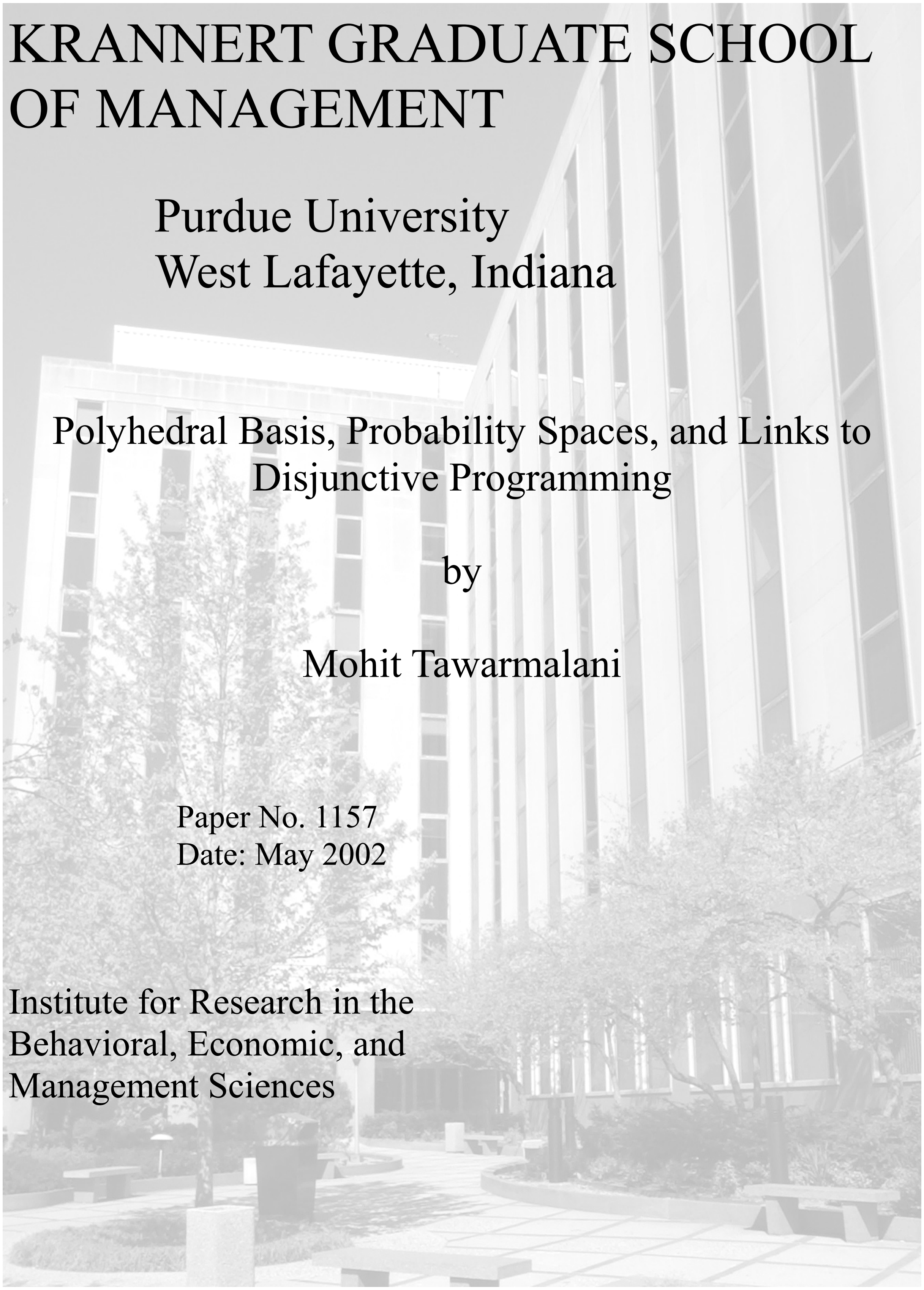

KRANNERT GRADUATE SCHOOL OF MANAGEMENT

Purdue University
West Lafayette, Indiana

Polyhedral Basis, Probability Spaces, and Links to Disjunctive Programming

by

Mohit Tawarmalani

Paper No. 1157
Date: May 2002

Institute for Research in the
Behavioral, Economic, and
Management Sciences

Mohit Tawarmalani

Polyhedral Basis, Finite Probability Spaces and Links to Disjunctive Programming

May, 2002

Abstract. Motivated by our earlier study of convex extensions [25], we define a polyhedral basis over a convex set. We explore properties of polyhedral functions that lead us into studying their relations to convexification and disjunctive programming. In particular, we show that a polyhedral basis is easy to identify for cartesian products of simplices. A special case, that of an n-dimensional hypercube, is of particular interest in this study. In this case, we show that convex and concave envelopes of multilinear sets are derived as a consequence of disjunctive programming and Reformulation Linearization Techniques. We answer in negative an open question whether there exist polynomial functions that will provide convexification processes for general integer programs just as multilinear functions are used to convexify 0-1 programs in the reformulation linearization technique and the lift and project algorithm.

The polyhedral functions also correspond to nonlinear rounding ideas and we will explore these in the article. This interpretation allows us to study probabilistic mathematical programs and explore their relation to multilinear programs. More concretely, we demonstrate that a probabilistic mathematical program is equivalent to the lagrangian relaxation of a multilinear program. Consequently, we study approximation schemes and demonstrate that rounding schemes often hint at polyhedral basis functions. Some negative results about Lagrangian relaxation are presented.

This is a preliminary set of proofs and provides many directions for further work in convexification techniques. Some of the results that follow easily are stated without proofs.

Mohit Tawarmalani: Krannert School of Management, Purdue University.

1. Introduction

This working paper presents the proofs of many results of interest in convexification and their relation to disjunctive programming [4, 3, 5, 23], reformulation linearization technique [17, 20, 18, 16, 19], semidefinite relaxations [9]. The work presented herein is related to Lasserre's hierarchy in that the probability measures are associated with variables, but is distinct in the fact that we consider subsets of the face lattice of the hypercube (and there is no face lattice for semi-algebraic sets considered in Lasserre's work) and also that we go beyond the general definition of probabilistic measures to provide specific types of probabilistic measures that allow us to reconstruct whatever can be said about all the probabilistic measures. For example independent events allow us through convexification operation to determine if there exists a probabilistic measure satisfying certain prespecified probabilities.

2. Polyhedral Function Basis

We often deal with finite collection $\mathcal{C} = \{C_i \mid i \in I\}$ of sets. With a slight abuse of notation, but with the view of simplifying presentation, we denote $\bigcup_{i\in I} C_i$ by $\mathcal{C}$ itself. For example, $\text{conv}(\mathcal{C})$ denotes $\text{conv}(\cup_{i\in I} C_i)$ and $x \in \mathcal{C}$ is equivalent to $x \in \cup_{i\in I} C_i$.

Definition 1. *Given a finite collection of disjoint convex sets* $\mathcal{C} = \{C_i \mid i \in I\}$, *and a set* X *such that* $\bigcup_{i\in I} C_i \subseteq X \subseteq \text{conv}(\mathcal{C})$, *the polyhedral basis is a collection* *of functions* $\mathcal{F} = \{f_i \mid i \in I\}$ *such that for any* $x \in X$ *and the collection* $\mathcal{F}$*, the following hold:*

- $f_i(x) \geq 0$
- $\sum_{f_i \in \mathcal{F}} f_i(x) = 1$
- *for every* $f_i(x) \neq 0$*, there exists* $y_i \in C_i$ *such that* $x = \sum_{i \in I} f_i(x) y_i$

and for every $x \in C_j$*,* $f_j(x) = 1$*.*

If the set X is not explicitly specified, it is to be assumed as $\text{conv}(C)$.

