann reformulation for GDP

Lee & Grossmann [2000] proposed a valid mixed-integer representation of a generalized disjunctive program. For the linear generalized disjunctive program in (1.3), their reformulation translates into the following mixed-integer program (see also Sawaya and Grossmann [2005]):

$$\begin{aligned}
&Min\ Z = \sum_{k\in K}\sum_{j\in J_k}\gamma_{jk}y_{jk} + d^T x \\
&s.t. \qquad Bx \ge b \\
&x = \sum_{j\in J_k} v^{jk} && k\in K \\
&A^{jk}v^{jk} \ge a^{jk}y_{jk} && j\in J_k, k\in K \\
&x^L y_{jk} \le v^{jk} \le x^U y_{jk} && j\in J_k, k\in K \\
&\sum_{j\in J_k} y_{jk} = 1 && k\in K \\
&Hy \ge h \\
&y_{jk}\in\{0,1\} && j\in J_k, k\in K
\end{aligned} \tag{3.11}$$

In order to reconstruct the above formulation using disjunctive programming theory, we make use of the following theorems from Balas [1985]. The first theorem expresses the convex hull of a disjunctive set as the projection of a higher dimensional polyhedron onto $\mathbf{R}^n$.

Theorem 3.2. (Theorem 3.3 and Theorem 3.4 in [Balas, 1985]). Let $F = \bigcup_{i\in Q} P_i$, $P_i = \left\{z\in\mathbf{R}^n : \tilde{A}^i z \ge \tilde{a}_0^{\,i}\right\}$, $i\in Q$, where Q is an arbitrary set and each $(\tilde{A}^i, \tilde{a}_0^{\,i})$

is an $m_i \times (n+1)$ matrix. Furthermore, let the set $\zeta(Q)$ be the set of all those $z \in \mathbf{R}^n$ such that there exist vectors $(v^i, y_i) \in \mathbf{R}^{n+1},\ i \in Q$, satisfying

$$
\begin{aligned}
& z - \sum_{i \in Q} v^i = 0 && \\
& \tilde{A}^i v^i - \tilde{a}_0{}^i y_i \geq 0 && i \in Q \\
& y_i \geq 0 && i \in Q, \\
& \sum_{i \in Q} y_i = 1 && i \in Q
\end{aligned}
$$

where the cone $C_q = \{\overline{z} \in \mathbf{R}^n \mid \tilde{A}^i \overline{z} \geq 0\}$ satisfies the condition $C_q \subseteq \sum_{i \in Q^*} C_i,\ \forall q \in Q \setminus (Q^* \equiv \{i \in Q \mid P_i \neq \varnothing\})$. Then $cl\ conv\ F = \zeta(Q)$.

The next theorem expresses the claim that any disjunctive program that satisfies a technical condition on the recession directions of its polyhedra can be represented as the mixed-integer program $\zeta_I(Q) := \left\{ z \in \zeta(Q) : y_i \in \{0,1\},\ i \in Q \right\}$.

Theorem 3.3. (Theorem 3.6 in [Balas, 1985]). Let $F = \cup_{i \in Q} P_i,\ Q^* = \{i \in Q \mid P_i \neq \varnothing\},\ Q^{**} = \{i \in Q^* \mid P_i \subseteq P_j, \forall j \in Q^* \setminus \{i\}\}$. If F satisfies

$$C_i = C_j \quad \forall i, j \in Q^{**}$$

and

$$C_q = C_i \quad \forall q \in Q \setminus Q^*, i \in Q^{**}$$

then

$$\zeta_I(Q) = F.$$

We are now ready to reconstruct Lee and Grossmann's formulation. In the following proposition, we show that the formulation in (3.11) obtains from the application of basic and parallel basic steps to disjunctive form (2.6).

Proposition 3.4: The Lee & Grossmann formulation presented in (3.11) for the linear generalized disjunctive program in (1.3) obtains from the application of a series of $2|I_X|$ *parallel basic steps* and $2\sum_{k\in K}|J_k|$ *basic steps* to the disjunctive form in (2.6).

Proof: Consider the disjunctive form in (2.6). Then a series of $2|I_X|$ parallel basic steps and $2\sum_{k\in K}|J_k|$ basic steps to (2.6) results in the intersection of relevant bounds on x and λ with $\bar{A}^{jk}z \geq \bar{a}^{jk}$, $j \in J_k, k \in K$, as follows:

$$F = \left\{ z \in \mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|} : \left(\begin{array}{l} \bigcap_{i\in I_B} \begin{pmatrix} (b^i)^T & 0 & 0 \end{pmatrix} z \geq b_0^{\,i} \\ \bigcap_{k\in K_S} \begin{pmatrix} 0 & (e^{\forall j\in J_k})^T & 0 \end{pmatrix} z \geq 1_{|K_S|} \\ \bigcap_{k\in K_S} \begin{pmatrix} 0 & (-e^{\forall j\in J_k})^T & 0 \end{pmatrix} z \geq -1_{|K_S|} \\ \bigcap_{i\in I_H} \begin{pmatrix} 0 & (h^i)^T & 0 \end{pmatrix} z \geq h_0^{\,i} \\ \bigcap_{k\in K} \bigcup_{j\in J_k} \left(\begin{array}{l} \bar{A}^{jk} z \geq \bar{a}^{jk} \\ \bigcap_{i\in I_X} \begin{pmatrix} (\hat{e}^i)^T & 0 & 0 \end{pmatrix} z \geq x_i^L \\ \bigcap_{i\in I_X} \begin{pmatrix} (-\hat{e}^i)^T & 0 & 0 \end{pmatrix} z \geq -x_i^U \\ \bigcap_{\hat{j}\in J_k} \begin{pmatrix} 0 & (e^{\hat{j}k})^T & 0 \end{pmatrix} z \geq 0_{|J_k|} \\ \bigcap_{\hat{j}\in J_k} \begin{pmatrix} 0 & (-e^{\hat{j}k})^T & 0 \end{pmatrix} z \geq -1_{|J_k|} \end{array} \right) \end{array} \right) \right\}, \qquad (3.12)$$

Since every term $j \in J_k$ of disjunction $k \in K$ in (3.12) corresponds to a non-empty and bounded polyhedron, we can apply Theorem 3.3. This results in the following MIP:

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

$$s.t. \qquad Bx \geq b$$

$$\lambda_{jk} = \sum_{\hat{j} \in J_k} u^{\hat{j}}{}_{jk} \qquad j \in J_k,\ k \in K \tag{3.13}$$

$$x = \sum_{j \in J_k} v^{jk} \qquad k \in K \tag{3.14}$$

$$c_k = \sum_{j \in J_k} w^{jk} \qquad k \in K \tag{3.15}$$

$$u^{\hat{j}}{}_{jk} = y_{\hat{j}k} \qquad \hat{j}, j = \hat{j} \in J_k, k \in K \tag{3.16}$$

$$A^{jk} v^{jk} \geq a^{jk} y_{jk} \qquad j \in J_k, k \in K \tag{3.17}$$

$$w^{jk} = \gamma_{jk} y_{jk} \qquad j \in J_k, k \in K \tag{3.18}$$

$$0 \leq u^{\hat{j}}{}_{jk} \leq y_{\hat{j}k} \qquad \hat{j}, j \in J_k, k \in K \tag{3.19}$$

$$x^L y_{jk} \leq v^{jk} \leq x^U y_{jk} \qquad j \in J_k, k \in K \tag{3.20}$$

$$\sum_{j \in J_k} \lambda_{jk} = 1 \qquad k \in K \tag{3.21}$$

$$\sum_{\hat{j} \in J_k} y_{\hat{j}k} = 1 \qquad k \in K \tag{3.22}$$

$$H\lambda \geq h$$

$$y_{jk} \in \{0,1\} \qquad j \in J_k, k \in K$$

Here, $\hat{j} \in J_k$ is an alias for $j \in J_k$. Also, as in the case of the big-M reformulation, the above model can be significantly simplified if we observe that $y_{\hat{j}k} = \lambda_{\hat{j}k},\ \forall \hat{j} \in J_k, \forall k \in K$. Indeed,

