Lift-and-project for 0–1 programming via algebraic geometry

Luis F. Zuluaga* Juan C. Vera† Javier F. Peña‡
Carnegie Mellon University

December 4, 2003

Abstract

Recently, tools from algebraic geometry have been successfully applied to develop solution schemes for new classes of optimization problems. A central idea in these constructions is to express a polynomial that is positive on a given domain in terms of polynomials of higher degree so that its positivity is readily revealed. This resembles the "lifting" phase of the lift-and-project procedures for 0–1 programming.

We propose an enhancement to these solution schemes via a construction that is reminiscent of the "projecting" phase of the lift-and-project procedures. Our construction applies to domains that can be represented as the intersection of a set and an affine variety.

To illustrate the power of our approach, we provide novel derivations of some of the lift-and-project procedures for 0–1 programming due to Balas, Ceria and Cornuéjols; Sherali and Adams; Lovász and Schrijver; and Lasserre. These derivations add new insight into this interesting subject, and suggest a number of variations and extensions.

*Supported by NSF grants CCR-0092655 and DMI-0098427. lzuluaga@andrew.cmu.edu
†Supported by NSF grant CCR-0092655. jvera@andrew.cmu.edu
‡Supported by NSF grant CCR-0092655. jfp@andrew.cmu.edu

1 Introduction

The motivation for this work is to provide a connection between two major trends in optimization. One of these trends is the lift-and-project procedures for 0–1 programming. The other one is the recent research activity in optimization that draws on tools from real algebraic geometry to devise solution schemes for *polynomial optimization problems*, whose objective and constraints are multivariate polynomials. The first trend includes the relaxation schemes of Balas-Ceria-Cornuéjols [1], Sherali-Adams [20], Lovász-Schrijver [12], and in more generality, the so-called *Reformulation Linearization Techniques* and *Convexification Techniques* (for surveys see [21] and [25]). The second trend includes the recent work by de Klerk [4], Kojima et al. [7], Laurent [11], Lasserre [8, 9, 10] and Parrilo [16, 17, 18], along with previous work by Shor [22, 23] and Nesterov [13], among others.

The first trend, lift-and-project procedures, originally aims to describe the convex hull of the 0–1 points inside a given polyhedron. The fundamental ideas underlying these procedures are: 1) *lift* the problem to a higher dimensional space where the structure of the 0–1 polyhedron is more clearly revealed, and valid inequalities can be inferred; 2) *project* the problem back to the original space.

The second trend relies on representation theorems from real algebraic geometry to derive solution schemes for polynomial optimization problems. Such representation theorems characterize the set of polynomials that are positive on a given domain. This typically involves writing a given polynomial $p(x)$ in terms of polynomials of higher degree, so that the positivity of $p(x)$ is readily revealed. This resembles the *lifting* phase of the lift-and-project procedures mentioned above.

Our central goal is to enhance the solution schemes for polynomial optimization problems by incorporating a construction that is reminiscent of the *projecting* phase of the lift-and-project procedures. Specifically, we consider polynomial optimization problems over a domain S of the form $S = D \cap V$, where D is a set and V is an affine variety (i.e., a set defined by a finite system of polynomial equations). We formalize the following basic idea: use information about the (conceivable simpler) set of polynomials that are positive over D to infer information about the (more complicated) set of polynomials that are positive over $D \cap V$.

Our results yield novel derivations from a dual viewpoint of the lift-and-project procedures of Balas-Ceria-Cornuéjols [1]; Sherali-Adams [20]; Lovász-Schrijver [12]; and Lasserre [8]. The unified derivation of these results is similar in spirit to the work by Ceria [2] and Laurent [11]. However, in contrast to Ceria's lift-and-project approach [2] and Laurent's combinatorial approach [11], we follow an algebraic-geometric approach similar to the one introduced by Parrilo in [18]. This adds new insight into this interesting subject, and suggests a number of variations and extensions.

The rest of the paper is organized as follows. In Section 2, we present the relevant problem, terminology, and notation. We also formally state our central objective; namely, to obtain information about the set of polynomials that are positive over a domain S, by exploiting the particular decomposition structure $S = D \cap V$ of such domain. Section 3 presents the first formalization of the above idea. We show that under suitable assumptions, a *modulo-ideal* construction takes information about the set of polynomials that are positive over D, and yields information about the set of polynomials that are positive over $D \cap V$. Sections 4 and 5 refine the results of Section 3. In particular, we obtain new derivations for certain known hierarchies of relaxations for 0–1 programming via the modulo-ideal construction. Finally, in Section 6 we discuss an issue of crucial algorithmic relevance: We show that if a set of polynomials has a *computable description,* e.g., as a polyhedron or as a system of linear matrix inequalities (LMI), then the modulo-ideal constructions yield a set of polynomials with the same property.

2 Preliminaries

2.1 Problems and notation

The following optimization problem is our central object of study:

$$\begin{array}{rll} \rho = & \inf & f(x) \\ & \text{s.t.} & x \in S, \end{array} \tag{1}$$

where $f(x)$ is a given polynomial in n variables and $S \subseteq \mathbb{R}^n$ is a given set. We can rephrase this problem as

$$\begin{array}{rll} \rho = & \sup & \lambda \\ & \text{s.t.} & f(x) - \lambda \geq 0 \quad \text{for all } x \in S. \end{array} \tag{2}$$

As it has been recognized by several researchers (see, e.g., [7, 8, 13, 16, 22]), the latter formulation suggests working with a dual object; namely, the set of polynomials that are non-negative in the domain S. Specifically, let $\mathcal{H}_n := \mathbb{R}[x_1, \ldots, x_n]$ be the set of polynomials in n variables with real coefficients, and consider the cones of polynomials $\mathcal{P}_n(S)$, $\mathcal{P}_n^o(S)$, Σ_n defined as follows.

Let $\mathcal{P}_n(S)$ be the cone of polynomials in $\mathcal{H}_n$ that are *positive semidefinite* (non-negative) in the domain $S \subseteq \mathbb{R}^n$; that is,

$$\mathcal{P}_n(S) = \{p \in \mathcal{H}_n : p(x) \geq 0 \text{ for all } x \in S\}.$$

Let $\mathcal{P}_n^o(S)$ be the cone of polynomials in $\mathcal{H}_n$ that are *positive definite* in the domain $S \subseteq \mathbb{R}^n$; that is,

$$\mathcal{P}_n^o(S) = \{p \in \mathcal{H}_n : p(x) > 0 \text{ for all } x \in S\}.$$

Finally, let Σ_n be the cone of polynomials in $\mathcal{H}_n$ that are sum of squares of polynomials; that is,

$$\Sigma_n = \text{conv}\{q(x)^2 : q \in \mathcal{H}_n\}.$$

We will also work with the subsets of $\mathcal{H}_n$, $\mathcal{P}_n(S), \mathcal{P}_n^o(S)$ and Σ_n obtained by considering only polynomials of bounded degree: For a positive integer m, let $\mathcal{H}_{n,m}$ be the set of polynomials in $\mathcal{H}_n$ with degree at most m, $\mathcal{P}_{n,m}(S) := \mathcal{P}_n(S) \cap \mathcal{H}_{n,m}$, $\mathcal{P}_{n,m}^o(S) := \mathcal{P}_n^o(S) \cap \mathcal{H}_{n,m}$ and $\Sigma_{n,m} := \Sigma_n \cap \mathcal{H}_{n,m}$.

