POSITIVSTELLENSATZ CERTIFICATES FOR CONTAINMENT OF 
\(H\)-POLYTOPES IN \(V\)-POLYTOPES

KAI KELLNER AND THORSTEN THEOBALD

Abstract. Given an \(H\)-polytope \(P\) and a \(V\)-polytope \(Q\), the decision problem whether \(P\) is contained in \(Q\) is co-NP-complete. This hardness remains if \(P\) is restricted to be a standard cube and \(Q\) is restricted to be the affine image of a cross polytope. While this hardness classification by Freund and Orlin dates back to 1985, there seems to be only limited progress on that problem so far.

Here, we study the \(H\)-in-\(V\) containment problem from the viewpoint of a bilinear feasibility problem and in connection with linear and semidefinite relaxations. Handelman’s and Putinar’s Positivstellensatz yield hierarchies of linear programs and semidefinite programs, respectively, to decide the containment problem. We study their geometric properties and Positivstellensatz certificates.

As a main result, we show that under mild and explicitly known preconditions the semidefinite hierarchy converges in finitely many steps. In particular, this is the case if \(P\) is the standard cube and \(Q\) is the standard cross polytope.

1. Introduction

Polytopes can be represented as the convex hull of finitely many points ("\(V\)-polytopes") or as the intersection of finitely many halfspaces ("\(H\)-polytopes"). For \(a \in \mathbb{R}^k\), \(A \in \mathbb{R}^{k \times d}\), and \(B = [b^{(1)}, \ldots, b^{(l)}] \in \mathbb{R}^{d \times l}\) let
\[
P = P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\} \quad \text{and} \quad Q = Q_B = \text{conv}(B) = \text{conv}(b^{(1)}, \ldots, b^{(l)})
\]
be an \(H\)-polytope and a \(V\)-polytope, respectively. The subscript in the notion of \(P\) and \(Q\) indicates the dependency on the specific representation of the polytopes involved. However, if there is no risk of confusion, we often state \(P\) and \(Q\) without subscript. The following two problems are prominent problems in algorithmic polytope theory (see Kaibel and Pfetsch [15]). We always assume that the polytope data is given in terms of rational numbers.

Polytope verification:
\textbf{Input:} \(d \in \mathbb{N}\), an \(H\)-polytope \(P \subseteq \mathbb{R}^d\) and a \(V\)-polytope \(Q \subseteq \mathbb{R}^d\).
\textbf{Task:} Decide whether \(P = Q\).

Polytope containment (or \(H\)-in-\(V\) containment):
\textbf{Input:} \(d \in \mathbb{N}\), an \(H\)-polytope \(P \subseteq \mathbb{R}^d\) and a \(V\)-polytope \(Q \subseteq \mathbb{R}^d\).
\textbf{Task:} Decide whether \(P \subseteq Q\).

While the complexity status of the first problem is open, the second problem is co-NP-complete (see Freund and Orlin [6]); note that it is trivial to decide the converse question \(Q \subseteq P\). It is well known that the problem of enumerating all facets of a polytope given by a finite set of points (or, equivalently, enumerating all vertices of a polytope given by a
finite number of halfspaces) can be polynomially reduced to the Polytope verification problem (see Avis et. al. [2], Kaibel and Pfetsch [15]). Note that enumerating the vertices of an (unbounded) polyhedron is hard [18]. Moreover, Joswig and Ziegler [14] showed that the Polytope verification problem is polynomially equivalent to a geometric polytope completeness problem. For the Polytope Containment problem, however, despite of its fundamental nature, there seems to be only limited progress on that problem so far.

In this paper we study the Polytope Containment problem. Our main focus is to consider the problem from the viewpoint of the transition from linear/polyhedral problems to low-degree semialgebraic problems. To that end, we formulate the problem as a disjointly constrained bilinear feasibility problem and consider semialgebraic Positivstellensatz relaxations.

The main idea – which is meanwhile common in polynomial optimization, but whose understanding of the particular potential on low-degree geometric problems is still a major challenge – can be explained as follows. One point of view towards linear programming is as an application of Farkas’ Lemma which characterizes the (non-)solvability of a system of linear inequalities. The affine form of Farkas’ Lemma [34, Corollary 7.1h] characterizes linear polynomials which are nonnegative on a given polyhedron. By omitting the linearity condition, one gets a polynomial nonnegativity question, leading to so called Positivstellensätze (or, more precisely, Nichtnegativstellensätze). These Positivstellensätze provide a certificate for the positivity of a polynomial function, in terms of a polynomial identity. As in the linear case, the Positivstellensätze are the foundation of polynomial optimization and relaxation methods (see [21, 23, 24]).

In our context, using an optimization version of the bilinear formulation of the Polytope Containment problem allows to apply linear relaxations based on Handelman’s Positivstellensatz [12] respectively semidefinite relaxations based on Putinar’s Positivstellensatz [33]; see also Laurent’s extensive survey [24] on these techniques. First we treat Handelman’s Positivstellensatz which deals with positivity of polynomials on polytopes and provides hierarchies of linear programs. Subsequently, we consider Putinar’s Positivstellensatz. That theorem deals with more general, semialgebraic constraint sets and it provides hierarchies of semidefinite programs. The general machinery from polynomial optimization automatically implies convergence results, but often these results come with restrictions or technical assumptions. A main purpose of the present paper is to initiate further detailed research on the capabilities of semialgebraic relaxations on low-degree, combinatorial problems, such as (non-linear) problems of polyhedra.

