Don Coppersmith · Oktay Günlük · Jon Lee · Janny Leung

A Polytope for a Product of Real Linear Functions in 0/1 Variables

Original: 29 September 1999 as IBM Research Report RC21568
Revised version: 30 November 2003

Abstract. In the context of integer programming, we develop a polyhedral method for linearizing a product of a pair of real linear functions in 0/1 variables. As an example, by writing a pair of integer variables in binary expansion, we have a technique for linearizing their product. We give a complete linear description for the resulting polytope, and we provide an efficient algorithm for the separation problem. Along the way to establishing the complete description, we also give a complete description for an extended-variable formulation, and we point out a generalization.

Introduction

We assume familiarity with polyhedral methods of linear integer programming (see [NW88], for example). There is a well-known method of linear integer programming for modeling the product (i.e., logical and) of a *pair* of binary variables. Specifically, the 0/1 solutions of $y = x^1 x^2$ are, precisely, the extreme points of the polytope in $\mathbb{R}^3$ that is the solution set of

$$y \geq 0 \; ; \tag{1}$$
$$y \geq x^1 + x^2 - 1 \; ; \tag{2}$$
$$y \leq x^1 \; ; \tag{3}$$
$$y \leq x^2 \; . \tag{4}$$

There has been extensive interest in the development of linear integer programming methods for handling and/or exploiting products of *many* 0/1 variables. In particular, quite a lot is known about the facets of the convex hull of the 0/1 solutions to: $y^{ij} = x^i x^j$, $1 \leq i < j \leq n$ (the *boolean quadric polytope*). See [Pad89]. This polytope is related to other well-studied structures such as the cut polytope [DS90] and the correlation polytope [Pit91]. Also see [BB98], [DL97] and references therein. Of course, we know less and less about the totality of the facets of the polytope as n increases, because optimization over these 0/1 solutions is NP-hard.

Don Coppersmith, Oktay Günlük, Jon Lee: Department of Mathematical Sciences, T.J. Watson Research Center, IBM. e-mail: {dcopper,gunluk,jonlee}@us.ibm.com

Janny Leung: Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong. e-mail: janny@se.cuhk.edu.hk

Mathematics Subject Classification (1991): 52B12, 90C10

Rather than considering pairwise products of many 0/1 variables, we consider a single product of a pair of real linear functions in 0/1 variables. For $l = 1, 2$, let k_l be a positive integer, let $K_l = \{1, 2, \ldots, k_l\}$, let $\mathbf{a}^l$ be a k_l-vector of positive real numbers a_i^l, and let $\mathbf{x}^l$ be a k_l-vector of binary variables x_i^l. Note that if we had any $a_i^l = 0$, then we could just delete such x_i^l, and if we had any $a_i^l < 0$, then we could just complement such x_i^l and apply our methods to the nonlinear part of the resulting product.

Now, we let $P(\mathbf{a}^1, \mathbf{a}^2)$ be the convex hull of the solutions in $\mathbb{R}^{k_1+k_2+1}$ of

$$y = \left(\sum_{i \in K_1} a_i^1 x_i^1\right)\left(\sum_{j \in K_2} a_j^2 x_j^2\right) = \sum_{i \in K_1}\sum_{j \in K_2} a_i^1 a_j^2 x_i^1 x_j^2 \; ; \tag{5}$$

$$x_i^l \in \{0,1\}, \text{ for } i \in K_l, \;\; l = 1, 2 \; . \tag{6}$$

We note that $P((1),(1))$ is just the solution set of (1–4). Our goal is to investigate the polytope $P(\mathbf{a}^1, \mathbf{a}^2)$ generally.

In Section 1, we describe an application to modeling a product of a pair of nonnegative integer variables using binary expansion. In Section 2, we describe a linear integer formulation of $P(\mathbf{a}^1, \mathbf{a}^2)$ In Section 3, we investigate which of our inequalities are facet describing. In Section 4, we determine a complete polyhedral characterization of $P(\mathbf{a}^1, \mathbf{a}^2)$. In establishing this characterization, we also find an inequality characterization of a natural extended-variable formulation. In Section 5, we demonstrate how to solve the separation problem for the facet describing inequalities of $P(\mathbf{a}^1, \mathbf{a}^2)$. In Section 6, we investigate some topological properties of *real* points in the $P(\mathbf{a}^1, \mathbf{a}^2)$ that satisfy the product equation (5). In Section 7, we briefly describe a generalization of our results.

1. Products of Nonnegative Integer Variables

Let x^1 and x^2 be a pair of nonnegative integer variables. We can look directly at the convex hull of the integer solutions of $y = x^1 x^2$. This convex set, in $\mathbb{R}^3$, contains integer points that do not correspond to solutions of $y = x^1 x^2$. For example, $x^1 = x^2 = y = 0$ is in this set, as is $x^1 = x^2 = 2, y = 4$, but the average of these two points is not a solution of $y = x^1 x^2$. More concretely, the convex hull of the integer solutions to:

$$\begin{aligned} &y = x^1 x^2 \; ; \\ &0 \le x^1 \le 2 \; ; \\ &0 \le x^2 \le 2 \end{aligned}$$

is precisely the solution set of

$$\begin{aligned} &y \ge 0 \; ; \\ &y \ge 2x^1 + 2x^2 - 4 \; ; \\ &y \le 2x^1 \; ; \\ &y \le 2x^2 \; . \end{aligned}$$

If we use these latter linear inequalities to model the product y, and then seek to maximize y subject to these constraints and a side constraint $x^1 + x^2 \leq 2$, we find the optimal solution $x^1 = 1$, $x^2 = 1$, $y = 2$, which does not satisfy $y = x^1 x^2$. Therefore, this naïve approach is inadequate in the context of linear integer programming.

We adopt an approach that avoids the problem above. Specifically, we assume that, for practical purposes, x^1 and x^2 can be bounded above. So we can write the x^l in binary expansion: $x^l = \sum_{i \in K_l} 2^{i-1} x_i^l$, for $l = 1, 2$. That is, we let $a_i^l = 2^{i-1}$. The only integer points in $P(\mathbf{a}^1, \mathbf{a}^2)$ are the solutions of (5–6). Therefore, we avoid the problem that we encountered when we did not use the binary expansions of x^1 and x^2.

2. Linear Integer Formulation

Obviously, for $l = 1, 2$, the *simple bound inequalities* are valid for $P(\mathbf{a}^1, \mathbf{a}^2)$:

$$x_i^l \geq 0, \quad \text{for } i \in K_l \ ; \tag{7}$$

$$x_i^l \leq 1, \quad \text{for } i \in K_l \ . \tag{8}$$

We view the remaining inequalities that we present as lower and upper bounds on the product variable y. In the sequel, H is a subset of $K_l \times K_{\bar{l}}$, where l is either 1 or 2, and $\bar{l} = 3 - l$.