2.2 Approximation schemes

Consider problem (2) above and let $m = \deg(f)$. Notice that for any K satisfying

$$\mathcal{P}_{n,m}^o(S) \subseteq K \subseteq \mathcal{P}_{n,m}(S), \tag{3}$$

we have

$$\begin{aligned} \rho = \quad & \sup \quad \lambda \\ & \text{s.t.} \quad f(x) - \lambda \in K. \end{aligned} \tag{4}$$

Also, given a sequence K_r, $r = 0, 1, \dots$ satisfying

$$K_r \subseteq K_{r+1} \subseteq \mathcal{P}_{n,m}(S),\ r = 0, 1, \dots \text{ and } \mathcal{P}_{n,m}^o(S) \subseteq \bigcup_{r=0}^{\infty} K_r \subseteq \mathcal{P}_{n,m}(S), \tag{5}$$

consider the sequence ρ_r, $r = 0, 1, \dots$, defined by

$$\begin{aligned} \rho_r = \quad & \sup \quad \lambda \\ & \text{s.t.} \quad f(x) - \lambda \in K_r. \end{aligned} \tag{6}$$

Observe that $\rho_r \uparrow \rho$. In other words, for a fixed r, (6) yields an approximation to (2).

Throughout the sequel, we use the following notational convention. We write $K \approx \mathcal{P}_{n,m}(S)$ to indicate that K satisfies (3), and $K_r \uparrow \mathcal{P}_{n,m}(S)$ to indicate that the sequence K_r, $r = 0, 1, \dots$ satisfies (5). Likewise for $K \approx \mathcal{P}_n(S)$ and $K_r \uparrow \mathcal{P}_n(S)$.

As the examples in Section 2.3 below show, for certain domains $S \subseteq \mathbb{R}^n$, representation results from real algebraic geometry can be used to obtain $K \approx \mathcal{P}_{n,m}(S)$ or $K_r \uparrow \mathcal{P}_{n,m}(S)$ with *computable descriptions*. In such cases problem (4) or (6) can be cast as a linear, second order cone or semidefinite program, and hence be amenable to modern optimization technology. These types of *approximation schemes* underlie some of the main constructions in [4, 8, 9, 16].

Our central goal is to enhance these approximation schemes by exploiting the particular structure of the domain S. Specifically, we consider domains of the form $S = D \cap V$, where D is a set and V is an *affine variety,* that is,

$$V = V(f_1, \dots, f_v) := \{x \in \mathbb{R}^n : f_k(x) = 0,\ k = 1, \dots, v\} \tag{7}$$

for some given $f_k \in \mathcal{H}_n$, $k = 1, \dots, v$.

A particular case of a domain with this form, and central to integer programming, is

$$S = \{x \in \mathbb{R}^n : Ax \geq b\} \cap \{0, 1\}^n,$$

which can be written as $D \cap V$ for

$$D = \{x \in \mathbb{R}^n : Ax \geq b\},\ V = \{x \in \mathbb{R}^n : x_i(x_i - 1) = 0,\ i = 1, \ldots, n\}.$$

In the sequel we develop several constructions that formalize the following basic idea: use information about the (conceivable simpler) set $\mathcal{P}_{n,m}(D)$ (or $\mathcal{P}_n(D)$) to describe $\mathcal{P}_{n,m}(D \cap V)$. We illustrate these constructions by providing alternative derivations of some known hierarchies of relaxations for 0–1 programming.

2.3 Computable descriptions

Observe that a polynomial $p(x) \in \mathcal{H}_{n,m}$ can be identified with its (finite-dimensional) vector of coefficients. We shall use p to denote this vector of coefficients, assuming some ordering of the monomials in $\mathcal{H}_{n,m}$ is set. With this identification in mind, we shall say that $K \subseteq \mathcal{H}_{n,m}$ has a *computable description* if K can be written in the form

$$K = \{p \in \mathcal{H}_{n,m} : \mathcal{A}p = \mathcal{B}z,\ z \in C\}, \tag{8}$$

where $\mathcal{A}, \mathcal{B}$ are suitable linear maps, and C is a cone amenable to modern optimization technology, e.g., modern interior-point methods. This is known to be the case if C is the non-negative orthant, the positive semidefinite cone, or the second order cone. These are indeed special cases of the broader class of *self-scaled cones* (also known as symmetric cones), for which there are interior-point algorithms (see [14, 15]). We use the term "computable description" because such a description for $K \approx \mathcal{P}_{n,m}(S)$ yields conceptual algorithms for problem (4). This is illustrated by Examples 1, 2, and 3 below.

Example 1 Let $S = \{x : Ax \geq b\}$ be a non-empty polyhedron. Then by Farkas Lemma it follows that

$$p(x) = c^{\mathrm{T}}x + d \in \mathcal{P}_{n,1}(S) \text{ if and only if } \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} A^{\mathrm{T}} & 0 \\ -b^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} z \\ \tau \end{bmatrix} \text{ for some } \begin{bmatrix} z \\ \tau \end{bmatrix} \geq 0. \tag{9}$$

Furthermore, if in addition $\{x : Ax \geq -b\} = \emptyset$ then, again by Farkas Lemma

$$p(x) = c^{\mathrm{T}}x + d \in \mathcal{P}_{n,1}(S) \text{ if and only if } \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} A^{\mathrm{T}} \\ -b^{\mathrm{T}} \end{bmatrix} z \text{ for some } z \geq 0. \tag{10}$$

Example 2 Let $S = \{x \in \mathbb{R}^n : x^{\mathrm{T}}Mx + 2d^{\mathrm{T}}x + e \geq 0\}$, where $M \in \mathbb{S}^n$, $d \in \mathbb{R}^n$, $e \in \mathbb{R}$. Then by the S-lemma [26, 24]

$$p(x) = x^{\mathrm{T}}Qx + 2b^{\mathrm{T}}x + c \in \mathcal{P}_{n,2}(S) \text{ if and only if}$$
$$\begin{bmatrix} Q & b \\ b^{\mathrm{T}} & c \end{bmatrix} - t \begin{bmatrix} M & d \\ d^{\mathrm{T}} & e \end{bmatrix} \succeq 0 \text{ for some } t \geq 0.$$

($X \succeq 0$ denotes that X is a positive semidefinite matrix.)

Example 3 Let $S = \mathbb{R}$. Using the fact that an univariate polynomial is non-negative if and only if it is a sum of squares (see, e.g., [13]) it follows that

$$p(x) = a_0 + a_1 x + \ldots + a_{2m} x^{2m} \in \mathcal{P}_{1,2m}(S) \text{ if and only if}$$

$$a_k = \sum_{i+j=k} y_{ij},\ k = 0, \ldots, 2m \text{ for some } Y = \begin{bmatrix} y_{00} & y_{01} & \cdots & y_{0m} \\ y_{01} & y_{11} & \cdots & y_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ y_{0m} & y_{1m} & \cdots & y_{mm} \end{bmatrix} \succeq 0.$$

In a similar fashion, the following example illustrates a case where there exists an inner approximation $K_r \uparrow \mathcal{P}_{n,m}(S)$, where each K_r has a computable description.

Example 4 Let $S = \{x \in \mathbb{R}^n_+ : x_1 + \cdots + x_n = 1\}$. Then by Pólya's Theorem, $K_r \uparrow \mathcal{P}_{n,m}(S)$, where

$$K_r = \{p \in \mathcal{H}_{n,m} : (x_1 + \cdots + x_n)^r p(x) \text{ has non-negative coefficients}\}.$$

3 Approximation schemes for $\mathcal{P}_n(D \cap V)$

Throughout the sequel we shall make the following assumption.

Blanket Assumption: $D \subseteq \mathbb{R}^n$ is a given set, and V is an affine variety as in (7).

We next present a *modulo-ideal* construction that yields approximations for $\mathcal{P}_n(D \cap V)$, given approximations for $\mathcal{P}_n(D)$.