**Our contributions:**

1. Based on a formulation of the Polytope containment problem in terms of a polynomial (in fact, bilinear) feasibility problem (Proposition 3.1), we characterize geometric properties of a natural bilinear programming reformulation; see Lemma 3.5.

2. We study linear relaxations of the bilinear programming formulation, as based on Handelman’s Positivstellensatz. Beyond a standard convergence result (Theorem 4.2), we provide some characterizations on the (widely open) question of degree bounds, including some explicit Handelman certificates in certain specific symmetric examples (such as cubes and cross polytopes). Based on various geometric properties of the relaxation
scheme in the $\mathcal{H}$-in-$\mathcal{V}$-setting, we provide a necessary criterion for the set of $\mathcal{H}$-polytopes $P$ whose containment in a fixed $\mathcal{V}$-polytope $Q$ is certified by the $t$-th relaxation step of the Handelman hierarchy (Theorem 4.5).

3. We study the application of Putinar’s Positivstellensatz on the bilinear programming formulation. For a fixed $\mathcal{V}$-polytope $Q$, we provide a necessary characterization for the set of $\mathcal{H}$-polytopes $P$ whose containment in $Q$ is certified in the $t$-th relaxation step (Theorem 5.7). This characterization is given in terms of a projection of a spectrahedron.

4. An important point is whether the hierarchies always converge in finitely many steps. While in the case of strict containment, this property comes out from the general theory (Theorem 5.3), in the case of non-strict containment this is a critical issue. As a main result of this paper, we can give a positive answer for Putinar’s Positivstellensatz. We show that under mild and explicitly known conditions, the Putinar relaxation converges in finitely many steps (Theorem 5.8), based on recent results by Nie [28]. For the situation of Handelman’s Positivstellensatz the question remains open.

While in this paper, we concentrate on the main problem of deciding whether an $\mathcal{H}$-polytope is contained in a $\mathcal{V}$-polytope, let us briefly mention some related problems. Finding the largest simplex in a $\mathcal{V}$-polytope is an NP-hard problem [11]. However for that problem Packer has given a polynomial-time approximation [30]. Recently, Gouveia et. al. have studied the question which nonnegative matrices are slack matrices [9], and they establish equivalence of the decision problem to the polyhedral verification problem. For containment of polytopes and spectrahedra see [16, 17].

The paper is structured as follows. After introducing the relevant notation and presenting some application examples in Section 2 we study geometric properties of a natural bilinear programming formulation in Section 3. In Section 4 we discuss linear relaxations of the Polytope containment problem, as based on Handelman’s Positivstellensatz. Section 5 deals with semidefinite relaxations, as based on Putinar’s Positivstellensatz. Finally, Section 6 lists several open questions.

2. Preliminaries

We recall relevant notation, existing computational results on the containment of $\mathcal{H}$-polytopes in $\mathcal{V}$-polytopes, and sketch some application scenarios.

2.1. Polyhedra and polytopes. A polyhedron $P$ is the intersection of finitely many affine halfspaces in $\mathbb{R}^d$, and a bounded polyhedron is called a polytope. As a general reference we refer to Ziegler’s book [35]. Denote by $V(P)$ the set of vertices of a polytope $P$, and by $F(P)$ the set of facets. By McMullen’s Upper bound Theorem [26], any $d$-polytope with $k$ vertices (resp. facets) has at most

$$\left( k - \left\lfloor \frac{1}{2}(d+1) \right\rfloor \right) + \left( k - \left\lfloor \frac{1}{2}(d+2) \right\rfloor \right)$$

facets (resp. vertices). This bound is sharp for neighborly polytopes, such as cyclic polytopes.
If \( P \subseteq \mathbb{R}^d \) and \( Q \subseteq \mathbb{R}^e \) are polytopes, then their product \( P \times Q \subseteq \mathbb{R}^{d+e} \) is a polytope of dimension \( \dim P + \dim Q \), whose nonempty faces are the products of nonempty faces of \( P \) and nonempty faces of \( Q \); see for example [35, Page 10].

2.2. Containment of polytopes in polytopes. It is well-known that the complexity of deciding containment of one polytope in another one strongly depends on the type of input representations (see [6] and [10, Theorems 2.3 and 2.4]).

Proposition 2.1. Deciding whether a polytope \( P \) is contained in a polytope \( Q \) can be done in polynomial time in the following cases:

1. Both \( P \) and \( Q \) are \( \mathcal{H} \)-polytopes,
2. both \( P \) and \( Q \) are \( \mathcal{V} \)-polytopes, or
3. \( P \) is a \( \mathcal{V} \)-polytope while \( Q \) is an \( \mathcal{H} \)-polytope.

However, deciding whether an \( \mathcal{H} \)-polytope is contained in a \( \mathcal{V} \)-polytope is co-NP-complete. This hardness remains if \( P \) is restricted to be a standard cube and \( Q \) is restricted to be the affine image of a cross polytope.

If the dimension is fixed, then the problem of deciding whether an \( \mathcal{H} \)-polytope is contained in a \( \mathcal{V} \)-polytope can be decided in polynomial time.

Note that the result for fixed dimension can be strengthened slightly.