$y_{\hat{j}'k'} = 1, \text{ for some } \hat{j}' \in J_{k'}, k' \in K$

$$\Rightarrow \begin{cases} u^{\hat{j}'}{}_{jk'} = 1, \ \hat{j}', j = \hat{j}' \in J_{k'}, k' \in K \text{ from (3.16)} & (3.23) \\ y_{\hat{j}k'} = 0, \ \hat{j} \setminus \hat{j}' \in J_{k'}, k' \in K \text{ from (3.22)} \\ \quad \Rightarrow 0 \le u^{\hat{j}}{}_{jk'} \le 0, \ \hat{j} \setminus \hat{j}', j \in J_{k'}, k' \in K \text{ from (3.19)} \\ \quad\quad \Rightarrow u^{\hat{j}}{}_{jk'} = 0, \ \hat{j} \setminus \hat{j}', j \in J_{k'}, k' \in K & (3.24) \end{cases}$$

We can rewrite (3.13) as

$$\lambda_{jk'} = u^{\hat{j}'}{}_{jk'} + \sum_{\hat{j} \setminus \hat{j}' \in J_{k'}} u^{\hat{j}}{}_{jk'}, \ j = \hat{j}' \in J_{k'}, k' \in K \quad (3.25)$$

$$\lambda_{jk'} = u^{\hat{j}}{}_{jk'} + \sum_{\hat{j} \setminus \hat{j}' \in J_{k'}} u^{\hat{j}}{}_{jk'}, \ j \ne \hat{j}' \in J_{k'}, k' \in K$$

For (3.25)

$\lambda_{jk'} = 1 + 0 = 1, \ j = \hat{j}' \in J_{k'}, k' \in K$ from (3.23) and (3.24)

$\Rightarrow \lambda_{jk'} = 0, \ j \ne \hat{j}' \in J_{k'}, k' \in K$ from (3.21)

Similarly, the above can be shown for $y_{\hat{j}'k'} = 0, \text{ for some } \hat{j}' \in J_{k'}, k' \in K.$

Thus, we can replace all instances of λ with y, and remove redundant constraints (3.13), (3.16), (3.19) and (3.21). Furthermore, we note that (3.15) and (3.18) can be combined into the equivalent expression

$$c_k = \sum_{j \in J_k} \gamma_{jk} y_{jk}, \ k \in K,$$

which we can then substitute into the objective function, thus eliminating $c_k, k \in K$ from the program. These modifications lead exactly to the reformulation in (3.11).

■

3.3. A family of MIP reformulations for GDP

It is clear from section 3.2 that the Lee and Grossmann reformulation for the linear generalized disjunctive program in (1.3) belongs to a larger family of MIP reformulations for linear GDP. Indeed, its feasible region was shown to be a valid MIP representation of the disjunctive program in (3.12), the latter having been obtained from (2.6) by a series of basic and parallel basic steps that amounted, in essence, to intersecting every term J_k of every disjunction $k \in K$ with the relevant bounds on $\lambda_{jk}, j \in J_k, k \in K$ and *x*. Thus, it is possible to develop a family of MIP reformulations for linear GDP by applying, one at a time, (parallel) basic steps as in (2.3) to the disjunctive form in (2.6), and subsequently applying Theorem 3.3 to the resulting form. Equivalently, since these basic and parallel basic steps result in the instantiation of certain index sets of the formulation in (2.8) (which serves as a template for all possible equivalent forms of GDP), it is possible to develop a family of MIP reformulations for linear GDP by transforming (2.8) into a MIP, and then instantiating, one at a time, the index sets of (2.8). Doing so, however, raises two issues which warrant mentioning.

Firstly, when transforming a disjunctive program into a MIP, we need to ensure that the constraints within every term $j \in J_k$ and $m \in J_n$ of disjunctions $n \in N$ represent polyhedra that satisfy the conditions (on recession cones) presented in Theorem 3.3. Secondly, in cases where $|M_{mn'}| > 1,\ m \in J_{n'}, n' \in N$, applying Theorem 3.3 to disjunction $n' \in N$ in (2.8) often results in more total binary variables being used than in cases where $|M_{mn''}| = 1,\ m \in J_{n''}, n'' \neq n' \in N$. Indeed, the former scenario is a result of intersecting *proper* disjunctive sets in (2.6) through the application of some basic step(s) (we call such a basic step a proper basic step), which often leads to the creation of new disjunction(s) with many more terms than those of the originally intersected disjunctions. In addition, the number of terms of these newly created disjunctions increases exponentially with every recursive proper basic step applied to them, resulting in the need of an ever increasing number of binary variables to germanely represent the original disjunctive program as a MIP. As performing proper basic steps can have substantial benefits (as examined in the following section on relaxations of disjunctive programs), there is a need to address this issue. Fortunately, Theorem 4.4 in [Balas, 1985] deals precisely with this problem. We state it as follows:

Theorem 3.5. (Theorem 4.4 in [Balas, 1985]): Let the disjunctive set $F = \bigcap_{j \in T} S_j$, where each S_j is a union of polyhedra, be in regular form, and assume that *F* satisfies the conditions of Theorem 3.3. Furthermore, let F_0 be the disjunctive set in CNF consisting of those $z \in \mathbb{R}^n$ satisfying

$$\bigvee_{s \in Q_r} (a^s z \ge a^s{}_0), \quad r \in T_0 \tag{3.26}$$

and let *F* be the same set in regular form obtained from F_0 by some sequence of basic steps, given as the set of $z \in \mathbb{R}^n$ satisfying

$$\bigvee_{i \in Q_j} (A^i z \ge a^i{}_0), \quad j \in T, \tag{3.27}$$

where every $j \in T$ would then correspond to some subset T_{0j} of T_0, with $T_0 = \bigcup_{j \in T} T_{0j}$, such that the disjunction in (3.27) indexed by *j* is the disjunctive normal form of the set of disjunctions in (3.26) indexed by T_{0j}. Finally, let M_i be the index set of the inequalities $a^s z \ge a^s{}_0$ making up the system $A^i z \ge a^i{}_0$. Then the constraint set (3.27) is equivalent to the following constraint set

$$\begin{aligned}
& z - \sum_{i \in Q_j} v^i = 0 && j \in T, \\
& \tilde{A}^i v^i - \tilde{a}_0{}^i y_i \ge 0 && i \in Q_j, j \in T, \\
& y_i \ge 0 && i \in Q_j, j \in T, \\
& \sum_{i \in Q_j} y_i = 1 && j \in T, \\
& \sum_{i \in Q_j | s \in M_i} y_i = \delta_s{}^r && s \in Q_r, r \in T_0, \\
& \sum_{s \in Q_r} \delta_s{}^r = 1 && r \in T_0, \\
& \delta_s{}^r \in \{0,1\} && s \in Q_r, r \in T_0
\end{aligned} \tag{3.28}$$

in that for every solution to (3.27), there exist vectors $(v^i, y_i) \in \mathbf{R}^{n+1}$, $i \in Q_j$, $j \in T$ and scalars $\delta_s{}^r$, $s \in Q_r, r \in T_0$ that together with *z* satisfy (3.28); and conversely, the *z*-component of any solution to (3.28) is a solution to (3.27).

