ELSEVIER

Discrete Applied Mathematics 52 (1994) 83–106

DISCRETE APPLIED MATHEMATICS

A hierarchy of relaxations and convex hull characterizations for mixed-integer zero–one programming problems

Hanif D. Sherali[a],*, Warren P. Adams[b]

[a]*Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0118, USA*

[b]*Department of Math Sciences, Clemson University, Clemson, SC 29634-1907, USA*

(Received 21 October 1991; revised 18 August 1992)

Abstract

This paper is concerned with the generation of tight equivalent representations for mixed-integer zero–one programming problems. For the linear case, we propose a technique which first converts the problem into a nonlinear, polynomial mixed-integer zero–one problem by multiplying the constraints with some suitable d-degree polynomial factors involving the n binary variables, for any given $d \in \{0, \dots, n\}$, and subsequently linearizes the resulting problem through appropriate variable transformations. As d varies from zero to n, we obtain a hierarchy of relaxations spanning from the ordinary linear programming relaxation to the convex hull of feasible solutions. The facets of the convex hull of feasible solutions in terms of the original problem variables are available through a standard projection operation. We also suggest an alternate scheme for applying this technique which gives a similar hierarchy of relaxations, but involving fewer "complicating" constraints. Techniques for tightening intermediate level relaxations, and insights and interpretations within a disjunctive programming framework are also presented. The methodology readily extends to multilinear mixed-integer zero–one polynomial programming problems in which the continuous variables appear linearly in the problem.

Key words: Mixed-integer zero–one problems; Tight relaxations; Convex hull representations; Facetial inequalities; Disjunctive programming

1. Introduction

Recently, Sherali and Adams [7] have proposed a new technique for generating a hierarchy of relaxations for linear and polynomial zero–one programming problems, spanning the spectrum from the continuous relaxation to the convex hull representation. The present paper provides an extension of this approach to the

* Corresponding author.

SSDI 0166-218X(92)00190-W

important case of mixed-integer problems, which arise more commonly in practice. Similar to the pure zero–one case, we multiply the problem constraints with d-degree polynomial factors composed of the n binary variables and their complements, for some fixed $d \in \{0, 1, \ldots, n\}$, where the zero-degree factors are taken as unity. We then linearize the resulting nonlinear program through a suitable redefinition of variables. However, in contrast with the pure zero–one situation, because of the presence of the continuous variables, an additional set of variables are defined, and the simple nonnegativity restrictions on the factor expressions are replaced by variable upper bounding types of constraints. This necessitates a different analytical approach to show that as d varies from zero to n, a hierarchy of tighter relaxations are generated between the continuous and the convex hull representations at the two extremes. We characterize all the facets of the convex hull of feasible solutions in terms of the original problem variables through a projection operation on the explicitly available final relaxation. Moreover, we demonstrate an alternate technique which is peculiar to the mixed-integer situation, and which generates a similar hierarchy of relaxations using fewer "complicating" constraints and having a different structure. We also provide additional strategies for tightening intermediate level relaxations, and present certain extreme point characterizations that relate particularly to the mixed-integer case. The overall methodology is also applicable to mixed-integer zero–one polynomial programming problems in which the continuous variables appear linearly, for which a similar hierarchy of linear relaxations is obtained.

Our approach here is in the same spirit as that of Balas [3], in which it is shown how a hierarchy of relaxations spanning the spectrum from the linear programming relaxation to the convex hull representation can be obtained for linear mixed-integer zero–one programming problems. However, the methodology employed by Balas is based on constructing the convex hull of the union of certain polyhedra defining the feasible region using disjunctive programming techniques, and the inductive process used can generate a variety of possible relaxations between the two extreme representations. In contrast, our approach generates a sequence of $(n + 1)$ relaxations, where n is the number of binary variables, with each relaxation being precisely defined in closed form. Moreover, our approach is different also in that it is designed to accommodate zero–one polynomial programming problems. However, as we show in the sequel, our relaxations can indeed be generated via Balas' "hull-relaxations" through a formulation of the feasible region as a conjunction of certain suitable disjunctions. Because of the nonstandard nature of this disjunctive formulation, the demonstration of this relationship lends further insights into both the approaches.

Since the writing of this paper (June, 1989), three somewhat related papers have emerged. The first is a paper by Boros et al. [5] which deals with the unconstrained quadratic pseudo-Boolean programming problem. For this case, they construct a standard linear programming relaxation which coincides with our relaxation at level $d = 1$, and then show in an existential fashion how a hierarchy of relaxations indexed by $d = 1, \ldots, n$ leading up to the convex hull representation at level n can be generated. This is done by including at level d, constraints corresponding to the extreme directions of the cone of nonnegative quadratic pseudo-Boolean functions which involve at most d of the n-variables. Each such relaxation can be viewed as the projection of one of our *explicitly* stated higher-order relaxations onto the first level variable space. Moreover, in contrast, our approach also permits one to consider general pseudo-Boolean polynomials, constrained problems, as well as mixed-integer situations.

In a second related approach, Lovász and Schrijver [6] address linear, pure zero–one programming problems. For this case, they generate a hierarchy of n relaxations by transforming the representation in the original n variables at each stage to a suitable representation in an n^2-variable space via an operation which amounts essentially to the application of our technique at the first level. Projecting back onto the original variable space (implicitly) yields the next level representation. Repeating n times, they recover the convex hull of feasible solutions. Again, no explicit algebraic characterization of the relaxations is readily available, and extensions to nonlinear zero–one or mixed-integer situations are not evident.

Following a similar lifting and projection scheme, Balas et al. [4] have proposed another hierarchy leading to the convex hull representation for linear, mixed-integer zero–one problems. Here, at the first level, the "lifting" operation is done using the same technique as in the present paper, but treating only one variable as binary valued, say, x_1. Projecting the resulting formulation onto the original variable space, produces the convex hull of solutions feasible to the original linear programming relaxation with the added restriction that x_1 is binary valued. This result follows from our development in the sequel by treating only x_1 as binary valued, and the remaining variables as continuous, so that the first level relaxation itself produces the corresponding convex hull representation. However, although the projection operation yields only an implicit representation, the variable x_1 is now binary valued at all vertices of the resulting polytope as argued above. This enables Balas et al. to repeat the foregoing process with the remaining variables $x_2, \ldots, x_n$ in turn, each time applying the foregoing technique to the most recent projected polytope, to (implicitly) generate a hierarchy of relaxations leading to the convex hull representation. Based on the first level "lift-and-project" scheme, an interesting cutting plane algorithm is also described, and some encouraging computational results have been presented.

The remainder of this paper is organized as follows. Section 2 presents the technique proposed for generating the sharper representations of the linear mixed-integer zero–one programming problem. Section 3 establishes the validity of this scheme and exhibits the hierarchy among the relaxations generated, along with the convex hull property for the final relaxation obtained. The relationship with disjunctive programming is explored in Section 4, and Section 5 presents certain additional characterizations for the intermediate relaxations along with strategies for further tightening them. Section 6 addresses the characterization of the resulting facets when the set is projected onto the space of the original variables. Finally, Section 7 presents some sample computational results, discusses an alternate, more compact, representation for each of the intermediate relaxations, and extends the methodology and results to the multilinear mixed-integer zero–one polynomial programming problem.

2. Generation of a sequence of sharper representations

Consider a linear mixed-integer zero–one programming problem whose feasible region is given as follows:

$$X = \left\{(x, y)\colon \sum_{j=1}^{n} \alpha_{rj} x_j + \sum_{k=1}^{m} \gamma_{rk} y_k \geqslant \beta_r \text{ for } r = 1, \ldots, R,\right.$$

$$\left. 0 \leqslant x \leqslant e_n,\ x \text{ integer},\ 0 \leqslant y \leqslant e_m \right\}, \tag{1}$$

where e_n and e_m are, respectively, column vectors of n and m entries of 1, and where the continuous variables y_k are assumed to be bounded and appropriately scaled to lie in the interval $[0, 1]$ for $k = 1, \ldots, m$. Note that any equality constraints present in the formulation can be accommodated in a similar manner as are the inequalities in the following derivation, and we omit writing them explicitly in (1) only for simplifying the presentation. However, we will show later that the equality constraints can be treated in a special manner, which in fact, encourages the writing of the R inequalities in (1) as equalities by using slack variables.

Now, for any $d \in \{1, \ldots, n\}$, let us define the (nonnegative) polynomial *factors of degree d* as

$$F_d(J_1, J_2) = \left[\prod_{j \in J_1} x_j\right]\left[\prod_{j \in J_2} (1 - x_j)\right] \quad \text{for each } J_1, J_2 \subseteq N \equiv \{1, \ldots, n\}$$

$$\text{such that } J_1 \cap J_2 = \emptyset, \text{ and } |J_1 \cup J_2| = d. \tag{2}$$

Any (J_1, J_2) satisfying the conditions in (2) will be said to be of *order d*. For example, for $n = 3$ and $d = 2$, these factors are $x_1x_2, x_1x_3, x_2x_3, x_1(1 - x_2), x_1(1 - x_3), x_2(1 - x_1), x_2(1 - x_3), x_3(1 - x_1), x_3(1 - x_2), (1 - x_1)(1 - x_2), (1 - x_1)(1 - x_3)$, and $(1 - x_2)(1 - x_3)$. In general, there are $\binom{n}{d}2^d$ such factors. For convenience, we will consider the single factor of degree zero to be $F_0(\emptyset, \emptyset) \equiv 1$, and accordingly assume products over null sets to be unity. Using these factors, let us construct a relaxation X_d of X, for any given $d \in \{0, \ldots, n\}$, using the following two steps that comprise our proposed *reformulation-linearization technique* (RLT).