Theorem 2.2. If the dimension of \( P \) or the dimension of \( Q \) is fixed, then containment of an \( \mathcal{H} \)-polytope \( P \) in a \( \mathcal{V} \)-polytope \( Q \) can be decided in polynomial time.

Proof. If the dimension of \( P \) is fixed, then first compute the affine hull of \( P \). This can be done in polynomial time. Taking that affine hull as ambient space in fixed dimension, \( P \) can be transformed into a \( \mathcal{V} \)-representation in polynomial time. It remains to decide containment of a \( \mathcal{V} \)-polytope in a \( \mathcal{V} \)-polytope, which can be done in polynomial time.

Similarly, if the dimension of \( Q \) is fixed, an \( \mathcal{H} \)-representation of \( Q \) can be computed in polynomial time, and the resulting problem of deciding whether an \( \mathcal{H} \)-polytope is contained in an \( \mathcal{H} \)-polytope can be decided in polynomial time. \( \square \)

2.3. Application examples. While it is a fundamental geometric problem by itself, we sketch some exemplary application scenarios in which the POLYTOPE CONTAINMENT problem occurs. Generally, many applications in data analysis or shape analysis of point clouds involve the convex hull of point sets (see, e.g., [4]), and a POLYTOPE CONTAINMENT problem can be used to answer questions about certain (polyhedral) properties on the set.

A textbook type example in diet realization. We consider \( d \) different food types and \( k \) underlying basic nutrients. Assume that each unit of food \( j \) contains \( A_{ij} \) units of the \( i \)-th nutrient \( a_i \). A dietary requirement is described by linear inequalities of the form

\[
\sum_{j=1}^{d} A_{ij} x_j \geq a_i, \quad 1 \leq i \leq k,
\]

with positive \( A_{i1}, \ldots, A_{id} \) for a minimum of a requirement and negative \( A_{i1}, \ldots, A_{id} \) for a maximum of a requirement. Thus the requirements define an \( \mathcal{H} \)-polytope \( P \).
Moreover, assume we are given \( l \) fixed combinations of can food, each can containing one unit. Type \( s \) consists of an amount of \( b^{(s)} \) of food \( t \). The convex combinations of the vectors \( b^{(1)}, \ldots, b^{(l)} \) correspond to the food which can be combined from the can food, where the convex combination signifies that the resulting food has also the size of one unit. The question is if every food combination satisfying the dietary constraints can be assembled from the can food in that way. This is a Polytope containment problem.

Similar scenarios occur, e.g., in the mixing of liquids (such as oil).

**Theorem proving in linear real arithmetic.** A sub-branch in theorem proving is based on formulas in linear real arithmetic; see, e.g., [5, 27]. Given a set of (quantifier-free) linear inequalities of the form

\[
L_i(x_1, \ldots, x_d) \geq 0
\]

in the real variables \( x_1, \ldots, x_d \) specifying the assumptions of a certain theorem, one may ask whether all these solutions satisfy a certain property \( Q \). If \( Q \) is described as the convex hull of a finite number of points, then the theorem proving problem corresponds to a Polytope containment problem.

3. A bilinear programming approach to the Polytope containment problem

We first collect some geometric properties of the Polytope containment problem. Here and throughout the paper, we assume boundedness and nonemptiness of \( P \), which both can be tested in polynomial time [19]. W.l.o.g. we assume \( 0 \in Q \). If this is not the case, one can translate \( Q \) and \( P \) by the centroid of the vertices of \( Q \). The centroid is clearly contained in \( Q \), and it is an interior point of \( Q \) whenever \( Q \) has an interior point. Recall that the polar polyhedron of \( Q \) is

\[
Q^o = \{ z \in \mathbb{R}^d \mid 1_l - B^T z \geq 0 \}
\]

where \( 1_l \) is the all-1-vector in \( \mathbb{R}^l \). Our starting point is the following reformulation of the Polytope containment problem as a bilinear feasibility problem.

**Proposition 3.1.** Let \( a \in \mathbb{R}^k \), \( A \in \mathbb{R}^{k \times d} \), and \( B = [b^{(1)}, \ldots, b^{(l)}] \in \mathbb{R}^{d \times l} \) such that the \( H \)-polytope \( P = P_A = \{ x \in \mathbb{R}^d \mid a - Ax \geq 0 \} \) is nonempty and the \( V \)-polytope \( Q = Q_B = \text{conv}(B) = \text{conv}(b^{(1)}, \ldots, b^{(l)}) \) contains the origin.

1. \( P \) is contained in \( Q \) if and only if

\[
x^T z \leq 1 \quad \text{for all } (x, z) \in P \times Q^o.
\]

2. We have

\[
\sup \{ x^T z \mid (x, z) \in P \times Q^o \} = 1
\]

if and only if \( P \subseteq Q \) and \( \partial P \cap \partial Q \neq \emptyset \).

**Proof.** To (1): If \( P \subseteq Q \) then, clearly, for any \( x \in P \) we have \( x^T z \leq 1 \) for all \( z \in Q^o \). Conversely, if \( x^T z \leq 1 \) holds for all \( z \in Q^o \), then for any \( x \in P \) we have \( P \subseteq Q^{oo} = Q \).