The motivation behind our definition of polyhedral function basis should be apparent. For every point in $\text{conv}(\mathcal{C})$ but not in $\mathcal{C}$, the polyhedral basis provides a convex representation of x in terms of points in $\mathcal{C}$, or in other words, a certificate of its inclusion in $\text{conv}(\mathcal{C})$. Indeed, for the definition to be meaningful, it is necessary that $\mathcal{C}$ be a collection of disjoint sets, or else, the polyhedral basis can not possibly exist (consider a point $x \in C_i \cap C_j$). However, it may be noted that the disjointness does not impose a restriction since any collection of convex sets can always be written as a collection of disjoint convex sets using the meet irreducibles of the canonical distributive lattice of the sets.

Example 1. Consider the canonical simplex Δ_3 in $\mathbb{R}^3$ given by:

$$\{(x_1, x_2, x_3) \mid x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_1 + x_2 + x_3 \leq 1\}.$$

Let $\mathcal{C}$ be the collection of extreme points of Δ_3. Then, it is easy to verify that $x_1, x_2, 1 - x_1 - x_2$ forms a polyhedral basis for $\mathcal{C}$. It is in fact clear in this example that the above functions are the only functions that satisfy the requirements of a polyhedral basis since the convex multipliers are determined uniquely (see for example exercise 2.28(b) in [15]). In general the unique polyhedral basis for Δ_n is given by $x_1, \ldots, x_n, 1 - x_1 - \cdots - x_n$.

The next result motivates our definition of polyhedral basis.

Theorem 1. *Consider a polyhedral basis $\mathcal{F}$ of $\mathcal{C}$ over X. The generating set of the epigraph (hypograph) of the convex (concave) envelope of any $f_i \in \mathcal{F}$ over X is a subset of $\mathcal{C}$.*

Proof. Any $x \notin \mathcal{C}$ can be expressed as: $x = \sum_{i \in I} f_i(x) y_i$. By definition, $f_i(x)$ equals 1 for every $x \in C_i$ and 0 for every $x \in C_j$ where $j \neq i$. Further, $f_i(x) = \sum_{i \in I} f_i(x) f_i(y_i) = f_i(x)$. It follows then as a simple corollary of Theorem 7 in [25] that x does not belong to the generating set.

In particular, if $\mathcal{C}$ is a collection of a finite number of points, then Theorem 1 implies that in order to construct the convex hull of any function in $\mathcal{F}$, we need to consider only the points in $\mathcal{C}$.

From Definition 1, and the definition of convex hull, it is clear that there always exists a polyhedral basis. In fact, Theorem 1 can be generalized significantly. In order to do so, we first define a lifting of a point/subset of $\mathbb{R}^n$.

Definition 2. *Consider a polyhedral basis $\mathcal{F}$ of $\mathcal{C}$ as in Definition 1. The lifting of a point, x is a mapping, $\phi(x) : \mathbb{R}^n \mapsto \mathbb{R}^{n+|I|}$ such that $\phi(x) = (x, f_1(x), \ldots, f_m(x))$. The lifting of a subset, S, of $\mathbb{R}^n$ is denoted as $\phi(S)$ and represents $\bigcup_{x \in S} \phi(x)$. The lifting of a collection of sets $\mathcal{S} = \{S_i \mid i \in I\}$ is defined as $\phi(\mathcal{S}) = \phi\left(\cup_{i \in I} S_i\right)$.*

Theorem 2. *Consider a polyhedral basis $\mathcal{F}$ of $\mathcal{C}$ over X and the corresponding lifting ϕ. Then,* $\operatorname{conv}(\phi(\mathcal{C})) = \operatorname{conv}(\phi(X))$.

Proof. ($\subseteq$) Since $\bigcup_{i\in I} C_i \subseteq X$, $\phi(\mathcal{C}) \subseteq \phi(X)$. It is therefore apparent that that

$$\text{conv}(\phi(\mathcal{C})) \subseteq \text{conv}(\phi(X)).$$

($\supseteq$) Consider the representation of a point $x \in X$ as $x = \sum_{i\in I} f_i(x)y_i$, where each $y_i \in C_i$ as guaranteed by the definition of polyhedral basis. Restrict attention to the y_i used in the above representation. It can be easily verified that $\phi(x) = \sum_{i\in I} f_i(x)\phi(y_i)$ since $\phi(y_i) = (y_i, e_i)$ where e_i is the i^{th} unit vector of $R^{|I|}$. Therefore, $\phi(x) \in \text{conv}(\phi(\mathcal{C}))$. In other words, $\phi(X) \subseteq \text{conv}(\phi(\mathcal{C}))$. Taking the convex hull on both sides, we have $\text{conv}(\phi(X)) \subseteq \text{conv}(\phi(\mathcal{C}))$.

Lemma 1. *Consider a set X and a linear transformation A. Then $A\,\text{conv}(X) = \text{conv}(AX)$.*

Proof. ($\supseteq$) Follows easily from $AX \subseteq A\,\text{conv}(X)$ and that $A\,\text{conv}(X)$ is convex (see Theorem 3.4 [14]).

($\subseteq$) For an arbitrary $x \in A\,\text{conv}(X)$, there exists a y in $\text{conv}(X)$ such that $x = Ay$. In particular, if $X \subseteq \mathbb{R}^n$, there exist $n+1$ points in X, say $y_1, \ldots, y_{n+1}$ such that $y = \sum_{i=1}^{n+1} \lambda_i y_i$ (see Theorem 2.29 [15]). Then, $x = Ay = \sum_{i=1}^{n+1} \lambda_i A y_i \in \text{conv}(AX)$.

The following generalization of Theorem 2 and Theorem 1 follows easily.

Theorem 3. *Consider a polyhedral basis $\mathcal{F}$ of $\mathcal{C}$ (see definition 1), the corresponding lifting ϕ (see definition 2) and a linear transformation A with range $\mathcal{C}$. Then:*

$$\text{conv}(A\phi(\mathcal{C})) = \text{conv}(A\phi(X)).$$

Proof. It follows from Theorem 2 and Lemma 1 that:

$$\text{conv}(A\phi(\mathcal{C})) = A\,\text{conv}(\phi(\mathcal{C})) = A\,\text{conv}(\phi(X)) = \text{conv}(A\phi(X)).$$

We shall return to Theorem 3 occasionally during this paper as it embodies the motivation for our work on polyhedral basis functions.

Corollary 1. *Consider a polyhedral basis $\mathcal{F}$ of $\mathcal{C}$, the corresponding lifting ϕ and the function $g(x) = \sum_{i\in I} a_i f_i(x)$, where a_i, $i \in I$ are real constants. The generating set of the epigraph (hypograph) of the convex (concave) envelope of $g(x)$ is a subset of $\mathcal{C}$.*

Proof. Follows as a direct application of Theorem 3 by considering the linear transformation, A, that maps $\phi(x) = (x, f_1(x), \dots, f_{|I|}(x))$ to $(x, \sum_{i\in I} a_i f_i(x))$.

It follows easily from Definition 1 that the functions in a polyhedral basis are linearly independent. In particular, assume that $f_i(x)$ can be expressed as a linear combination of some functions in $\mathcal{F}$ indexed by a subset of $I\backslash\{i\}$. Then,

$$f_i(x) = \sum_{\substack{j\in I\\ j\neq i}} \lambda_j f_j(x). \tag{1}$$

We arrive at a contradiction by evaluating (1) at any point x in C_i. Corollary 1 along with the independence of the functions in $\mathcal{F}$ motivates our definition of a vector space of polyhedral functions.