Thus, in order to obtain a valid MIP representation that requires as many binary variables in cases where $|M_{mn'}|>1,\ m\in J_{n'}, n'\in N$ as in cases where $|M_{mn''}|=1,\ m\in J_{n''}, n''\neq n'\in N$, we apply Theorem 3.5 to disjunctions $n\in N$. Note that in cases where $|M_{mn''}|=1,\ m\in J_{n''}, n''\neq n'\in N$, Theorem 3.5 reduces to Theorem 3.3.

In light of the issues discussed above, we present the following MIP reformulation of (2.8):

$$
\begin{aligned}
& Min\ Z=\sum_{k\in K}\sum_{j\in J_k}\gamma_{jk}y_{jk}+d^T x \\
& s.t. \\
& b^i x\geq b_0^{\,i} && i\in I_{B_1} \\
& h^i y\geq h_0^{\,i} && i\in I_{H_1} \\
& x_i^L\leq x_i\leq x_i^U && i\in I_{X_1} \\
& y_{jk}=\sum_{m\in J_n}\hat{u}_{jk}^{\,mn} && (j,k)\in L_{2_n}\cup K_{S_{2_n}}\cup I_{H_{2_n}}, n\in N \\
& x=\sum_{m\in J_n}\hat{v}^{mn} && n\in N \\
& b^i\hat{v}^{mn}\geq b_0^{\,i}\hat{y}_{mn} && i\in I_{B_{2_n}}, m\in J_n, n\in N \\
& \sum_{j\in J_k}\hat{u}_{jk}^{\,mn}=\hat{y}_{mn} && k\in K_{S_{2_n}}, m\in J_n, n\in N \\
& h^i\hat{u}^{mn}\geq h_0^{\,i}\hat{y}_{mn} && i\in I_{H_{2_n}}, m\in J_n, n\in N \\
& \hat{u}_{jk}^{\,mn}=\hat{y}_{mn} && (j,k)\in M_{mn}, m\in J_n, n\in N \\
& A^{jk}\hat{v}^{mn}\geq a^{jk}\hat{y}_{mn} && (j,k)\in M_{mn}, m\in J_n, n\in N \\
& x^L\hat{y}_{mn}\leq\hat{v}^{mn}\leq x^U\hat{y}_{mn} && m\in J_n, n\in N \\
& 0\leq\hat{u}_{jk}^{\,mn}\leq\hat{y}_{mn} && (j,k)\in L_{3_n}, m\in J_n, n\in N \\
& \sum_{m\in J_n}\hat{y}_{mn}=1 && n\in N \\
& \sum_{m\in Q_{n_{jk}}}\hat{y}_{mn_{jk}}=y_{jk} && n_{jk}\in N, j\in J_k, k\in K \\
& \sum_{j\in J_k}y_{jk}=1 && k\in K \\
& \hat{y}_{mn}\geq 0 && m\in J_n, n\in N \\
& y_{jk}\in\{0,1\} && j\in J_k, k\in K
\end{aligned}
\tag{3.29}
$$

We note that in the constraints $\sum_{m\in Q_{n_{jk}}} \hat{y}_{mn_{jk}} = y_{jk}$, $n_{jk} \in N, j \in J_k, k \in K$, the index $n_{jk} \in N$ represents that disjunction $n \in N$ that contains the intersected constraints of term $j \in J_k$, $k \in K$ in some of its terms, while the index set $Q_{n_{jk}}$ refers to these latter terms. For instance, let $S_1 = (P_{11} \cup P_{21})$ and $S_2 = (P_{12} \cup P_{22} \cup P_{32})$ correspond to two disjunctions of set K, where P_{11} and P_{21} represent the constraints of disjunction $S_1 \in K$, and P_{12}, P_{22} and P_{32} represent the constraints of disjunction $S_2 \in K$. The application of a proper basic step to S_1 and S_2 results in the new disjunction $S_{12} = (P_{11} \cap P_{12}) \cup (P_{11} \cap P_{22}) \cup (P_{11} \cap P_{32}) \cup (P_{21} \cap P_{12}) \cup (P_{21} \cap P_{22}) \cup (P_{21} \cap P_{32})$, where $S_{12} \in N$. In this case, the set of constraints $\sum_{m\in Q_{n_{jk}}} \hat{y}_{mn_{jk}} = y_{jk}$, $n_{jk} \in N, j \in J_k, k \in K$ are such that

$$
\begin{aligned}
&\hat{y}_{1S_{12}} + \hat{y}_{2S_{12}} + \hat{y}_{3S_{12}} = y_{1S_1} \\
&\hat{y}_{4S_{12}} + \hat{y}_{5S_{12}} + \hat{y}_{6S_{12}} = y_{2S_1} \\
&\hat{y}_{1S_{12}} + \hat{y}_{4S_{12}} = y_{1S_2} \\
&\hat{y}_{2S_{12}} + \hat{y}_{5S_{12}} = y_{2S_2} \\
&\hat{y}_{3S_{12}} + \hat{y}_{6S_{12}} = y_{3S_2},
\end{aligned}
$$

where $J_{S_1} = \{1,2\}$, $J_{S_2} = \{1,2,3\}$, $\hat{n}_{\hat{j}\hat{k}} = S_{12}$, $Q_{S_{12\,1S_1}} = \{1,2,3\}$, $Q_{S_{12\,2S_1}} = \{4,5,6\}$, $Q_{S_{12\,1S_2}} = \{1,4\}$, $Q_{S_{12\,2S_2}} = \{2,5\}$, $Q_{S_{12\,3S_2}} = \{3,6\}$. Finally, we note that the above MIP formulation in (3.29) is in simplified form, in the sense that the variables λ were replaced by the variables y, and redundant constraints $\sum_{j\in J_k} \lambda_{jk} = 1,\ k \in K_{S_1}$ were eliminated from the formulation while all relevant constraints containing c were substituted into the objective function.

Section 4. LP relaxations for linear GDP

A key property of MIP reformulations of GPD problems is the quality of their LP relaxations. This section examines a hierarchy of relaxations that can be obtained from different MIP reformulations.

Having derived different MIP reformulations for the original GDP problem in (1.3) from Section 1, it is straightforward to obtain their LP relaxations. In this section, we first present relaxations for the Big-M and Lee & Grossmann reformulations in (3.1) and (3.11), respectively. We then propose a hierarchy of relaxations based on the concept of *hull-relaxation* that mirror those of Balas [1985], and establish theoretical properties regarding the relative tightness of these relaxations. Finally, we discuss the inherent trade-offs between size and tightness of reformulations.

4.1. Big-M and Lee & Grossmann relaxations for GDP

The Big-M and Lee and Grossmann relaxations are obtained by relaxing the binary variables in models (3.1) and (3.11), respectively. The Big-M relaxation is then as follows [Raman & Grossmann, 1994]:

$$
\begin{aligned}
& Min\ Z = \sum_{k\in K}\sum_{j\in J_k}\gamma_{jk}y_{jk} + d^T x \\
& s.t. \quad Bx \ge b \\
& A^{jk}x \ge a^{jk} - M^{jk}(1-y_{jk}) \qquad j\in J_k, k\in K \\
& \sum_{j\in J_k} y_{jk} = 1 \qquad k\in K \\
& Hy \ge h \\
& x^L \le x \le x^U \\
& 0 \le y_{jk} \le 1 \qquad j\in J_k, k\in K
\end{aligned}
\tag{4.1}
$$

The Lee & Grossmann relaxation is then as follows [Lee & Grossmann, 2000]:

$$
\begin{aligned}
& Min\ Z = \sum_{k \in K} \sum_{j \in J_k} \gamma_{jk} y_{jk} + d^T x \\
& s.t. \qquad Bx \ge b \\
& x = \sum_{j \in J_k} v^{jk} \qquad k \in K \\
& A^{jk} v^{jk} \ge a^{jk} y_{jk} \qquad j \in J_k, k \in K \\
& x^L y_{jk} \le v^{jk} \le x^U y_{jk} \qquad j \in J_k, k \in K \\
& \sum_{j \in J_k} y_{jk} = 1 \qquad k \in K \\
& Hy \ge h \\
& 0 \le y_{jk} \le 1 \qquad j \in J_k, k \in K
\end{aligned} \tag{4.2}
$$

In comparing the relaxations in (4.1) and (4.2), the following trade-offs can be observed. On the one hand, the size of the formulation in (4.2) is considerably larger than the size of that in (4.1) because of the addition of disaggregated variables $v^{jk},\ j \in J_k, k \in K$ and convex hull constraints. On the other hand, the projected feasible region of the formulation in (4.2) onto the *(x,y)* space is at least as tight, if not tighter, than the feasible region of the formulation in (4.1). The latter is proven in Grossmann & Lee [2003] for the general (nonlinear) case. We re-state this claim for the linear case in the following proposition.