To (2): Let \( P \subseteq Q \) and \( \partial P \cap \partial Q \) be nonempty. Then there exists a vertex \( v \in V(P) \) and a facet \( F \in F(Q) \) such that \( v \in F \). \( F \) defines a vertex \( f \) of the polar \( Q^o \). Further, \( f^T v = 1 \) implies that the supremum is at least one. By part (1) of the statement, the supremum must be exactly one.
Conversely, if the supremum is one, then \( x^Tz \leq 1 \) for all \((x,z) \in P \times Q^\circ\) and since the set \( P \times Q^\circ \) is closed, there exists a point \((\bar{x}, \bar{z}) \in P \times Q^\circ\) such that \( x^T\bar{z} = 1 \). Hence \( \bar{x}^Tz \leq 1 \) for all \( z \in Q^\circ \) and \( \bar{x}^T\bar{z} = 1 \), i.e., \( \bar{x} \) defines a supporting hyperplane of \( Q^\circ \). Thus \( \bar{x} \) is a boundary point of \( Q \). Similarly, \( x^T\bar{z} \leq 1 \) for all \( x \in P \) and \( \bar{x}^T\bar{z} = 1 \), implying \( \bar{x} \in \partial P \). Consequently, \( \bar{x} \in \partial Q \cap \partial P \). \( \square \)

We record the following slightly more general version for the case that the precondition 0 \( \in \text{int} \, Q \) is not satisfied.

**Corollary 3.2.** Let \( Q_B = \text{conv}(B) \) be an irredundant representation and let the interior of \( Q_B \) be nonempty. Denote by \( c = \frac{1}{d} \sum_{i=1}^d b^{(i)} \) the centroid of \( Q_B \). Then \( P_A \subseteq Q_B \) if and only if

\[
x^Tz \leq 1 \quad \text{for all} \quad (x, z) \in (P_A - c) \times Q^\circ_{B-c},
\]

where \( P_A - c = \{ x \in \mathbb{R}^d \mid a - A(x + c) \geq 0 \} \) and \( B - c = B - c\mathbb{1}_{d \times 1} \).

Proposition 3.1 suggests to formulate the **Polytope containment** problem via a disjointly constrained bilinear program

\[
\mu^* := \sup \quad x^Tz \\
\quad \text{s.t.} \quad (x, z) \in P \times Q^\circ.
\]

Since 0 \( \in Q^\circ \), the supremum is nonnegative whenever \( P \) is nonempty. By Proposition 3.1, \( P \subseteq Q \) if and only if the supremum of (3.1) is at most 1. The following characterization of the optimal solutions is shown in Konno’s work [20] on bilinear programming.

**Proposition 3.3.** Let \( P \) be a nonempty polytope. If the origin is an interior point of \( Q \), then the set of optimal solutions is contained in the boundary \( \partial P \times \partial Q^\circ \), and the supremum is finite and attained at a pair of vertices of \( P \) and \( Q^\circ \).

For the convenience of the reader, we recall the short proof.

**Proof.** The objective function attains its maximum since it is continuous and the feasible region is nonempty, compact by 0 \( \in \text{int} \, Q \).

Let \((\bar{x}, \bar{z}) \in P \times Q^\circ\) be an optimal solution. The linear program \( \max\{x^Tz \mid z \in Q^\circ\}\) has a finite optimal value since its feasible region is nonempty and bounded. Hence there exists a vertex \( \bar{z} \) of \( Q^\circ \) at which the optimal value of the LP is attained. Since \( \bar{z} \in Q^\circ \), we get \( \bar{x}^T\bar{z} \geq \bar{x}^T\bar{z} \). Analogously, there exists a vertex \( \hat{x} \in P \) at which the optimal value of the LP \( \max\{x^Tz \mid x \in P\} \) is attained. Consequently, \( \hat{x}^T\hat{z} \geq \bar{x}^T\bar{z} \geq \bar{x}^T\bar{z} \) and, by the optimality of \((\bar{x}, \bar{z})\), the point \((\hat{x}, \bar{z})\) is an optimal solution for (3.1). \( \square \)

There is a nice geometric interpretation of the latter proposition. Since, in the case 0 \( \in \text{int} \, Q \), each vertex of \( Q^\circ \) corresponds to a facet of \( Q \) and vice versa, an optimal solution \((x, z) \in V(P) \times V(Q^\circ)\) of (3.1) yields a pair of a vertex of \( P \) and a facet defining normal vector of \( Q \) which either certify containment or non-containment. However, since computing the set of vertices \( V(Q^\circ) \) is an NP-hard problem, it is not reasonable to reduce the problem to the set of vertices, in general.
The (sign-)oriented euclidean distance of a point \(v\) and a hyperplane \(H\) given by \(H = \{x \mid 1 - h^T x = 0\}\) is defined as
\[
\text{dist}(v, H) = \frac{1 - h^T v}{\|h\|}.
\]
If the hyperplane defines a facet of a certain polytope, we call this the distance between the point and the facet. Given two polytopes \(P\) and \(Q\), we denote the minimum oriented distance of the vertices of \(P\) and the facets of \(Q\) as
\[
d(P, Q) := \min\{\text{dist}(v, F) \mid x \in V(P), \ F \in F(Q)\}.
\]

**Corollary 3.4.** Let the origin be an interior point of \(Q\). Denote by \((\bar{x}, \bar{z}) \in V(P) \times V(Q^*)\) an optimal solution of (3.1). Then the minimum (sign-)oriented euclidean distance between the vertices of \(P\) and the facets of \(Q\) is given by the distance of \(\bar{x}\) and \(F_{\bar{z}} = Q \cap \{x \in \mathbb{R}^d \mid 1 - \bar{z}^T x = 0\}\), i.e., \(d(P, Q) = \text{dist}(\bar{x}, F_{\bar{z}}) = \frac{1 - \bar{z}^T \bar{x}}{\|\bar{z}\|}\).