Definition 3. *Consider a polyhedral basis $\mathcal{F}$ of $\mathcal{C}$. The vector space of polyhedral functions $\mathcal{V}(\mathcal{F})$ is defined as linear combinations of the functions in the polyhedral function basis. $\mathcal{F}$ forms an orthogonal basis of $\mathcal{V}(\mathcal{F})$.*

Definition 4. *Consider two functions $g(x)$ and $h(x)$ that belong to $\mathcal{V}(\mathcal{F})$. We define their polyhedral product, $g \diamond h$ as the function corresponding to the vector formed by taking the termwise product of their vector representations.*

The motivation for the polyhedral product arises from the fact that for any two polyhedral basis functions $f_i(x)f_j(x) = 0$ for $i \neq j$ and $f_i^2(x) = f_i(x)$ everywhere on $\mathcal{C}$. It is easy to verify that $(g \diamond h)(x) = g(x)h(x)$ for every x in $\mathcal{C}$. Further, from Corollary 1, the generating set of the epigraph and hypograph of $g \diamond h$ is a subset of $\mathcal{C}$.

When can polynomial functions be used to describe polyhedral basis functions? We try to answer this question partially in the following few results:

Theorem 4. *Consider a collection of convex sets $\mathcal{C}$ indexed by the set I and its polyhedral basis $\mathcal{F}$. $\mathcal{F}$ is not composed of polynomial functions if there exist i and j in I such that $\operatorname{aff}(C_i) \cap \operatorname{aff}(C_j) \neq \emptyset$.*

Proof. Either C_i is a single point or contains an infinite number of points. In either case, it follows from the fundamental theorem of algebra that if $f_i(x)$ is a polynomial then $f_i(x) = 1$ over the affine hull of C_i. Similarly, $f_i(x) = 0$ for any point in C_j, $j \neq i$. Therefore, $\operatorname{aff}(C_j) \cap \operatorname{aff}(C_i) = \emptyset$.

Theorem 5. *Consider a collection of convex sets $\mathcal{C}$ indexed by the set I and let its polyhedral basis be $\mathcal{F}$. If there exists a set $C_i \in \mathcal{C}$ and a point $\bar{x} \in C_i$ such that there exists a neighborhood of $\bar{x}$ where the points y_j in the representation $x = \sum_{j \in I} f_j(x) y_j$ can be expressed as $y_j = g_j(x)$ such that each component of g_j is continuous and differentiable and $\bar{x}$ can be expressed as a convex combination* *of the points $g_{i_1},\ldots,g_{i_m}$ where $\{i_1,\ldots,i_m\}\subseteq I\backslash\{i\}$, then $\mathcal{F}$ is not composed of polynomial functions.*

Proof. Consider the representation $x=\sum_j f_j(x)g_j(x)$ in the neighborhood of x. Then, taking the derivative on both sides with respect to the l^{th} coordinate we get:

$$e_l=\sum_j \nabla_l f_j(x)g_j(x)+\sum_j f_j(x)(\nabla_l g_j(x))^t, \tag{2}$$

where ∇_l denotes the differential operator with respect to the l^{th} co-ordinate. Let $P\subseteq I\backslash\{i\}$ be the set of points $g_p(x)$ which have non-zero multipliers in the representation of $\bar{x}$. We denote the convex hull of $g_p(x)$, $p\in P$ by C_P. Clearly, $\bar{x}\in\operatorname{ri}(C_P)$ (see Theorem 6.9 in [14]), $C_P\subseteq\operatorname{conv}(\mathcal{C})$, and each function in $\mathcal{F}$ is bounded between 0 and 1 over $\operatorname{conv}(\mathcal{C})$. Therefore, for any $j\in I$, $\nabla f_j(\bar{x})$ is orthogonal to $\operatorname{aff}(C_P)$. Consider any $d\in\operatorname{aff}(C_P)$. Then:

$$\begin{aligned}
d&=\sum_l d_l\sum_j \nabla_l f_j(\bar{x})g_j(\bar{x})+\sum_l d_l\sum_j f_j(\bar{x})(\nabla_l g_j i)^t(\bar{x})\\
&=\sum_j g_j(\bar{x})\sum_l d_l\nabla_l f_j(\bar{x})+\sum_j f_j(\bar{x})\sum_l d_l(\nabla_l g_j(\bar{x}))^t\\
&=\sum_j f_j(\bar{x})\sum_l d_l(\nabla_l g_j(\bar{x}))^t\\
&=\sum_j f_j(\bar{x})(\nabla g_j(x))^t d\\
&=(\nabla g_i(\bar{x}))^t d
\end{aligned} \tag{3}$$

where d_l is the l^{th} co-ordinate of d. (3) shows that the derivative of $g_i(\bar{x}+\lambda d)$ with respect to λ is d. Restricting attention to some $\bar{p}\in P$, and choosing $d=(g_{\bar{p}}(\bar{x})-\bar{x})$, the directional derivative is then $g_{\bar{p}}(\bar{x})-\bar{x}$. This implies that $\bar{x}+g_{\bar{p}}(\bar{x})-\bar{x}\in\operatorname{aff}(C_P)$. Theorem 4, however asserts that $g_{\bar{p}}(\bar{x})\notin\operatorname{aff}(C_i)$ leading to a contradiction of the assumption that a polynomial polyhedral basis can be constructed.

The above result easily provides the following corollary.

Corollary 2. *Consider a collection of points* $P = \{x_1, \ldots, x_m\}$*. If* x_1 *can be expressed as a convex combination of points in* $P \backslash \{x_1\}$*, then there does not exist a polynomial polyhedral basis for* P*.*

Proof. Since there is only one point in each convex set, the functions $g_j(x)$ are constant. Therefore, the current result follows directly from Theorem 5.

Example 2. Consider a collection of three points in a one-dimensional space coordinatized by the values 0, 1 and 2. Even for this small example, there does not exist a polynomial polyhedral basis as shown in Corollary 2. In the following, we construct $f_0(x)$, $f_1(x)$ and $f_2(x)$ that can be easily verified to form a polyhedral basis for this small example:

$$f_0(x) = \frac{1}{2}(1 - x + |1 - x|)$$
$$f_1(x) = 1 - |1 - x|$$
$$f_2(x) = \frac{1}{2}(x - 1 + |1 - x|).$$

In fact, define if we define $h_{i^-}(x) = \frac{1}{2}(i - x + |i - x|)$, $h_i(x) = 1 - |i - x|$ and $h_{i^+}(x) = \frac{1}{2}(x - i + |i - x|)$, then a polyhedral basis for $0, 1, \ldots, n$ can be formed by the functions defined over $[0, n]$:

$$f_0(x) = h_{1^-}(x)$$
$$f_i(x) = h_i(x) + h_{i-1^-}(x) + h_{i+1^+}(x) \qquad i = 1, \ldots, n-1$$
$$f_n(x) = h_{n-1^+}(x).$$

Theorem 6. *Consider a collection of convex sets* $\mathcal{C}$ *indexed by the set* I *and let be* $\mathcal{F}$ *be a polyhedral basis for* $\mathcal{C}$*. Assume that* d *is a unit vector along a recession direction for* $\mathrm{ri}(\mathrm{conv}(\mathcal{C}))$*. Then, for every polynomial function* $f \in \mathcal{F}$ *and every point* $\bar{x} \in \mathrm{aff}(\mathrm{conv}(\mathcal{C}))$*,* $f(\bar{x} + \lambda d) = f(\bar{x})$ *for all* $\lambda \in \mathbb{R}$ *(*i.e, f *is constant along the direction* d*).*