Proposition 4.1. Let the feasible region of the big-M relaxation in (4.1) be defined as FR_{BM} and that of the Lee & Grossmann relaxation in (4.2) as $FR_{L\&G}$. Furthermore, let us define the projection of $FR_{L\&G}$ onto the (x,y) space as $FR_{PL\&G} := \{(x,y) : \exists v : (x,v,y) \in FR_{L\&G}\}$. Then $FR_{PL\&G} \subseteq FR_{BM}$.

4.2. A hierarchy of relaxations for GDP

We now present a hierarchy of relaxations for the most general disjunctive form in (2.8), based on the concept of *hull-relaxation.* According to Balas [1985], the hull-relaxation of a disjunctive set $F = \bigcap_{j \in T} S_j$ in regular form, where each S_j is a union of polyhedra, is denoted as $h-rel\ F$ and defined as

$$h-rel\ F := \bigcap_{j \in T} clconv\ S_j .$$

The hull-relaxation of F is then the formulation that results from intersecting the convex hulls of *every* union of polyhedra $S_j,\ j \in T$.

Following the above concept, we obtain the hull-relaxation of the most general disjunctive form in (2.8) by applying Theorem 3.2 to every disjunction in (2.8) – under the assumption that every disjunction in (2.8) satisfies the conditions of Theorem 3.2. After simplification, we obtain the following:

$$
\begin{aligned}
&Min\ Z = \sum_{k\in K}\sum_{j\in J_k}\gamma_{jk}y_{jk} + d^T x \\
&s.t. \\
&b^i x \geq b_0{}^i && i \in I_{B_1} \\
&h^i y \geq h_0{}^i && i \in I_{H_1} \\
&x_i^L \leq x_i \leq x_i^U && i \in I_{X_1} \\
&y_{jk} = \sum_{m\in J_n}\hat{u}_{jk}{}^{mn} && (j,k)\in L_{2_n} \cup K_{S_{2_n}} \cup I_{H_{2_n}}, n\in N \\
&x = \sum_{m\in J_n}\hat{v}^{mn} && n \in N \\
&b^i\hat{v}^{mn} \geq b_0{}^i\hat{y}_{mn} && i\in I_{B_{2_n}}, m\in J_n, n\in N \\
&\sum_{j\in J_k}\hat{u}_{jk}{}^{mn} = \hat{y}_{mn} && k\in K_{S_{2_n}}, m\in J_n, n\in N \\
&h^i\hat{u}^{mn} \geq h_0{}^i\hat{y}_{mn} && i\in I_{H_{2_n}}, m\in J_n, n\in N \\
&\hat{u}_{jk}{}^{mn} = \hat{y}_{mn} && (j,k)\in M_{mn}, m\in J_n, n\in N \\
&A^{jk}\hat{v}^{mn} \geq a^{jk}\hat{y}_{mn} && (j,k)\in M_{mn}, m\in J_n, n\in N \\
&x^L\hat{y}_{mn} \leq \hat{v}^{mn} \leq x^U\hat{y}_{mn} && m\in J_n, n\in N \\
&0 \leq \hat{u}_{jk}{}^{mn} \leq \hat{y}_{mn} && (j,k)\in L_{3_n}, m\in J_n, n\in N \\
&\sum_{m\in J_n}\hat{y}_{mn} = 1 && n\in N \\
&\sum_{m\in Q_{n_{jk}}}\hat{y}_{mn_{jk}} = y_{jk} && n_{jk}\in N, j\in J_k, k\in K \\
&\sum_{j\in J_k} y_{jk} = 1 && k\in K \\
&\hat{y}_{mn} \geq 0 && m\in J_n, n\in N \\
&0 \leq y_{jk} \leq 1 && j\in J_k, k\in K
\end{aligned}
\tag{4.3}
$$

We can now exploit the following theorem by Balas in order to generate a hierarchy of relaxations for GDP:

Theorem 4.2. (Theorem 4.3 in [Balas, 1985]): For $i = 0,1,\ldots,t,$ let $F_i = \bigcap_{j\in T_i} S_j$ be a sequence of regular forms of a disjunctive set, such that

i) F_0 is in CNF, with $P_0 = \bigcap_{j\in T_0} S_j$;

ii) F_t is in DNF;

iii) for $i = 1, \ldots, t$, F_i is obtained from F_{i-1} by a basic step.

Then,

$$P_0 = h-rel\ F_0 \supseteq h-rel\ F_1 \supseteq \cdots \supseteq h-rel\ F_t = clconv\ F_t.$$

It was previously shown that the disjunctive form in (2.6) was in regular form and in an intermediary form between the CNF and DNF. Since any form in (4.3) is obtained through the application of some (parallel) basic steps(s) which instantiates the various index sets previously described, it is clear that every form of (4.3) is in regular form. The following corollary thus holds:

Corollary 4.3. For $i = 0,1,...,|\overline{T}|+|K|-1$, let F_{GDP_i} be a sequence of regular forms of the disjunctive set in (2.8), such that

i) F_{GDP_0} corresponds to the disjunctive form in (2.6);

ii) $F_{GDP_{|\overline{T}|+|K|-1}} := F_t$ is in DNF;

iii) for $i = 1, \ldots, t$, F_{GDP_i} is obtained from $F_{GDP_{i-1}}$ by a basic step.

Then,

$$h-rel\ F_{GDP_0} \supseteq FR_{L\&G} = h-rel\ F_{GDP_{(2|I_X|+2\sum_{k\in K}|J_k|)}} \supseteq \cdots \supseteq h-rel\ F_{GDP_{|\overline{T}|+|K|-1}} = clconv\ F_{GDP_{|\overline{T}|+|K|-1}} = clconv\ F_t.$$

■

4.3 Trade-offs between size and tightness of relaxations

As we have seen in the previous section, the relaxations of a disjunctive program become tighter as the disjunctive form approaches the DNF through the application of (parallel) basic steps. This tightening, however, comes at the cost of an increase in the number of variables and constraints in the algebraic reformulation. Although this observation holds in the most general of cases, there may be particular cases where we can exploit the explicit logical structure of the GDP model such that the problem size of the relaxation can be significantly reduced without compromising tightness. We illustrate this in the following example.