**Proof.** Let \((v, f) \in V(P) \times V(Q^*)\). Since \(0 \in \text{int } Q\), \(f\) defines a facet \(F = Q \cap \{x \in \mathbb{R}^d \mid 1 - f^T x = 0\}\) of \(Q\). The oriented euclidean distance \(\text{dist}(v, F)\) of \(v\) and \(F\) is \(\text{dist}(v, F) = \frac{1 - f^T v}{\|f\|}\).

For a given \(F\), the distance has a nonnegative value if and only if \(v\) is contained in the positive halfspace, i.e., \(v \in \{x \mid 1 - f^T v \geq 0\}\).

By Proposition 3.1, \(P\) is contained in \(Q\) if and only if \(v^T f \leq 1\) for every \((v, f) \in V(P) \times V(Q^*)\). Thus \(P \subseteq Q\) if and only if the oriented euclidean distance \(\text{dist}(v, F)\) is nonnegative for all vertices of \(P\) and all facets of \(Q\).

Consider an optimal solution \((\bar{x}, \bar{z}) \in V(P) \times V(Q^*)\), which exists by Proposition 3.3. Then the distance \(\text{dist}(\bar{x}, F_{\bar{z}}) = \frac{1 - \bar{z}^T \bar{x}}{\|\bar{z}\|}\) is minimal over the set of vertices \(V(P)\) and the set of facets \(F(Q)\).

The optimal value of problem (3.1) might be attained by other boundary points than vertices, and, moreover, there might be infinitely many optimal solutions. From a geometric point of view, this only occurs in somewhat degenerate cases.

**Lemma 3.5.** Let the interior of \(P\) be nonempty and let \(0 \in \text{int } Q\). The following statements are equivalent.

1. Problem (3.1) has finitely many optimal solutions.
2. Every optimal solution of Problem (3.1) is a pair of vertices of \(P\) and \(Q^*\).
3. Let \(\alpha > 0\) be the unique minimal factor so that \(\partial P \cap \partial(\alpha Q) \neq \emptyset\). \(\partial P \cap \partial(\alpha Q)\) is zero-dimensional and every \(v \in \partial P \cap \partial(\alpha Q)\) lies in the relative interior of a facet of \(\alpha Q\).
4. Let \(\alpha > 0\) be the unique minimal factor so that \(\partial P \cap \partial(\alpha Q) \neq \emptyset\). \(\partial P \cap \partial(\alpha Q)\) is zero-dimensional and for every \(v \in \partial P \cap \partial(\alpha Q)\) the outer normal cone of \(v\) with respect to \(\alpha Q\) is 1-dimensional.

**Proof.** We consider the equivalence of the first two statements. Clearly, if the set of optimal solutions is a subset of \(V(P) \times V(Q)\), then there are only finitely many solutions. For the converse, assume there exists an optimal solution \((\bar{x}, \bar{z}) \in P \times Q^*\) which is not a pair of vertices. Since \(\bar{z}\) is an optimal solution of the LP \(\max\{\bar{z}^T z \mid z \in Q^*\}\), every point in
the minimal face $F(\bar{z})$ of $Q^o$ containing $\bar{z}$ is an optimal solution of the LP with the same optimal value $\mu^* = \bar{x}^T \bar{z}$. Thus the number of optimal solutions is unbounded.

Equivalence of the other statements can be shown in a similar way and is left to the reader. □

**Remark 3.6.** To some extent all in this section holds for the unbounded case as well, i.e., given a nonempty $\mathcal{H}$-presented polyhedron $P$ and a $\mathcal{V}$-presented polyhedron $Q = \text{conv}(B) + \text{cone}(C)$ containing the origin, $P$ is contained in $Q$ if and only if

$$\sup\{x^T z \mid x \in P, \ 1 - B^T z \geq 0, -C^T z \geq 0\} \leq 1.$$ 

Proposition 3.3 indicates to consider vertex tracking algorithms. There is rich literature of cutting plane and branch-and-bound algorithms for bilinear programming following this approach; see, e.g., [7, 20]. Unfortunately, for bilinear programming problems Lagrange duality fails. So far, no converging algorithm is known based on this approach. In the subsequent sections, we study the bilinear reformulation of the POLYTOPE CONTAINMENT from the viewpoint of algebraic certificates and (linear and semidefinite) relaxations.

4. **Linear relaxations based on Handelman’s Positivstellensatz**

In this section, we discuss LP-relaxations for POLYTOPE CONTAINMENT, based on Handelman’s Positivstellensatz

4.1. **Handelman’s Positivstellensatz.** For a polytope $P = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$, define the cone

$$(4.1) \qquad H(P) = \left\{ p \in \mathbb{R}[x] \mid p = \sum_{\alpha \in \mathbb{N}^k} c_\alpha \prod_{j=1}^k (a - Ax)^{\alpha_j}, \ c_\alpha \in \mathbb{R}_+ \right\}$$

in the ring $\mathbb{R}[x]$ of real polynomials in $x = (x_1, \ldots, x_d)$. Here we assume implicitly that only finitely many scalar multipliers $c_\alpha$ are nonzero.