Proof. $\mathrm{aff}(\mathrm{ri}(\mathrm{conv}(\mathcal{C}))) = \mathrm{aff}(\mathrm{conv}(\mathcal{C}))$ (Theorem 6.2, [14]). For any point $\hat{x} \in \mathrm{ri}(\mathrm{conv}(\mathcal{C}))$, $\hat{x} + \lambda d \in \mathrm{ri}(\mathrm{conv}(\mathcal{C}))$ for all $\lambda \geq 0$ and $f(x)$ is bounded between 0 and 1. It then follows easily that $f(\hat{x} + \lambda d) = f(\hat{x})$ for all λ. Now, consider a point $\bar{x} \in \mathrm{aff}(\mathrm{conv}(\mathcal{C}))\backslash\mathrm{ri}(\mathrm{conv}(\mathcal{C}))$ and a point $\hat{x} \in \mathrm{ri}(\mathrm{conv}(\mathcal{C}))$. From the definition of the relative interior, there exists a ϵ larger than zero such that $\hat{x} + \delta(\bar{x} - \hat{x}) \in \mathrm{ri}(\mathrm{conv}(\mathcal{C}))$ for every δ such that $0 \leq \delta < \epsilon$. We have shown that function $f(\hat{x} + \delta(\bar{x} - \hat{x}) + \lambda d) - f(\hat{x} + \delta(\bar{x} - \hat{x}))$ is zero and therefore from fundamental theorem of algebra $f(\hat{x} + \gamma(\bar{x} - \hat{x}) + \lambda d) - f(\hat{x} + \gamma(\bar{x} - \hat{x})) = 0$ for any $\gamma \in \mathbb{R}$. Choose $\gamma = 1$ to prove the current result.

Example 3. Consider the collection of three convex sets $A = (0, 0)$, $B = (0, 1)$ and $C = \{(1, a) | a \in \mathbb{R}\}$. Denote the basis function corresponding to the three sets by f_A, f_B and f_C respectively and assume f_A is polynomial. Then, f_A is independent of the second co-ordinate since $(0, 1)$ is a recession direction of $\mathrm{ri}(\mathrm{conv}(A \cup B \cup C))$. By noting that $f_A(0, 1) = 0$ and $f_A(0, 0) = 1$, we derive a contradiction. A similar argument shows that f_B is also not polynomial.

Definition 5. *Consider two arbitrary collections* $\mathcal{C}$ *and* $\mathcal{D}$ *of convex sets. Let* $C = \{C_1, \ldots, C_n\}$ *and* $\mathcal{D} = \{D_1, \ldots, D_m\}$*. We define* $\mathcal{C} + \mathcal{D}$ *as the collection of all Minkowski sums of the form* $C_i + D_j$ *where* $i \in \{1, \ldots, n\}$ *and* $j \in \{1, \ldots, m\}$*.*

Definition 6. *Consider two arbitrary collectons $\mathcal{F}$ and $\mathcal{G}$ of functions. Let $\mathcal{F} = \{f_1, \ldots, f_n\}$ and $\mathcal{G} = \{g_1, \ldots, g_m\}$ We define $\mathcal{F} \times \mathcal{G}$ as the collection of functions $f_i g_j$ where $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$.*

Lemma 2. *For any two sets S_1 and S_2, $\operatorname{conv}(S_1 + S_2) = \operatorname{conv}(S_1) + \operatorname{conv}(S_2)$.*

Proof. ($\subseteq$) Clearly, any $x \in \operatorname{conv}(S_1 + S_2)$ is expressible as a convex combination of a finite number of points in $S_1 + S_2$ (see Theorem 2.27 and 2.29 in [15]) or $x = \sum_j \lambda_j (y_j + z_j)$ where $y_j \in S_1$ and $z_j \in S_2$ for every j. Then, $x \in \sum_j \lambda_j y_j + \sum_j \lambda_j z_j \in \operatorname{conv}(S_1) + \operatorname{conv}(S_2)$.

($\supseteq$) If $x \in \operatorname{conv}(S_1) + \operatorname{conv}(S_2)$, then $x = y + z$ where $y = \sum_{j=1}^{l} \lambda_j y_j$ and $z = \sum_{k=1}^{m} \gamma_k z_k$. We can therefore express x as follows:

$$\begin{aligned} x &= \sum_{j=1}^{n} \lambda_j \sum_{k=1}^{m} \gamma_k y_j + \sum_{k=1}^{m} \gamma_k \sum_{j=1}^{n} \lambda_j z_k \\ &= \sum_{j=1}^{n} \sum_{k=1}^{m} \lambda_j \gamma_k (y_j + z_k) \end{aligned}$$

Lemma 3. *For sets $S_1 \subseteq \mathbb{R}^n$, $S_2 \subseteq \mathbb{R}^m$ and linear transformations $A_1 : \mathbb{R}^n \mapsto \mathbb{R}^p$ and $A_2 : \mathbb{R}^m \mapsto \mathbb{R}^p$, $\operatorname{conv}(A_1 S_1 + A_2 S_2) = A_1 \operatorname{conv}(S_1) + A_2 \operatorname{conv}(S_2)$.*

Proof. Follows from Lemma 1 and Lemma 2.

We denote by $A\mathcal{C}$ the collection $\{AC_1, \ldots, AC_n\}$ where $C = \{C_1, \ldots, C_n\}$.

Theorem 7. *Let $\mathcal{C}_1, \ldots, \mathcal{C}_n$ be n collections of convex sets and let $\mathcal{F}_1, \ldots, \mathcal{F}_n$ be their polyhedral basis respectively. and $A_1, \ldots, A_n$ be linear transformations. Assume $C_i \in \mathbb{R}^{m_i}$ and $A_i : \mathbb{R}^{m_i} \mapsto \mathbb{R}^m$. Then,*

$$\operatorname{conv}(A_1 \mathcal{C}_1 + \cdots + A_n \mathcal{C}_n) = A_1 \operatorname{conv}(\mathcal{C}_1) + \cdots + A_n \operatorname{conv}(\mathcal{C}_n). \tag{4}$$

There are (vector) functions $k_1, \ldots, k_n$ *such that for every point* x *in* $\operatorname{conv}(A_1\mathcal{C}_1 + \ldots + A_n\mathcal{C}_n)$*:*

1. $k_i(x) \in \operatorname{conv}(\mathcal{C}_i)$, *for* $i = 1, \ldots, n$
2. $x = \sum_{i=1}^n A_i k_i(x)$
3. *If* $x \in A_1\mathcal{C}_1 + \cdots + A_n\mathcal{C}_n$, *then* $k_i(x) \in \mathcal{C}_i$.

Let $\mathcal{G}_i$ *be a collection of functions defined over* $\operatorname{conv}(A_1\mathcal{C}_1 + \ldots + A_n\mathcal{C}_n)$ *by composing each function in* $\mathcal{F}_i$ *with* k_i*. If the sets in the collection* $A_1\mathcal{C}_1 + \ldots + A_n\mathcal{C}_n$ *are disjoint, then* $\mathcal{G}_1 \times \cdots \times \mathcal{G}_n$ *forms a polyhedral basis for* $A_1\mathcal{C}_1 + \ldots + A_n\mathcal{C}_n$.

Proof. Observe that

$$\bigcup_{C \in A_i\mathcal{C}_i + A_j\mathcal{C}_j} C = \bigcup_{C \in \mathcal{C}_i} A_i C + \bigcup_{C \in \mathcal{C}_j} A_j C.$$

Applying Lemma 3 (tail) recursively to derive (4).