Step 1 (*Reformulation step*): Multiply each of the inequalities in (1), including $0 \leqslant x \leqslant e_n$ and $0 \leqslant y \leqslant e_m$, by each of the factors $F_d(J_1, J_2)$ of degree d as defined in (2). Upon using the identity $x_j^2 \equiv x_j$ (and so $x_j(1 - x_j) = 0$) for each binary variable x_j, $j = 1, \ldots, n$, this gives the following set of additional, implied, nonlinear constraints where $D \equiv \min\{d + 1, n\}$:

$$\left[\sum_{j\in J_1} \alpha_{rj} - \beta_r\right] F_d(J_1, J_2) + \sum_{j\in N-(J_1\cup J_2)} \alpha_{rj} F_{d+1}(J_1 + j, J_2) + \sum_{k=1}^{m} \gamma_{rk} y_k F_d(J_1, J_2) \geqslant 0$$

$$\text{for } r = 1, \ldots, R \text{ and for each } (J_1, J_2) \text{ of order } d, \tag{3a}$$

$$F_D(J_1, J_2) \geqslant 0 \quad \text{for each } (J_1, J_2) \text{ or order } D, \tag{3b}$$

$$F_d(J_1, J_2) \geqslant y_k F_d(J_1, J_2) \geqslant 0$$

$$\text{for } k = 1, \ldots, m, \text{ and for each } (J_1, J_2) \text{ of order } d. \tag{3c}$$

Step 2 (*Linearization step*): Viewing the constraints in (3) in expanded form as a sum of monomials, linearize them by substituting the following variables for the corresponding nonlinear terms for each $J \subseteq N$:

$$w_J = \prod_{j\in J} x_j \quad \text{and} \quad v_{Jk} \equiv y_k \prod_{j\in J} x_j, \text{ for } k = 1, \ldots, m, \tag{4a}$$

where we will assume the notation that

$$w_j \equiv x_j \text{ for } j = 1, \ldots, n, \quad w_\emptyset \equiv 1, \quad v_{\emptyset k} \equiv y_k \text{ for } k = 1, \ldots, m. \tag{4b}$$

Furthermore, denoting by $f_d(J_1, J_2)$ and $f_d^k(J_1, J_2)$ the respective linearized forms of the polynomial expressions $F_d(J_1, J_2)$ and $y_k F_d(J_1, J_2)$ under such a substitution, we obtain the following polyhedral set X_d whose projection onto the (x, y) space is claimed to yield a relaxation for X:

$$X_d = \Bigg\{(x, y, w, v):$$

$$\left[\sum_{j\in J_1} \alpha_{rj} - \beta_r\right] f_d(J_1, J_2) + \sum_{j\in N-(J_1\cup J_2)} \alpha_{rj} f_{d+1}(J_1 + j, J_2)$$

$$+ \sum_{k=1}^{m} \gamma_{rk} f_d^k(J_1, J_2) \geqslant 0 \quad \text{for } r = 1, \ldots, R$$

$$\text{and for each } (J_1, J_2) \text{ of order } d, \tag{5a}$$

$$f_D(J_1, J_2) \geqslant 0 \text{ for each } (J_1, J_2) \text{ of order } D \equiv \min\{d + 1, n\}, \tag{5b}$$

$$f_d(J_1, J_2) \geqslant f_d^k(J_1, J_2) \geqslant 0 \text{ for } k = 1, \ldots, m,$$

$$\text{and for each } (J_1, J_2) \text{ of order } d\Bigg\}. \tag{5c}$$

Example 2.1. Consider the set

$$X = \{(x, y): \alpha_1 x_1 + \alpha_2 x_2 + \gamma_1 y_1 + \gamma_2 y_2 \geqslant \beta, 0 \leqslant x \leqslant e_2, x \text{ integer}, 0 \leqslant y \leqslant e_2\}.$$

Hence, we have $n = m = 2$. Let us consider $d = 2$, so that $D \equiv \min\{n + 1, d\} = 2$ as well. The various sets (J_1, J_2) of order 2 and the corresponding factors are given below:

(J_1, J_2)	$(\{1,2\},\emptyset)$	$(\{1\},\{2\})$	$(\{2\},\{1\})$	$(\emptyset,\{1,2\})$
$F_2(J_1, J_2)$	x_1x_2	$x_1(1-x_2)$	$x_2(1-x_1)$	$(1-x_1)(1-x_2)$
$f_2(J_1, J_2)$	w_{12}	$x_1 - w_{12}$	$x_2 - w_{12}$	$1-(x_1+x_2)+w_{12}$
$f_2^k(J_1, J_2)$, $k = 1,2$	v_{12k}	$v_{1k} - v_{12k}$	$v_{2k} - v_{12k}$	$y_k - (v_{1k} + v_{2k}) + v_{12k}$

Hence, we obtain the following constraints (6a)–(6c) corresponding to (5a)–(5c), respectively, where v_{Jk} has been written as $v_{J,k}$ for clarity:

$$X_2 = \{(x, y, w, v):$$

$$\left.\begin{aligned}&(\alpha_1+\alpha_2-\beta)w_{12}+\gamma_1 v_{12,1}+\gamma_2 v_{12,2}\geqslant 0,\\&(\alpha_1-\beta)[x_1-w_{12}]+\gamma_1(v_{1,1}-v_{12,1})+\gamma_2(v_{1,2}-v_{12,2})\geqslant 0,\\&(\alpha_2-\beta)[x_2-w_{12}]+\gamma_1(v_{2,1}-v_{12,1})+\gamma_2(v_{2,2}-v_{12,2})\geqslant 0,\\&\beta(-1+(x_1+x_2)-w_{12})+\gamma_1[y_1-(v_{1,1}+v_{2,1})+v_{12,1}]\\&\quad+\gamma_2[y_2-(v_{1,2}+v_{2,2})+v_{12,2}]\geqslant 0,\end{aligned}\right\}\tag{6a}$$

$$w_{12}\geqslant 0,\ x_1-w_{12}\geqslant 0,\ x_2-w_{12}\geqslant 0,\ 1-(x_1+x_2)+w_{12}\geqslant 0,\tag{6b}$$

$$\left.\begin{aligned}&w_{12}\geqslant v_{12,k}\geqslant 0,\ (x_1-w_{12})\geqslant(v_{1,k}-v_{12,k})\geqslant 0,\\&(x_2-w_{12})\geqslant(v_{2,k}-v_{12,k})\geqslant 0,\\&\text{and } (1-(x_1+x_2)+w_{12})\geqslant(y_k-(v_{1,k}+v_{2,k})+v_{12,k})\geqslant 0\\&\quad\text{for } k=1,2\}.\end{aligned}\right\}\tag{6c}$$

Some comments and illustrations are in order at this point. First, note that we could symmetrically have employed the terms y_k and $(1 - y_k)$, for $k = 1, \ldots, m$, as factors to multiply each of the constraints in (1) which involve only the x-variables, so that linearity would be preserved upon using the substitution (4). While the additional inequalities thus generated would possibly yield a tighter relaxation, such inequalities would be present in all the sets X_d for $d = 1, \ldots, n$, and by our convex hull assertion in Theorem 3.5, these constraints would be implied by those defining X_n above. Hence, our hierarchy results remain unaffected, and so for simplicity we omit such constraints. Nonetheless, we address the issue of the validity of including such constraints, among others, at the end of Section 5, and note that one may include them in a computational scheme employing the sets X_d, $d < n$. Note that as far as multiplying the constraints $0 \leqslant x \leqslant e_n$ by such factors is concerned, the resulting inequalities are explicitly present in (5c) when $d = 1$, while for $d > 1$, these constraints are implied as Lemma 3.1 below establishes.

Second, note that for the case $d = 0$, using the fact that $f_0(\emptyset, \emptyset) \equiv 1$, that $f_0^k(\emptyset, \emptyset) \equiv y_k$ for $k = 1, \ldots, m$, and that $f_1(j, \emptyset) \equiv x_j$ and $f_1(\emptyset, j) \equiv (1 - x_j)$ for $j = 1, \ldots, n$, it follows that X_0 given by (5) is precisely the continuous relaxation of X in which the integrality restrictions on the x-variables are dropped. Finally, for $d = n$, note that the inequalities (5b) are implied by (5c) and can therefore be omitted from the representation X_n. The following section establishes the fact that for $d = 0, 1, \ldots, n$, the sets X_d represent a sequence of nested, valid relaxations leading up to the convex hull representation.

3. Validity and the hierarchy of relaxations leading to the convex hull representation

The main result of this section is that $\mathrm{conv}(X) \equiv X_{\mathrm{P}n} \subseteq X_{\mathrm{P}(n-1)} \subseteq \cdots \subseteq X_{\mathrm{P}1} \subseteq X_{\mathrm{P}0} \equiv X_0$, where $\mathrm{conv}(X)$ denotes the convex hull of X, and where

$$X_{\mathrm{P}d} = \{(x, y)\colon (x, y, w, v) \in X_d\} \quad \text{for } d = 0, 1, \ldots, n \tag{7}$$

denotes the projection of the set X_d onto the space of the original variables (x, y). The lemma given below first sets up a hierarchy of implications with respect to constraints (5b) and (5c) via a surrogation process.

Lemma 3.1. *For any $d \in (0, \ldots, n-1\}$, the constraints $f_{d+1}(J_1, J_2) \geqslant 0$ for all (J_1, J_2) of order $(d+1)$ imply that $f_d(J_1, J_2) \geqslant 0$ for all (J_1, J_2) of order d. Similarly, the constraints $f_{d+1}(J_1, J_2) \geqslant f^k_{d+1}(J_1, J_2) \geqslant 0$ for all $k = 1, \ldots, m$, and (J_1, J_2) of order $(d+1)$ imply that $f_d(J_1, J_2) \geqslant f^k_d(J_1, J_2) \geqslant 0$ for all $k = 1, \ldots, m$, and (J_1, J_2) of order d.*

Proof. Consider any (J_1, J_2) of order d with $0 \leqslant d < n$ and any $k \in \{1, \ldots, m\}$, and let $t \in N - (J_1 \cup J_2)$. Then we have,

$$\begin{aligned} &F_{d+1}(J_1 + t, J_2) + F_{d+1}(J_1, J_2 + t) = F_d(J_1, J_2), \\ &y_k F_{d+1}(J_1 + t, J_2) + y_k F_{d+1}(J_1, J_2 + t) = y_k F_d(J_1, J_2). \end{aligned} \tag{8}$$

It is readily seen that these equations are preserved upon using the substitution (4), so that we also have,

$$\begin{aligned} &f_{d+1}(J_1 + t, J_2) + f_{d+1}(J_1, J_2 + t) = f_d(J_1, J_2), \\ &f^k_{d+1}(J_1 + t, J_2) + f^k_{d+1}(J_1, J_2 + t) = f^k_d(J_1, J_2). \end{aligned} \tag{9}$$

The required result now follows from (9), and the proof is complete. □

The equivalence of X to X_d for any $d \in \{0, \ldots, n\}$ under integrality restrictions on the x-variables, and the hierarchy among the relaxations are established next.

Theorem 3.2. *Let $X_{\mathrm{P}d}$ denote the projection of the set X_d onto the space of the (x, y) variables as defined by (7), for $d = 0, 1, \ldots, n$. Then*

$$\mathrm{conv}(X) \subseteq X_{\mathrm{P}n} \subseteq X_{\mathrm{P}(n-1)} \subseteq \cdots \subseteq X_{\mathrm{P}1} \subseteq X_{\mathrm{P}0} \equiv X_0. \tag{10}$$

In particular, $X_{\mathrm{P}d} \cap \{(x, y)\colon x \text{ binary}\} \equiv X$ for all $d = 0, 1, \ldots, n$.