Consider the following *lower bound inequalities* for y:

$$y \geq \sum_{(i,j) \in H} a_i^1 a_j^2 \left(x_i^1 + x_j^2 - 1 \right) \ . \tag{9}$$

Proposition 1. *The inequalities (9) are valid for $P(\mathbf{a}^1, \mathbf{a}^2)$.*

Proof.

$$\begin{aligned} y &= \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^1 x_j^2 \\ &\geq \sum_{(i,j) \in H} a_i^1 a_j^2 x_i^1 x_j^2 \\ &\geq \sum_{(i,j) \in H} a_i^1 a_j^2 \left(x_i^1 + x_j^2 - 1 \right) \ . \end{aligned}$$

□

We derive upper bounds on y by considering certain transformations ϕ^l, $l = 1, 2$, of the polytope $P(\mathbf{a}^1, \mathbf{a}^2)$:

$$\begin{aligned} \tilde{x}_i^l &= 1 - x_i^l \ , \quad \text{for } i \in K_l \ ; \\ \tilde{x}_j^{\bar{l}} &= x_j^{\bar{l}} \ , \quad \text{for } j \in K_{\bar{l}} \ ; \\ \tilde{y} &= \sum_{i \in K_l} \sum_{j \in K_{\bar{l}}} a_i^l a_j^{\bar{l}} \tilde{x}_i^l \tilde{x}_j^{\bar{l}} \ . \end{aligned}$$

Proposition 2. *The transformation ϕ^l, $l = 1, 2$, is an affine involution.*

Proof. Clearly $\phi^l \circ \phi^l$ is the identity transformation. To see that it is affine, we need just check that $\tilde{y}$ is an affine function of the original variables.

$$\begin{aligned}\tilde{y} &= \sum_{i\in K_l}\sum_{j\in K_{\bar{l}}} a_i^l a_j^{\bar{l}} \tilde{x}_i^l \tilde{x}_j^{\bar{l}} \\ &= \sum_{i\in K_l}\sum_{j\in K_{\bar{l}}} a_i^l a_j^{\bar{l}} (1 - x_i^l) x_j^{\bar{l}} \\ &= \sum_{i\in K_l}\sum_{j\in K_{\bar{l}}} a_i^l a_j^{\bar{l}} x_j^{\bar{l}} - \sum_{i\in K_l}\sum_{j\in K_{\bar{l}}} a_i^l a_j^{\bar{l}} x_i^l x_j^{\bar{l}} \\ &= \sum_{i\in K_l}\sum_{j\in K_{\bar{l}}} a_i^l a_j^{\bar{l}} x_j^{\bar{l}} - y \ .\end{aligned}$$

□

Our *upper bound inequalities* take the form:

$$y \le \sum_{i\in K_l}\sum_{j\in K_{\bar{l}}} a_i^l a_j^{\bar{l}} x_j^{\bar{l}} + \sum_{(i,j)\in H} a_i^l a_j^{\bar{l}} \left(x_i^l - x_j^{\bar{l}} \right) \ . \qquad (10)$$

Proposition 3. *The inequalities (10) are valid for $P(\mathbf{a}^1, \mathbf{a}^2)$.*

Proof. We apply the lower bound inequalities (9) to the variables transformed by ϕ^l, to obtain

$$\sum_{i\in K_l}\sum_{j\in K_{\bar{l}}} a_i^l a_j^{\bar{l}} x_j^{\bar{l}} - y = \tilde{y} \ge \sum_{(i,j)\in H} a_i^l a_j^{\bar{l}} \left(\tilde{x}_i^l + \tilde{x}_j^{\bar{l}} - 1 \right) \ .$$

Solving for y, and rewriting this in terms of x^l and $x^{\bar{l}}$, we obtain the upper bound inequalities (10). □

Note that the sets of inequalities (10) with $l = 1$ and $l = 2$ are equivalent — this follows by checking that changing l is equivalent to complementing H.

The transformation ϕ^l corresponds to the "switching" operation used in the analysis of the cut polytope (see [DL97], for example). Specifically, (2) and (3) are switches of each other under ϕ^1, and (2) and (4) are switches of each other under ϕ^2.

Proposition 4. *The points satisfying (6) and (9–10) for all cross products H are precisely the points satisfying (5–6).*

Proof. By Propositions 1 and 3, we need only show that every point satisfying (6) and (9–10) for all cross products H also satisfies (5). Let $(\mathbf{x}^1, \mathbf{x}^2, y)$ be a point satisfying (6) and (9–10). Letting

$$H = \{i \in K_1 \ : \ x_i^1 = 1\} \times \{j \in K_2 \ : \ x_j^2 = 1\} \ ,$$

we obtain a lower bound inequality (9) that is satisfied as an equation by the point $(\mathbf{x}^1, \mathbf{x}^2, y)$. Similarly, letting

$$H = \{i \in K_1 \ : \ x_i^1 = 0\} \times \{j \in K_2 \ : \ x_j^2 = 1\} \ ,$$

we obtain an upper bound inequality (10) that is satisfied as an equation by the point $(\mathbf{x}^1, \mathbf{x}^2, y)$. So $\mathbf{x}^1$ and $\mathbf{x}^2$ together with (9–10) for cross products H determine y. But by the definition of $P(\mathbf{a}^1, \mathbf{a}^2)$, the point $\left(\mathbf{x}^1, \mathbf{x}^2, \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^1 x_j^2\right)$ is in $P(\mathbf{a}^1, \mathbf{a}^2)$, so (5) must be satisfied. □

3. Facets

For $i \in K_l$, $l = 1, 2$, we let $\mathbf{e^i} \in \mathbb{R}^{k_l}$ denote the i-th standard unit vector. For simplicity, we say that a point is *tight* for an inequality if it satisfies the inequality as an equation.

Proposition 5. *For $k_1, k_2 > 0$, the polytope $P(\mathbf{a}^1, \mathbf{a}^2)$ is full dimensional.*

Proof. We prove this directly. The following $k_1 + k_2 + 2$ points in $P(\mathbf{a}^1, \mathbf{a}^2)$ are easily seen to be affinely independent:

- $(\mathbf{x}^1, \mathbf{x}^2, y) = (\mathbf{0}, \mathbf{0}, 0)$;
- $(\mathbf{x}^1, \mathbf{x}^2, y) = (\mathbf{e^j}, \mathbf{0}, 0)$, for $j \in K_1$;
- $(\mathbf{x}^1, \mathbf{x}^2, y) = (\mathbf{0}, \mathbf{e^j}, 0)$, for $j \in K_2$;
- $(\mathbf{x}^1, \mathbf{x}^2, y) = (\mathbf{e^1}, \mathbf{e^1}, a_1^1 a_1^2)$.