**Proposition 4.1** (Handelman’s Positivstellensatz [12]). Given a nonempty polytope $P = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$, a polynomial $f = f(x) \in \mathbb{R}[x]$ is positive on $P$ if and only if $f \in H(P)$.

Handelman’s original statement deals with subfields of the real numbers. This, in particular, implies that whenever $a \in \mathbb{Q}^k$, $A \in \mathbb{Q}^{k \times d}$, and $f \in \mathbb{Q}[x]$ are rational, then there is a (rational) certificate for membership of $f$ in the cone $H(P) \subseteq \mathbb{Q}[x]$ and thus for positivity of $f$ on $P$.

We denote by $H_t(P)$ the truncated cone

$$H_t(P) := \left\{ p \in \mathbb{R}[x] \mid p = \sum_{\alpha \in \mathbb{N}^k, |\alpha| \leq t} c_\alpha \prod_{j=1}^k (a - Ax)^{\alpha_j}, \ c_\alpha \in \mathbb{R}_+ \right\},$$

where $|\alpha| := \alpha_1 + \cdots + \alpha_k$. Observe that membership of a polynomial to some fixed $H_t(P)$ can be decided by comparing coefficients, which is a linear programming problem. For the
maximization problem \( \max \{ f(x) \mid x \in P \} \), this leads to a hierarchy of linear programming relaxations

\[
\max \{ f(x) \mid x \in P \} = \min \{ \nu \mid \nu - f \in H(P) \} \leq \min \{ \nu \mid \nu - f \in H_t(P) \}.
\]

Since in the latter program all information is encoded in the coefficients of the polynomials involved, it reduces to a (finite dimensional) linear program. Clearly, it does not make sense to consider a relaxation order \( t \) less than the degree of \( f \). Thus we always assume \( t \geq \deg(f) \) and call \( t = \deg(f) \) the initial relaxation order.

There are some advantages and disadvantages of the LP-approach based on Handelman’s Theorem (as compared, e.g., with the semidefinite approach in the subsequent section). The relaxation leads to a hierarchy of linear programs (rather than semidefinite programs). Since, both in theoretical issues (such as exactness of duality theory) and in practical speed, linear programming has advantages compared to semidefinite programming, this makes the Handelman approach attractive.

On the other hand, if for one of the global maximizers the set of active constraints is empty, i.e., a global maximizer lies in the interior of the polytope, then no Handelman representation exists; see [22, Theorem 3.1]. Since in our specific problem, all maximizers are part of the boundary (by Proposition 3.1), this obstacle does not occur in the \( \mathcal{H}\)-in-\( \mathcal{V} \) setting. However, because of the large binomial coefficients involved, the LP relaxation is ill conditioned in general. In practice, setting up the problem, i.e., extracting the linear program from the input, takes a lot more time than actually solving the resulting LP.

4.2. Handelman certificates for Polytope containment. We study certificates coming from Handelman’s Positivstellensatz. To keep notation simple, we denote the (truncated) cone generated by the linear constraints \( a - Ax \) and \( 1 - B^T z \) by \( H_t(P, Q) \). The Handelman relaxation for the bilinear program (3.1) is

\[
\nu(t) = \min \{ \nu \mid \nu - x^T z \in H_t(P, Q) \}.
\]

Since the objective function \( x^T z \) has degree two, the initial relaxation step is \( t = 2 \).

Asymptotic convergence of the relaxation in the general case and finite convergence in the strict containment case are direct consequences of Handelman’s Positivstellensatz.

**Theorem 4.2.** Let \( P \) be an \( \mathcal{H} \)-polytope and \( Q \) be a \( \mathcal{V} \)-polytope with \( 0 \in \text{int } Q \).

1. If \( \nu(t) \leq 1 \) for some integer \( t \geq 2 \), then \( P \subseteq Q \).
2. The relaxation (4.2) converges asymptotically from above to the optimal value \( \mu^* \) of problem (3.1).
3. If \( P \) is strictly contained in \( Q \), i.e., \( P \subseteq Q \) and \( \partial P \cap \partial Q = \emptyset \), then the relaxation (4.2) converges in finitely many steps to the optimal value of problem (3.1), i.e., there exists an integer \( t \geq 2 \) such that \( \nu(t) = \mu^* \).

**Proof.** The first statement is clear by construction of the relaxation.

For the third statement, let \( P \) be strictly contained in \( Q \). Then the optimal value \( \mu^* \) of problem (3.1) is less than one by part (2) of Proposition 3.1 and thus the polynomial \( 1 - x^T z \) is positive on \( P \times Q^c \). Handelman’s Positivstellensatz [4.1] implies the claim.

The second statement follows from part (3) by blowing up \( Q \) such that \( P \subseteq Q \) and \( \partial P \cap \partial Q = \emptyset \).
In the following, we deduce a necessary condition for the set of $\mathcal{H}$-polytopes whose containment in a fixed $\mathcal{V}$-polytope is certified by the $t$-th relaxation step of the Handelman hierarchy (see Theorem 4.5 below). Before proving this criterion, we collect some relevant properties of relaxation [4,2] in the forthcoming lemmas. These properties also show that the relaxation behaves geometrically in a natural way.