Proof. Consider any $d \in \{1, \ldots, n\}$, and let $(x, y, w, v) \in X_d$. We will show that this same solution (using the components which appear in X_{d-1}) satisfies X_{d-1}, hence implying that $X_{Pd} \subseteq X_{P(d-1)}$. By Lemma 3.1, we have that the constraints (5b) and (5c) defining X_{d-1} are satisfied, and hence let us show by a similar surrogation process that the constraints (5a) are also satisfied. Toward this end, consider any (J_1, J_2) or order $(d-1)$, and any $r \in \{1, \ldots, R\}$. For any $t \notin (J_1 \cup J_2)$, by summing the two inequalities in (5a) corresponding to the sets $(J_1, J_2 + t)$ and $(J_1 + t, J_2)$ of order d, and using (9), we obtain the constraint (5a) for X_{d-1} corresponding to the set (J_1, J_2) of order $(d-1)$. Hence, $X_{Pn} \subseteq X_{P(n-1)} \subseteq \cdots \subseteq X_0$.

Next, let us show that $\mathrm{conv}(X) \subseteq X_{Pn}$. If $X = \emptyset$, this is trivial. Otherwise, given any $(x, y) \in X$, define w_J and v_{Jk} for all $J \subseteq N$, $k = 1, \ldots, m$, as in (4). Then, by construction, $(x, y, w, v) \in X_n$. Hence $X \subseteq X_{Pn}$, and since X_{Pn} is convex, we have $\mathrm{conv}(X) \subseteq X_{Pn}$, and so (10) holds. Finally, since $X \equiv \mathrm{conv}(X) \cap \{(x, y)\colon x \text{ binary}\} \equiv X_0 \cap \{(x, y)\colon x \text{ binary}\}$, it follows from (10) that $X_{Pd} \cap \{(x, y)\colon x \text{ binary}\} \equiv X$ for all $d = 0, 1, \ldots, n$, and this completes the proof. $\square$

Hence, by Theorem 3.2, we see that for any $d \in \{0, 1, \ldots, n\}$, the set X_{Pd} is a polyhedral relaxation of X in that it contains X and is equivalent to X if the x-variables are enforced to be binary valued. Moreover, the sets X_{Pd} are all nested, one within the previous set as d varies from zero to n, initializing with the ordinary continuous relaxation $X_{P0} \equiv X_0$. In fact, as shown later in Theorem 3.5, the final relaxation X_{Pn} coincides with $\mathrm{conv}(X)$. But first, let us introduce the following transformation which we shall find useful throughout this paper.

Lemma 3.3. *Consider the affine transformation:* $\{w_J, J \subseteq N\} \to \{U_J^0, J \subseteq N\}$ *defined by*

$$U_J^0 = f_n(J, \bar{J}) \equiv \sum_{J' \subseteq \bar{J}} (-1)^{|J'|} w_{J \cup J'} \quad \text{for all } J \subseteq N, \tag{11a}$$

where $\bar{J} = J - N$ *for* $J \subseteq N$. *This transformation is nonsingular with inverse*

$$w_J = \sum_{J' \subseteq \bar{J}} U_{J \cup J'}^0 \quad \text{for all } J \subseteq N, \tag{11b}$$

where as defined in (4b), $w_\emptyset \equiv 1$, *and* $w_j \equiv x_j$ *for* $j = 1, \ldots, n$. *Similarly, for each* $k = 1, \ldots, m$, *consider the linear transformation:* $\{v_{Jk}, J \subseteq N\} \to \{U_J^k, J \subseteq N\}$ *defined by*

$$U_J^k = f_n^k(J, \bar{J}) \equiv \sum_{J' \subseteq \bar{J}} (-1)^{|J'|} v_{(J \cup J')k} \quad \text{for all } J \subseteq N. \tag{12a}$$

Then this defines a nonsingular transformation with inverse

$$v_{Jk} = \sum_{J' \subseteq \bar{J}} U_{J \cup J'}^k \quad \text{for all } J \subseteq N, \tag{12b}$$

where as defined in (4b), $v_{\emptyset k} \equiv y_k$, *for* $k = 1, \ldots, m$. *In particular, under* (11) *and* (12), *we have*

$$x_j = \sum_{J \subseteq N:\, j \in J} U_J^0 \text{ for } j = 1, \ldots, n \quad \text{and} \quad y_k = \sum_{J \subseteq N} U_J^k \text{ for } k = 1, \ldots, m. \tag{13}$$

Proof. Note that from (11a), where the expression for $f_n(J, \bar{J})$ follows easily from (2), the sum in (11b) is given by

$$\sum_{J' \subseteq \bar{J}} U^0_{J \cup J'} = \sum_{J' \subseteq \bar{J}} \sum_{K \subseteq J \cup J'} (-1)^{|K|} w_{J \cup J' \cup K}$$

$$= \sum_{H \subseteq \bar{J}} \left[\sum_{K \subseteq H} (-1)^{|K|} \right] w_{J \cup H} = w_J. \tag{14}$$

The last step follows because the sum $\sum_{K \subseteq H}(-1)^{|K|}$ in (14) equals zero whenever $H \neq \emptyset$, and equals 1 when $H = \emptyset$. Hence, given the system (11a), we see from (14) that the system (11b) must be satisfied, yielding a unique solution. This proves the assertion involving (11).

In an identical fashion, the system (12a) is equivalent to (12b). Finally, noting that (13) simply rewrites (11b) for $J = \{j\}, j = 1, \ldots, n$, and (12b) for $J = \emptyset, k = 1, \ldots, m$, the proof is complete. □

Example 3.4. To illustrate Lemma 3.3, consider a situation with $n = 3$ and let us verify the transformation (12), for example. Then for any $k \in \{1, \ldots, m\}$, the system (12a) is of the form

$$U^k_{123} = v_{123k}, \quad U^k_{12} = v_{12k} - v_{123k}, \quad U^k_{13} = v_{13k} - v_{123k}, \quad U^k_{23} = v_{23k} - v_{123k},$$

$$U^k_1 = v_{1k} - (v_{12k} + v_{13k}) + v_{123k}, \quad U^k_2 = v_{2k} - (v_{12k} + v_{23k}) + v_{123k},$$

$$U^k_3 = v_{3k} - (v_{13k} + v_{23k}) + v_{123k},$$

$$U^k_\emptyset = y_k - (v_{1k} + v_{2k} + v_{3k}) + (v_{12k} + v_{13k} + v_{23k}) - v_{123k}.$$

The inverse transformation (12b) is as follows:

$$v_{123k} = U^k_{123}, \quad v_{12k} = U^k_{12} + U^k_{123}, \quad v_{13k} = U^k_{13} + U^k_{123}, \quad v_{23k} = U^k_{23} + U^k_{123},$$

$$v_{1k} = U^k_1 + U^k_{12} + U^k_{13} + U^k_{123}, \quad v_{2k} = U^k_2 + U^k_{12} + U^k_{23} + U^k_{123},$$

$$v_{3k} = U^k_3 + U^k_{13} + U^k_{23} + U^k_{123},$$

$$v_{\emptyset k} \equiv y_k = U^k_\emptyset + U^k_1 + U^k_2 + U^k_3 + U^k_{12} + U^k_{13} + U^k_{23} + U^k_{123}.$$

The following theorem now provides the desired convex hull characterization.

Theorem 3.5. *Let the polyhedral relaxation X_{P_n} of X be as defined by* (5) *and* (7). *Then* $X_{P_n} \equiv \operatorname{conv}(X)$.

Proof. By Theorem 3.2, we need to show that $X_{P_n} \subseteq \operatorname{conv}(X)$. If $X_{P_n} = \emptyset$, then this result is trivial, and so we assume that $X_{P_n} \neq \emptyset$. Since X_{P_n} is bounded, and since $X_{P_n} \subseteq X_0$ by Theorem 3.2, we only need to show that x is binary valued at all extreme points (x, y) of X_{P_n}. Equivalently, we need to show that the linear program

$$\text{LP:} \quad \text{maximize} \left\{ \sum_{j=1}^{n} c_j x_j + \sum_{k=1}^{m} d_k y_k : (x, y) \in X_{P_n} \right\} \tag{15a}$$

has an optimal solution at which x is binary for any objective function $(cx + dy)$. Noting the definition of X_{P_n} given via (5) and (7) with $d = n$, we may write (15a) as follows:

$$\text{LP:} \quad \text{maximize} \quad \sum_{j=1}^{n} c_j x_j + \sum_{k=1}^{m} d_k y_k,$$

$$\text{subject to} \quad \left[\sum_{j \in J} \alpha_{rj} - \beta_r\right] f_n(J, \bar{J}) + \sum_{k=1}^{m} \gamma_{rk} f_n^k(J, \bar{J}) \geqslant 0$$

$$\text{for all } r = 1, \ldots, R, \quad J \subseteq N, \tag{15b}$$

$$f_n(J, \bar{J}) \geqslant f_n^k(J, \bar{J}) \geqslant 0 \quad \text{for all } k = 1, \ldots, m, \ J \subseteq N.$$

Now, consider the nonsingular linear transformation given by (11) and (12). Noting (13), and using Lemma 3.3, the linear program LP given in (15b) gets equivalently transformed into the following problem:

$$\text{LP:} \quad \text{maximize} \quad \sum_{J \subseteq N} c_J^0 U_J^0 + \sum_{J \subseteq N} \sum_{k=1}^{m} d_k U_J^k, \tag{16a}$$

$$\text{subject to} \quad \sum_{k=1}^{m} \gamma_{rk} U_J^k \geqslant \delta_{Jr} U_J^0 \quad \text{for all } r = 1, \ldots, R, \ J \subseteq N, \tag{16b}$$

$$\sum_{J \subseteq N} U_J^0 = 1, \tag{16c}$$

$$0 \leqslant U_J^k \leqslant U_J^0 \quad \text{for all } k = 1, \ldots, m, \ J \subseteq N, \tag{16d}$$

where

$$c_J^0 \equiv \sum_{j \in J} c_j \quad \text{for all } J \subseteq N,$$

$$\delta_{Jr} \equiv \beta_r - \sum_{j \in J} \alpha_{rj} \quad \text{for all } r = 1, \ldots, R, \ J \subseteq N, \tag{16e}$$

and where (11b) and (12b) give the optimal solution to (15) corresponding to the optimal solution to (16). (Above, note that (16c) corresponds to (11b) for $J = \emptyset$.)