□

Proposition 6. *For $l = 1, 2$, the inequalities (7) describe facets of $P(\mathbf{a}^1, \mathbf{a}^2)$ when $k_l > 1$.*

Proof. Again, we proceed directly. For both $l = 1$ and $l = 2$, the following $k_1 + k_2 + 1$ points in $P(\mathbf{a}^1, \mathbf{a}^2)$ are tight for (7) and are easily seen to be affinely independent:

- $(\mathbf{x}^1, \mathbf{x}^2, y) = (\mathbf{0}, \mathbf{0}, 0)$;
- $(\mathbf{x}^1, \mathbf{x}^2, y)$ defined by $\mathbf{x}^l = \mathbf{e^j}$, $\mathbf{x}^{\bar{l}} = \mathbf{0}$, $y = 0$, for $j \in K_l \setminus \{i\}$;
- $(\mathbf{x}^1, \mathbf{x}^2, y)$ defined by $\mathbf{x}^l = \mathbf{0}$, $\mathbf{x}^{\bar{l}} = \mathbf{e^j}$, $y = 0$, for $j \in K_{\bar{l}}$;
- $(\mathbf{x}^1, \mathbf{x}^2, y)$ defined by $\mathbf{x}^l = \mathbf{e^m}$, for some $m \in K_l \setminus \{i\}$, $\mathbf{x}^{\bar{l}} = \mathbf{e^1}$, $y = a_m^l a_1^{\bar{l}}$.

□

Proposition 7. *For $l = 1, 2$, the inequalities (8) describe facets of $P(\mathbf{a}^1, \mathbf{a}^2)$ when $k_l > 1$.*

Proof. The affine transformation ϕ^l is an involution from the face of $P(\mathbf{a}^1, \mathbf{a}^2)$ described by the simple lower bound inequality $x_i^l \geq 0$ to the face of $P(\mathbf{a}^1, \mathbf{a}^2)$ described by the simple upper bound inequality $x_i^l \leq 1$. Since the affine transformation is invertible, it is dimension preserving. Therefore, by Proposition 6, the result follows. □

One might suspect from Proposition 4 that the only inequalities of the form (9–10) that yield facets have H being a cross product, but this is not the case. Let $\mathcal{H}(k_1, k_2)$ denote a particular set of subsets of $K_1 \times K_2$. Each H in $\mathcal{H}(k_1, k_2)$ arises by choosing a permutation of the variables

$$\{x_i^1 \ : \ i \in K_1\} \cup \{x_j^2 \ : \ j \in K_2\} \ .$$

If x_i^1 precedes x_j^2 in the permutation, which we denote by $x_i^1 \prec x_j^2$, then (i, j) is in H.

For example, if $S_1 \subset K_1$ and $S_2 \subset K_2$, then we get the cross product $H = S_1 \times S_2 \in \mathcal{H}(k_1, k_2)$ by choosing any permutation of the variables having $\{x_j^2 \ : \ j \in K_2 \setminus S_2\}$ first, followed by $\{x_i^1 \ : \ i \in S_1\}$, followed by $\{x_j^2 \ : \ j \in S_2\}$, followed by $\{x_i^1 \ : \ i \in K_1 \setminus S_1\}$.

As another example, let $k_1 = 3$ and $k_2 = 2$, and consider the permutation: $x_2^1, x_2^2, x_1^1, x_3^1, x_1^2$, which yields

$$H = \{(2, 2), (2, 1), (1, 1), (3, 1)\} \ .$$

This choice of H is not a cross product. However, this H yields the lower bound inequality:

$$\begin{aligned} y \geq\ & a_2^1 a_2^2 (x_2^1 + x_2^2 - 1) \\ & + a_2^1 a_1^2 (x_2^1 + x_1^2 - 1) \\ & + a_1^1 a_1^2 (x_1^1 + x_1^2 - 1) \\ & + a_3^1 a_1^2 (x_3^1 + x_1^2 - 1) \ . \end{aligned}$$

We demonstrate that this inequality describes a facet by displaying $k_1 + k_2 + 1 = 6$ affinely independent points in $P(\mathbf{a}^1, \mathbf{a}^2)$ that are tight for the inequality: We display the points as rows of the matrix below, where we have permuted the columns, in a certain manner, according to the permutation of the variables that led to H, and we have inserted a column of 1's. It suffices to check that the square "caterpillar matrix" obtained by deleting the y column is nonsingular.

x_2^2	x_1^2	1	x_2^1	x_1^1	x_3^1	y
0	0	1	1	1	1	$0 = (a_3^1 + a_1^1 + a_2^1)(0)$
0	1	1	1	1	1	$(a_3^1 + a_1^1 + a_2^1)(a_1^2)$
0	1	1	1	1	0	$(a_1^1 + a_2^1)(a_1^2)$
0	1	1	1	0	0	$(a_2^1)(a_1^2)$
1	1	1	1	0	0	$(a_2^1)(a_1^2 + a_2^2)$
1	1	1	0	0	0	$0 = (0)(a_1^2 + a_2^2)$

For $m \geq 3$, an order-m *caterpillar matrix* is obtained by choosing $k_1, k_2 \geq 1$ with $k_1 + k_2 + 1 = m$. The first row of such a matrix is $(\mathbf{0} \in \mathbb{R}^{k_2}, 1, \mathbf{1} \in \mathbb{R}^{k_1})$, and the last row is $(\mathbf{1} \in \mathbb{R}^{k_2}, 1, \mathbf{0} \in \mathbb{R}^{k_1})$. Each row is obtained from the one above it by either flipping the right-most 1 to 0, or flipping the 0 immediately to the left of the left-most 1 to 1. So the rows depict snapshots of a "caterpillar" of 1's that moves from right to left by either pushing its head forward a bit or pulling its tail forward a bit. We note that besides translating affine to linear dependence, the column of 1's precludes the unsettling possibility of a 1-bit caterpillar disappearing by pulling its tail forward a bit, and then reappearing by pushing its head forward a bit. Since caterpillar matrices have their 1's in each row consecutively appearing, they are totally unimodular (see [HK56]). That is, the determinant of a caterpillar matrix is in $\{0, \pm 1\}$. In fact, we have the following result.

Proposition 8. *The determinant of a caterpillar matrix is in* $\{\pm 1\}$.

Proof. The proof is by induction on m. The only order-3 caterpillar matrices are

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} .$$

It is easy to check that these have determinant in $\{\pm 1\}$.