4.3.1. Illustrative example (cont'd): synthesis of process network with fixed charges

We rewrite the GDP model in section 1.2.1 as the following disjunctive program (DP1):

$$Min\ Z = c_1 + c_2 + c_3 + d^T x$$
$$s.t.$$
$$x_1 = x_2 + x_4$$
$$x_6 = x_3 + x_5$$
$$\begin{bmatrix} \lambda_{11} = 1 \\ x_3 = p_1 x_2 \\ c_1 = \gamma_1 \end{bmatrix} \vee \begin{bmatrix} \lambda_{21} = 1 \\ x_3 = x_2 = 0 \\ c_1 = 0 \end{bmatrix} \tag{4.4}$$
$$\begin{bmatrix} \lambda_{12} = 1 \\ x_5 = p_2 x_4 \\ c_2 = \gamma_2 \end{bmatrix} \vee \begin{bmatrix} \lambda_{22} = 1 \\ x_5 = x_4 = 0 \\ c_2 = 0 \end{bmatrix} \tag{4.5}$$
$$\begin{bmatrix} \lambda_{13} = 1 \\ x_7 = p_3 x_6 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \lambda_{23} = 1 \\ x_7 = x_6 = 0 \\ c_3 = 0 \end{bmatrix} \tag{4.6}$$
$$\lambda_{11} + \lambda_{21} = 1$$
$$\lambda_{12} + \lambda_{22} = 1$$
$$\lambda_{13} + \lambda_{23} = 1$$
$$\lambda_{13} \ge \lambda_{11}$$
$$\lambda_{13} \ge \lambda_{12}$$
$$\lambda_{11} + \lambda_{12} \ge \lambda_{13}$$
$$\lambda_{21} + \lambda_{22} \ge 1$$
$$0 \le x \le x^U$$
$$\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \lambda_{13}, \lambda_{23} \ge 0$$
$$c_1, c_2, c_3 \in \mathbf{R}^1$$

If we intersect disjunctions (4.4), (4.5) and (4.6) through the application of two consecutive proper basic steps, we obtain the following disjunctive program (DP2):

$$Min\ Z = c_1 + c_2 + c_3 + d^T x$$
$$s.t.$$
$$x_1 = x_2 + x_4$$
$$x_6 = x_3 + x_5$$

$$\begin{bmatrix} \lambda_{11}=1 \\ \lambda_{12}=1 \\ \lambda_{13}=1 \\ x_3 = p_1 x_2 \\ x_5 = p_2 x_4 \\ x_7 = p_3 x_6 \\ c_1 = \gamma_1 \\ c_2 = \gamma_2 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \lambda_{11}=1 \\ \lambda_{12}=1 \\ \lambda_{23}=1 \\ x_3 = p_1 x_2 \\ x_5 = p_2 x_4 \\ x_7 = x_6 = 0 \\ c_1 = \gamma_1 \\ c_2 = \gamma_2 \\ c_3 = 0 \end{bmatrix} \vee \begin{bmatrix} \lambda_{11}=1 \\ \lambda_{22}=1 \\ \lambda_{13}=1 \\ x_3 = p_1 x_2 \\ x_5 = x_4 = 0 \\ x_7 = p_3 x_6 \\ c_1 = \gamma_1 \\ c_2 = 0 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \lambda_{11}=1 \\ \lambda_{22}=1 \\ \lambda_{23}=1 \\ x_3 = p_1 x_2 \\ x_5 = x_4 = 0 \\ x_7 = x_6 = 0 \\ c_1 = \gamma_1 \\ c_2 = 0 \\ c_3 = 0 \end{bmatrix}$$

$$\vee \begin{bmatrix} \lambda_{21}=1 \\ \lambda_{12}=1 \\ \lambda_{13}=1 \\ x_3 = x_2 = 0 \\ x_5 = p_2 x_4 \\ x_7 = p_3 x_6 \\ c_1 = 0 \\ c_2 = \gamma_2 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \lambda_{21}=1 \\ \lambda_{12}=1 \\ \lambda_{23}=1 \\ x_3 = x_2 = 0 \\ x_5 = p_2 x_4 \\ x_7 = x_6 = 0 \\ c_1 = 0 \\ c_2 = \gamma_2 \\ c_3 = 0 \end{bmatrix} \vee \begin{bmatrix} \lambda_{21}=1 \\ \lambda_{22}=1 \\ \lambda_{13}=1 \\ x_3 = x_2 = 0 \\ x_5 = x_4 = 0 \\ x_7 = p_3 x_6 \\ c_1 = 0 \\ c_2 = 0 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \lambda_{21}=1 \\ \lambda_{22}=1 \\ \lambda_{23}=1 \\ x_3 = x_2 = 0 \\ x_5 = x_4 = 0 \\ x_7 = x_6 = 0 \\ c_1 = 0 \\ c_2 = 0 \\ c_3 = 0 \end{bmatrix} \qquad (4.7)$$

$$\lambda_{11} + \lambda_{21} = 1 \qquad (4.8)$$
$$\lambda_{12} + \lambda_{22} = 1 \qquad (4.9)$$
$$\lambda_{13} + \lambda_{23} = 1 \qquad (4.10)$$
$$\lambda_{13} \ge \lambda_{11} \qquad (4.11)$$
$$\lambda_{13} \ge \lambda_{12} \qquad (4.12)$$
$$\lambda_{11} + \lambda_{12} \ge \lambda_{13} \qquad (4.13)$$
$$\lambda_{21} + \lambda_{22} \ge 1 \qquad (4.14)$$
$$0 \le x \le x^U$$
$$\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \lambda_{13}, \lambda_{23} \ge 0$$
$$c_1, c_2, c_3 \in \mathbf{R}^1$$

If the relaxation of (DP2) is taken as in (4.3), then from Corollary 4.3, it is clearly at least as tight as that of (DP1), if not much tighter. On the other hand, it requires a significant number of additional variables and constraints relative to that of (DP1) (see Table 1).

However, if we use logic constraints (4.8)-(4.14) to eliminate those terms of disjunction (4.7) that are infeasible, the size of (DP2)'s relaxation can be significantly reduced. Indeed, the first and second terms violate constraint (4.14) because $\lambda_{21} = 0$ and $\lambda_{22} = 0$ from (4.8) and (4.9); the fourth term violates constraint (4.11) and the sixth term violates constraint (4.12), because $\lambda_{13} = 0$ from (4.10); and the seventh term violates constraint (4.13) because $\lambda_{11} = 0$ and $\lambda_{12} = 0$ from (4.8) and (4.9). Thus, only the third, fifth and eighth terms of disjunction (4.7) are feasible, which leads to the following disjunctive program (DP3):

$$
\begin{aligned}
& Min\ Z = c_1 + c_2 + c_3 + d^T x \\
& s.t. \\
& x_1 = x_2 + x_4 \\
& x_6 = x_3 + x_5 \\
& \begin{bmatrix} \lambda_{11} = 1 \\ \lambda_{22} = 1 \\ \lambda_{13} = 1 \\ x_3 = p_1 x_2 \\ x_5 = x_4 = 0 \\ x_7 = p_3 x_6 \\ c_1 = \gamma_1 \\ c_2 = 0 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \lambda_{21} = 1 \\ \lambda_{12} = 1 \\ \lambda_{13} = 1 \\ x_3 = x_2 = 0 \\ x_5 = p_2 x_4 \\ x_7 = p_3 x_6 \\ c_1 = 0 \\ c_2 = \gamma_2 \\ c_3 = \gamma_3 \end{bmatrix} \vee \begin{bmatrix} \lambda_{21} = 1 \\ \lambda_{22} = 1 \\ \lambda_{23} = 1 \\ x_3 = x_2 = 0 \\ x_5 = x_4 = 0 \\ x_7 = x_6 = 0 \\ c_1 = 0 \\ c_2 = 0 \\ c_3 = 0 \end{bmatrix} \\
& \lambda_{11} + \lambda_{21} = 1 \\
& \lambda_{12} + \lambda_{22} = 1 \\
& \lambda_{13} + \lambda_{23} = 1 \\
& 0 \le x \le x^U \\
& \lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \lambda_{13}, \lambda_{23} \ge 0 \\
& c_1, c_2, c_3 \in \mathbf{R}^1
\end{aligned}
$$

The above disjunctive program's relaxation has been significantly reduced in size. Indeed, the number of variables for (DP3)'s relaxation is only marginally larger than that of (DP1), while the number of constraints is marginally smaller (see Table 1).