Now projecting onto the space of the variables U_J^0, $J \subseteq N$, and defining

$$S_0 = \left\{ U^0 \equiv (U_J^0, J \subseteq N) \colon \sum_{J \subseteq N} U_J^0 = 1, \ U^0 \geqslant 0 \right\} \tag{17a}$$

and for all $J \subseteq N$, letting

$$\Delta_J \equiv \max \left\{ \sum_{k=1}^{m} d_k U_J^k \colon \sum_{k=1}^{m} \gamma_{rk} U_J^k \geqslant \delta_{Jr} \text{ for } r = 1, \ldots, R, \right.$$
$$\left. 0 \leqslant U_J^k \leqslant 1 \text{ for } k = 1, \ldots, m \right\}, \tag{17b}$$

where $\Delta_J \equiv -\infty$ if (17b) is infeasible, it is readily seen that problem (16) is equivalent to the problem

$$\text{maximize}\left\{\sum_{J \subseteq N} (c_J^0 + \Delta_J)U_J^0\colon U^0 \in S_0\right\}. \tag{18}$$

(The equivalence follows by noting that for a fixed $U^0 \in S_0$, the problem (16) decomposes into separable problems over $J \subseteq N$, with each such problem being given by (17b) in which all the right-hand sides are multiplied by the corresponding scalar U_J^0.)

Now since $X_{P_n} \neq \emptyset$ by assumption, $\Delta_J > -\infty$ for at least some $J \subseteq N$ in (17b). Noting (17a), we have at optimality in (18), that $U_{J^*}^0 = 1$ for some $J^* \subseteq N$, and $U_J^0 = 0$ for $J \subseteq N$, $J \neq J^*$. Accordingly, from (16), $U_J^k = 0$ for $k = 1, \ldots, m$ for all $J \subseteq N$, $J \neq J^*$, while $U_{J^*}^k$, $k = 1, \ldots, m$, are given at optimality by the solution $U_{J^*}^{*k}$, $k = 1, \ldots, m$, to the problem in (17b) for $J = J^*$. Hence, from (13), we obtain at optimality for LP that

$$x_j = \begin{cases} 1 & \text{if } j \in J^* \\ 0 & \text{otherwise,} \end{cases} \quad \text{for } j = 1, \ldots, n, \quad \text{and} \quad y_k = U_{J^*}^{*k} \quad \text{for } k = 1, \ldots, m. \tag{19}$$

Since x is binary valued at optimality, this completes the proof. □

4. Relationships with disjunctive programming

In this section we explore the connections between our approach and that of Balas [3] by demonstrating that the sets X_{Pd}, $d = 0, 1, \ldots, n$, can be viewed as "hull-relaxations" of certain disjunctive formulations of the set X. This provides insights by not only putting our development in the framework of disjunctive programming, but also, by noting the peculiar manner in which the equivalent disjunctions are constructed, it sheds light on formulating disjunctions in order to obtain tighter representations. This insight may prove to be useful in devising partial applications of either technique for generating computationally useable, tight linear programming relaxations.

Toward this end, consider the set S given by the union of the polyhedra P_i, $i \in Q$, where we assume that $P_i \equiv \{z\colon A^i z \geqslant a_0^i\}$ is bounded for each $i \in Q$. Then, as shown by Balas [3], the convex hull of S is given by

$$\text{conv}(S) = \left\{z\colon A^i\xi^i - a_0^i\xi_0^i \geqslant 0\ \forall i \in Q,\ \sum_{i \in Q} \xi_0^i = 1,\right.$$
$$\left.\xi_0^i \geqslant 0\ \forall i \in Q, \quad \text{and } z = \sum_{i \in Q} \xi^i\right\}. \tag{20}$$

Accordingly, if the feasible region X of a given problem is represented in the *conjunctive normal form* (CNF) given by $X \equiv \bigcap_{j \in T} S_j$ for some index set T, where each S_j, $j \in T$, is the union of certain (bounded) polyhedra, then Balas [3] considers the *hull-relaxation* of X defined as follows:

$$\text{h-rel}(X) \equiv \bigcap_{j \in T} \text{conv}(S_j). \tag{21}$$

Clearly, we have $\text{conv}(X) \subseteq \text{h-rel}(X)$. In particular, for X given by (1), if we represent X as $X_0 \cap \{(x, y)\colon x_1 \leqslant 0 \text{ or } x_1 \geqslant 1\} \cap \{(x, y)\colon x_2 \leqslant 0 \text{ or } x_2 \geqslant 1\} \cap \cdots \cap \{(x, y)\colon x_n \leqslant 0 \text{ or } x_n \geqslant 1\}$, where X_0 is the linear programming relaxation of X, then $\text{h-rel}(X) \equiv X_0$. On the other hand, if we write $X \equiv S_1$, where $T \equiv \{1\}$, and where $S_1 \equiv \bigcup_{J \subseteq N} P_{(J, \bar{J})}$ with $P_{(J, \bar{J})}$ being the polytope $X_0 \cap \{(x, y)\colon x_j = 1\ \forall j \in J,\ x_j = 0\ \forall j \in \bar{J}\}$, then we have $\text{h-rel}(X) \equiv \text{conv}(X) \equiv X_{\text{P}n}$. Hence, the foregoing two CNF representations of X produce as hull-relaxations the two relaxations at the extreme ends of our hierarchy. By the same argument, suppose that we were to consider some $J \subseteq N$, $|J| = d \in \{1, \ldots, n\}$, and that we were to construct all factors $F_d(J_1, J_2)$ of order d where $J_1 \subseteq J$ and $J_2 = J - J_1$, and use these factors in the spirit of our approach to generate a set X'_d. Then the projection $X'_{\text{P}d}$, say, of X'_d onto the (x, y) variable space would precisely be the convex hull of $X_0 \cap \{(x, y)\colon x_j \text{ binary for } j \in J\}$. This follows simply by treating x_j, $j \in \bar{J}$ in addition to the y-variables as continuous variables and constructing the corresponding final relaxation in the hierarchy. The question of principal interest, however, is whether our intermediate relaxations $X_{\text{P}d}$, $d \in \{1, \ldots, n-1\}$ are also recoverable as hull-relaxations of suitable CNF representations of X. This is indeed the case, as stated in Theorem 4.1 below. Notationally, any $J \subseteq N$, $|J| = d$ will also be referred to as being *of order d*, and given a set J of order d, we will denote $J(d) \equiv \{(J_1, J_2)\colon J_1 \subseteq J,\ J_2 = J - J_1\}$.

Theorem 4.1. *Given any $d \in \{1, \ldots, n-1\}$, the feasible region X in* (1) *can be written as*

$$X = \left\{(x, y)\colon (x, y, w, v) \in \bigcap_{\substack{J \subseteq N \\ J \text{ of order } d}} S_J\right\}, \tag{22a}$$

where

$$S_J \equiv \bigcup_{(J_1, J_2) \in J(d)} P_{(J_1, J_2)} \tag{22b}$$

and where for each $(J_1, J_2) \in J(d)$, by fixing $x_j = 1\ \forall j \in J_1$ and $x_j = 0\ \forall j \in J_2$ in X_0, and including the resulting consequence on the w and v variables, we construct

$$\begin{aligned} P_{(J_1, J_2)} = \Bigg\{(x, y, w, v)\colon & \sum_{j \in \bar{J}} \alpha_{rj} x_j + \sum_{k=1}^{m} \gamma_{rk} y_k \geqslant \Bigg(\beta_r - \sum_{j \in J_1} \alpha_{rj}\Bigg)\ \forall r = 1, \ldots, R, \\ & 0 \leqslant y_k \leqslant 1\ \forall k = 1, \ldots, m,\ 0 \leqslant x_j \leqslant 1\ \forall j \in \bar{J}, \\ & x_j = 1\ \forall j \in J_1 \textit{ and } x_j = 0\ \forall j \in J_2, \\ & v_{J'k} = y_k\ \forall k = 1, \ldots, m,\ \forall J' \subseteq J,\ J' \neq \emptyset,\ J' \subseteq J_1, \\ & v_{J'k} = 0\ \forall k = 1, \ldots, m,\ \forall J' \subseteq J,\ J' \neq \emptyset,\ J' \not\subseteq J_1, \end{aligned}$$

$$\left.\begin{aligned}
&w_{J'} = 1\ \forall J' \subseteq J,\ |J'| \geqslant 2,\ J' \subseteq J_1,\\
&w_{J'} = 0\ \forall J' \subseteq J,\ |J'| \geqslant 2,\ J' \not\subseteq J_1,\\
&w_{j \cup J'} = x_j\ \forall j \in \bar{J},\ \forall J' \subseteq J,\ J' \neq \emptyset,\ J' \subseteq J_1,\\
&w_{j \cup J'} = 0\ \forall j \in \bar{J},\ \forall J' \subseteq J,\ J' \neq \emptyset,\ J' \not\subseteq J_1
\end{aligned}\right\}. \tag{22c}$$

Then we have

$$\text{h-rel}(X) \equiv \left\{(x, y)\colon (x, y, w, v) \in \text{h-rel}\left[\bigcap_{\substack{J \subseteq N \\ J \text{ of order } d}} S_J\right]\right\} = X_{Pd}. \tag{23}$$

Proof. Note that $(x, y) \in X$ if and only if for each $J \subseteq N$ of order d, we have that for at least one (actually exactly one) $(J_1, J_2) \in J(d)$, the restrictions $(x, y) \in X_0$, $x_j = 1\ \forall j \in J_1$, and $x_j = 0\ \forall j \in J_2$ hold. We can append to this statement the inconsequential identities based on (4a) that $w_{J'} = 1$ if $x_j = 1\ \forall j \in J'$, and 0 otherwise, for each $J' \subseteq N$, $|J'| \geqslant 2$, and that for all $j \notin J'$, $w_{j \cup J'} = x_j$ if $x_k = 1\ \forall k \in J'$, and 0 otherwise, for each $J' \subseteq N$, $J' \neq \emptyset$, and also that for $k = 1, \ldots, m$, $v_{J'k} = y_k$ if $x_j = 1\ \forall j \in J'$, and 0 otherwise, for each $J' \subseteq N$, $J' \neq \emptyset$. Noting that each $P_{(J_1, J_2)}$ asserts that $(x, y) \in X_0$, $x_j = 1\ \forall j \in J_1$, and $x_j = 0\ \forall j \in J_2$, and includes the latter identities for all relevant $J' \subseteq J \equiv J_1 \cup J_2$, we obtain (22a).