We need to verify that the functions $k_1(x), \ldots, k_n(x)$ exist. Consider any point x in $\operatorname{conv}(A_1\mathcal{C}_1 + \cdots + A_n\mathcal{C}_n)$. By (4), $x \in A_1 \operatorname{conv}(\mathcal{C}_1) + \cdots + A_n \operatorname{conv}(\mathcal{C}_n)$. In other words, there exist $x_1, \ldots, x_n$ such that $x_i \in \operatorname{conv}(\mathcal{C}_i)$, $i = 1, \ldots, n$ and $x = A_1 x_1 + \cdots + A_n x_n$. Define $k_i(x) = x_i$. If $x \in A_1\mathcal{C}_1 + \cdots + A_n\mathcal{C}_n$, then by definition $k_i(x)$ can be defined such that $k_i(x) \in \mathcal{C}_i$. We have thus argued that the functions $k_i(x)$ as provided by the statement of the thoerem exist.

Now, we verify that $\mathcal{G}_1 \times \cdots \times \mathcal{G}_n$ is indeed a polyhedral basis for $\mathcal{C}_1 + \ldots + \mathcal{C}_n$. Let the cardinality of $\mathcal{C}_i$ be I_i. Property 1. Each function in $\mathcal{G}_1 \times \cdots \times \mathcal{G}_n$ is a product of non-negative functions and is therefore non-negative. Property 2. Since $\sum_{f \in \mathcal{F}_i} f(k_i(x)) = 1$, $\prod_{i=1}^n \sum_{f \in \mathcal{F}_i} f(k_i(x)) = 1$. Property 3. $k_i(x) =$

$\sum_{f_{ij}\in\mathcal{F}_i} f_{ij}(k_i(x))y_{ij}$ where $y_{ij} \in C_{ij} \in \mathcal{C}_i$. Therefore,

$$\begin{aligned}
x &= \sum_{j=1}^{n} \sum_{i_j \in I_j} f_{ji_j}(k_j(x)) A_j y_{ji_j} \\
&= \sum_{j=1}^{n} \sum_{i_j \in I_j} f_{ji_j}(k_j(x)) \left(\prod_{l \neq j} \sum_{i_l \in I_l} f_{li_l}(k_l(x)) \right) A_j y_{ji_l} \\
&= \sum_{j=1}^{n} \sum_{i_j \in I_j} f_{ji_j}(k_j(x)) \left(\sum_{i_1 \in I_1} \cdots \sum_{i_{j-1} \in I_{j-1}} \sum_{i_{j+1} \in I_{j+1}} \cdots \sum_{i_n \in I_n} \right. \\
&\qquad \left. \prod_{\substack{l=1 \\ l \neq j}}^{n} f_{li_l}(k_l(x)) \right) A_j y_{ji_l} \\
&= \sum_{i_1 \in I_1} \cdots \sum_{i_n \in I_n} \prod_{l=1}^{n} f_{li_l}(k_l(x_j)) \sum_{l=1}^{n} A_l y_{li_l}
\end{aligned}$$

Clearly, $\sum_{l=1}^{n} A_l y_{li_l} \in \sum_{l=1}^{n} C_{li_l}$ where $\mathcal{C}_i = \{C_{i1}, \ldots, C_{i|I_i|}\}$.

Now, we only need to show that for every $x \in A_1\mathcal{C}_1 + \cdots + A_n\mathcal{C}_n$, one of the polyhedral functions takes a value 1 and the rest take a value 0. By the disjointness of sets in $A_1\mathcal{C}_1 + \cdots + A_n\mathcal{C}_n$, if $x \in \sum_{l=1}^{n} C_{li_l}$, then $k_l(x) \in C_{li_l} \in \mathcal{C}_l$. Then, $\prod_{l=1}^{n} f_{li_l} = 1$ and the rest of the functions in $\mathcal{G}_1 \times \cdots \times \mathcal{G}_n$ take a value 0 at x.

Definition 7. *Consider two arbitrary collections $\mathcal{C}$ and $\mathcal{D}$ of convex sets. Let $C = \{C_1, \ldots, C_n\}$ and $\mathcal{D} = \{D_1, \ldots, D_m\}$. We define $\mathcal{C} \times \mathcal{D}$ as the collection of all sets of the form $C_i \times D_j$ where $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$.*

Corollary 3. *Let $\mathcal{C}_1, \ldots, \mathcal{C}_n$ be n collections of convex sets and let $\mathcal{F}_1, \ldots, \mathcal{F}_n$ be their polyhedral basis respectively. Assume $\mathcal{C}_i \subseteq \mathbb{R}^{m_i}$ Then,*

$$\operatorname{conv}(\mathcal{C}_1 \times \cdots \times \mathcal{C}_n) = \operatorname{conv}(\mathcal{C}_1) \times \cdots \times \operatorname{conv}(\mathcal{C}_n).$$

Let $k_i(x)$ be the projection of x into $\mathbb{R}^{m_i}$. Define $\mathcal{G}_i$ to be a collection of functions defined over $\operatorname{conv}(\mathcal{C}_1 \times \ldots \times \mathcal{C}_n)$ *by composing each function in $\mathcal{F}_i$ with k_i. $\mathcal{G}_1 \times \cdots \times \mathcal{G}_n$ forms a polyhedral basis for $\mathcal{C}_1 \times \ldots \times \mathcal{C}_n$.*

Proof. Let 0_i be the collection $\big\{\{0\}\big\}$ in $\mathbb{R}^{m_i}$ (*i.e,* it contains only the origin). Apply Theorem 7 to the collections $\mathcal{C}_l \times \prod_{\substack{j=1 \\ j\neq l}}^{n} 0_j$. It is easy to verify that the functions k_i satisfy the requirements of Theorem 7 and the disjointness of the sets in $\mathcal{C}_1 \times \ldots \times \mathcal{C}_n$ follows from the disjointness of sets in $\mathcal{C}_1, \ldots, \mathcal{C}_n$.

3. Multilinear Sets

In this section, we explore multilinear sets and show that the convex hull of the solutions of multilinear functions over a hypercube is polyhedral as a simple corollary of Theorem 3.

Indeed, this allows us to consider arbitrary subsets of the face lattice of the hypercube to develop relaxations for 0-1 programs. In fact these correspond to the different multilinear functions that can be generated as a result of Theorem 10 in [25]. Indeed any 2^n linearly independent columns, where each column corresponds to a multilinear function and the entries correspond to the function evaluated at the extreme points of the hypercube, provides alternate basis for the Reformulation Linearization Hierarchy. For a binary program, it is clear from the proof of Theorem 10 in [25] that there is a single multilinear function which if multiplied with all the constraints reduces all the constraints to redundant inequalities and the convex envelope of the function intersected with the hyperplane $f(x) = 0$ is the convex hull of the solutions to the IP. In fact, it can be argued that multiplying an inequality with the basis function of a face and then linearizing it is equivalent to defining polyhedral functions for the sub-faces and then posing the inequality on the face using the polyhedral functions to define the point on the face.

4. Finite Probability Spaces and Probabilistic Rounding

Consider a finite probability space with a universal set U an arbitrary collection of events $\mathcal{A} = \{A_1, \ldots, A_n\}$ where each $A_i \subseteq U$. The set of events generated by $\mathcal{A}$ is the set of events that are obtained by taking unions or intersections of the sets in $\mathcal{A}$ or their complements. The generated set of events will be denoted by $G(\mathcal{A})$.

Definition 8. *Consider a collection $\mathcal{A}$ of events. A probability assignment assigns probabilities to events in $G(\mathcal{A})$ obeying the axioms of probability.*

Definition 9. *A convex probabilistic inequality for $\mathcal{A}$ is an inequality that must be obeyed by any probability assignment to $\mathcal{A}$ and is convex in the probabilities of the events in $G(\mathcal{A})$.*

One of the simplest convex (in fact linear) probabilistic inequalities is $P(A_1 \cap A_2) \leq P(A_1)$. A more interesting example is $P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) \leq 1$ which follows easily from the inclusion-exclusion principle.