Table 1.

Number of variables and constraints for relaxation of (DP1), (DP2) and (DP3)

	Number of Variables	Number of Constraints
DP1	24	42
DP2	57	67
DP3	27	41

Section 5. Disjunctive cutting planes for linear GDP: facets of the hull-relaxation

Having derived a hierarchy of relaxations for linear GDP in Section 4, we are now interested in generating valid cutting planes for linear GDP by exploiting these relaxations. As the hull-relaxations were produced in a higher-dimensional space that often requires many additional variables and constraints, we seek to circumvent this drawback by generating valid cutting planes in the original space of the problem. We begin by describing the family of inequalities implied by the constraint set of GDP as described in its most general form in (2.8) from Section 2, and then identify the strongest ones amongst them (i.e. facets of the hull-relaxations).

5.1 Valid inequalities

As previously remarked, we are interested in the family of inequalities implied by (2.8), which represents the constraint set of GDP in its most general form. An inequality $\beta x \geq \beta_0$ is said to be implied by, or is a consequence of, an inequality $\alpha x \geq \alpha_0$ if every x that satisfies $\alpha x \geq \alpha_0$ also satisfies $\beta x \geq \beta_0$. Thus, all valid cutting planes for (2.8) belong to this family of inequalities implied by (2.8). A characterization of this family is given in the following proposition, which mirrors Balas' Theorem 3.3.1 in [Balas, 1979]:

Proposition 5.1. The inequality $\alpha z \geq \alpha_0$ is a consequence of

$$F = \left\{ z := (x, \lambda, c) \in \mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|} : \bigcap_{i\in \widehat{T}} \overline{b}^i z \geq \overline{b}_0^{\,i} \bigcap_{n\in N} \bigcup_{m\in J_n} (\hat{A}^{mn} z \geq \hat{a}^{mn}) \right\},$$

if and only if there exists a set of $\hat{\theta}^{mn} \geq 0,\ m \in J_n, n \in N;$ and $\Gamma^i,\ i \in \widehat{T}$, satisfying

$$\alpha \geq \sum_{i\in \widehat{T}} \Gamma^i \overline{b}^i + \sum_{n\in N} \hat{\theta}^{mn} \hat{A}^{mn},\ m \in J_n$$

$$\alpha_0 \leq \sum_{i\in \widehat{T}} \Gamma^i \overline{b}_0^{\,i} + \sum_{n\in N} \hat{\theta}^{mn} \hat{a}^{mn},\ m \in J_n$$

Proof: Let $n^{'} \in N$. Then from Theorem 3.3.1 in Balas [1979], the inequality $\beta^{n^{'}} z \geq \beta_0^{\,n^{'}}$ is a consequence of the constraint

$$\bigcup_{m\in J_{n'}} (\hat{A}^{mn'} z \geq \hat{a}^{mn'}),$$

if and only if there exists a set of $\hat{\delta}^{mn'} \geq 0,\ m \in J_{n'}, n' \in N$ satisfying

$$\beta^{n'} \geq \hat{\delta}^{mn'} \hat{A}^{mn'},\ m \in J_{n'}, n' \in N$$

$$\beta_0^{\,n'} \leq \hat{\delta}^{mn'} \hat{a}^{mn'},\ m \in J_{n'}, n' \in N.$$

Clearly then, this implies that $\beta^{n^{'}} z \geq \beta_0^{\,n^{'}}$ is a consequence of

$$clconv \bigcup_{m\in J_{n'}} (\hat{A}^{mn'} z \geq \hat{a}^{mn'}),$$

since $\beta^{n^{'}} z \geq \beta_0^{\,n^{'}}$ is a consequence of every term of the disjunction, and thus, must be a consequence of the smallest convex set containing all the terms of the disjunction (i.e. the convex hull).

This implies that the system of constraints $\beta^n z \geq \beta_0^{\,n},\ n \in N$ is a consequence of the constraints

$$clconv \bigcup_{m\in J_n} (\hat{A}^{mn} z \geq \hat{a}^{mn}),\ n \in N$$

if and only if there exists a set of $\hat{\delta}^{mn} \geq 0,\ m \in J_n, n \in N$ satisfying

$$\beta^n \geq \hat{\delta}^{mn} \hat{A}^{mn},\ m \in J_n, n \in N$$

$$\beta_0^{\,n} \leq \hat{\delta}^{mn} \hat{a}^{mn},\ m \in J_n, n \in N.$$

From Theorem 22.3 in Rockafellar [1970], the inequality $\Delta z \ge \Delta_0$ is a consequence of the system of constraints

$$\bigcap_{i\in \hat{T}} \bar{b}^i z \ge \bar{b}_0^i$$

if and only if there exists a set of $\xi^i \ge 0,\ i \in \hat{T}$ satisfying

$$\Delta \ge \sum_{i\in \hat{T}} \xi^i \bar{b}^i$$
$$\Delta_0 \le \sum_{i\in \hat{T}} \xi^i \bar{b}_0^i.$$

Finally, $\alpha z \ge \alpha_0$ is a consequence of the constraints

$$\bigcap_{i\in \hat{T}} \bar{b}^i z \ge \bar{b}_0^i \bigcap_{n\in N} clconv \bigcup_{m\in J_n} (\hat{A}^{mn} z \ge \hat{a}^{mn}) := h-rel\ F_{GDP_i},\ i = 0,1,...,|\bar{T}|+|K|-1$$

if and only if it is a consequence of any

$$\Delta z \ge \Delta_0 \text{, where } (\Delta, \Delta_0) \in \left\{ \begin{array}{l} \Delta \ge \sum_{i\in \hat{T}} \xi^i \bar{b}^i \\ \Delta_0 \le \sum_{i\in \hat{T}} \xi^i \bar{b}_0^i \\ \xi^i \ge 0,\ i \in \hat{T} \end{array} \right\},$$

and of any

$$\beta^n z \ge \beta_0^n,\ n \in N \text{, where } (\beta^n, \beta_0^n) \in \left\{ \begin{array}{l} \beta^n \ge \hat{\delta}^{mn} \hat{A}^{mn},\ m \in J_n, n \in N \\ \beta_0^n \le \hat{\delta}^{mn} \hat{a}^{mn},\ m \in J_n, n \in N \\ \hat{\delta}^{mn} \ge 0,\ m \in J_n, n \in N \end{array} \right\}.$$

Thus, $\alpha z \ge \alpha_0$ is a consequence of the constraints

$$\bigcap_{i\in \hat{T}} \bar{b}^i z \ge \bar{b}_0^i \bigcap_{n\in N} clconv \bigcup_{m\in J_n} (\hat{A}^{mn} z \ge \hat{a}^{mn}) := h-rel\ F_{GDP_i},\ i = 0,1,...,|\bar{T}|+|K|-1$$

if and only if there exists a set of $\Lambda \ge 0,\ \hat{\lambda}_n \ge 0, n \in N$; $\xi^i \ge 0,\ i \in \hat{T}$; and $\hat{\delta}^{mn} \ge 0,\ m \in J_n, n \in N$ satisfying

$$\alpha \geq \Lambda\Delta + \sum_{n\in\hat{K}} \hat{\lambda}_n \beta^n \tag{5.1}$$

$$\alpha_0 \leq \Lambda\Delta_0 + \sum_{n\in\hat{K}} \hat{\lambda}_n \beta_0{}^n \tag{5.2}$$

$$\Delta \geq \sum_{i\in\hat{T}} \xi^i \overline{b}^i \tag{5.3}$$

$$\Delta_0 \leq \sum_{i\in\hat{T}} \xi^i \overline{b}_0{}^i \tag{5.4}$$

$$\beta^n \geq \hat{\delta}^{mn} \hat{A}^{mn},\ m \in J_n, n \in N \tag{5.5}$$

$$\beta_0{}^n \leq \hat{\delta}^{mn} \hat{a}^{mn},\ m \in J_n, n \in N \tag{5.6}$$