To prove (23), it is sufficient to show from (21) that

$$X_d = \bigcap_{\substack{J \subseteq N \\ J \text{ of order } d}} \operatorname{conv}(S_J). \tag{24}$$

Applying (20), we obtain with obvious notation that

$$\bigcap_{\substack{J \subseteq N \\ J \text{ of order } d}} \operatorname{conv}(S_J) = \{(x, y, w, v)\colon \text{for each } J \subseteq N \text{ of order } d,$$
$$\text{constraints (25a–j) hold}\},$$

where

$$\left(\sum_{j \in J_1} \alpha_{rj} - \beta_r\right) z_{(J_1, J_2)} + \sum_{j \in \bar{J}} \alpha_{rj} x_{j,(J_1, J_2)}$$
$$+ \sum_{k=1}^{m} \gamma_{rk} y_{k,(J_1, J_2)} \geqslant 0\ \forall r,\ \forall (J_1, J_2) \in J(d), \tag{25a}$$

$$0 \leqslant y_{k,(J_1, J_2)} \leqslant z_{(J_1, J_2)}\ \forall k,\ \forall (J_1, J_2) \in J(d), \tag{25b}$$

$$0 \leqslant x_{j,(J_1, J_2)} \leqslant z_{(J_1, J_2)}\ \forall j \in \bar{J},\ \forall (J_1, J_2) \in J(d), \tag{25c}$$

$$y_k = \sum_{(J_1, J_2) \in J(d)} y_{k,(J_1, J_2)}\ \forall k, \tag{25d}$$

$$x_j = \sum_{(J_1, J_2) \in J(d)} x_{j,(J_1, J_2)}\ \forall j \in \bar{J}, \tag{25e}$$

$$x_j = \sum_{(J_1, J_2) \in J(d):\, j \in J_1} z_{(J_1, J_2)} \;\forall j \in J, \tag{25f}$$

$$\sum_{(J_1, J_2) \in J(d)} z_{(J_1, J_2)} = 1, \tag{25g}$$

$$v_{J'k} = \sum_{(J_1, J_2) \in J(d):\, J' \subseteq J_1} y_{k,(J_1, J_2)} \;\forall k,\; \forall J' \subseteq J,\; J' \neq \emptyset, \tag{25h}$$

$$w_{J'} = \sum_{(J_1, J_2) \in J(d):\, J' \subseteq J_1} z_{(J_1, J_2)} \;\forall J' \subseteq J,\; |J'| \geqslant 2, \tag{25i}$$

$$w_{j \cup J'} = \sum_{(J_1, J_2) \in J(d):\, J' \subseteq J_1} x_{j,(J_1, J_2)} \;\forall j \in \bar{J},\; \forall J' \subseteq J, J' \neq \emptyset. \tag{25j}$$

Above, $z_{(J_1,J_2)}$ plays the role of ξ_0^i in (20), and $x_{j,(J_1,J_2)}$ and $y_{k,(J_1,J_2)}$ play the role of ξ^i in (20). In this spirit, note that the constraints $x_j = 1\ \forall j \in J_1$ and $x_j = 0\ \forall j \in J_2$ in (22c) actually translate under (20) to the constraints $x_{j,(J_1,J_2)} = z_{(J_1,J_2)}\ \forall j \in J_1$ and $x_{j,(J_1,J_2)} = 0\ \forall j \in J_2$, $\forall (J_1, J_2) \in J(d)$, and that $x_j = \sum_{(J_1,J_2) \in J(d)} x_{j,(J_1,J_2)}\ \forall j \in J$. Eliminating $x_{j,(J_1,J_2)}\ \forall j \in J \equiv J_1 \cup J_2$ by substitution, yields the equivalent set of constraints (25f). Similarly, the two sets of constraints involving $v_{J'k}$ in (22c) translate under (20) to the form

$$v_{J'k,(J_1,J_2)} = \begin{cases} y_{k,(J_1,J_2)} & \forall k,\; \forall J' \subseteq J,\; J' \neq \emptyset,\; J' \subseteq J_1, \\ 0 & \forall k,\; \forall J' \subseteq J,\; J' \neq \emptyset,\; J' \not\subseteq J_1, \end{cases} \quad \forall (J_1, J_2) \in J(d),$$

$$v_{J'k} = \sum_{(J_1,J_2) \in J(d)} v_{J'k,(J_1,J_2)} \quad \forall k,\; \forall J' \subseteq J,\; J' \neq \emptyset.$$

Again by substituting out the $v_{J'k,(J_1,J_2)}$ variables from the above constraints, we obtain the equivalent set of constraints (25h). In a likewise manner, applying (20) and simplifying the last two pairs of constraints in (22c) equivalently yields (25i) and (25j), respectively.

Now, for each $k = 1, \ldots, m$ and for each $J \subseteq N$ of order d, consider the constraints

$$y_k = \sum_{(J_1,J_2) \in J(d)} y_{k,(J_1,J_2)} \quad \text{and} \quad v_{J'k} = \sum_{(J_1,J_2) \in J(d):\, J' \subseteq J_1} y_{k,(J_1,J_2)} \;\forall J' \subseteq J,\; J' \neq \emptyset.$$

From (12) in Lemma 3.3, this equation system is equivalent via elementary row operations to

$$y_{k,(J_1,J_2)} = f_d^k(J_1, J_2) \quad \forall (J_1, J_2) \in J(d).$$

Hence, the equations (25d) and (25h) written for all $J \subseteq N$ of order d are equivalent to the system of equations

$$y_{k,(J_1,J_2)} = f_d^k(J_1, J_2) \quad \forall k,\; \forall (J_1, J_2) \in J(d),\; \forall J \subseteq N \text{ of order } d. \tag{26a}$$

Similarly, examining the Eqs. (25f), (25g) and (25i) written for any $J \subseteq N$ of order d, and applying (11) of Lemma 3.3, we see that these equations are collectively equivalent to the equations

$$z_{(J_1,J_2)} = f_d(J_1, J_2) \quad \forall (J_1, J_2) \in J(d),\ \forall J \subseteq N \text{ of order } d. \tag{26b}$$

In a likewise fashion, for each $J \subseteq N$ of order d, and for each $j \in \bar{J}$, examine the equation system comprised of (25e) written for the particular $j \in \bar{J}$, and the Eqs. (25j) written for all $J' \subseteq J$, $J' \neq \emptyset$, for the particular J and $j \in \bar{J}$. Then applying (12) of Lemma 3.3 to this set of equations, we see that the system of equations in (25e) and (25j) for all $J \subseteq N$ of order d are collectively equivalent to the equations

$$x_{j,(J_1,J_2)} = f_{d+1}(J_1 + j, J_2) \quad \forall (J_1, J_2) \in J(d),\ \forall j \in \bar{J},\ \forall J \subseteq N \text{ of order } d. \tag{26c}$$

Hence, applying Lemma 3.3 to the set of Eqs. (25d)–(25j) written for all $J \subseteq N$ of order d yields the equivalent system of equations given in (26a)–(26c). Finally, using (26) to substitute out the $y_{k,(J_1,J_2)}$, $z_{(J_1,J_2)}$, and $x_{j,(J_1,J_2)}$ variables from (25a)–(25c), transforms (25a) to (5a), (25b) to (5c), and (25c) to (5b), where the latter set for $d < n$ is equivalently written as $0 \leqslant f_{d+1}(J_1 + j,\ J_2) \leqslant f_d(J_1, J_2)$ $\forall j \in \bar{J}$, $(J_1, J_2) \in J(d)$, for each $J \subseteq N$ of order d. Hence, under (26), (25) transforms precisely to X_d as given by (5), and this completes the proof. $\square$

We remark here that Balas [3] constructs his hierarchy of relaxations by commencing from the hull-relaxation of the CNF representation which yields the usual linear programming relaxation, and then combines pairs of disjunctive sets in the conjunction to obtain tighter relaxations. No realization of this process will evidently result in the CNF representation of Theorem 4.1, although Theorem 4.1 casts our relaxations in the framework of hull-relaxations. Note that the CNF representation given by (22) says more than simply that for each $J \subseteq N$ of order d, we must have for at least one $(J_1, J_2) \in J(d)$ that $(x,\ y) \in X_0$, $x_j = 1$ $\forall j \in J_1$, and $x_j = 0$ $\forall j \in J_2$. The additional constraints involving the w and v variables in (22c), although being implied in the discrete representation of X in (22), serve to tighten the hull-relaxation by tying in the constraints obtained by applying (20) to some $S_{\hat{J}}$ to those obtained for another $S_{\tilde{J}}$, whenever $\hat{J} \cap \tilde{J} \neq \emptyset$, so that there exist some common subsets J'. This insight hints at a scheme which can be used in a more general form in order to tighten hull-relaxations obtained through disjunctive programming methods.

5. Further insights into the structure of the relaxations

In this section, we begin by presenting two lemmas which show that if x is binary in X_d, then the identity (4) holds precisely due to the constraints (5c) defining X_d, while the remaining constraints serve to contain such a corresponding solution within X. Along with Theorem 3.5, this enables us to present a more complete characterization of the vertices of X_n, and permits us to suggest further strategies for tightening the intermediate relaxations while maintaining the hierarchy.

Lemma 5.1. *For any $d \in \{1, \ldots, n\}$, define the following set composed from the constraints* (5c).

$$Z_d = \{(x, y, w, v)\colon f_d(J_1, J_2) \geqslant f_d^k(J_1, J_2) \geqslant 0 \text{ for } k = 1, \ldots, m \text{ and for each } (J_1, J_2) \text{ of order } d\}. \tag{27}$$

Let $\hat{x}$ be any binary vector. Then $(\hat{x}, \hat{y}, \hat{w}, \hat{v}) \in Z_d$ if and only if

$$0 \leqslant \hat{y} \leqslant e_m \tag{28a}$$

and

$$\hat{w}_J = \prod_{j \in J} \hat{x}_j \quad \text{and} \quad \hat{v}_{Jk} = \hat{y}_k \prod_{j \in J} \hat{x}_j, \quad \text{for all } k = 1, \ldots, m \text{ and } J \subseteq N \text{ with } |J| = 1, \ldots, d. \tag{28b}$$

Proof. If $(\hat{x}, \hat{y}, \hat{w}, \hat{v})$ satisfies (28), then the values of $f_d(J_1, J_2)$ and $f_d^k(J_1, J_2)$ match those of $F_d(J_1, J_2)$ and $y_k F_d(J_1, J_2)$, respectively, and so $(\hat{x}, \hat{y}, \hat{w}, \hat{v}) \in Z_d$. Conversely, let $(\hat{x}, \hat{y}, \hat{w}, \hat{v}) \in Z_d$. Let us show by induction on d that (28) holds. Note that for $d = 1$, the set Z_1 has constraints

$$x_j \geqslant v_{jk} \geqslant 0 \quad \text{and} \quad (1 - x_j) \geqslant (y_k - v_{jk}) \geqslant 0 \quad \text{for } j = 1, \ldots, n, \ k = 1, \ldots, m.$$

Hence, $0 \leqslant \hat{y}_k \leqslant 1$, and for any $j = 1, \ldots, n$, if $\hat{x}_j = 0$, then $\hat{v}_{jk} = 0 = \hat{y}_k \hat{x}_j$, for $k = 1, \ldots, m$, while if $\hat{x}_j = 1$, then $\hat{v}_{jk} = \hat{y}_k = \hat{y}_k \hat{x}_j$ for $k = 1, \ldots, m$. Moreover, since $\hat{w}_j \equiv \hat{x}_j, j = 1, \ldots, n$ from (4b), we have that (28) holds. Therefore, the result is true for Z_1. Now, assume that it is true for $Z_1, \ldots, Z_{d-1}$, and consider the set Z_d for any $d \in \{2, \ldots, n\}$.