**Lemma 4.3** (Redundant constraints). Let $P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\}$ and $Q_B = \text{conv}(B)$ be nonempty polytopes and let $a_{k+1} \in \mathbb{R}$, $A_{k+1} \in \mathbb{R}^{1 \times d}$, and $b^{(l+1)} \in \mathbb{R}^d$.

1. If $a_{k+1} - A_{k+1}x \geq 0$ is a redundant inequality in the $\mathcal{H}$-representation of $P_A$, then it is also redundant in the Handelman representation [4,3], i.e., the inclusion of $P_A$ in $Q_B$ is certified by a certain relaxation step if and only if it is certified by the same relaxation step considering $P_{[A,A_{k+1}]}$ instead.

2. If $b^{(l+1)}$ is a redundant point in the $\mathcal{V}$-representation of $Q_B$, then it is also redundant in the Handelman representation [4,3], i.e., $P_A \subseteq Q_B$ is certified by a certain relaxation step if and only if it is certified by the same relaxation step considering $Q_{[B,b^{(l+1)}]}$ instead.

**Proof.** We only prove statement (1), the proof of part (2) is analogous. Let

$$
\nu(t) - x^Tz = \sum_{|\alpha,\beta,\gamma| \leq t} c_{\alpha,\beta,\gamma}(a - Ax)^\alpha(1 - B^Tz)^\beta(a_{k+1} - A_{k+1}x)^\gamma \in H_t(P_{[A,A_{k+1}]}, Q)
$$

with nonnegative $c_{\alpha,\beta,\gamma}$ be a Handelman representation of $\nu(t) - x^Tz$ for some $t \geq 2$. Since $a_{k+1} - A_{k+1}x$ is redundant in the description of $P_{[A,A_{k+1}]}$, we can write it as a convex combination of the remaining linear polynomials,

$$
a_{k+1} - A_{k+1}x = \lambda^T(a - Ax), \quad \lambda^T1 = 1, \quad \lambda \in \mathbb{R}_+^k.
$$

The multinomial theorem implies

$$
(a_{k+1} - A_{k+1}x)^\gamma = \sum_{|\delta| = \gamma} \left(\gamma \atop \delta_1, \ldots, \delta_k\right) \prod_{j=1}^k (\lambda_j(a - Ax)_j)^{\delta_j}.
$$

Replacing $(a_{k+1} - A_{k+1}x)^\gamma$ in the Handelman representation by the above term for any $\gamma$ yields a Handelman representation of the form

$$
\nu(t) - x^Tz = \sum_{|\alpha',\beta'| \leq t} c_{\alpha',\beta'}(a - Ax)^{\alpha'}(1 - B^Tz)^{\beta'} \in H_t(P_A, Q)
$$

with $c_{\alpha',\beta'} \geq 0$. \qed

Removing redundant constraints may lead to faster computations. Note that removing redundant constraints is a polynomial time process; see, e.g., [10, Theorem 2.1] for a constructive proof.

**Lemma 4.4** (Transitivity).

1. Given a $\mathcal{V}$-polytope $Q$ and $\mathcal{H}$-polytopes $P$ and $P'$ such that $P' \subseteq P \subseteq Q$. If there is a Handelman representation of a certain degree $t \geq 2$ certifying containment of $P$ in $Q$, then it also certifies containment of $P'$ in $Q$. 

(2) Given $\mathcal{V}$-polytopes $Q$ and $Q'$, and an $\mathcal{H}$-polytope $P$ such that $P \subseteq Q \subseteq Q'$. If there is a Handelman representation of a certain degree $t \geq 2$ certifying containment of $P$ in $Q$, then it also certifies containment of $P$ in $Q'$.

Proof. Assume $\nu(t) - x^T z \in H_t(P, Q)$ for some $t \geq 2$. By Farkas’ Lemma, we can write the linear polynomials defining $P'$ as convex combinations of the one defining $P$. Using the multinomial theorem as in Lemma 4.3 yields a Handelman representation $\nu(t) - x^T z \in H_t(P', Q)$. This proves part (1) of the lemma. The proof of part (2) is analogous. □

To end this subsection, we state a necessary characterization for the set of $\mathcal{H}$-polytopes whose containment in a fixed $\mathcal{V}$-polytope is certified by the $t$-th relaxation step of the Handelman hierarchy. The underlying idea is to consider the $\mathcal{H}$-polytope as a union of points and interpreting each of these points as a (degenerated) polytope in $\mathcal{H}$-representation.

We define the (formal) natural $\mathcal{H}$-representation of a point $\bar{x}$ considered as $d$-dimensional cube with edge length 0,

\begin{equation}
C_d(\bar{x}) := \left\{ x \in \mathbb{R}^d \mid \left(\begin{array}{c} -\bar{x} \\ \bar{x} \end{array}\right) - I_d x \geq 0 \right\}, \quad I_d = \left[ \begin{array}{c} I_d \\ -I_d \end{array} \right],
\end{equation}

where $I_d$ is the $d \times d$-identity matrix. For $t \geq 2$ and a matrix $B \in \mathbb{R}^{d \times t}$, let $Q = \text{conv}(B)$ and define the set

\[ R_B^t = \{ \bar{x} \in \mathbb{R}^d \mid 1 \geq \min\{\nu : \nu - x^T z \in H_t(C_d(\bar{x}), Q)\} \} . \]

Clearly, $R_B^t$ is a subset of $Q$ for each $t \geq 2$.