The system of inequalities (5.1)-(5.6) can be simplified as follows:

$$\alpha \geq \Lambda\Delta + \sum_{n\in\hat{K}} \hat{\lambda}_n \beta^n \qquad \text{from (5.1)}$$

$$\Rightarrow \alpha \geq \Lambda \sum_{i\in\hat{T}} \xi^i \overline{b}^i + \sum_{n\in N} \hat{\lambda}_n \hat{\delta}^{mn} \hat{A}^{mn},\ j \in J_k, m \in J_n \qquad \text{from (5.3),(5.5)}$$

$$\Rightarrow \alpha \geq \sum_{i\in\hat{T}} \Gamma^i \overline{b}^i + \sum_{n\in N} \hat{\theta}^{mn} \hat{A}^{mn},\ j \in J_k, m \in J_n$$

where

$$\Gamma^i := \Lambda \xi^i,\ i \in \widehat{T}$$

$$\hat{\theta}^{mn} := \hat{\lambda}_n \hat{\delta}^{mn},\ m \in J_n, n \in N.$$

Similarly,

$$\alpha_0 \leq \Lambda\Delta_0 + \sum_{n\in\hat{K}} \hat{\lambda}_n \beta_0{}^n \qquad \text{from (5.2)}$$

$$\Rightarrow \alpha_0 \leq \Lambda \sum_{i\in\hat{T}} \xi^i \overline{b}_0{}^i + \sum_{n\in N} \hat{\lambda}_n \hat{\delta}^{mn} \hat{a}^{mn},\ j \in J_k, m \in J_n \qquad \text{from (5.4),(5.6)}$$

$$\Rightarrow \alpha_0 \leq \sum_{i\in\hat{T}} \Gamma^i \overline{b}_0{}^i + \sum_{n\in N} \hat{\theta}^{mn} \hat{a}^{mn},\ j \in J_k, m \in J_n$$

■

5.2. Facets of the hull-relaxation

Having described the family of valid inequalities for (2.8), we are interested in identifying the strongest amongst them: the facets of the hull-relaxation. Our presentation here follows that of Balas in [1979]. Let

$$F=\left\{z:=(x,\lambda,c)\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\bigcap_{i\in\widehat{T}}\overline{b}^i z\geq\overline{b}_0^i\bigcap_{n\in N}\bigcup_{m\in J_n}(\hat{A}^{mn}z\geq\hat{a}^{mn})\right\}$$

Then for a given scalar α_0, the family of inequalities $\alpha z\geq\alpha_0$ satisfied by all $z\in F$ is isomorphic to the family of vectors $\alpha\in F^{\#}{}_{(\alpha_0)}$, where

$$F^{\#}{}_{(\alpha_0)}=\left\{\psi\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\psi z\geq\alpha_0,\ \forall z\in F\right\},$$

in the sense that $\alpha z\geq\alpha_0$ is a valid inequality if and only if $\alpha\in F^{\#}{}_{(\alpha_0)}$. We can now describe the facets of the hull-relaxation of (2.8) as follows:

Proposition 5.2. $\alpha z\geq\alpha_0$ with $\alpha_0\neq 0$ is a facet of the $h-rel\ F_{GDP_i}$, $i=0,1,...,|\overline{T}|+|K|-1$ if and only if $\alpha\neq 0$ is a vertex of the polyhedron

$$\left\{\psi\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}\left|\begin{array}{l}\psi\geq\sum_{i\in\widehat{T}}\Gamma^i\overline{b}^i+\sum_{n\in N}\hat{\theta}^{mn}\hat{A}^{mn},\ j\in J_k,m\in J_n\\ \alpha_0\leq\sum_{i\in\widehat{T}}\Gamma^i\overline{b}_0^i+\sum_{n\in N}\hat{\theta}^{mn}\hat{a}^{mn},\ j\in J_k,m\in J_n\\ \hat{\theta}^{mn}\geq 0,\ m\in J_n,n\in N\\ \Gamma^i,\ i\in\widehat{T}\end{array}\right.\right\}. \qquad (5.7)$$

Proof: From Proposition 5.1, the set $F^{\#}{}_{(\alpha_0)}=\left\{\psi\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\psi z\geq\alpha_0,\ \forall z\in F\right\}$ is of the form claimed above. The rest is a direct application of Balas' Theorems 5.5 and 5.6 in [Balas, 1979].

■

5.2.1. An alternative way of describing facets of the hull-relaxation

It is possible to describe an alternative system of inequalities to (5.7) that corresponds to the facets of the hull-relaxation. We show this in the following proposition and corollary:

Proposition 5.3. Every facet of $h-rel\ F_{GDP_i},\ i=0,1,...,|\overline{T}|+|K|-1$ as described in (5.7) is a facet of $F_n=\left\{z:=(x,\lambda,c)\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\bigcap_{i\in\overline{T}}\ \overline{b}^i z\geq\overline{b}_0^{\,i}\cap clconv\bigcup_{m\in J_n}(\hat{A}^{mn}z\geq\hat{a}^{mn})\right\}$ for some $n\in N$.

Proof: Let us represent F_n as follows:

$$F_n=\left\{z:=(x,\lambda,c)\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\bigcap_{i\in\overline{T}}\ \overline{b}^i z\geq\overline{b}_0^{\,i}\cap clconv\bigcup_{m\in J_n}(\hat{A}^{mn}z\geq\hat{a}^{mn})\right\}$$
$$:=\left\{z:=(x,\lambda,c)\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\pi^i z\geq\pi^i{}_0,\ i\in\hat{\Theta}_n\right\},$$

where $\hat{\Theta}_n,\ n\in N$ indexes the facets of F_n. Therefore,

$$\bigcap_{n\in N}F_n=\left\{z:=(x,\lambda,c)\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\left\{\begin{array}{l}\bigcap_{i\in\overline{T}}\ \overline{b}^i z\geq\overline{b}_0^{\,i}\\ \bigcap_{n\in N}clconv\bigcup_{m\in J_n}(\hat{A}^{mn}z\geq\hat{a}^{mn})\end{array}\right\}\right\}$$
$$:=\left\{z:=(x,\lambda,c)\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\pi^i z\geq\pi^i{}_0,\ i\in\bigcup_{n\in N}\hat{\Theta}_n\right\}.$$

Thus,

$$h-rel\ F_{GDP_i},\ i=0,1,...,|\overline{T}|+|K|-1:=\left\{z:=(x,\lambda,c)\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}:\pi^i z\geq\pi^i{}_0,\ i\in\bigcup_{n\in N}\hat{\Theta}_n\right\}.$$

■

Corollary 5.4. $\alpha z\geq\alpha_0$ with $\alpha_0\neq 0$ is a facet of the $h-rel\ F_{GDP_i},\ i=0,1,...,|\overline{T}|+|K|-1$ if and only if $\alpha\neq 0$ is a vertex of the polyhedron

$$\left\{\psi\in\mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|}\ \middle|\ \begin{array}{l}\psi\geq\sum_{i\in\overline{T}}\Gamma^i\overline{b}^i+\hat{\theta}^{mn}\hat{A}^{mn},\ m\in J_n\\ \alpha_0\leq\sum_{i\in\overline{T}}\Gamma^i\overline{b}_0^{\,i}+\hat{\theta}^{mn}\hat{a}^{mn},\ m\in J_n\\ \hat{\theta}^{mn}\geq 0,\ m\in J_n,n\in N\\ \Gamma^i,\ i\in\overline{T}\end{array}\right\},\qquad(5.8)$$