First recall some basic terminology and notation from algebraic geometry (for an excellent reference on the subject see [3]). Given an affine variety V in $\mathbb{R}^n$, $I(V)$ is the *ideal of* V, in other words, $I(V) = \{f \in \mathcal{H}_n : f(x) = 0 \text{ for all } x \in V\}$. Also, given $g_1, \ldots, g_l \in \mathcal{H}_n$, let $\langle g_1, \ldots, g_l \rangle$ be the *ideal generated by* $\{g_1, \ldots, g_l\}$, i.e., $\langle g_1, \ldots, g_l \rangle = \{p \in \mathcal{H}_n : p(x) = \sum_{i=1}^l h_i(x) g_i(x), h_i \in \mathcal{H}_n, i = 1, \ldots, l\}$. Given $p, q \in \mathcal{H}_n$, the notation $p \equiv q \mod I(V)$ indicates that p and q are *congruent modulo* $I(V)$; that is, $p - q \in I(V)$.

Lemma 1 *Given $K \subseteq \mathcal{P}_n(D)$, let $Q = \{p \in \mathcal{H}_n : \exists q \in K,\ p \equiv q \mod I(V)\}$. Then $Q \subseteq \mathcal{P}_n(D \cap V)$.*

Proof. Let $p \in Q$. Then there exists $q \in K$ such that $p \equiv q \mod I(V)$. Thus $p(x) - q(x) = 0$ for any $x \in V$. Therefore, since $K \subseteq \mathcal{P}_n(D)$, we have $p(x) = q(x) \geq 0$ for all $x \in D \cap V$, i.e., $p \in \mathcal{P}_n(D \cap V)$.

□

The following theorem shows how to obtain a *one-step approximation* $Q \approx \mathcal{P}_n(D \cap V)$ given a one-step approximation $K \approx \mathcal{P}_n(D)$.

Theorem 2 (one-step approximation) *Assume D is compact. Given $K \approx \mathcal{P}_n(D)$, let $Q = \{p \in \mathcal{H}_n : \exists q \in K,\ p \equiv q \mod I(V)\}$. Then $Q \approx \mathcal{P}_n(D \cap V)$.*

Proof. If $D = \emptyset$ then $D = D \cap V = \emptyset$ and there is nothing to show. Hence assume $D \neq \emptyset$. The inclusion $Q \subseteq \mathcal{P}_n(D \cap V)$ follows from Lemma 1. For the other inclusion, take $p \in \mathcal{P}_n^o(D \cap V)$ and let $f(x) = \sum_{k=1}^{v} f_k(x)^2$ (recall (7)). It suffices to show that there exists $c \geq 0$ such that $p(x) + cf(x) \in \mathcal{P}_n^o(D) \subseteq K$. If $D \cap V = \emptyset$ take $c > |\min\{p(x) : x \in D\}| / \min\{f(x) : x \in D\}$. Otherwise let $\mu = \min\{p(x) : x \in D \cap V\} > 0$, and fix $\epsilon > 0$ such that $x \in D$ and $d(x, V) < \epsilon$ implies $p(x) > \mu/2$, where $d(x, V)$ is the distance from x to the set V. Notice that if $f(x) = 0$, then $x \in V$. Let $\sigma = \min\{f(x) : x \in D, d(x, V) \geq \epsilon\} > 0$. To finish, take $c > \frac{1}{\sigma} |\min\{p(x) : x \in D\}|$.

□

The following theorem shows how to obtain a *sequential approximation* $Q_r \uparrow \mathcal{P}_n(D \cap V)$, given a sequential approximation $K_r \uparrow \mathcal{P}_n(D)$,

Theorem 3 (sequential approximation) *Assume D is compact. Given $K_r \uparrow \mathcal{P}_n(D)$, let $Q_r = \{p \in \mathcal{H}_n : \exists q \in K_r,\ p \equiv q \mod I(V)\}$. Then $Q_r \uparrow \mathcal{P}_n(D \cap V)$.*

Proof. This follows by putting together Lemma 1 and Theorem 2. □

Given an approximation for the set $\mathcal{P}_n(D)$, Theorems 2 and 3 yield theoretical tools that yield an approximation for $\mathcal{P}_n(D \cap V)$. However, their statements concern polynomials in $\mathcal{H}_n$, which is an infinite-dimensional space, as the polynomials in $\mathcal{H}_n$ have unrestricted degree. In the subsequent sections we provide refinements of these results that concern the relevant subsets of the finite-dimensional space $\mathcal{H}_{n,m}$. Furthermore, as we show in Section 6, these refinements have an important property; namely, if the starting cone K (or sequence K_r) has a computable description, then the resulting cone Q (or sequence Q_r) has a computable description as well.

4 Refining the one-step approximation scheme

We next provide refinements of the one-step approximation in Theorem 2 to construct $Q \approx \mathcal{P}_{n,m}(D \cap V)$ given $K \approx \mathcal{P}_{n,m}(D)$ or, more generally, given $K \subseteq \mathcal{P}_n(D)$.

4.1 Restricted-degree modulo-ideal construction

The first natural attempt to obtain $Q \approx \mathcal{P}_{n,m}(D \cap V)$ given $K \approx \mathcal{P}_{n,m}(D)$ is to restrict the construction in Theorem 2 to polynomials of degree bounded by m. Although this simple idea does not always work, it does work when the degree m satisfies a suitable lower bound. The next corollary follows from the proof of Theorem 2.

Corollary 1 *Assume D is compact and $m \geq deg(f)$, where $f(x) = \sum_{k=1}^{v} f_k(x)^2$ (recall (7)). Given $K \approx \mathcal{P}_{n,m}(D)$, let $Q = \{p \in \mathcal{H}_{n,m} : \exists q \in K,\ p \equiv q \mod I(V)\}$. Then $Q \approx \mathcal{P}_{n,m}(D \cap V)$.*

Furthermore, if each $f_k \in \mathcal{P}_n(D)$, $k = 1, \ldots, v$ then the same holds under the weaker condition $m \geq \max\{\deg(f_k) : k = 1, \ldots, v\}$.

Corollary 1 yields a direct derivation, via the S-lemma [26, 24], of one of the main results presented by Sturm and Zhang in [24]: Let $V = \{x : r(x) = 0\}$ and $D = \{x : r(x) \geq 0\}$, where r is a strictly concave quadratic polynomial. The latter condition ensures that D is compact. By the S-lemma [26, 24], the set $K = \{\varphi(x) + tr(x) : t \geq 0, \varphi \in \Sigma_{n,2}\}$ satisfies $K \approx \mathcal{P}_{n,2}(D)$. Thus Corollary 1 yields $Q = \{p \in \mathcal{H}_{n,2} : \exists q \in K \text{ s.t. } p \equiv q \mod I(V)\} \approx \mathcal{P}_{n,2}(D \cap V)$. But $I(V) = \langle r(x) \rangle$, thus $Q = \{p \in \mathcal{H}_{n,2} : \exists q \in K \text{ s.t. } p(x) - q(x) = t'r(x), t' \in \mathbb{R}\} = \{\varphi(x) + t'r(x) : t' \in \mathbb{R}, \varphi \in \Sigma_{n,2}\}$, which is equivalent to [24, Thm. 2].

For a second application of Corollary 1, consider the affine variety

$$V_{a,b} = \{x \in \mathbb{R}^n : a^{\mathrm{T}}x - b = 0\},$$

where $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ are given. Define the operator $\mathcal{Z}_{a,b}$ as follows: Given $K \subseteq \mathcal{H}_{n,m}$, let

$$\mathcal{Z}_{a,b}(K) = \{p \in \mathcal{H}_{n,m} : p \equiv q \mod I(V_{a,b}), q \in K\}.$$

(Notice that $I(V_{a,b}) = \langle a^{\mathrm{T}}x - b \rangle$.)