Now, suppose that we have a caterpillar matrix of order $m \geq 4$. Depending on the bit flip that produces the last row from the one above it, the matrix has the form

$$\left(\begin{array}{c|c|c} 0\,0\cdots 0 & 1 & 1\,1\cdots 1\,1 \\ 0\,0\cdots 0 & 1 & 1\,1\cdots 1\,0 \\ & \vdots & \vdots \\ & 1 & 0 \\ & 1 & 0 \end{array}\right)$$

or

$$\left(\begin{array}{c|c|c} 0\,0\cdots 0\,0 & 1 & 1\,1\cdots 1 \\ 0\,0\cdots 0\,1 & 1 & 1\,1\cdots 1 \\ \vdots & \vdots & \\ 1 & 1 & \\ 1 & 1 & \end{array}\right) .$$

In the first case, we can expand the determinant along the last column and we obtain a caterpillar matrix of order $m-1$ with the same determinant as the original matrix, up to the sign. In the second case, we subtract the first row from the second row (which does not affect the determinant), and then we expand along the second row of the resulting matrix. Again, we obtain a caterpillar matrix of order $m-1$ with the same determinant as the original matrix, up to the sign. In either case, the result follows by induction. □

Proposition 9. *An inequality of the form (9) describes a facet of* $P(\mathbf{a}^1, \mathbf{a}^2)$ *when* $H \in \mathcal{H}(k_1, k_2)$.

Proof. By Proposition 1, these inequalities are valid for $P(\mathbf{a}^1, \mathbf{a}^2)$. It suffices to exhibit $k_1 + k_2 + 1$ affinely independent points of $P(\mathbf{a}^1, \mathbf{a}^2)$ that are tight for (9). Let Φ be the permutation that gives rise to H. It is easy to check that $(\mathbf{x}^1, \mathbf{x}^2, y) = (\mathbf{0}, \mathbf{1}, 0)$ is a point of $P(\mathbf{a}^1, \mathbf{a}^2)$ that is tight for (9). We generate the remaining $k_1 + k_2$ points by successively flipping bits in the order of the permutation Φ. We simply need to check that each bit flip preserves equality in (9). If a variable x_i^1 is flipped from 0 to 1, the increase in y (i.e., the left-hand side of (9)) and in $\sum_{(i,j)\in H} a_i^1 a_j^2 \left(x_i^1 + x_j^2 - 1\right)$ (i.e., the right-hand side of (9)) is precisely $\sum_{j:x_i^1 \prec x_j^2} a_i^1 a_j^2$. Similarly, if a variable x_j^2 is flipped from 1 to 0, the decrease in both of these quantities is precisely $\sum_{i:x_i^1 \prec x_j^2} a_i^1 a_j^2$.

Next, we arrange these $k_1 + k_2 + 1$ points, in the order generated, as the rows of a caterpillar matrix of order $k_1 + k_2 + 1$. A point $(\mathbf{x}^1, \mathbf{x}^2, y)$ yields the row $(\mathbf{x}_\Phi^2, 1, \mathbf{x}_\Phi^1)$, where $\mathbf{x}_\Phi^l$ is just $\mathbf{x}^l$ permuted according to the order of the x_i^l in Φ. Clearly this defines a caterpillar matrix, which is nonsingular by Proposition 8. Hence, the generated points are affinely independent, so (9) describes a facet when $H \in \mathcal{H}(k_1, k_2)$. □

Corollary 1. *Each inequality (9) with $H \in \mathcal{H}(k_1, k_2)$ admits a set of tight points in $P(\mathbf{a}^1, \mathbf{a}^2)$ that correspond to the rows of a caterpillar matrix.*

Proposition 10. *An inequality of the form (10) describes a facet of $P(\mathbf{a}^1, \mathbf{a}^2)$ when $H \in \mathcal{H}(k_1, k_2)$.*

Proof. Using the transformation ϕ^l, this follows from Proposition 9. □

Conversely, every caterpillar matrix of order $k_1 + k_2 + 1$ corresponds to a facet of the form (9). More precisely, we have the following result.

Proposition 11. *Let C be a caterpillar matrix of order $k_1 + k_2 + 1$ such that its first k_2 columns correspond to a specific permutation of $\{x_j^2 \ : \ j \in K_2\}$ and its last k_1 columns correspond to a specific permutation of $\{x_i^1 \ : \ i \in K_1\}$. Then there exists a facet of $P(a_1, a_2)$ of the form (9) such that the points corresponding to the rows of C are tight for it.*

Proof. It is easy to determine the permutation Ψ that corresponds to C, by interleaving the given permutations of $\{x_j^2 \ : \ j \in K_2\}$ and of $\{x_i^1 \ : \ i \in K_1\}$, according to the head and tail moves of the caterpillar. Then, as before, we form the H of (9) by putting (i,j) in H if $x_i^1 \prec x_j^2$ in the final permutation.

It is easy to see that each row of C corresponds to a point of $P(\mathbf{a}^1, \mathbf{a}^2)$ that is tight for the resulting inequality (9). □

4. Inequality Characterization

In this section, we demonstrate how every facet of $P(\mathbf{a}^1, \mathbf{a}^2)$ is one of the ones described in Section 3. We do this by projecting from an extended formulation in a higher-dimensional space.

Consider the system

$$y = \sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 \delta_{ij} \ ; \tag{11}$$

$$\delta_{ij} \le x^1_i \ , \quad \text{for all } i \in K_1, j \in K_2 \ ; \tag{12}$$

$$\delta_{ij} \le x^2_j \ , \quad \text{for all } i \in K_1, j \in K_2 \ ; \tag{13}$$

$$\delta_{ij} \ge x^1_i + x^2_j - 1 \ , \quad \text{for all } i \in K_1, j \in K_2 \ ; \tag{14}$$

$$\delta_{ij} \ge 0 \ , \quad \text{for all } i \in K_1, j \in K_2 \ , \tag{15}$$

and let

$$Q^\delta(\mathbf{a}^1, \mathbf{a}^2) = \text{conv}\Big\{\mathbf{x}^1 \in \mathbb{R}^{k_1},\ \mathbf{x}^2 \in \mathbb{R}^{k_2}\ , y \in \mathbb{R},\ \delta \in \mathbb{R}^{k_1 \times k_2}\ :\ \text{(7-8,11-15)}\Big\},$$

where we use conv(X) to denote the convex hull of X. Let $Q(\mathbf{a}^1, \mathbf{a}^2)$ be the projection of $Q^\delta(\mathbf{a}^1, \mathbf{a}^2)$ in the space of $\mathbf{x}^1, \mathbf{x}^2$ and y variables. We next show that $Q(\mathbf{a}^1, \mathbf{a}^2)$ is integral.