Definition 10. *We define a probabilistic mathematical program as a mathematical program that minimizes a convex function of probabilities of events in $G(\mathcal{A})$ over a set of inequalities relating the probability of the events.*

Example 4. We consider a simple (probably the most simple) probabilistic mathematical program to illustrate the concept:

$$
\text{(P)} \qquad \begin{aligned} \min \quad & P(A \cap B) \\ \text{s.t.} \quad & P(A) \geq 0.5 \\ & P(B) \geq 0.6. \end{aligned}
$$

Clearly, $P(A \cap B) \geq 0.1$ as is easily seen by the addition law of probabilities: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Setting $P(A) = 0.5$ and $P(B) = 0.6$ in a manner such that $P(A \cup B) = 1$, we get a valid probabilistic assignment obeying the constraints of $P(B)$.

The probabilistic mathematical program does not seem to have been used in practice. We imagine that it has a potential to serve as an important instrument in bounding probabilities as is often done in proof techniques based on the probabilistic method ([2]). From a more practical standpoint, probabilistic mathematical programs may be useful when the worst/best case scenario for the occurrence of certain event/events is desired possibly in the context of decision making under uncertainity. Unfortunately, probabilistic mathematical programs are not easy to solve and it will be clear as a by-product of our discussion that in general the probabilistic mathematical program is NP-Hard.

Interestingly, even though the probabilistic mathematical program does not require the events to be independent, we will show next that placing such a restriction provides many structural results about the unrestricted probabilistic mathematical program.

Consider an abitrary collection of events $\mathcal{A}$ indexed by the set I. An arbitrary event in $G(\mathcal{A})$ can be wrriten in disjunctive normal form as a union of intersections. In other words, any event $E \in G(\mathcal{A})$ can be written as:

$$E = \bigcup_{c=1}^{n} \bigcap_{j \in J_c} S_j$$

where each $J_c \subset I$ and S_j is either A_j or its complement. We associate with each event E, an expression multilinear in terms of the probabilities of the events A_i such that the expression provides the probability of E under the assumption that $A_1, \ldots, A_n$ are independent. The inclusion-exclusion formula provides such a description easily. Another way to derive the same formula is by expressing E as:

$$E = \bar{\bar{E}} = \overline{\bigcap_{c=1}^{n} \overline{\bigcap_{j \in J_c} S_j}}.$$

In the sequel, each term of the form $\prod_i P(A_i)$ will be referred to as a multilinear term.

Example 5. Consider an event $E = (A_1 \cap A_2) \cup A_3$. The multilinear formula we associate with $P(E)$ is obtained by expressing E as:

$$E = \overline{\overline{(A_1 \cap A_2)} \cap \overline{A}_3}.$$

Therefore, $P(E)$ assuming independence of A_1, A_2 and A_3 is:

$$\begin{aligned} P(E) &= 1 - (1 - P(A_1)P(A_2))(1 - P(A_3)) \\ &= P(A_1)P(A_2) + P(A_3) - P(A_1)P(A_2)P(A_3) \end{aligned}$$

which also follows easily from the addition law (a special case of inclusion-exclusion). The multilinear terms are $P(A_1)P(A_2)$, $P(A_3)$ and $P(A_1)P(A_2)P(A_3)$.

Following the idea of translating each of the event probabilities to multilinear expressions, we construct a mathematical program from the probabilistic mathematical program in the following manner:

Algorithm AssocMathProgram:

1. For each event E used in the probabilistic mathematical program introduce a variable P_E.
2. For each event $A_i \in \mathcal{A}$, introduce $0 \leq P_{A_i} \leq 1$.
3. For every event E introduced in Step 1 such that $E \notin \mathcal{A}$, introduce a constraint equating P_E to the associated multilinear expression.

Taking Example 4, we get the following mathematical program after simple preprocessing:

$$(P_M) \qquad \min\{P_A P_B \mid 0.5 \leq P_A \leq 1,\ 0.6 \leq P_B \leq 1\}$$

Clearly, the optimal solution to (P_M) is 0.3. The result is as expected since we are restricting to independent events A and B. So, in general the following is apparent.

Theorem 8. *Consider any probabilistic mathematical program P. The associated mathematical program derived using Algorithm AssocMathProgram provides an upper bound to P.*

The associated mathematical progam has the form that it includes inequalities relating multilinear functions of the probabilities. In fact, it is easy to show that every multilinear function can be expressed as a linear combination of probabilities of some events in $G(\mathcal{A})$ under the additional assumption that the events are independent. This ties in with the fact that multilinear functions form a polyhedral basis of the extreme points of a hypercube (see Definition 3, and the preceding discussion). As a result any mathematical program involving multilinear functions of variables is associated with a probabilistic mathematical program where the multilinear functions are replaced by the appropriate probabilities. In particular, a multilinear program is associated with a linear probabilistic mathematical program.

What makes the associated mathematical program interesting is that it can be easily relaxed to produce a reformulation of the probabilistic mathematical program. Before, we provide such a reformulation, we argue in Theorem 9 that the axioms of probability can be expressed in terms of convex hulls of certain multilinear equations over the unit hypercube.

The following result is hinted to in [9].

Theorem 9. *Consider a collection of events $\mathcal{A}$ indexed by I and another collection $E \subseteq G(\mathcal{A})$. A probability assignment is valid for E if and only if the probabilities lie in $\mathrm{conv}(F)$, where F is the set of feasible solutions to $0 \leq P(A_i) \leq 1$ and the multilinear inequalities relating each event in E to the multilinear function associated with it.*

Proof. Let $\mathcal{A}$ denote the collection of events over which P is formulated and let $\mathcal{A}$ be indexed by I. Now, consider the collection of events $\mathcal{B}$, consisting of events of the form $\bigcap_{i\in I} S_i$ where each S_i is either equal to A_i or $\overline{A_i}$. Note that the events in B are disjoint and cover the probability space in the sense that $\bigcup_{B_i\in B} B_i = U$ where U is the universal set. Now, consider an event E in $G(\mathcal{A})$. By expressing each event in the disjunctive normal form, it is easily seen that E can be expressed as a disjoint union of events in $\mathcal{B}$. Therefore, if the probabilities of events in $\mathcal{B}$ are known, probabilities for other events in $G(\mathcal{A})$ can be easily computed. Further, there exists a valid probability assignment to events in E if and only if there exists an assignment of probabilities to events in $\mathcal{B}$ such that for every $B_i \in \mathcal{B}$, $P(B_i) \geq 0$ and $\sum_{B_i\in\mathcal{B}} P(B_i) = 1$. This is the canonical simplex in $2^{|I|}$ dimensional space. Consider the linear transformation $M : \mathbb{R}^{2^{|I|}} \mapsto \mathbb{R}^{|E|}$ such that each point $(P(B_1), \ldots, P(B_{2^{|I|}}))$ is transformed to $\bigl(P(E_1), \ldots, P(E_{|E|})\bigr)$. We have already shown in Theorem 3 (as a very special case) that $\text{conv}(F_M^2)$ is precisely the above linear transformation of the canonical simplex.

Corollary 4. *A linear relation between the probabilities of the events holds if and only if the corresponding multilinear inequality holds over the unit hypercube.*

Proof. Follows directly from Theorem 9.

Note that linearity/convexity of the relation is necessary in Corollary 4. For general nonlinear expressions, the result does not hold. For example, $P(A_1 \cap A_2) \neq P(A_1)P(A_2)$, even though the corresponding multilinear expression holds.