Observe by Lemma 3.1 that the set Z_d enforces the constraints $f_{d-1}(J_1, J_2) \geqslant f_{d-1}^k(J_1, J_2) \geqslant 0$ for all (J_1, J_2) of order $(d - 1)$, and so by the induction hypothesis, (28a) holds and moreover, (28b) holds for all $J \subseteq N$ such that $|J| \in \{1, \ldots, d - 1\}$. Hence, consider any $J \subseteq N$ with $|J| = d$, and let us show that (28b) holds for this case as well.

Note that for any $s, t \in J$, since $F_d(J - t, t) = (1 - x_t) \prod_{j \in J - t} x_j$ and $F_d(J - s - t, \{s, t\}) = (1 - x_s - x_t + x_s x_t) \prod_{j \in J - \{s, t\}} x_j$, we have from the constraint $f_d(J - t, t) \geqslant 0$ that $w_J \leqslant w_{J-t}$, and from the constraint $f_d(J - s - t, \{s, t\}) \geqslant 0$ that $w_J \geqslant w_{J-s} + w_{J-t} - w_{J-s-t}$. Using the induction hypothesis, this means that

$$\hat{w}_J \leqslant \prod_{j \in J - t} \hat{x}_j \quad \text{and} \quad \hat{w}_J \geqslant \prod_{j \in J - s} \hat{x}_j + \prod_{j \in J - t} \hat{x}_j - \prod_{j \in J - s - t} \hat{x}_j \quad \text{for all } s, t \in J. \tag{29}$$

Now, suppose that $\prod_{j \in J} \hat{x}_j = 0$. Then from the first inequality in (29) and that $w_J \equiv f_d(J, \emptyset) \geqslant f_d^k(J, \emptyset) \equiv v_{Jk} \geqslant 0$, we have that $\hat{w}_J = \hat{v}_{Jk} = 0$ for $k = 1, \ldots, m$, and so (28b) holds. On the other hand, suppose that $\prod_{j \in J} \hat{x}_j = 1$. Then from (29), we have $\hat{w}_J = 1$. Moreover, for any $t \in J$ and $k \in \{1, \ldots, m\}$, the constraint $f_d(J - t, t) \geqslant f_d^k(J - t, t) \geqslant 0$ in (27) yields $(\hat{w}_{J-t} - \hat{w}_J) \geqslant (\hat{v}_{(J-t)k} - \hat{v}_{Jk}) \geqslant 0$. Since $\hat{w}_{J-t} = \hat{w}_J = 1$, we have $\hat{v}_{Jk} = \hat{v}_{(J-t)k}$. But by the induction hypothesis, since (28b) holds for $(J-t)$ as $|J-t| = d-1$, we have, $\hat{v}_{(J-t)k} = \hat{y}_k \prod_{j\in J-t} \hat{x}_j = \hat{y}_k$, and so $\hat{v}_{Jk} = \hat{y}_k$. Therefore, if $\prod_{j\in J} \hat{x}_j = 1$, then $\hat{w}_J = 1$ and $\hat{v}_{Jk} = \hat{y}_k$ for $k = 1,\ldots,m$, and so again (28b) holds. This completes the proof. □

Lemma 5.2. *Consider any $d \in \{1,\ldots,n\}$, and let $\hat{x}$ be any binary vector. Then $(\hat{x},\hat{y},\hat{w},\hat{v}) \in X_d$ if and only if:*

$$
\begin{aligned}
&(\hat{x},\hat{y}) \in X, \hat{w}_J = \prod_{j\in J} \hat{x}_j \quad \textit{for all } J \subseteq N \textit{ with } |J| = 1,\ldots,D \equiv \min\{d+1,n\},\\
&\hat{v}_{Jk} = \hat{y}_k \prod_{j\in J} \hat{x}_j \quad \textit{for all } k = 1,\ldots,m \textit{ and } J \subseteq N \textit{ with } |J| = 0,1,\ldots,d.
\end{aligned} \tag{30}
$$

Proof. For any $d \in \{1,\ldots,n\}$, and $\hat{x}$ binary, if $(\hat{x},\hat{y},\hat{w},\hat{v}) \in X_d$, then since $(\hat{x},\hat{y}) \in X_{\mathrm{P}d}$, $X_{\mathrm{P}d} \subseteq X_0$ by Theorem 3.2, and $X \equiv X_0 \cap \{(x,y)\colon x \text{ binary}\}$, we have that $(\hat{x},\hat{y}) \in X$. Moreover, noting (5b), we have from Lemma 5.1 that the other conditions in (30) hold as well. Conversely, if (30) holds, then $(\hat{x},\hat{y},\hat{w},\hat{v}) \in X_d$ by construction, and the proof is complete. □

Theorem 5.3. *The solution $(\hat{x},\hat{y},\hat{w},\hat{v})$ is a vertex of X_n if and only if $\hat{x}$ is binary valued, $\hat{w}_J = \prod_{j\in J} \hat{x}_j$ for all $J \subseteq N$, $J \neq \emptyset$, $\hat{v}_{Jk} = \hat{y}_k \prod_{j\in J} \hat{x}_j$ for all $J \subseteq N$, $J \neq \emptyset$, $k = 1,\ldots,m$, and $(\hat{y}_1,\ldots,\hat{y}_m)$ is an extreme point of the set $\hat{Y} \equiv \{y\colon (\hat{x},y) \in X\}$.*

Proof. Let $(\hat{x},\hat{y},\hat{w},\hat{v})$ be a vertex of X_n. Then there exists a linear objective function defined on the (x,y,w,v) space such that the maximum of this function over X_n occurs uniquely at $(\hat{x},\hat{y},\hat{w},\hat{v})$. Now, following the approach in Theorem 3.5, under the transformations (11) and (12) of Lemma 3.3, the foregoing linear program can be put into the form (16) with appropriately defined objective coefficients. Consequently, from (19), the unique optimum $\hat{x}$ must be binary valued. Hence, from Lemma 5.2, $\hat{w}_J = \prod_{j\in J} \hat{x}_j$ for all $J \subseteq N$, $J \neq \emptyset$, and $\hat{v}_{Jk} = \hat{y}_k \prod_{j\in J} \hat{x}_j$ for all $k = 1,\ldots,m$, $J \subseteq N$, $J \neq \emptyset$. Furthermore, from (17b) and (19), the (unique) optimum $\hat{y}$ is obtained as the solution $U^{*k}_{J^*}$, $k = 1,\ldots,m$ to (17b) for some $J = J^*$. Noting in (17b) that $\delta_{J^*r} \equiv \beta_r - \sum_{j\in J^*} \alpha_{rj} = \beta_r - \sum_{j=1}^n \alpha_{rj}\hat{x}_j$ from (19), we obtain $\hat{y}$ as the unique solution to a linear program over the polyhedron $\hat{Y} \equiv \{y\colon \sum_{k=1}^m \gamma_{rk} y_k \geqslant (\beta_r - \sum_{j=1}^n \alpha_{rj}\hat{x}_j)$, $0 \leqslant y \leqslant e_m\}$. Hence, $\hat{y}$ is a vertex of $\hat{Y}$.

Conversely, suppose that we are given $(\hat{x},\hat{y},\hat{w},\hat{v})$ satisfying the conditions stated in Theorem 5.3. By Lemma 5.2, $(\hat{x},\hat{y},\hat{w},\hat{v}) \in X_n$ and, in particular, $(\hat{x},\hat{y}) \in X_{\mathrm{P}n}$. It is sufficient to show that $(\hat{x},\hat{y})$ is an extreme point of $X_{\mathrm{P}n}$ since by Lemma 5.2, for feasibility to X_n, we must uniquely have $w = \hat{w}$ and $v = \hat{v}$ as the completion to this vertex. To accomplish this, we show that for $(\hat{x},\hat{y})$ to satisfy $(\hat{x},\hat{y}) = \lambda(\bar{x},\bar{y}) + (1-\lambda)(\tilde{x},\tilde{y})$ for some $\lambda \in (0,1)$, $(\bar{x},\bar{y}) \in X_{\mathrm{P}n}$, and $(\tilde{x},\tilde{y}) \in X_{\mathrm{P}n}$, we must have $(\hat{x},\hat{y}) = (\bar{x},\bar{y}) = (\tilde{x},\tilde{y})$. Observe that since $X_{\mathrm{P}n} \subseteq X_0$ by Theorem 3.2, we have $0 \leqslant \bar{x} \leqslant e_n$ and $0 \leqslant \tilde{x} \leqslant e_n$. Since $\hat{x}$ is binary, for any $\lambda \in (0,1)$ we get $\hat{x} = \bar{x} = \tilde{x}$. Using this result along with the supposition that $\hat{y}$ is an extreme point to $\hat{Y}$, we deduce that $\hat{y} = \bar{y} = \tilde{y}$, and the proof is complete. □

Note that Theorem 5.3 essentially asserts that in projecting the set X_n from the (x, y, w, v) space to the set X_{Pn} in the (x, y) space, all extreme points are preserved. If $(\hat{x}, \hat{y})$ is a vertex of X_{Pn}, then by Theorem 3.5 and Lemma 5.2, $(\hat{x}, \hat{y}, \hat{w}, \hat{v})$ as defined by the theorem is a vertex of X_n. Conversely, if $(\hat{x}, \hat{y}, \hat{w}, \hat{v})$ is a vertex of X_n, then it satisfies the conditions stated in the theorem and, as shown in the proof, yields $(\hat{x}, \hat{y})$ as a vertex of X_{Pn}.