Proposition 12. *$Q(\mathbf{a}^1, \mathbf{a}^2)$ is integral on $\mathbf{x}^1$ and $\mathbf{x}^2$.*

Proof. We will show that if $\mathbf{p}$ is fractional (on $\mathbf{x}^1$ and $\mathbf{x}^2$), then it is not an extreme point of $Q(\mathbf{a}^1, \mathbf{a}^2)$.

Assume that $\mathbf{p} = (\mathbf{x}^1, \mathbf{x}^2, y)$ is a fractional extreme point of $Q(\mathbf{a}^1, \mathbf{a}^2)$. For $v \in R$, let $(v)^+ = \max\{0, v\}$. For fixed $\mathbf{x}^1$ and $\mathbf{x}^2$, notice that $\delta \in \mathbb{R}^{k_1 \times k_2}$ is feasible to (12-15) if an only if it satisfies $\min\{x^1_i, x^2_j\} \ge \delta_{ij} \ge (x^1_i + x^2_j - 1)^+$ for all $i \in K_1$, $j \in K_2$. Therefore, if we define

$$y_{up} = \sum_{i \in K_1} \sum_{j \in K_2} a^1_i a^2_j \min\{x^1_i, x^2_j\},$$

$$y_{down} = \sum_{i \in K_1} \sum_{j \in K_2} a^1_i a^2_j (x^1_i + x^2_j - 1)^+$$

then the points $\mathbf{p}_{up} = (\mathbf{x}^1, \mathbf{x}^2, y_{up})$, and $\mathbf{p}_{down} = (\mathbf{x}^1, \mathbf{x}^2, y_{down})$ are in $Q(\mathbf{a}^1, \mathbf{a}^2)$, and $\mathbf{p}_{up} \ge \mathbf{p} \ge \mathbf{p}_{down}$. Furthermore, if $\mathbf{p}$ is an extreme point, it has to be one of $\mathbf{p}_{up}$ and $\mathbf{p}_{down}$.

Let $\bar{K}_1 \subseteq K_1$ and $\bar{K}_2 \subseteq K_2$ be the set of indices corresponding to fractional components of $\mathbf{x}^1$ and $\mathbf{x}^2$ respectively. Clearly, $\bar{K}_1 \cup \bar{K}_2 \neq \emptyset$. Let $\epsilon > 0$ be a small number so that $1 > x^l_i + \epsilon > x^l_i - \epsilon > 0$ for all $i \in \bar{K}_l$, $l = 1, 2$. Define $\mathbf{x}^{l+}$ where $x^{l+}_i := x^l_i + \epsilon$ if $i \in \bar{K}_l$ and $x^{l+}_i := x^l_i$, otherwise. Define x^{l-} similarly. We consider the following two cases and show that if $\mathbf{p}$ is fractional, then it can be represented as a convex combination of two distinct points in $Q(\mathbf{a}^1, \mathbf{a}^2)$.

Case 1. Assume $\mathbf{p} = \mathbf{p}_{up}$. Let $\mathbf{p}^a = (\mathbf{x}^{1+}, \mathbf{x}^{2+}, y^a)$ and $\mathbf{p}^b = (\mathbf{x}^{1-}, \mathbf{x}^{2-}, y^b)$ where

$$y^a = \sum_{i \in K_1} \sum_{j \in K_2} a^1_i a^2_j \min\{x^{1+}_i, x^{2+}_j\},$$
$$y^b = \sum_{i \in K_1} \sum_{j \in K_2} a^1_i a^2_j \min\{x^{1-}_i, x^{2-}_j\},$$

and note that $\mathbf{p}^a, \mathbf{p}^b \in Q(\mathbf{a}^1, \mathbf{a}^2)$ and $\mathbf{p}^a \neq \mathbf{p}^b$.

For $i \in K_1$ and $j \in K_2$, let $\delta_{ij} = \min\{x_i^1, x_j^2\}$, and define δ_{ij}^a and δ_{ij}^b similarly. Note that $y = \sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 \delta_{ij}$. Due to the construction, if $\min\{x_i^1, x_j^2\} = 0$, then we have $\min\{x_i^{1+}, x_j^{2+}\} = \min\{x_i^{1-}, x_j^{2-}\} = 0$, and therefore $\delta_{ij} = \delta_{ij}^a = \delta_{ij}^b = 0$. Similarly, if $\delta_{ij} = 1$, then we have $\delta_{ij}^a = \delta_{ij}^b = 1$ as well. On the other hand, if $\delta_{ij} \notin \{0, 1\}$, then $\delta_{ij}^a = \delta_{ij} + \epsilon$ and $\delta_{ij}^b = \delta_{ij} - \epsilon$. Therefore, $\delta^a + \delta^b = 2\delta$ and $\mathbf{p} = (1/2)\mathbf{p}^a + (1/2)\mathbf{p}^b$.

Case 2. Assume $\mathbf{p} = \mathbf{p}_{down}$. Let $\mathbf{p}^c = (\mathbf{x}^{1+}, \mathbf{x}^{2-}, y^c)$ and $\mathbf{p}^d = (\mathbf{x}^{1-}, \mathbf{x}^{2+}, y^d)$ where

$$y^c = \sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 (x_i^{1+} + x_j^{2-} - 1)^+,$$
$$y^d = \sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 (x_i^{1-} + x_j^{2+} - 1)^+,$$

and note that $\mathbf{p}^c, \mathbf{p}^d \in Q(\mathbf{a}^1, \mathbf{a}^2)$ and $\mathbf{p}^c \neq \mathbf{p}^d$.

Let $\delta_{ij} = (x_i^1 + x_j^2 - 1)^+$ for $i \in K_1$ and $j \in K_2$, and define δ_{ij}^c and δ_{ij}^d similarly. Note that $y = \sum_{i\in K_1} \sum_{j\in K_2} \delta_{ij}$.

If $\min\{x_i^1, x_j^2\} = 0$, then $\min\{x_i^{1+}, x_j^{2-}\} = \min\{x_i^{1-}, x_j^{2+}\} = 0$ and $\delta_{ij} = \delta_{ij}^c = \delta_{ij}^d = 0$. If $x_i^1 = 1$, then $\delta_{ij} = x_j^2$, implying $\delta_{ij}^c = x_j^{2-}$ and $\delta_{ij}^d = x_j^{2+}$. Similarly, if $x_i^2 = 1$, then $\delta_{ij} = x_j^1$, implying $\delta_{ij}^c = x_j^{1+}$ and $\delta_{ij}^d = x_j^{1-}$. Finally, if $1 > x_i^1, x_j^2 > 0$, then $\delta_{ij} = \delta_{ij}^c = \delta_{ij}^d$. Therefore, $\delta^c + \delta^d = 2\delta$ and $\mathbf{p} = (1/2)\mathbf{p}^c + (1/2)\mathbf{p}^d$. □

Now, let $R(\mathbf{a}^1, \mathbf{a}^2)$ be the real solution set of (7–10), and note that $P(\mathbf{a}^1, \mathbf{a}^2) \subseteq R(\mathbf{a}^1, \mathbf{a}^2)$. To prove that $P(\mathbf{a}^1, \mathbf{a}^2) = R(\mathbf{a}^1, \mathbf{a}^2)$, we will first argue that $Q(\mathbf{a}^1, \mathbf{a}^2) \subseteq P(\mathbf{a}^1, \mathbf{a}^2)$, and then we will show that $R(\mathbf{a}^1, \mathbf{a}^2) \subseteq Q(\mathbf{a}^1, \mathbf{a}^2)$.