Corollary 5. *Let $\mathcal{A}$ be a collection of events and consider a probabilistic mathematical program P that relates probabilities of events in a collection E^P, where $E_P \subseteq G(\mathcal{A})$. Let F be the feasible region of P. Apply Algorithm AssocMath-Program to generate an associated mathematical program P_M. Let F_M denote the feasible region of P_M. Further, let F_M^1 be the set of feasible solutions to the inequalities of P_M generated in 1^{st} step of AssocMathProgram and F_M^2 be the set of feasible solutions to the inequalities of 2^{nd} and 3^{rd} step. Then, P can be refomulated as the relaxation of P_M where F_M is relaxed to $F_M^1 \cap \text{conv}(F_M^2)$.*

Proof. Follows directly from Theorem 9.

We illustrate the ideas in Theorem 9 by returning to the Example 4. Using the well-known convex envelope of a bilinear term [10,1], Corollary 5 establishes that the following mathematical program:

$$\min\{\max(P(A) + P(B) - 1, 0) \mid P(A) \geq 0.5,\ P(B) \geq 0.6\}, \qquad (5)$$

is equivalent to (P). For this example, it is easily seen that the optimal value indeed matches.

The relation of probabilistic experiments to polyhedral basis is rather close. Consider a collection of convex sets $\mathcal{C}$ indexed by I. Inject the convex sets in a higher dimensional space such that the sets are located at the extreme points of the canonical simplex. More formally, construct the collection $\mathcal{S}$ indexed by I such that $S_i \in \mathcal{S}$ if: $S_i = \{(y_i, e_i) \mid y_i \in C_i, i \in I\}$, and e_i is the i^{th} unit vector. Clearly, by Theorem 2, if we have available to us a polyhedral basis for $\mathcal{C}$ over a superset X and ϕ is the lifting of $\mathcal{C}$, then $\text{conv}(\mathcal{S}) = \text{conv}(\phi(\mathcal{C})) = \text{conv}(\phi(X))$.

Then, constructing the polyhedral basis is in essence equivalent to constructing an experiment such that given a point $x \in X$, the polyhedral functions assign probabilities to some arbitrarily chosen point in C_i such that the expected value is x. In particular, if $f_i(x)$ is the polyhedral function associated with C_i and $f_j(x)$ is the polyhedral function associated with C_j, then treating f_i and f_j as probabilities $f_i(x) + f_j(x)$ can be considered as a probability associated with $C_i \cup C_j$. Corollary 3 provides the formal explanation of why independence of events is important in constructing such probabilistic experiments or polyhedral basis.

Now, we show that probabilistic rounding schemes have inadvertently though in rather inventive ways used the essential idea behind Theorem 9. Many noteworthy advances in the art of probabilistic rounding have interpreted, albeit this does not seem to have been realized, relaxations of mathematical programs as probabilistic mathematical programs and have developed experiments for sampling from the event space interpreting the optimal solution as a solution to probabilistic mathematical program. As pointed above, this often amounts to defining a polyhedral basis for the feasible region which is expressible as a collection of convex sets.

In the seminal paper on randomized rounding [12], the authors identified the optimal value of each variable x in the linear programming relaxation with its probability of being 1. Further, they rounded each variable independently of the other. This corresponds to using multilinear functions as a polyhedral basis for the underlying hypercube. In the presence of an inequality of the form

$\sum_{i=1}^{n} x_i \leq 1$, the LP solution is rounded by picking a random number r uniformly over $[0, 1]$ and assigning x_j to 1 if and only if $r \in [\sum_{i=1}^{j-1} x_i, \sum_{i=1}^{j} x_i]$. Such an experiment corresponds to the unique polyhedral basis for the simplex provided in Example 1.

A remarkable probabilistic rounding scheme was designed in the paper of Goemans and Willimason [8]. We argue that this experiment in fact defines a polyhedral basis for the extreme points of the boolean quadric polytope. We first present a brief overview of the probabilistic rounding scheme developed in [8]. The authors define a relaxation for the max-cut problem by employing a matrix $X = xx^t$ and relaxing the condition that X is rank-one to its positive-semidefiniteness following [9]. Then as guaranteed by incomplete Cholesky Factorization, they express $X = VV^t$, thereby associating each $x_i x_j$ with $\langle v_i, v_j \rangle$, where v_i is a unit vector and corresponds to the i^{th} row of V. Choosing another unit vector r randomly, a variable x_i is set to 1 if $\langle v_i, r \rangle \geq 0$ and -1 if $\langle v_i, r \rangle < 0$.

To see that the rounding scheme defines a polyhedral basis for the boolean quadric polytope, define the event A_i as the event that x_i is set to 1. Choose an additional unit-vector v_t and assign x_i to 1 if $\operatorname{sgn}(\langle v_i, r \rangle) = \operatorname{sgn}(\langle v_t, r \rangle)$ and to -1 if $\operatorname{sgn}(\langle v_i, r \rangle) = -\operatorname{sgn}(\langle v_t, r \rangle)$, where $\operatorname{sgn}(a)$ is 1 if $a \geq 0$ and -1 otherwise. We need to argue that there exists a superset, S_B, of the extreme points of the boolean quadric polytope such that for any $\bar{y} \in S_B$, a probabilistic experiment can be constructed where the probabilities of A_i is provided by the i^{th} co-ordinate, $\bar{y}_i$, and the probabilities of $A_i \cap A_j$ is provided by $\bar{y}_{ij}$ which corresponds to the co-ordinate $y_i y_j$. Note that $\bar{y}_{ij}$ does not necessarily equal $\bar{y}_i \bar{y}_j$

when $\bar{y}$ is not an extreme point of the quadric polytope. As shown in Lemma 2.2 of [8]:

$$P(A_i \cap A_j) + P(\overline{A_i} \cap \overline{A_j}) = \frac{\arccos(\langle v_i, v_j \rangle)}{\pi}. \tag{6}$$

It can be easily verfied (and also following from the multilinear equality: $x_1x_2 + (1-x_1)(1-x_2) = 1 - x_1 - x_2 + 2x_1x_2$ and Corollary 4) that:

$$P(A_i \cap A_j) + P(\overline{A_i} \cap \overline{A_j}) = 1 - P(A_i) - P(A_j) + 2P(A_i \cap A_j). \tag{7}$$

Consider a point $\bar{y} \in S_B$. Interpreting the co-ordinates of $\bar{y}$ as the probabilities of $P(A_i)$ and $P(A_i \cap A_j)$, we easily obtain:

$$x_{it} = \cos(\bar{y}_i\pi)$$
$$x_{ij} = \cos\big((1 - \bar{y}_i - \bar{y}_j + 2\bar{y}_{ij})\pi\big).$$

Provided the matrix $X = [x_{ij}]$ is positive semidefinite, the vectors v_i can be found by Incomplete Cholesky Factorization allowing the construction of a probabilistic experiment with the required probabilities since given x_{ij} and x_{it} it is easy to find the probability of $P(A_i \cap A_j)$ using (7). On closer inspection, the reader may notice that our argument implicitly employs the linear transformation between the cut polytope and the boolean quadric polytope originally provided in [22]. The remaining part is to show that the matrix X defined above is positive-semidefinite if $\bar{y}$ is an extreme point of the boolean quadric polytope. This is easy to verify by setting $v_i = v_t$ if $\bar{y}_i = 1$ and $v_i = -v_t$ otherwise. Interestingly, the constructed poolyhedral basis is not explicitly available in closed form but requires the application of Cholesky Factorization. Indeed, since the feasible region of the trignometric formulation in Theorem 2.9 in [8] is now expressed as a subset of the boolean quadric polytope it should be clear that the formulation is not just a relaxation but a reformulation of max-cut.