The following result follows from Corollary 1.

Corollary 2 *Let $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, D be a compact set and $m > 1$. If $K \approx \mathcal{P}_{n,m}(D)$, then $\mathcal{Z}_{a,b}(K) \approx \mathcal{P}_{n,m}(D \cap V_{a,b})$.*

Without the lower bound condition on m, the statement in Corollary 1 may not hold. For example, when $m = 1$ and $V = \{x \in \mathbb{R}^n : x_j \in \{0,1\}\}$ it is easy to see that for any given $K \subseteq \mathcal{H}_{n,1}$, $\{p \in \mathcal{H}_{n,1} : \exists q \in K,\ p \equiv q \mod I(V)\} = K$. Hence in this case this *restricted-degree* modulo-ideal construction does not yield anything new.

We next present several constructions that give one-step approximations for $\mathcal{P}_{n,m}(D \cap V)$ for any m. The tradeoff for dropping the condition on m is to apply a "lifting" step to $K \approx \mathcal{P}_{n,m}(D)$ to get a richer subset of $\mathcal{P}_n(D)$. The restricted-degree modulo-ideal operation is then applied to such set. The latter operation can be seen as a "projecting" step. The following lemma provides the core of the constructions presented in the subsequent sections.

Lemma 4 *Assume $K \subseteq \mathcal{P}_n(D)$ and let $Q = \{p \in \mathcal{H}_{n,m} : \exists q \in K,\ p \equiv q \mod I(V)\}$. Then $Q \subseteq P_{n,m}(D \cap V)$. Furthermore, if K is such that for all $p \in \mathcal{P}_{n,m}^o(D \cap V)$ there exists $q \in K$ with $p - q \in I(V)$, then $Q \approx \mathcal{P}_{n,m}(D \cap V)$.*

Proof. The inclusion $Q \subseteq \mathcal{P}_{n,m}(D \cap V)$ readily follows from Lemma 1. For the second part, let $p \in \mathcal{P}_{n,m}^o(D \cap V)$. Then $q - p \in I(V)$ for some $q \in K$. Thus $p \in Q$. This shows $\mathcal{P}_{n,m}^o(D \cap V) \subseteq Q$ as well. □

4.2 Some relaxations for 0–1 programming

We next present some refinements of the one-step approximation scheme for the following particular variety, which plays a central role in 0–1 programming: Given $j \in \{1, \ldots, n\}$, let

$$V_j = \{x \in \mathbb{R}^n : x_j \in \{0, 1\}\}.$$

The constructions in this section give alternative derivations and extensions of some well-known relaxation hierarchies for 0–1 programming.

4.2.1 Balas-Ceria-Cornuéjols

In the same spirit as the lift-and-project operator P_j of Balas, Ceria and Cornuéjols [1], the following construction yields a one-step approximation for $\mathcal{P}_{n,m}(D \cap V_j)$ given a one-step approximation for $\mathcal{P}_{n,m}(D)$. Given $j \in \{1, \ldots, n\}$ and $K \subseteq \mathcal{H}_{n,m}$, let

$$B_j(K) = \{p \in \mathcal{H}_{n,m} : \exists q, q' \in K \text{ s.t. } p(x) \equiv (x_j q(x) + (1 - x_j) q'(x)) \mod I(V_j)\}.$$

(Notice that $I(V_j) = \langle x_j(1 - x_j) \rangle$.)

Theorem 5 *Let $j \in \{1, \ldots, n\}$. Assume D is compact and $D \subseteq \{x \in \mathbb{R}^n : 0 \leq x_j \leq 1\}$. If $K \approx \mathcal{P}_{n,m}(D)$, then $B_j(K) \approx \mathcal{P}_{n,m}(D \cap V_j)$.*

Proof. By Lemma 4, it suffices to show that given $p \in \mathcal{P}^o_{n,m}(D \cap V_j)$ there exist $q, q' \in K$ such that $x_j q(x) + (1 - x_j) q'(x) - p(x) \in I(V_j)$. If $p \in \mathcal{P}^o_{n,m}(D)$ or $D = \emptyset$, then simply take $q = q' = p$. Otherwise let $\epsilon = \min\{1 - x_j : p(x) \leq 0, x \in D\} > 0$, $\epsilon' = \min\{x_j : p(x) \leq 0, x \in D\} > 0$, and $\mu = |\min\{p(x) : x \in D\}|$. Put $q(x) = p(x) + c(1 - x_j), q'(x) = p(x) + c' x_j$ where $c > \frac{\mu}{\epsilon}$ and $c' > \frac{\mu}{\epsilon'}$. It follows that $q, q' \in P_{n,m}(D)$ and $x_j q(x) + (1 - x_j) q'(x) - p(x) \in I(V_j)$. □

For $J = \{j_1, \ldots, j_k\} \subseteq \{1, \ldots, n\}$ and $K \subseteq \mathcal{H}_{n,m}$ define

$$B_J(K) = B_{j_k}(B_{j_{k-1}}(\cdots (B_{j_1}(K)) \cdots)),$$

and

$$V_J = \bigcap_{j \in J} V_j.$$

The next corollary follows by induction from Theorem 5.

Corollary 3 *Let $J \subseteq \{1, \ldots, n\}$. Assume D is compact and $D \subseteq \bigcap_{j \in J} \{x \in \mathbb{R}^n : 0 \leq x_j \leq 1\}$. If $K \approx \mathcal{P}_{n,m}(D)$, then $B_J(K) \approx \mathcal{P}_{n,m}(D \cap V_J)$.*

Notice that under the conditions of Corollary 3, we obtain a *finite* sequential approximation of $\mathcal{P}_{n,m}(D \cap V_J)$. That is, if $J_i = \{j_1, \ldots, j_i\}$, $i = 1, \ldots, |J|$, then

$$K \subseteq B_{J_1}(K) \subseteq B_{J_2}(K) \subseteq \cdots \subseteq B_J(K) \approx \mathcal{P}_{n,m}(D \cap V_J).$$

Remark 1 The compactness hypothesis on D in Theorem 5 is only used to ensure that ϵ, ϵ' and μ are attained and finite. When $m = 1$ and D is a polyhedron, this compactness hypothesis is not necessary, as in this case ϵ, ϵ' and μ are simply the optimal values of feasible bounded linear programs. As a consequence, in this case the operator B_j is precisely the dual of the lift-and-project operator P_j in [1]. We next discuss this in more detail.

The operator B_j defined above is closely related to the operator P_j in [1]. Indeed, for linear polynomials, i.e., for $m = 1$ and D a polyhedron, the operator B_j is the dual counterpart of P_j. Specifically, for $D = \{x : Ax \geq b\}$ with $D \subseteq \{x : 0 \leq x_j \leq 1\}$ and $K = \mathcal{P}_{n,1}(D)$, Theorem 5 together with Remark 1 imply that the polynomials in $B_j(K)$ are precisely the set of valid inequalities of the convex hull of $D \cap \{x : x_j \in \{0, 1\}\}$, which is denoted $P_j^*(D)$ in [1]. Observe that in this case, by (10),

$$\mathcal{P}_{n,1}(D) = \{c^{\mathrm{T}}x + d : \exists z \geq 0 \ \text{ s.t. } c = A^{\mathrm{T}}z, \ d = -b^{\mathrm{T}}z\}.$$

Hence for $K = \mathcal{P}_{n,1}(D)$, it follows that $\alpha^{\mathrm{T}}x + \beta \in B_j(K)$ if and only if there exist $u, v \geq 0$ such that

$$\alpha^{\mathrm{T}}x + \beta \equiv x_j((A^{\mathrm{T}}u)^{\mathrm{T}}x - b^{\mathrm{T}}u) + (1 - x_j)((A^{\mathrm{T}}v)^{\mathrm{T}}x - b^{\mathrm{T}}v) \mod I(V_j).$$