Proposition 13. $Q(\mathbf{a}^1, \mathbf{a}^2) \subseteq P(\mathbf{a}^1, \mathbf{a}^2)$.

Proof. As $P(\mathbf{a}^1, \mathbf{a}^2)$ is a bounded convex set, it is sufficient to show that all of the extreme points of $Q(\mathbf{a}^1, \mathbf{a}^2)$ are contained in $P(\mathbf{a}^1, \mathbf{a}^2)$. Using Proposition 12 and its proof, we know that if $\mathbf{p} = (\mathbf{x}^1, \mathbf{x}^2, y)$ is an extreme point of $Q(\mathbf{a}^1, \mathbf{a}^2)$, then $\mathbf{x}^1$ and $\mathbf{x}^2$ are integral and

$$\sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 \min\{x_i^1, x_j^2\} \ \geq \ y \ \geq \ \sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 (x_i^1 + x_j^2 - 1)^+.$$

Notice that for any $u, v \in \{0, 1\}$, $\min\{u, v\} = (u + v - 1)^+ = uv$. Therefore, $y = \sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 x_i^1 x_j^2$, and $\mathbf{p} \in P(\mathbf{a}^1, \mathbf{a}^2)$. □

Proposition 14. $R(\mathbf{a}^1, \mathbf{a}^2) \subseteq Q(\mathbf{a}^1, \mathbf{a}^2)$.

Proof. Assume not, and let $\mathbf{p} = (\mathbf{x}^1, \mathbf{x}^2, y) \in R(\mathbf{a}^1, \mathbf{a}^2) \backslash Q(\mathbf{a}^1, \mathbf{a}^2)$. As in the proof of Proposition 12, let $\mathbf{p}_{up} = (\mathbf{x}^1, \mathbf{x}^2, y_{up})$ and $\mathbf{p}_{down} = (\mathbf{x}^1, \mathbf{x}^2, y_{down})$, where $y_{up} = \sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 \min\{x_i^1, x_j^2\}$ and $y_{down} = \sum_{i\in K_1} \sum_{j\in K_2} a_i^1 a_j^2 (x_i^1 + x_j^2 -$

$1)^+$. Note that, $\mathbf{p}_{up}, \mathbf{p}_{down} \in Q(\mathbf{a}^1, \mathbf{a}^2)$. We next show that $y_{up} \geq y \geq y_{down}$, and therefore $\mathbf{p} \in \mathrm{conv}\{\mathbf{p}_{up}, \mathbf{p}_{down}\} \subseteq Q(\mathbf{a}^1, \mathbf{a}^2)$.

Let $H_1 = \{(i,j) \in K_1 \times K_2 \ : \ x_i^1 > x_j^2\}$ and $H_2 = \{(i,j) \in K_1 \times K_2 \ : \ x_j^2 + x_i^1 > 1\}$. Applying (10) with $H = H_1$ gives

$$\begin{aligned} y &\leq \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^1 + \sum_{(i,j) \in H_1} a_i^1 a_j^2 (x_j^2 - x_i^1) \\ &= \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^1 + \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 \min\{0, x_j^2 - x_i^1\} \\ &= \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 (x_i^1 + \min\{0, x_j^2 - x_i^1\}) \\ &= \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 \min\{x_i^1, x_j^2\} \quad = \quad y_{up} \end{aligned}$$

Applying (9) with $H = H_2$ gives

$$\begin{aligned} y &\geq \sum_{(i,j) \in H_2} a_i^1 a_j^2 (x_i^1 + x_j^2 - 1) \\ &= \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 \max\{0, x_i^1 + x_j^2 - 1\} \quad = \quad y_{down} \end{aligned}$$

□

As a consequence, we have our main theorem:

Theorem 15 $P(\mathbf{a}^1, \mathbf{a}^2) = R(\mathbf{a}^1, \mathbf{a}^2) = Q(\mathbf{a}^1, \mathbf{a}^2)$.

Although our main goal was to establish the inequality description (7-10) of $P(\mathbf{a}^1, \mathbf{a}^2)$, we have established that from a mathematical point of view, the extended formulation (11-15) has the same power as a description. Which formulation will be preferable in an application will likely depend on implementation details.

5. Separation

We can efficiently include all facet describing inequalities of $P(\mathbf{a}^1, \mathbf{a}^2)$ implicitly in a linear programming formulation, provided that we can separate on them in polynomial time (see [GLS84, GLS81, GLS93]). That is, provided we have a polynomial-time algorithm that determines whether a given point is in $P(\mathbf{a}^1, \mathbf{a}^2)$ and provides a violated facet describing inequality if the point is not in $P(\mathbf{a}^1, \mathbf{a}^2)$.

Separation for the simple lower and upper bound inequalities (7–8) is easily handled by enumeration. For a point $(\mathbf{x}^1, \mathbf{x}^2, y)$ satisfying (7–8), separation for the lower and upper bound inequalities (9–10) is also rather simple. For the lower bound inequalities (9), we simply let

$$H_0 = \left\{(i,j) \in K_1 \times K_2 \ : \ x_i^1 + x_j^2 > 1\right\} \ ,$$

and then we just check whether $(\mathbf{x}^1, \mathbf{x}^2, y)$ violates the lower bound inequality (9) for the choice of $H = H_0$. Similarly, for the upper bound inequalities (10), we let

$$H_0 = \left\{(i,j) \in K_l \times K_{\bar{l}} \ : \ x_i^l - x_j^{\bar{l}} < 0\right\} \ ,$$

and then we just check whether $(\mathbf{x}^1, \mathbf{x}^2, y)$ violates the upper bound inequality (10) for the choice of $H = H_0$. Note that for any $H \subseteq K_1 \times K_2$,

$$\sum_{(i,j)\in H} a_i^1 a_j^2 (x_i^1 + x_j^2 - 1) \le \sum_{(i,j)\in H_0} a_i^1 a_j^2 (x_i^1 + x_j^2 - 1).$$

Therefore, $(\mathbf{x}^1, \mathbf{x}^2, y)$ satisfies the lower bounds (9) for all sets $H \subseteq K_1 \times K_2$ if and only if it satisfies (9) for $H = H_0$.