In concluding this section, let us comment on the situation in which there exist certain constraints from the first set of inequalities in (1) which involve only the x-variables. As mentioned in Section 2, one can multiply such constraints with the factors y_k and $(1 - y_k)$ for $k = 1, \ldots, m$, and then linearize the resulting constraints using (4) as with the other constraints (3). By Lemma 5.2, these constraints are implied when x is binary in any feasible solution, and moreover, since they serve to tighten the continuous relaxation, Theorem 3.2 continues to hold. Furthermore, since X_n has x binary for all vertices by Theorem 5.3, and X_n is bounded, these constraints are implied by the other constraints defining X_n by Lemma 5.2. Consequently, Theorems 3.5 and 5.3 also continue to hold with the inclusion of such constraints. In a likewise fashion, such inequalities defining X which involve only x-variables can be used in the same spirit as the factors $x_j \geqslant 0$ and $(1 - x_j) \geqslant 0$ to generate products of inequalities up to a given order level in order to further tighten intermediate level relaxations. Of course, by virtue of Lemma 5.2 and Theorem 5.3, such constraints are again all implied by the constraints defining X_n.

6. Characterization of the facets of the convex hull of feasible solutions

We will now derive a characterization for the facets of $X_{Pn} \equiv \operatorname{conv}(X)$ using a projection operation. Since under a nonsingular linear transformation, the extreme points, facets, and the boundedness of a polyhedron are preserved, we will conveniently use the form (16b)–(16d) obtained under the transformation (11) and (12) of Lemma 3.3 to represent the set X_{Pn}. Noting (13), we may therefore write

$$X_{Pn} = \Bigg\{(x, y):$$

$$x_j = \sum_{J \subseteq N:\, j \in J} U_J^0 \text{ for } j = 1, \ldots, n, \tag{31a}$$

$$y_k = \sum_{J \subseteq N} U_J^k \text{ for } k = 1, \ldots, m, \tag{31b}$$

$$\sum_{k=1}^{m} \gamma_{rk} U_J^k \geqslant \delta_{Jr} U_J^0 \text{ for } r = 1, \ldots, R,\ J \subseteq N, \tag{31c}$$

$$\sum_{J \subseteq N} U_J^0 = 1, \tag{31d}$$

$$0 \leqslant U_J^k \leqslant U_J^0 \text{ for } k = 1, \ldots, m,\ J \subseteq N\Bigg\}. \tag{31e}$$

Associating Lagrange multipliers π_j, $j = 1, \dots, n$, λ_k, $k = 1, \dots, m$, θ_{Jr}, $r = 1, \dots, R$, $J \subseteq N$, π_0, and ψ_{Jk}, $k = 1, \dots, m$, $J \subseteq N$ with respect to the constraints (31a), (31b), (31c), (31d), and the variable upper bounding constraints in (31e), respectively, we have by linear programming duality that $(x, y) \in X_{Pn}$ if and only if

$$0 = \max\{\pi x + \lambda y + \pi_0 \colon (\pi, \lambda, \theta, \pi_0, \psi) \in \mathrm{PC}\} \tag{32}$$

where PC is a polyhedral cone defined as follows:

$$\mathrm{PC} = \Bigg\{(\pi, \lambda, \theta, \pi_0, \psi):$$

$$\sum_{j \in J} \pi_j - \sum_{r=1}^{R} \delta_{Jr}\theta_{Jr} + \pi_0 + \sum_{k=1}^{m} \psi_{Jk} \leqslant 0 \text{ for } J \subseteq N, \tag{33a}$$

$$\lambda_k + \sum_{r=1}^{R} \gamma_{rk}\theta_{Jr} - \psi_{Jk} \leqslant 0 \text{ for } k = 1, \dots, m,\ J \subseteq N, \tag{33b}$$

$$\theta_{Jr} \geqslant 0,\ r = 1, \dots, R,\ J \subseteq N,\ \psi_{Jk} \geqslant 0,\ k = 1, \dots, m,\ J \subseteq N\Bigg\}. \tag{33c}$$

Now, consider the following result.

Theorem 6.1. *The set PC defined in* (33) *is an unbounded polyhedral cone with vertex at the origin and has some L distinct extreme directions or generators* $(\pi^l, \lambda^l, \theta^l, \pi_0^l, \psi^l)$, $l = 1, \dots, L$, $L \geqslant 1$, *with* $\pi_0^l = 0, +1$, *or* -1. *Moreover,*

$$X_{Pn} = \{(x, y) \colon \pi^l x + \lambda^l y \leqslant -\pi_0^l,\ l = 1, \dots, L\}. \tag{34}$$

Proof. Noting that π, λ, and π_0 are unrestricted in sign in (33), PC is clearly unbounded. Furthermore, enforcing all the defining inequalities to be binding yields $\theta \equiv 0$ and $\psi \equiv 0$ from (33c), which implies that $\lambda \equiv 0$ from (33b), and from (33a) for $J = \emptyset, \{1\}, \{2\}, \dots, \{n\}$, we, respectively, obtain $\pi_0 = 0$, $\pi_1 = 0, \dots, \pi_n = 0$. Hence, this produces the origin as the unique feasible solution, and so there exist some q linearly independent defining hyperplanes in (33) which are binding at the origin, where q is the dimension of $(\pi, \lambda, \theta, \pi_0, \psi)$. Consequently, PC is a pointed polyhedral cone with the vertex at the origin, and has some L distinct extreme directions as stated in the theorem, each produced by some $(q - 1)$ linearly independent hyperplanes binding from (33). Moreover, (32) holds if and only if $\pi^l x + \lambda^l y + \pi_0^l \leqslant 0$ for $l = 1, \dots, L$, and hence, X_{Pn} is given by (34). This completes the proof. $\square$

Corollary 6.2. *Alternately,* X_{Pn} *is given by* (34) *where* $(\pi^l, \lambda^l, \theta^l, \pi_0^l, \psi^l)$, $l = 1, \dots, L$, *are the extreme points of the set*

$$\overline{\mathrm{PC}} = \mathrm{PC} \cap \Bigg\{(\pi, \lambda, \theta, \pi_0, \psi) \colon \sum_{J \subseteq N} \Bigg[\sum_{j \in J} \pi_j + \sum_{k=1}^{m} \lambda_k + \pi_0 - \sum_{k=1}^{m} \psi_{Jk} + \sum_{r=1}^{R} \theta_{Jr} \Bigg(\sum_{k=1}^{m} \gamma_{rk} - \delta_{Jr} - 1 \Bigg) = -1 \Bigg] \Bigg\}. \tag{35}$$

Proof. Follows from the fact that the constraint imposed on PC in (35) is simply a regularization constraint on the generators of PC. $\square$

Note that if $X_{\mathrm{P}n} \neq \emptyset$, then the facet defining inequalities for $X_{\mathrm{P}n}$ are among the constraints in (34) if $X_{\mathrm{P}n}$ is full dimensional, and are given through appropriate intersections of these defining hyperplanes otherwise. Of course, a total enumeration of such constraints is prohibitive. However, the structure of $X_{\mathrm{P}n}$ and PC given in (31) and (33), respectively, can be possibly exploited to generate various classes of facets via Theorem 6.1. This might be achievable by characterizing certain classes (not necessarily all) generators of PC for special types of problems. This is the principal potential utility of this characterization. Furthermore, we can also generate specific valid inequalities for the problem by noting that any inequality $\pi x + \lambda y + \pi_0 \leqslant 0$ is valid for $X_{\mathrm{P}n}$ if and only if it can be obtained by surrogating the constraints (31) using a solution feasible to the (dual) system (33) defining PC. (Here, the surrogate multipliers to be used on the two families of nonnegativity restrictions on U_j^0 and U_j^k in (31e) are the slacks in (33a) and (33b), respectively.)

7. Preliminary computational results and extensions

We have presented in this paper a RLT for generating tight linear programming relaxations for linear mixed-integer zero–one programming problems. These relaxations span the spectrum from the ordinary continuous relaxation to the convex hull of feasible solutions, in a hierarchy of sharper representations. Our main objective in this paper has been to lay the theoretical foundation for this approach. The framework developed has the potential for deriving strong valid inequalities and characterizing facets for various classes of special problems. From a practical computational viewpoint, one may work with only the relaxation X_1, or one may devise techniques for generating tight valid inequalities implied by higher-order relaxations, or one may explicitly generate convex hull representations, as in this paper, separately for various subsets of sparse constraints which involve a manageable number of variables.

To provide some computational evidence, we present a sample of test results from [1, 2], wherein an algorithm is developed to solve a mixed-integer bilinear programming problem of the form

$$\text{minimize}\{c^{\mathrm{t}}x + d^{\mathrm{t}}y + y^{\mathrm{t}}Cx\colon x \in X,\ x \in Y,\ x \text{ binary}\}, \tag{36}$$

where X and Y are nonempty polytopes. The test problems relate to an application in which the set Y is comprised of transportation constraints representing a flow of products between certain origin-destination pairs, X is comprised of set covering types of constraints representing certain discrete decisions dealing with geographical or technological coverage, and the cross-product terms in $y^{\mathrm{t}}Cx$ subsidize shipment costs associated with implemented discrete decisions. The strategy used here was to generate a linear programming relaxation for (36) in the spirit of the first level

Table 1
Sample test results on the strength of a first level relaxation

Problem	X (rows, columns) (density $= 0.5$)	Y (rows, columns)	LP value / MIP value	LB / MIP value
1	(10, 15)	(14, 49)	0.985	0.479
2	(10, 20)	(14, 49)	0.993	0.510
3	(10, 25)	(14, 49)	0.997	0.510
4	(10, 15)	(20, 100)	0.986	0.542
5	(10, 20)	(20, 100)	0.995	0.483
6	(10, 25)	(20, 100)	0.998	0.452

relaxation along with the comments given toward the end of Section 5, by first multiplying the constraints in Y with factors x_j and $(1 - x_j)$ $\forall j$, and also by multiplying the constraints in X with factors y_i and $(y_i^+ - y_i)$ $\forall i$, where $y_i^+ = \max\{y_i: y \in Y\}$ $\forall i$, and then using (4) to linearize the resulting problem. Table 1 gives a typical set of results. Note that the linear programming based bound (*LP value*) is within 1–2% of optimality (*the MIP value*), whereas a standard bound given by

$$\mathrm{LB} = \min\{c^{\mathrm{t}}x: x \in X\} + \min\left\{\sum_i (d_i + C_i^-)y_i: y \in Y\right\}, \tag{37}$$

where $C_i^- \equiv \min\{\sum_j C_{ij}x_j: x \in X\}$ $\forall i$ performs quite poorly in comparison.