In addition to the connection between probabilistic rounding schemes and polyhedral basis Theorem 9 points to some interesting relations between relaxations of multilinear programs. In fact, we next argue that the relaxation constructed in Corollary 5 is a natural Lagrangian relaxation of the associated mathematical program. This in turn implies that most relaxations for multilinear programs are further relaxations the associated probabilistic mathematical program.

Theorem 10. *Consider a mathematical program with multilinear expressions of variables restricted to lie in the unit hypercube. Substitute new variables for the multilinear expressions. Form a Lagrangian relaxation by dualizing all constraints except those enforcing that the variables lie in the unit hypercube and the equality constraints between the newly introduced variables and the multilinear expressions they represent The Lagrangian relaxation thus formed is a reformulation of the associated probabilistic mathematical program.*

Proof. Since the feasible region of the Lagrangian subproblem is bounded, it follows from Theorem 1.17 in [15] that the perturbation function is proper and lower semicontinuous. Therefore, the Lagrangian dual and the relaxation constructed in Corollary 5 are duals of each other with no duality gap.

Returning to Example 4, Theorem 10 states that the following Lagrangean relaxation is a reformulation of the probabilistic mathematical program:

$$(L)\ \max_{\lambda_1,\lambda_2\geq 0}\ \min_{P(A_1),P(A_2)\in[0,1]}\ P(A_1)P(A_2)-\lambda_1(P(A_1)-0.5)-\lambda_2(P(A_2)-0.6)$$

The optimum value of a multilinear function over a hypercube is attained at one of the extreme points [7,13]. It is easy to see that (L) can be rewritten as:

$$\begin{aligned} &\max\ 0.5\lambda_1+0.6\lambda_2+a\\ &\text{s.t.}\quad a\leq\min\{1-\lambda_1-\lambda_2,-\lambda_1,-\lambda_2,0\}\\ &\qquad\ \ \lambda_1,\lambda_2\geq 0. \end{aligned}$$

The optimum value of the above program is indeed 0.1 which is attained at $(\lambda_1^*,\lambda_2^*,a^*)=(1,1,-1)$. Note that λ_1^* and λ_2^* are respectively the optimum dual multipliers of $P(A_1)\geq 0.5$ and $P(A_2)\geq 0.6$ in (5).

Lemma 4. *Consider a Lagrangian relaxation* $\max_\lambda \min_x L(\lambda,x)$ *and two real numbers* α, β *greater than or equal to zero. Assume that there exists an algorithm A that given a* λ *calculates* $z_A(\lambda)$, *such that* $\alpha\min_x L(\lambda,x)\geq z_A(\lambda)\geq \beta\min_x L(\lambda,x)$. *Then,* $\alpha\max_\lambda\min_x L(\lambda,x)\geq\max_\lambda z_A(\lambda)\geq\beta\max_\lambda\min_x L(\lambda,x)$.

Proof. If λ^* is the optimal solution to $\max_\lambda\min_x L(\lambda,x)$, then $\max_\lambda z_A(\lambda)\geq z_A(\lambda^*)\geq\beta\min_x L(\lambda^*,x)=\beta\max_\lambda\min_x L(\lambda,x)$. If λ' is the optimal solution to $\max_\lambda z_A(\lambda)$, then $\max_\lambda z_A(\lambda)=z_A(\lambda')\leq\alpha\min_x L(\lambda',x)\leq\alpha\max_\lambda\min_x L(\lambda,x)$.

Some related references. [20] [21][11][24][6]

References

1. F. A. Al-Khayyal and J. E. Falk. Jointly Constrained Biconvex Programming. *Mathematics of Operations Research*, 8:273–286, 1983.

2. N. Alon and J. H. Spencer. *The Probabilistic Method. With an Appendix on Open Problems by Paul Erdös.* Wiley-Intersceince Series in Discrete Mathematics and Optimization, New York, 1992.

3. E. Balas. Disjunctive programming and hierarchy of relaxations for discrete optimization problems. *SIAM J. Alg. Disc. Meth.*, 6:466–486, 1985.

4. E. Balas. Disjunctive programming: Properties of the convex hull of feasible points. *Discrete Applied Mathematics*, 89(1-3):3–44, 1998. Original manuscript was published as a technical report in 1974.

5. E. Balas, S. Ceria, and G. Cornuèjols. A Lift and Project Cutting Plane Algorithm for Mixed 0-1 Programs. *Mathematical Programming*, 58:295–324, 1993.

6. S. Ceria and J. Soares. Convex programming for disjunctive convex optimization. *Mathematical Programming*, 86A:595–614, 1999.

7. Y. Crama. Concave Extensions for Non-linear 0-1 Maximization Problems. *Mathematical Programming*, 61:53–60, 1993.

8. M. X. Goemans and D. P. Williamson. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *Journal of ACM*, 42:1115–1145, 1995.

9. L. Lovàsz and A. Schrijver. Cones of Matrices and Set-functions and 0-1 Optimization. *SIAM Journal on Optimization*, 1:166–190, May 1991.

10. G. P. McCormick. Computability of Global Solutions to Factorable Nonconvex Programs: Part I – Convex Underestimating Problems. *Mathematical Programming*, 10:147–175, 1976.

11. Y. Nesterov. Semidefinite Relaxation and Nonconvex Quadratic Optimization. *Optimization Methods and Software*, 9:141–160, 1997.

12. P. Raghavan and C. D. Thompson. Randomized Rounding. *Combinatorica*, 7:365–374, 1987.

13. A. D. Rikun. A Convex Envelope Formula for Multilinear Functions. *Journal of Global Optimization*, 10:425–437, 1997.

14. R. T. Rockafellar. *Convex Analysis*. Princeton Mathematical Series. Princeton University Press, 1970.

15. R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. A Series of Comprehensive Studies in Mathematics. Springer, Berlin, 1998.

16. H. D. Sherali. Convex Envelopes of Multilinear Functions over a Unit Hypercube and over Special Discrete Sets. *Acta Mathematica Vietnamica*, 22:245–270, 1997.

17. H. D. Sherali and W. P. Adams. A Hierarchy of Relaxations between the Continuous and Convex Hull Representations for Zero-One Programming Problems. *SIAM Journal of Discrete Mathematics*, 3:411–430, 1990.

18. H. D. Sherali and W. P. Adams. A hierarchy of relaxations and convex hull characterizations for mixed- integer zero-one programming problems. *Discrete Applied Mathematics*, 52(1):83–106, 1994.

19. H. D. Sherali and W. P. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, volume 31 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, Dordrecht, 1999.

20. H. D. Sherali and A. Alameddine. A New Reformulation–Linearization Technique for Bilinear Programming Problems. *Journal of Global Optimization*, 2:379–410, 1992.

21. H. D. Sherali and C. H. Tuncbilek. A Reformulation-Convexification Approach for Solving Nonconvex Quadratic Programming Problems. *Journal of Global Optimization*, 7:1–31, July 1995.

22. C. De Simeone. The Cut Poytope and the Boolean Quadric Polytope. *Discrete Mathematics*, 79:71–75, 1990.

23. R. A. Stubbs and S. Mehrotra. A Branch-and-Cut Method for 0-1 Convex Programming. Technical report, Northwestern University, January 1996.

24. M. Tawarmalani and N. V. Sahinidis. Semidefinite Relaxations of Fractional Programs via Novel Techniques for Constructing Convex Envelopes of Nonlinear Functions. *Journal of Global Optimization*, 20:137–158, 2001.

25. M. Tawarmalani and N. V. Sahinidis. Convex Extensions and Convex Envelopes of l.s.c. Functions. *Mathematical Programming*, (accepted 2002).