Using Propositions 9 and 10, we can see how this separation method yields *facet describing* violated inequalities. We develop a permutation of the variables

$$\{x_i^1 \ : \ i \in K_1\} \cup \{x_j^2 \ : \ j \in K_2\} \ ,$$

according to their values. Let $\delta_1 > \delta_2 > \cdots > \delta_p$ denote the distinct values of the $\{x_i^1 \ : \ i \in K_1\}$. For convenience, let $\delta_0 = 1$ and $\delta_{p+1} = 0$. We define the partition via $2p+1$ blocks, some of which may be empty. For, $t = 1, 2, \ldots, p$, block $2t$ consists of

$$\{x_i^1 \ : \ i \in K_1, \ x_i^1 = \delta_t\} \ .$$

For, $t = 1, 2, \ldots, p$, block $2t+1$ consists of

$$\{x_j^2 \ : \ j \in K_2, \ 1 - \delta_t < x_j^2 \le 1 - \delta_{t+1}\} \ ,$$

and block 1 consists of

$$\{x_j^2 \ : \ j \in K_2, \ 1 - \delta_0 = 0 \le x_j^2 \le 1 - \delta_{t+1}\}.$$

This permutation of the variables determines a subset H of $K_1 \times K_2$ as described in Section 3. This choice of H yields a facet-describing lower-bound inequality (9).

Similarly, for the upper bound inequalities (10), we let $\delta_1 < \delta_2 < \cdots < \delta_p$ denote the distinct values of the $\{x_i^l \ : \ i \in K_l\}$. As before, let $\delta_0 = 0$ and $\delta_{p+1} = 1$, and we define a partition via $2p+1$ blocks, some of which may be empty. For, $t = 1, 2, \ldots, p$, block $2t$ consists of

$$\{x_i^l \ : \ i \in K_l, \ x_i^l = \delta_t\} \ .$$

For, $t = 0, 1, 2, \ldots, p$, block $2t+1$ consists of

$$\{x_j^{\bar{l}} \ : \ j \in K_{\bar{l}}, \ \delta_t < x_j^{\bar{l}} \le \delta_{t+1}\} \ .$$

This permutation of the variables determines a subset H of $K_1 \times K_2$ as described in Section 3. This choice of H yields a facet describing upper bound inequality (10).

6. Ideal Points

For many combinatorial polytopes, it is natural to investigate adjacency of extreme points via edges (i.e., 1-dimensional faces). One motivation is that this notion of adjacency may prove useful in some local-search heuristics. In this section, we investigate a different notion of adjacency for extreme points — one that seems more natural for $P(\mathbf{a}^1, \mathbf{a}^2)$.

The point $(\mathbf{x}^1, \mathbf{x}^2, y) \in P(\mathbf{a}^1, \mathbf{a}^2)$ is *ideal* if it satisfies (5). Clearly, the extreme points of $P(\mathbf{a}^1, \mathbf{a}^2)$ are ideal. Also, $P(\mathbf{a}^1, \mathbf{a}^2)$ contains points that are not ideal. For example,

$$(x_1^1, x_1^2, y) = (\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}) = \frac{1}{2}(0,0,0) + \frac{1}{2}(1,1,1)$$

is in $P((1),(1))$, but it is not ideal because

$$\sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^1 x_j^2 = 1 \cdot 1 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \neq \frac{1}{2} \; .$$

Proposition 16. *Every pair of distinct extreme points of $P(\mathbf{a}^1, \mathbf{a}^2)$ is connected by a curve of ideal points of $P(\mathbf{a}^1, \mathbf{a}^2)$. Moreover, such a curve can be taken to be either a line segment or two line segments joined at another extreme point of $P(\mathbf{a}^1, \mathbf{a}^2)$.*

Proof. Let $(\mathbf{x}^{11}, \mathbf{x}^{12}, y^{11})$ and $(\mathbf{x}^{21}, \mathbf{x}^{22}, y^{22})$ be a pair of distinct extreme points of $P(\mathbf{a}^1, \mathbf{a}^2)$. If $\mathbf{x}^{11} = \mathbf{x}^{21}$ then we consider the curve obtained by letting

$$(\mathbf{z}^1, \mathbf{z}^2, y) = \lambda(\mathbf{x}^{11}, \mathbf{x}^{12}, y^{11}) + (1-\lambda)(\mathbf{x}^{21}, \mathbf{x}^{22}, y^{22}),$$

as λ ranges between 0 and 1. For $\lambda = 1$ we have $(\mathbf{z}^1, \mathbf{z}^2, y) = (\mathbf{x}^{11}, \mathbf{x}^{12}, y^{11})$, and for $\lambda = 0$ we have $(\mathbf{z}^1, \mathbf{z}^2, y) = (\mathbf{x}^{21}, \mathbf{x}^{22}, y^{22})$. Clearly, the curve is a line segment entirely contained in the convex polytope $P(\mathbf{a}^1, \mathbf{a}^2)$, because we have defined each point on the curve as a convex combination of the pair of extreme points of P. So it remains to demonstrate that each point on the curve is ideal:

$$\begin{aligned}
&\sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 z_i^1 z_j^2 \\
&\quad = \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 \left(\lambda x_i^{11} + (1-\lambda) x_i^{21}\right) \cdot \left(\lambda x_j^{12} + (1-\lambda) x_j^{22}\right) \\
&\quad = \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^{11} \left(\lambda x_j^{12} + (1-\lambda) x_j^{22}\right) \\
&\quad = \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 \lambda x_i^{11} x_j^{12} + (1-\lambda) x_i^{21} x_j^{22} \\
&\quad = \lambda y^{11} + (1-\lambda) y^{22} \\
&\quad = y \; .
\end{aligned}$$

Therefore the points on the curve are ideal.

Similarly, if $\mathbf{x}^{12} = \mathbf{x}^{22}$, we use the same line segment above to connect the points.

Suppose now that $\mathbf{x}^{11} \neq \mathbf{x}^{21}$ and $\mathbf{x}^{12} \neq \mathbf{x}^{22}$. We define a third point (x^{11}, x^{22}, y^{12}), where

$$y^{12} = \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^{11} x_j^{22} \ .$$

Then we connect this third point, in the manner above, to each of the points of the original pair. □

The curve of ideal points given to us by Proposition 16 is entirely contained in a 2-dimensional polytope, but it is not smooth in general. By allowing the curve to be contained in a 3-dimensional polytope, we can construct a smooth curve of ideal points connecting each pair of extreme points of $P(\mathbf{a}^1, \mathbf{a}^2)$.