We conclude this paper by presenting two important extensions. (For the development of a RLT to solve *continuous*, nonconvex, polynomial programming problems, see [9], and for an application, along with computational results, related to solving continuous bilinear programming problems, and certain location–allocation problems, see [8, 10].

The first extension presented herein concerns multilinear mixed-integer zero–one *polynomial* programming problems in which the continuous variables $0 \leqslant y \leqslant e_m$ appear linearly in the constraints and the objective function. This is discussed below.

Extension 1: Multilinear mixed-integer zero–one polynomial programming problems

Consider the set

$$X = \left\{(x, y): \sum_{t \in T_{r0}} \alpha_{rt}\, p(J_{1t}, J_{2t}) + \sum_{k=1}^{m} y_k \sum_{t \in T_{rk}} \gamma_{rkt}\, p(J_{1t}, J_{2t}) \geqslant \beta_r,\ r = 1, \ldots, R,\right.$$
$$\left. 0 \leqslant x \leqslant e_n \text{ and integer},\ 0 \leqslant y \leqslant e_m \right\}, \tag{36}$$

where for all t, $p(J_{1t}, J_{2t}) \equiv [\prod_{j \in J_{1t}} x_j][\prod_{j \in J_{2t}} (1 - x_j)]$ are polynomial terms for the various sets (J_{1t}, J_{2t}) in (36). For $d = 0, 1, \ldots, n$, we can construct a polyhedral relaxation X_d for X by using the factors $F_d(J_1, J_2)$ to multiply the first set of constraints as before, where (J_1, J_2) are of order d. However, denoting δ_1 as the maximum degree of the polynomial terms in x not involving the y-variables, and δ_2 as the maximum degree of the polynomial terms in x which are associated with products involving y-variables, in lieu of (3b), we now use $F_{D_1}(J_1, J_2) \geqslant 0$ for (J_1, J_2) of order $D_1 = \min\{d + \delta_1, n\}$, and in lieu of (3c), we employ the constraints $F_{D_2}(J_1, J_2) \geqslant y_k F_{D_2}(J_1, J_2) \geqslant 0$, $k = 1, \ldots, m$, for all (J_1, J_2) of order $D_2 = \min\{d + \delta_2, n\}$. Note that in computing δ_1 and δ_2 in an optimization context, we consider the terms in the objective function as well, and that for the linear case, we have $\delta_1 = 1$ and $\delta_2 = 0$. Now, linearizing the resulting constraints under the substitution (4) produces the desired set X_d. Because of Lemma 5.1, when the integrality on the x-variables is enforced, each such set X_d is equivalent to the set X.

Moreover, by Lemma 3.1 and Eq. (8), the proof of Theorem 3.2 continues to hold. In particular, because of (8), each constraint from the first set of inequalities in X_d for any $d < n$ is obtainable by surrogating two appropriate constraints from X_{d+1} as in the proof of Theorem 3.2. Hence, we again obtain the hierarchy of relaxations $\mathrm{conv}(X) \subseteq X_{\mathrm{P}n} \subseteq X_{\mathrm{P}(n-1)} \subseteq \cdots \subseteq X_{\mathrm{P}1} \subseteq X_0$. Furthermore, by Lemma 5.1, Lemma 5.2 also holds for this situation.

Now, consider the set $X_{\mathrm{P}n}$. The constraints (3b) and (3c) for this set are

$$f_n(J, \bar{J}) \geqslant f_n^k(J, \bar{J}) \geqslant 0, \quad \text{for all } k = 1, \ldots, m,\ J \subseteq N. \tag{37a}$$

Furthermore, multiplying the first set of constraints defining X in (36) by the factors $F_n(J, \bar{J})$ for all $J \subseteq N$ produces, upon linearization through the substitution (4), the following inequalities:

$$\sum_{k=1}^{m} \bar{\gamma}_{rk} f_n^k(J, \bar{J}) \geqslant \delta_{Jr} f_n(J, \bar{J}) \quad \text{for } r = 1, \ldots, R,\ J \subseteq N, \tag{37b}$$

where $\delta_{Jr} = \beta_r - \{\sum_t \alpha_{rt}\colon t \in T_{r0},\ J_{1t} \subseteq J, \text{ and } J_{2t} \subseteq \bar{J}\}$, for $r = 1, \ldots, R$, $J \subseteq N$, and where $\bar{\gamma}_{rk} = \{\sum_t \gamma_{rkt}\colon t \in T_{rk},\ J_{1t} \subseteq J, \text{ and } J_{2t} \subseteq \bar{J}\}$ for $r = 1, \ldots, R$ and $k = 1, \ldots, m$. Noting that (37) is precisely of the same form as the inequalities of $X_{\mathrm{P}n}$ given by (15b) for the linear case, the proof of Theorem 3.5 implies that $X_{\mathrm{P}n} \equiv \mathrm{conv}(X)$ for this case as well. Moreover, Theorem 5.3 and the characterization of facets for $\mathrm{conv}(X)$ continue to hold as for the linear case.

Extension 2: Construction of relaxations using equality constraint representations

Note that given any set X of the form (1), by adding slack variables to the first R constraints, determining upper bounds on these slacks as the sum of the positive constraint coefficients minus the right-hand side constant, and accordingly scaling these slack variables onto the unit interval, we may equivalently write the set X as

$$X = \left\{(x, y)\colon \sum_{j=1}^{n} \alpha_{rj} x_j + \sum_{k=1}^{m} \gamma_{rk} y_k + \gamma_{r(m+r)} y_{m+r} = \beta_r \text{ for } r = 1, \ldots, R,\right.$$
$$\left. 0 \leqslant x \leqslant e_n,\ x \text{ integer},\ 0 \leqslant y \leqslant e_{m+R}\right\}. \tag{38}$$

Now, for any $d \in \{1, \ldots, n)$, observe that the factor $F_d(J_1, J_2)$ for any (J_1, J_2) of order d is a linear combination of the factors $F_p(J, \emptyset)$ for $J \subseteq N$, $p \equiv |J| = 0, 1, \ldots, d$. Hence, the constraint derived by multiplying an *equality* from (38) by $F_d(J_1, J_2)$ and then linearizing it via (4) is obtainable via an appropriate surrogate (with mixed-sign multipliers) of the constraints derived similarly, but using the factors $F_p(J, \emptyset)$ for $J \subseteq N$, $p \equiv |J| = 0, 1, \ldots, d$. Hence, these latter factors produce constraints which can generate the other constraints, and so X_d defined by (5) corresponding to X as in (38) is *equivalent* to the following, where $f_p(J, \emptyset) \equiv w_J$, and $f_p^k(J, \emptyset) \equiv v_{Jk}$ for all $p = 0, 1, \ldots, d+1$ as in (4):

$$X_d = \Bigg\{(x, y, w, v)\text{: constraints of type (5b) and (5c) hold, and for } r = 1, \ldots, R,$$

$$\left[\sum_{j \in J} \alpha_{rj} - \beta_r\right] w_J + \sum_{i \notin J} \alpha_{rj}\, w_{J+j} + \sum_{k=1}^{m} \gamma_{rk}\, v_{Jk} + \gamma_{r(m+r)} v_{J(m+r)} = 0$$

$$\text{for all } J \subseteq N \text{ with } |J| = 0, 1, \ldots, d\Bigg\}. \tag{39}$$

Note that the savings in the number of constraints in (39) over that in (5) corresponding to the set X as in (38) is given by

$$R\left[2^d \binom{n}{d} - \sum_{i=0}^{d} \binom{n}{i}\right].$$

Also, observe that for $J = \emptyset$, the equalities in (39) are precisely the original equalities defining X in (38). Hence, using Lemma 3.1, the assertion of Theorem 3.2 is directly seen to be true for (39). Of course, because (39) is equivalent to the set of the type (5) which would have been derived using the factors $F_d(J_1, J_2)$ of degree d, all the foregoing results continue to hold for (39). However, establishing that $X_{Pn} = \mathrm{conv}(X)$ and characterizing the facets of X_{Pn} is more conveniently managed using the constructs of Sections 3 and 6.

While the approach in Sections 3 and 6 for the inequality constrained case avoids the manipulation of surrogates of the equalities in (39) for theoretical purposes, note that from a computational viewpoint, when $d < n$, the representation in (39) has fewer type (5a) "complicating" constraints and variables (including slacks in (5a)) than does (5) as given by the above savings expression, but has $R \times 2^d \binom{n}{d}$ additional constraints of the type (5c), counting the nonnegativity restrictions on the slacks in (5a) for the inequality constrained cases. Hence, depending on the structure which is more convenient, either form of the representation of these relaxations may be employed.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant Nos ECS-8807090 and DDM-9121419, and the Air Force Office of Scientific Research under Grant No. AFOSR-90-0191.

References

[1] W.P. Adams and H.D. Sherali, Linearization strategies for a class of zero–one mixed integer programming problems, Oper. Res. 38 (2) 1990 217–226.

[2] W.P. Adams and H.D. Sherali, Mixed-integer bilinear programming problems, Math. Programming 59 (1993) 279–305.

[3] E. Balas, Disjunctive programming and a hierarchy of relaxations for discrete optimization problems, SIAM J. Algebraic Discrete Methods 6 (1985) 466–486.

[4] E. Balas, S. Ceria and G. Cornuejols, A lift-and-project cutting plane algorithm for mixed 0–1 programs, Math. Programming 58 (1993) 295–324.

[5] E. Boros, Y. Crama and P.L. Hammer, Upper bounds for quadratic 0–1 maximization problems, RUTCOR, Report RRR #14–89, Rutgers University, New Brunswick, NJ 08903 (1989).

[6] L. Lovász and A. Schrijver, Cones of matrices and setfunctions, and 0–1 optimization, Department of Operations Research, Statistics, and System Theory, Report BS-R8925, Center of Mathematics and Computer Science, P.O.B. 4079, 1009 AB Amsterdam, Netherlands (1989).

[7] H.D. Sherali and W.P. Adams, A hierarchy of relaxations between the continuous and convex hull representations for zero–one programming problems, SIAM J. Discrete Math. 3 (1990) 411–430.

[8] H.D. Sherali and A. Allameddine, A new reformulation-linearization technique for bilinear programming problems, J. Global Optim., 2 (1992) 379–410.

[9] H.D. Sherali and C.H. Tuncbilek, A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique, J. Global Optim. 2 (1992) 101–112.

[10] H.D. Sherali and C.H. Tuncbilek, A squared-euclidean distance location–allocation problem, Naval Res. Logist. Quarterly 39 (1992) 447–469.