This in turn holds if and only if $\beta = -b^{\mathrm{T}}v$, $\alpha_i = (A^{\mathrm{T}}v)_i = (A^{\mathrm{T}}u)_i$, $i \neq j$, and $\alpha_j = (A^{\mathrm{T}}u)_j + b^{\mathrm{T}}u - b^{\mathrm{T}}v$. In other words, $\alpha^{\mathrm{T}}x + \beta \in B_j(K)$ if and only if there exist $u, v \geq 0$ and $u_0, v_0 \in \mathbb{R}$ such that

$$\begin{array}{rcl} \alpha & = & A^{\mathrm{T}}u + u_0 e_j \\ \alpha & = & A^{\mathrm{T}}v + v_0 e_j \\ \beta & = & -b^{\mathrm{T}}u - u_0 \\ \beta & = & -b^{\mathrm{T}}v. \end{array}$$

So we recover the characterization of $P_j^*(D)$ in [1, Thm 2.10].

In the case $D \subseteq \{x \in \mathbb{R}^n : 0 \leq x_j \leq 1, \, j \in J\}$, Theorem 5 and Corollary 3 extend some of the central properties of the lift-and-project constructions in [1, 2]. Furthermore, when $m = 1$ and D is a polyhedron, the identity (9) readily yields a polyhedral representation for $\mathcal{P}_{n,m}(D)$. In contrast, for $m > 1$, imposing the linear constraints $0 \leq x_j \leq 1$, $j \in J$ on D poses a difficulty as there is no current result analogous to (9) that yields a computable description for $\mathcal{P}_{n,m}(D)$. Fortunately, when $m > 1$ this extra condition on D is no longer necessary.

Theorem 6 *Let $j \in \{1, \ldots, n\}$, and D be a compact set. Assume $m > 1$. If $K \approx \mathcal{P}_{n,m}(D)$, then $B_j(K) \approx \mathcal{P}_{n,m}(D \cap V_j)$.*

Proof. Modify the proof of Theorem 5 by using x_j^2 instead of x_j and $(1 - x_j)^2$ instead of $(1 - x_j)$ in the construction of q, q', ϵ, and ϵ'. □

It thus follows that the condition $D \subseteq \{x \in \mathbb{R}^n : 0 \leq x_j \leq 1, \, j \in J\}$ can also be omitted in Corollary 3 when $m > 1$.

To see how Theorem 6 and Corollary 3 extend the range of application of the ideas developed in [1], consider the following example.

Example 5 For $Q \in \mathbb{S}^{n+1}$, consider the following pure 0–1 quadratic program:

$$\begin{array}{rl} \varrho = \min & (1, x^{\mathrm{T}}) Q (1, x^{\mathrm{T}})^{\mathrm{T}} \\ \text{s.t.} & x \in \{0,1\}^n. \end{array}$$

Let $D = \{x \in \mathbb{R}^n : (1, x^{\mathrm{T}}) M (1, x^{\mathrm{T}})^{\mathrm{T}} \geq 0\}$, where

$$M = \begin{bmatrix} 0 & \frac{1}{2}e \\ \frac{1}{2}e & -I \end{bmatrix}.$$

Here, $e = (1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^n$ and $I \in \mathbb{R}^{n \times n}$ is the identity matrix.

Let $J_i = \{1, \ldots, i\}$, $i = 1, \ldots, n$. Since D is the n-dimensional sphere of radius $\sqrt{n}/2$ centered at $\frac{1}{2}e$, it follows that

$$\begin{array}{rl} \varrho = \sup & \lambda \\ \text{s.t.} & (1, x^{\mathrm{T}})(Q - \lambda \delta_{00})(1, x^{\mathrm{T}})^{\mathrm{T}} \in \mathcal{P}_{n,2}(D \cap V_{J_n}), \end{array}$$

where $\delta_{00} \in \mathbb{S}^{n+1}$ is the matrix with all zeros except for a 1 in position $(0, 0)$. Notice that D is compact, and by the S-lemma [26, 24], $K = \{(1, x^{\mathrm{T}}) Z (1, x^{\mathrm{T}})^{\mathrm{T}} : Z - tM \succeq 0, t \geq 0\} \approx \mathcal{P}_{n,2}(D)$. Therefore, the sequence

$$\begin{array}{rl} \varrho_{J_i} = \sup & \lambda \\ \text{s.t.} & (1, x^{\mathrm{T}})(Q - \lambda \delta_{00})(1, x^{\mathrm{T}})^{\mathrm{T}} \in B_{J_i}(K), \end{array}$$

for $i = 1, \ldots, n$ satisfies $\varrho_{J_1} \geq \cdots \geq \varrho_{J_n} = \varrho$. Furthermore, since K has an LMI description, each problem in the sequence ϱ_{J_i}, $i = 1, \ldots, n$ is a semidefinite program.

Notice that in Example 5 we have used an important feature common to both the one-step and sequential approximation schemes; namely, that if we are interested in approximating $\mathcal{P}_n(D' \cap V)$, there is a lot of freedom to choose a compact set D such that $D' \cap V = D \cap V$. In turn, this provides room towards obtaining the necessary K or sequence $K_r, r = 0, 1, \ldots$ in Theorems 2 and 3, or their corresponding refinements.

Other authors have already recognized that the lift-and-project procedure of Balas, Ceria and Cornuéjols can be generalized in various ways (see, e.g., [6] and [25]). Our goal in this presentation is to emphasize a novel perspective to these generalizations via simple algebraic-geometric tools.

4.2.2 Sherali-Adams

The lift-and-project procedure of Balas, Ceria, Cornuéjols is related to the hierarchy of relaxations by Sherali and Adams [20]. The latter inspires the following sequential approximation scheme for $\mathcal{P}(D \cap V_J)$.

Assume $J \subseteq \{1, \ldots, n\}$ and $K \subseteq \mathcal{H}_{n,m}$ are given. For $t = 0, 1, \ldots, |J|$, put $\Gamma(t) = \{(J_1, J_2) : J_1, J_2 \subseteq J, J_1 \cap J_2 = \emptyset, |J_1 \cup J_2| \leq t\}$ and let

$$W_J^t(K) = \Bigg\{ p \in \mathcal{H}_{n,m} : \exists\, q_{J_1,J_2} \in K,\, (J_1, J_2) \in \Gamma(t) \text{ s.t.}$$

$$p(x) \equiv \sum_{(J_1,J_2)\in\Gamma(t)} q_{J_1,J_2}(x) \prod_{j_1\in J_1} x_{j_1} \prod_{j_2\in J_2} (1 - x_{j_2}) \mod I(V_J) \Bigg\},$$

with the convention $\prod_{j_1\in\emptyset} x_{j_1} = \prod_{j_2\in\emptyset}(1 - x_{j_2}) = 1$.

Notice that for $t = 0, 1, \ldots, |J|$, if $(J_1, J_2) \in \Gamma(t)$ and $j \in J \setminus (J_1 \cup J_2)$ then

$$\prod_{j_1\in J_1} x_{j_1} \prod_{j_2\in J_2} (1-x_{j_2}) = \prod_{j_1\in J_1\cup\{j\}} x_{j_1} \prod_{j_2\in J_2} (1-x_{j_2}) + \prod_{j_1\in J_1} x_{j_1} \prod_{j_2\in J_2\cup\{j\}} (1-x_{j_2}).$$

Therefore, $W_J^t(K) \subseteq W_J^{t+1}(K)$ for $t = 0, 1, \ldots, |J|$. Also, $K \subseteq W_J^0(K)$.