Remark 5.1. The above corollary defines the facets of every individual disjunction $n \in N$. Thus, although $h-rel\ F_{GDP_i},\ i=0,1,...,|\bar{T}|+|K|-1 := \left\{ z := (x,\lambda,c) \in \mathbf{R}^{n+\sum_{k\in K}|J_k|+|K|} : \pi^i z \geq \pi^i{}_0,\ i \in \bigcup_{n\in N} \hat{\Theta}_n \right\}$, it is clear that every facet of every individual disjunction is not necessarily a facet of $h-rel\ F_{GDP_i},\ i=0,1,...,|\bar{T}|+|K|-1$. In other words, there may exist (as is often the case) redundant constraints in the description of the hull-relaxation given above.

Section 6. Conclusion

In this paper, we established novel connections between disjunctive programming and linear GDP by providing the disjunctive programming equivalent of a linear GDP. We extended Balas' theory of equivalent forms to linear GDP by making use of the above transformation, which allows us to obtain equivalent linear GDP formulations to the original linear GDP problem. We developed a family of MIP reformulations for linear GDP, and showed that the Lee & Grossmann formulation is a particular instance of this family. We then developed a hierarchy of relaxations for linear GDP that mirror those developed by Balas for disjunctive programs, and showed that a subset of these hull-relaxations yields tighter relaxations than the traditional big-M and Lee & Grossmann relaxations. We subsequently described the family of inequalities implied by the constraint set of linear GDP in its most general form (as presented in section 2.3). We then identified the strongest amongst these inequalities (i.e. facets of the constraint set of the hull-relaxations), and showed that every facet of the constraint set of the hull-relaxation can be obtained from the convex hull of some *individual* disjunction.

In a subsequent paper, we develop a novel algorithm that efficiently applies the theory developed in this paper to challenging problems in Operations Research and Chemical Engineering.

Acknowledgments

The authors would like to gratefully acknowledge Egon Balas, whose seminal work on disjunctive programming paved the way for generalized disjunctive programming and inspired this particular work. His help and advice on this work, in particular regarding Proposition 5.3, which is due to him, were essential and invaluable. We are deeply indebted.

Nicolas Sawaya would like to thank Ashish Agarwal for the long discussions regarding generalized disjunctive programming and its connection to disjunctive programming. In particular, Remark 1.1 is due to him, and section 2.2 of chapter 2 was inspired by his work on logical modeling frameworks [Agarwal, 2006].

References

Agarwal A., "Logical Modeling Frameworks for the Optimization of Discrete-Continuous Systems", PhD Thesis, Carnegie Mellon University, 2006.

Balas E., "Disjunctive Programming: Properties of the Convex Hull of Feasible Points", *MSRR #348*, Carnegie Mellon University, Pittsburgh, PA, July 1974.

Balas E., "Disjunctive Programming", *Annals of Discrete Mathematics*, 5, 3-51, 1979.

Balas E., "Disjunctive Programming and a Hierarchy of Relaxations for Discrete Continuous Optimization Problems", *SIAM J. Alg. Disc. Meth.*, Vol. 6, No. 3, 1985.

Balas E., Tama J., Tind J., "Sequential Convexification in Reverse Convex & Disjunctive Programming", *Math. Programming*, 44, 337-350, 1989.

Balas E., Ceria S., Cornuejols G., "A Lift-and-Project Cutting Plane Algorithm for Mixed 0-1 Programs", *Math. Programming.*, 58, 295-324, 1993.

Balas E., "Disjunctive Programming: Properties of the Convex Hull of Feasible Points", Invited paper, with a foreword by G. Cornuejols and W. Pulleybank, *Discrete Applied Mathematics*, 89, 3-44, 1998. (Originally MSRR #348, Carnegie Mellon University, July 1974).

Balas E., Perregaard M., "Lift-and-Project for Mixed 0-1 Programming: Recent Progress", *Discrete Applied Mathematics*, 123, 129-154, 2002.

Blair C., "Facial Disjunctive Programs and Sequences of Cutting Planes", *Discrete Applied Mathematics,* 2: 173-179, 1980.

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

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

Grossmann, I.E. and Lee S., "Generalized Disjunctive Programming: Nonlinear Convex Hull Relaxation and Algorithms", *Computational Optimization and Applications*, **26**, 83-100, 2003.

Hooker J., “Logic-based methods for optimization: combining optimization and constraint satisfaction”, John Wiley & Sons, 2000.

Jackson J. and Grossmann I.E., “High-Level optimization model for the retrofit planning of process networks”, *Industrial Engineering and Chemistry Research*, **41**, 2002, 3762-3770.

Jackson J. and Grossmann I.E., “A Disjunctive Programming Approach for the Optimal Design of Reactive Distillation Columns”, *Computers and Chemical Engineering*, **25**, 1661-1673, 2001.

Jeroslow R.G., “Cutting Plane Theory: Disjunctive Methods”, *Annals of Discrete Mathematics*, 1, 293-330, 1977.

Jeroslow R.G., “Representability in Mixed Integer Programming, I: Characterization Results”, *Discrete Applied Mathematics*, 17, 223-243, 1987.

Jeroslow R.G., “Logic-Based Decision Support: Mixed-Integer Model Formulation”, *Annals of Discrete Mathematics*, 40, North Holland (Amsterdam), 1989.

Jeroslow R.G. & Lowe J.K., “Modeling with Integer Variables”, *Mathematical Programming Studies*, 22, 167-184, 1984.

Lee S. and Grossmann I.E., “New Algorithms for Nonlinear Generalized Disjunctive Programming”, *Computers and Chemical Engineering*, **24**, 2125-2141, 2000.

Raman R. and Grossmann I.E., “Modeling and Computational Techniques for Logic Based Integer Programming”, *Computers and Chemical Engineering*, **18** (7), 563-578, 1994.

Rockefeller, R.T., “Convex Analysis”, Princeton University Press, 1970.

Sawaya N.W., Grossmann I.E., “A Cutting Plane Method for Solving Linear Generalized Disjunctive Programming Problems”, *Computers and Chemical Engineering*, 29, 1891-1913, 2005.

Sherali, H. D. and Shetty, C. M. Optimization with Disjunctive Constraints, Springer-Verlag, Berlin-Heidelberg-New York, 1980, 156 p. (Series in Economics and Mathematical Systems, Publication Number 181).

Stubbs R. and Mehrotra S., “A Branch-and-Cut Method for 0-1 Mixed Convex Programming”, *Mathematical Programming* **86**, 515-532, 1999.

Turkay M., Grossmann I.E., “Logic-Based MINLP Algorithms for the Optimal Synthesis of Process Networks”, *Computers and Chemical Engineering*, **20** (8), 959-978, 1996.

Vecchietti A., Lee S., Grossmann I.E., “Modeling of Discrete/Continuous Optimization Problems: Characterization and Formulation of Disjunctions and their Relaxations”, *Computers and Chemical Engineering*, 27, 433-448, 2003.

Yeomans, H, Grossmann I.E., “Disjunctive Programming Model for the Optimal Design of Distillation Columns and Separation Sequences”, *Industrial Engineering and Chemistry Research*, 39, 1637-1648, 2000.

Williams H.P., “Model Building in Mathematical Programming”, Wiley, 1985.