Proposition 17. *Every pair of extreme points of $P(\mathbf{a}^1, \mathbf{a}^2)$ is connected by a smooth curve of ideal points of $P(\mathbf{a}^1, \mathbf{a}^2)$.*

Proof. Let $(\mathbf{x}^{11}, \mathbf{x}^{12}, y^{11})$ and $(\mathbf{x}^{21}, \mathbf{x}^{22}, y^{22})$ be a pair of distinct extreme points of $P(\mathbf{a}^1, \mathbf{a}^2)$. Our goal is to connect these points with a smooth curve of ideal points of $P(\mathbf{a}^1, \mathbf{a}^2)$. Toward this end, we consider two other points of $P(\mathbf{a}^1, \mathbf{a}^2)$, $(\mathbf{x}^{11}, \mathbf{x}^{22}, y^{12})$ and $(\mathbf{x}^{21}, \mathbf{x}^{12}, y^{21})$, which we obtain from the original pair by letting

$$y^{12} = \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^{11} x_j^{22}$$

and

$$y^{21} = \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 x_i^{21} x_j^{12} \ .$$

Now, we consider the curve obtained by letting

$$\begin{aligned}(\mathbf{z}^1, \mathbf{z}^2, y) = \lambda^2(\mathbf{x}^{11}, \mathbf{x}^{12}, y^{11}) + (1-\lambda)^2(\mathbf{x}^{21}, \mathbf{x}^{22}, y^{22}) \\ +\lambda(1-\lambda)(\mathbf{x}^{11}, \mathbf{x}^{22}, y^{12}) + \lambda(1-\lambda)(\mathbf{x}^{21}, \mathbf{x}^{12}, y^{21}) \ ,\end{aligned}$$

as λ ranges between 0 and 1. For $\lambda = 1$ we have $(\mathbf{z}^1, \mathbf{z}^2, y) = (\mathbf{x}^{11}, \mathbf{x}^{12}, y^{11})$, and for $\lambda = 0$ we have $(\mathbf{z}^1, \mathbf{z}^2, y) = (\mathbf{x}^{21}, \mathbf{x}^{22}, y^{22})$. Clearly, the curve is entirely contained in the convex polytope $P(\mathbf{a}^1, \mathbf{a}^2)$, because we have defined each point on the curve as a convex combination of extreme points of P. So it remains to demonstrate that each point on the curve is ideal:

$$\begin{aligned}&\sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 z_i^1 z_j^2 \\ &\quad = \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 \left(\lambda^2 x_i^{11} + (1-\lambda)^2 x_i^{21} + \lambda(1-\lambda) x_i^{11} + \lambda(1-\lambda) x_i^{21}\right) \\ &\qquad \cdot \left(\lambda^2 x_j^{12} + (1-\lambda)^2 x_j^{22} + \lambda(1-\lambda) x_j^{22} + \lambda(1-\lambda) x_j^{12}\right) \\ &\quad = \sum_{i \in K_1} \sum_{j \in K_2} a_i^1 a_j^2 \left(\lambda x_i^{11} + (1-\lambda) x_i^{21}\right) \cdot \left(\lambda x_j^{12} + (1-\lambda) x_j^{22}\right)\end{aligned}$$

$$
\begin{aligned}
&= \lambda^2 \sum_{i\in K_1}\sum_{j\in K_2} a_i^1 a_j^2 x_i^{11} x_j^{12} + (1-\lambda)^2 \sum_{i\in K_1}\sum_{j\in K_2} a_i^1 a_j^2 x_i^{21} x_j^{22} \\
&\quad +\lambda(1-\lambda) \sum_{i\in K_1}\sum_{j\in K_2} a_i^1 a_j^2 x_i^{11} x_j^{22} + \lambda(1-\lambda) \sum_{i\in K_1}\sum_{j\in K_2} a_i^1 a_j^2 x_i^{21} x_j^{22} \\
&= \lambda^2 y^{11} + (1-\lambda)^2 y^{22} + \lambda(1-\lambda) y^{12} + \lambda(1-\lambda) y^{21} \\
&= y \ .
\end{aligned}
$$

Therefore the points on the curve are ideal. □

7. A Generalization

Although we do not have an application for this, our results generalize. Let $\mathbf{A}$ be a $k_1 \times k_2$ matrix with positive components, and let $P(\mathbf{A})$ be the convex hull of solutions of

$$
\begin{aligned}
y &= \sum_{i\in K_1}\sum_{j\in K_2} a_{ij} x_i^1 x_j^2 \ ; \\
x_i^l &\in \{0,1\}, \text{ for } i \in K_l,\ l = 1,2 \ .
\end{aligned}
$$

The reader can easily check that everything that we have done applies to $P(\mathbf{A})$ by making the substitution of $a_i^1 a_j^2$ by a_{ij} throughout.

Acknowledgements. The authors are grateful to Komei Fukuda for making his program `cdd` available. Evidence collected with the use of `cdd` led us to conjecture some of our results. The authors thank the anonymous referees for their careful reading of the paper and for their suggestions which greatly improved the paper. The research of Jon Lee was supported in part by the Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong. The work of Janny Leung was partially supported by the Hong Kong Research Grants Council.

References

[BB98] T. Badics and E. Boros. Minimization of half-products. *Math. Oper. Res.*, 23(3):649–660, 1998.

[DL97] Michel Deza and Monique Laurent. *Geometry of cuts and metrics*. Springer-Verlag, Berlin, 1997.

[DS90] Caterina De Simone. The cut polytope and the Boolean quadric polytope. *Discrete Math.*, 79(1):71–75, 1989/90.

[GLS81] Martin Grötschel, László Lovász, and Alexander Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.

[GLS84] Martin Grötschel, László Lovász, and Alexander Schrijver. Corrigendum to our paper: "The ellipsoid method and its consequences in combinatorial optimization". *Combinatorica*, 4(4):291–295, 1984.

[GLS93] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*. Springer-Verlag, Berlin, second edition, 1993.

[HK56] Alan J. Hoffman and Joseph B. Kruskal. Integral boundary points of convex polyhedra. In *Linear inequalities and related systems*, pages 223–246. Princeton University Press, Princeton, N. J., 1956. Annals of Mathematics Studies, no. 38.

[NW88] George L. Nemhauser and Laurence A. Wolsey. *Integer and combinatorial optimization*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., New York, 1988. A Wiley-Interscience Publication.

[Pad89] Manfred W. Padberg. The Boolean quadric polytope: Some characteristics, facets and relatives. *Math. Programming, Ser. B*, 45(1):139–172, 1989.

[Pit91] Itamar Pitowsky. Correlation polytopes: their geometry and complexity. *Math. Programming*, 50(3, (Ser. A)):395–414, 1991.