Theorem 7 *Let $J \subseteq \{1, \ldots, n\}$ and D be a compact set. If $K \approx \mathcal{P}_{n,m}(D)$, then*

$$K \subseteq W_J^0(K) \subseteq \cdots \subseteq W_J^{|J|}(K) \approx \mathcal{P}_{n,m}(D \cap V_J).$$

Proof. From the discussion above and by Lemma 4, it suffices to show that given $p \in \mathcal{P}^o_{n,m}(D \cap V_J)$ there exist $q_{J_1,J_2} \in \mathcal{P}^o_{n,m}(D)$, $(J_1, J_2) \in \Gamma(|J|)$ such that

$$p - \sum_{(J_1,J_2)\in\Gamma(|J|)} q_{J_1,J_2} \prod_{j_1\in J_1} x_{j_1} \prod_{j_2\in J_2} (1 - x_{j_2}) \in I(V_J). \tag{11}$$

Fix $(J_1, J_2) \in \Gamma(|J|)$. Let $q_{J_1,J_2}(x)$ be the polynomial obtained by fixing the variables x_j, $j \in J$ in $p(x)$ as follows: $x_{j_1} = 1, j_1 \in J_1$ and $x_{j_2} = 0, j_2 \in J_2$. Since $p \in \mathcal{P}^o_{n,m}(D \cap V_J)$, it follows that each $q_{J_1,J_2} \in \mathcal{P}^o_{n,m}(D)$. Furthermore, this construction ensures that (11) holds. □

For the special case $m = 1$, and proceeding as in Section 4.2.1, it can be shown that the operator W_J^t defined above is the dual counterpart of the operator in the hierarchy of relaxations for 0-1 programming by Sherali and Adams [20].

4.2.3 Lovász-Schrijver

The lift-and-project construction by Balas, Ceria and Cornuéjols [1] and the hierarchy of relaxations by Sherali and Adams [20] are related to a third class of relaxations for 0–1 linear programming proposed by Lovász and Schrijver [12]. Again, the latter inspires a third finite sequential approximation of $\mathcal{P}_{n,m}(D \cap V_J)$.

Given $J \subseteq \{1,\ldots,n\}$ and $K \subseteq \mathcal{H}_{n,m}$, let

$$L_J(K) = \Bigg\{ p \in \mathcal{H}_{n,m} : \exists q_j, q'_j \in K,\, j \in J \ \text{ s.t.} \\ p(x) \equiv \sum_{j\in J} x_j q_j(x) + (1-x_j)q'_j(x) \mod I(V_J) \Bigg\},$$

and for $t = 2, 3, \ldots$, define

$$L^t_J(K) = L_J(L^{t-1}_J(K)).$$

Notice that $K \subseteq L_J(K)$; therefore L^t_J defines an increasing sequence; that is, $L^t_J(K) \subseteq L^{t+1}_J(K)$ for $t \geq 1$. Also, from their explicit constructions, it is easy to see that for $J' \subseteq J$ with $|J'| = t$

$$B_{J'}(K) \subseteq L^t_J(K) \subseteq W^t_J(K).$$

Hence the following result readily follows.

Theorem 8 *Let $J \subseteq \{1,\ldots,n\}$ and D be a compact set. Assume either $D \subseteq \bigcap_{j\in J}\{x \in \mathbb{R}^n : 0 \leq x_j \leq 1\}$ or $m > 1$. If $K \approx \mathcal{P}_{n,m}(D)$, then*

$$K \subseteq L_J(K) \subseteq \cdots \subseteq L^{|J|}_J(K) \approx \mathcal{P}_{n,m}(D \cap V_J).$$

Once again, proceeding as in Section 4.2.1 it can be seen that for the case when $m = 1$ and D is a polyhedron, the operator L_J is the dual counterpart of the operator N in [12].

We can also strengthen the construction of L_J to get an operator that corresponds to the dual counterpart of the operator N^+ in [12]. This can be done by defining L^+_J as follows

$$L^+_J(K) = \{p \in \mathcal{H}_{n,m} : p \equiv q + \varphi \mod I(V_J), q \in L_J(K), \varphi \in \Sigma_{n,m+1}\}. \quad (12)$$

The operator L^+_J clearly dominates L_J, consequently it also satisfies the statement of Theorem 8.

4.3 Other varieties

Although we have concentrated on refining the one-step approximation scheme for 0–1 programs, the same approach can be used to obtain refinements for other affine varieties. For example, given $j \in \{1,\ldots,n\}$, let

$$\widetilde{V}_j = \{x \in \mathbb{R}^n : x_j \in \{-1, 0, 1\}\}.$$

Define $\widetilde{B}_j$ as follows: Given $K \subseteq \mathcal{H}_{n,m}$, let

$$\widetilde{B}_j(K) = \{p \in \mathcal{H}_{n,m} : \exists q_0, q_{-1}, q_1 \in K \text{ s.t. } p(x) \equiv (2(1-x_j)(1+x_j)q_0(x)+ \\ x_j^2(1-x_j)q_{-1}(x) + x_j^2(1+x_j)q_1(x)) \mod I(\widetilde{V}_j)\}.$$

(Notice that $I(\widetilde{V}_j) = \langle x_j - x_j^3 \rangle$.)

Proceeding as in the proof of Theorem 6, we obtain the following result.

Theorem 9 *Let $j \in \{1, \dots, n\}$. Assume D is compact and $m > 1$. If $K \approx \mathcal{P}_{n,m}(D)$, then $\widetilde{B}_j(K) \approx \mathcal{P}_{n,m}(D \cap \widetilde{V}_j)$.*

5 Refining the sequential approximation scheme

We next show that under suitable conditions on D and V, the statement in Theorem 3 can be strengthened to obtain a *finite* sequential approximation $Q_0 \subseteq \cdots \subseteq Q_N \approx \mathcal{P}_{n,m}(D \cap V)$, given a sequential approximation $K_r \uparrow P_n(D)$.

We start with the following corollary of Theorem 3.

Corollary 4 *Assume D is compact. Given $K_r \uparrow \mathcal{P}_n(D)$, let $Q_r = \{p \in \mathcal{H}_{n,m} : p \equiv q \mod I(V), q \in K_r\}$. Then $Q_r \uparrow \mathcal{P}_{n,m}(D \cap V)$.*

We show below that for suitable D, V, and $K_r \uparrow \mathcal{P}_n(D)$, the sequence Q_r, $r \in \mathbb{N}$ in Corollary 4 eventually becomes constant, and in consequence Q_r yields a finite sequential approximation to $\mathcal{P}_{n,m}(D \cap V)$. The crux of these constructions is the specific choice of $K_r \uparrow \mathcal{P}_n(D)$.

5.1 A finite approximation scheme for D polyhedral and $V = \{0,1\}^n$

Our next construction relies on a theorem due to Handelman [5]. We present this result below in a format appropriate for our exposition.

Theorem 10 (Handelman) *Let $a_j \in \mathbb{R}^n$, $b_j \in \mathbb{R}$, $j = 1, \dots, d$ and assume*

$$D = \{x \in \mathbb{R}^n : a_j^{\mathrm{T}} x - b_j \geq 0, j = 1, \dots, d\} \tag{13}$$

is a bounded polyhedron. If $p \in \mathcal{P}_n^o(\mathcal{D})$, then for some positive integer N there exist $\lambda_\alpha \geq 0$, $\alpha \in \mathbb{N}^d$, $\|\alpha\|_1 \leq N$ such that

$$p(x) = \sum_{\alpha \in \mathbb{N}^n, \|\alpha\|_1 \leq N} \lambda_\alpha \prod_{j=1}^{d} (a_j^{\mathrm{T}} x - b_j)^{\alpha_j}.$$

Based on Theorem 10, we can construct $K_r \uparrow \mathcal{P}_n(D\cap[0,1]^n)$ for D polyhedral as follows. Assume $D = \{x \in \mathbb{R}^n : a_j^{\mathrm{T}} x - b_j \geq 0, j = 1, \ldots, d\}$. For $r \in \mathbb{N}$ put

$$K_r = \big\{p \in \mathcal{H}_{n,r} : \exists \lambda_{(\alpha,\beta,\beta')} \geq 0,\ (\alpha,\beta,\beta') \in \Gamma(r) \text{ s.t.}$$
$$p = \sum_{\Gamma(r)} \lambda_{(\alpha,\beta,\beta')} \prod_{j=1}^{d} (a_j^{\mathrm{T}} x - b_j)^{\alpha_j} \prod_{i=1}^{n} x_i^{\beta_i}(1-x_i)^{\beta'_i}\big\}, \quad (14)$$

where $\Gamma(r) = \{(\alpha,\beta,\beta') \in \mathbb{N}^d \times \mathbb{N}^n \times \mathbb{N}^n : ||(\alpha,\beta,\beta')||_1 \leq r\}$. By Theorem 10, $K_r \uparrow \mathcal{P}_{n,m}(D \cap [0,1]^n)$.

Remark 2 As we shall discuss in Section 6, each K_r above is a polyhedral cone.

Theorem 11 *Assume* $D = \{x \in \mathbb{R}^n : a_j^{\mathrm{T}} x - b_j \geq 0, j = 1, \ldots, d\}$ *and* K_r *is as in* (14)*. Let*

$$Q_r = \{p \in \mathcal{H}_{n,m} : \exists q \in K_r \ \text{ s.t. } p \equiv q \mod I(\{0,1\}^n)\}.$$

Then $Q_0 \subseteq Q_1 \subseteq \cdots \subseteq Q_{n+d} \approx \mathcal{P}_{n,m}(D \cap \{0,1\}^n)$.

Proof. Since $K_r \uparrow \mathcal{P}_n(D \cap [0,1]^n)$, by Corollary 4 it suffices to show that $Q_r = Q_{n+d}$ for $r \geq n+d$. This in turn is an immediate consequence of the following observations:

(i) $x_i^2 \equiv x_i \mod I(\{0,1\}^n)$ and $(1-x_i)^2 \equiv (1-x_i) \mod I(\{0,1\}^n)$.

(ii) Let $a \in \mathbb{R}^n, b \in \mathbb{R}$. Then $(a^{\mathrm{T}}x - b)^2 \equiv \sum_{\beta\in\{0,1\}^n} \lambda_\beta \prod_{i=1}^n x_i^{\beta_i}(1-x_i)^{1-\beta_i} \mod I(\{0,1\}^n)$ for some $\lambda_\beta \in \mathbb{R}_+$, $\beta \in \{0,1\}^n$,

both of which follow from the identity

$$f(x) \equiv \sum_{\beta\in\{0,1\}^n} f(\beta) \prod_{i=1}^n x_i^{\beta_i}(1-x_i)^{1-\beta_i} \mod I(\{0,1\}^n).$$

□

5.2 A finite approximation scheme for D semialgebraic and $V = \{0,1\}^n$

Another finite sequential approximation for $\mathcal{P}(D \cap \{0,1\}^n)$ can be obtained by using other representation theorems from real algebraic geometry to construct $K_r \uparrow \mathcal{P}(D)$. Lemma 12 below is one of such representation theorems stated in a format appropriate for our exposition. This result is a direct consequence of [19, Thm. 6.3.4].

Assume $D = \{x \in \mathbb{R}^n : g_j(x) \geq 0, j = 1, \ldots, d\}$ where $g_j \in \mathcal{H}_n, j = 1, \ldots, d$ and either all g_j have even degree or all g_j have odd degree. Assume also that g_j, $j = 1, \ldots, d$ satisfy the following technical condition:

$$\text{For all } x \in \mathbb{R}^n \setminus \{0\} \text{ there exists } i \in \{1, \ldots, d\} \ \text{ s.t. } \tilde{g}_i(x) < 0,$$

where $\tilde{g}_i(x)$ is the homogeneous component of $g_i(x)$ of highest degree. It is easy to see that the conditions above imply that D is compact.

The following lemma is a special case of [19, Thm 6.3.4].

Lemma 12 *Let $D = \{x \in \mathbb{R}^n : g_j(x) \geq 0, j = 1, \dots, d\}$ be such that above conditions hold. If $p \in \mathcal{P}_n^o(D)$, then $p(x) = \varphi_0(x) + \sum_{j=1}^d g_j(x)\varphi_j(x)$, for some $\varphi_j \in \Sigma_n$, $j = 0, \dots, d$.*

Based on this result, we can construct $K_r \uparrow \mathcal{P}_n(D)$ in a number of ways. One possible construction is the following. Let $m_j = \deg(g_j)$ and for $r \in \mathbb{N}$ put

$$K_r := \left\{ p \in \mathcal{H}_{n,2r} : \exists \varphi_0 \in \Sigma_{n,2r}, \varphi_j \in \Sigma_{n,2(r-m_j)} \text{ s.t. } p(x) = \varphi_0(x) + \sum_{j=1}^d g_j(x)\varphi_j(x) \right\}. \quad (15)$$

Remark 3 As we shall discuss in Section 6, each K_r above can be defined in terms of the cone of positive semidefinite matrices.

Theorem 13 *Assume D satisfies the conditions of Lemma 12, and let K_r be as in* (15)*. Let*

$$Q_r = \{p \in \mathcal{H}_{n,m} : p \equiv q \mod I(\{0,1\}^n), q \in K_r\}.$$

Then $Q_0 \subseteq Q_1 \subseteq \cdots \subseteq Q_N \approx \mathcal{P}_{n,m}(D \cap \{0,1\}^n)$ for $N = n + \max\{m_j : j = 1, \dots, d\}$.

Proof. Since $K_r \uparrow \mathcal{P}_n(D)$, by Corollary 4 it suffices to show that $Q_r = Q_N$ for $r \geq N$. This in turn is an immediate consequence of the following observation: if $p \in \Sigma_{n,2m}$ then $p \equiv q \mod I(\{0,1\}^n)$ for some $q \in \Sigma_{n,2n}$. □

In parallel to the connection between the operators B_j, L_J of Section 4 and the operators P_j, N of [1, 12], there is a connection between the sequence Q_r defined in Theorem 13 and the sequence of relaxation for 0–1 programming proposed by Lasserre [8]. Indeed, with a suitable choice of the sequence K_r in (15), it can be shown that the problems $(\mathbb{Q}_r^*)$ in [8] are precisely those obtained when the cone $\mathcal{P}_{n,m}(D \cap \{0,1\}^n)$ is approximated by the sequence Q_r given by Theorem 13 in problem (2). Since each Q_r has an LMI description (see Section 6), then each problem $(\mathbb{Q}_r^*)$ can be cast as a semidefinite program, as Lasserre showed [8].

However, observe that the statement in Theorem 13 applies to any sequence Q_r derived from a given $K_r \uparrow \mathcal{P}_n(D)$, as long as the specific construction of K_r ensures that the sequence Q_r eventually becomes constant. This readily suggests a number of variations of Lasserre's scheme. For example, we could conceivable exploit other algebraic geometry machinery such as the Positivstellensatz (see, e.g., [18]) to construct other sequences $K_r \uparrow \mathcal{P}_n(D)$.

6 Computable descriptions revisited

We now turn our attention to some details concerning the computable description of the cones obtained in our constructions. The two main goals of this section are: First, to show that if the starting cone K (or sequence K_r) has a computable description, then the constructions in Sections 4 and 5 yield cones Q (or sequences Q_r) with computable descriptions. Second, to show that the sequences defined in (14) and in (15) have computable descriptions. More precisely, the first one is a polyhedral cone, and the second one can be described in terms of the positive semidefinite cone.

We begin with some simple observations. Recall that for $p(x) \in \mathcal{H}_{n,m}$, we use p to denote its (finite-dimensional) vector of coefficients under a given ordering of the monomials in $\mathcal{H}_{n,m}$.

Observation 1 *Let $t_1, \dots, t_l \in \mathcal{H}_{n,m'}$ be given. Then there exists a linear map $M : \mathcal{H}_{n,m} \times \cdots \times \mathcal{H}_{n,m} \to \mathcal{H}_{n,m+m'}$ such that*

$$p(x) = \sum_{j=1}^{l} t_j(x) q_j(x),\ q_j \in \mathcal{H}_{n,m} \text{ if and only if } p = M(q_1, \dots, q_l).$$

Observation 1 shows that the cones K_r, $r \in \mathbb{N}$ defined in (14) have a computable description; and in particular, that each K_r is a polyhedral cone.

Observation 2 *Assume n, m, m' are positive integers, $p, q \in \mathcal{H}_{n,m}$, $f \in \mathcal{H}_{n,m'}$ and $I = \langle f \rangle$. Then*

$$p \equiv q \mod I \text{ if and only if } p(x) = q(x) + f(x)h(x) \text{ for some } h \in \mathcal{H}_{n,m}.$$

Since $I(V_{a,b}) = \langle a^{\mathrm{T}}x - b \rangle$, $I(V_j) = \langle x_j - x_j^2) \rangle$, and $I(\widetilde{V}_j) = \langle x_j - x_j^3 \rangle$, Observations 1 and 2 show that if a cone K has a computable description, then so do the cones $\mathcal{Z}_{a,b}(K)$, $B_j(K)$, and $\widetilde{B}_j(K)$.

Observation 3 *Assume n, m are positive integers and $p, q \in \mathcal{H}_{n,m}$. Let $J \subseteq \{1, \dots, n\}$ and $I = I(V_J) = \langle x_j(x_j - 1) : j \in J \rangle$. Then*

$$p \equiv q \mod I \text{ if and only if } \tilde{p}(x) = \tilde{q}(x),$$

where $t \mapsto \tilde{t}$ is the map defined by putting $\tilde{t}(x) = \sum_{\alpha \in \mathbb{N}^n : |\alpha| \le m} t_\alpha x^{\tilde{\alpha}}$ for

$$\tilde{\alpha}_i = \begin{cases} 1 & \text{if } \alpha_i \ge 1 \\ 0 & \text{otherwise,} \end{cases} \quad i = 1, \dots, n.$$

Observation 1 and 3 show that if a cone K has a computable description, then so do the cones $W_J^t(K)$ and $L_J(K)$. To show that the same property is satisfied by the remaining constructions in Sections 4 and 5 (i.e., those in (12), and in Theorems 11 and 13), as well as to show that each cone in the sequence (15) has a computable description, it suffices to show that that $\Sigma_{n,2m}$ has a computable description. This fact readily follows from the following observation.

Observation 4 *Let $p \in \mathcal{H}_{n,2m}$. Then*

$$p \in \Sigma_{n,2m} \Leftrightarrow \exists \Phi \succeq 0 \text{ s.t. } p = \mathcal{L}\Phi,$$

for a suitable linear map $\mathcal{L}$ (see, e.g., [16, 27]).

References

[1] E. Balas, S. Ceria, and G. Cornuéjols, *A lift-and-project cutting plane algorithm for mixed* 0-1 *programs*, Math. Program., 58:295-324, 1993.

[2] S. Ceria, *Lift-and-Project Methods for Mixed 0–1 Programs,* Ph.D. Dissertation, Carnegie Mellon University, 1993.

[3] D. Cox, J. Little and D. O'Shea, *Ideals, varieties, and algorithms. An introduction to computational algebraic geometry and commutative algebra*, second edition, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1998.

[4] E. de Klerk and D. Pasechnik, *Approximating the stability number of a graph via copositive programming*, SIAM J. Optim., 12(4):875-892, 2002.

[5] D. Handelman, *Representing polynomials by positive linear functions on compact convex polyhedra*, Pacific J. Math., 132:35-62, 1988.

[6] G. Iyengar and M. T. Cezik, *lift-and-project cutting planes for mixed 0-1 semidefinite programming*, working paper, 2001.

[7] M. Kojima, S. Kim and H. Waki, *A General Framework for Convex Relaxation of Polynomial Optimization Problems over Cones,* J. Oper. Res. Soc. Japan 46 (2003) 125–144.

[8] J. B. Lasserre, *An explicit equivalent positive semidefinite program for nonlinear 0-1 programs*, SIAM J. Optim., 12(3):756-769, 2001.

[9] J. B. Lasserre, *Bounds on Measures Satisfying Moment Conditions*, Ann. Appl. Probab., 12:1114–1137, 2002.

[10] J. B. Lasserre, *Global Optimization Problems with Polynomials and the Problem of Moments*, SIAM J. Optim., 11(3):796-817, 2001.

[11] M. Laurent, *A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for* 0-1 *programming*, Math. Oper. Res., 28(3):470–496, 2003.

[12] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and* 0-1 *optimization*, SIAM J. Optim., 1(2):166-190, 1991.

[13] Y. Nesterov, *Structure of non-negative polynomials and optimization problems*, CORE Discussion Paper No. 9749, 1997.

[14] Y. Nesterov and M. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Math. Oper. Res., 22:1-42, 1997.

[15] Y. Nesterov and M. Todd, *Primal-dual interior-point methods for self-scaled cones*, SIAM J. Optim., 8(2):324-364, 1998.

[16] P. Parrilo, *Structured semidefinite programs and semi-algebraic geometry methods in robustness and optimization*, Ph.D. Dissertation, California Institute of Technology, Pasadena, CA, 2000.

[17] P. Parrilo and B. Sturmfels, *Minimizing Polynomial Functions*, In Algorithmic and quantitative real algebraic geometry , DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 60:83–99, 2003.

[18] P. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program. B, 96(2):293-320, 2003.

[19] A. Prestel and C. Delzell, *Positive Polynomials: From Hilbert's 17th Problem to Real Algebra*, Springer Verlag, 2001.

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

[21] H. D. Sherali and W. P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, Kluwer Academic Publishers, 1998.

[22] N. Z. Shor, *Class of global minimum bounds of polynomial functions*, Cybernetics, 23(6):731-734, 1987.

[23] N. Z. Shor and P. I. Stetsyuk, *The use of a modification of the r-algorithm for finding the global minimum of polynomial functions*, Cybernet. Systems Anal., 33:482-497, 1997.

[24] J. S. Sturm and S. Zhang, *On Cones of Non-negative Quadratic Functions*, Math. Oper. Res., 28:246–267, 2003.

[25] M. Tawarmalani and N. V. Sahinidis, *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming. Theory, Algorithms, Software, and Applications* , Kluwer Academic Publishers , 2002.

[26] V. A. Yakubovich, *S-procedure in nonlinear control theory*, Vestnik Leninggradskogo Universiteta, Ser. Matematika, 62–77, 1971.

[27] L. Zuluaga, J. C. Vera and J. Peña, *LMI approximations for cones of positive semidefinite forms*, submitted to SIAM J. of Optim., available at `http://www.optimization-online.org/DB_HTML/2003/05/652.html`.