**Theorem 4.5.** Let $Q = \text{conv} B$ be a fixed $\mathcal{V}$-polytope. If the containment of an $\mathcal{H}$-polytope $P$ in $Q$ is certified by the $t$-th relaxation step of the Handelman hierarchy \((\ref{4.4})\), then $P$ is a subset of the polytope $R_B^t$.

Proof. Assume that the inclusion $P \subseteq Q$ is certified by the $t$-th relaxation step, i.e.,

\begin{equation}
1 \geq \min\{\nu : \nu - x^T z \in H_t(P, Q)\} ,
\end{equation}

and assume $P \nsubseteq S_B^t$. Then there exists $\bar{x} \in P \setminus S_B^t$. Considering $\bar{x}$ as fixed, we have

\[ 1 < \alpha := \min\{\nu : \nu - x^T z \in H_t(C_d(\bar{x}), Q)\} . \]

But, by Lemma 5.6, this implies a contradiction to \((\ref{4.4})\).

In order to show that $R_B^t$ is a polytope, observe that the set $\{ (\bar{x}, \nu) \in \mathbb{R}^d \times \mathbb{R} : \nu - x^T z \in H_t(C_d(\bar{x}), Q) \}$ is a polyhedron, and $R_B^t$ is the projection of this set on the $\bar{x}$-variables. Since $R_B^t$ is bounded, it is a polytope. □

### 4.3. Degree bounds.

The computational efforts to compute (good) certificates depends on the degree of a Handelman representation for the polynomial $1 - x^T z$. We are not aware of an explicit degree bound for a Handelman representation of this polynomial. However, a quantitative treatment of Averkov’s proof of Handelman’s Theorem \((11, \text{cf. also } [32])\) allows at least to provide an upper bound related to the Pólya exponent of a suitable polynomial. Here, for a homogeneous polynomial $f : \mathbb{R}^d \to \mathbb{R}$ positive on a simplex $\{ x \in \mathbb{R}^d_+ \mid \sum_{i=1}^d x_i = \alpha \}$ the Pólya exponent $\text{Pólya}(f)$ of $f$ is defined as the minimum $N$ such that $(x_1 + \cdots + x_d)^N f(x)$ has only nonnegative coefficients. The existence of such an
is a Handelman representation (4.1) of order $t$.

In order to state the connection in an appropriate way, assume that we apply a translation on $P$ and $Q^0$ so that they are contained in the positive orthant. Let $\tau$ be sufficiently large such that

$$\tau - \sum_{i=1}^{k} (a - Ax)_i - \sum_{i=1}^{l} (1 - B^T z)_i - \sum_{i=1}^{d} x_i - \sum_{i=1}^{d} z_i > 0$$

on $P \times Q^c$, and set $h : \mathbb{R}^{d+k} \times \mathbb{R}^{d+l} \to \mathbb{R}$,

$$h(x,z) = 1 - x^T z + c \sum_{i=1}^{k} (x_{d+i} - (a - Ax)_i) + c \sum_{i=1}^{l} (z_{d+i} - (1 - B^T z)_i)$$

for some constant $c > 0$. Let $\bar{h}(x,z,w)$ be the homogenization of $h$ with respect to $\sum_{i=1}^{d+k} x_i + \sum_{i=1}^{d+l} z_i$. Note that $h(x,z,w)$ is strictly positive on the simplex $\Delta = \{(x,z,w) \in \mathbb{R}^+ \mid w + \sum_{i=1}^{d+k} x_i + \sum_{i=1}^{d+l} z_i = 1\}$.

**Theorem 4.6.** Let $P$ and $Q^0$ be in the positive orthant and let $P$ be strictly contained in $Q$. Then there exists a Handelman representation of the polynomial $1 - x^T z$ whose degree is bounded by $2 + \text{Pólya}(h(x,z,w))$.

**Proof.** This follows from Averkov’s proof of Handelman’s Theorem in connection with the observation that in our situation, the re-substitutions $x_{d+i} \mapsto (a - Ax)_i$, $z_{d+i} \mapsto (1 - B^T z)_i$, $w \mapsto 1$ do not increase the degree. $\square$

In [32, Theorem 1], Powers and Reznick stated a bound on the Pólya exponent in dependence of the minimum objective value (on the underlying ground simplex).

### 4.4. Examples

To illustrate the behavior of the approach, we discuss some (structured) examples.

**Example 4.7.** Let $P = \{-1 \leq x_i \leq 1, \ i = 1, \ldots, d\}$ be the $d$-dimensional unit cube and let $Q^0 = \{-1 \leq e z_i \leq 1, \ i = 1, \ldots, d\}$, i.e., $Q$ is the $d$-dimensional unit cross polytope scaled by $e$. Then

$$\frac{d}{e} - x^T z = \frac{1}{2e} \sum_{i=1}^{d} (1 - x_i)(1 + e z_i) + \frac{1}{2e} \sum_{i=1}^{d} (1 + x_i)(1 - e z_i) \in H_2(P,Q)$$

is a Handelman representation (4.1) of order $t = 2$ certifying the containment $P \subseteq Q$ for $e \geq d$. Indeed, if $e \geq d$, then $1 - x^T z \geq \frac{d}{e} - x^T z \geq 0$, certifying the inclusion $P \subseteq Q$ (with strictness if $e > d$). If $e < d$, then $1 - x^T z < \frac{d}{e} - x^T z$. This is not a certificate for non-containment, since there might be a different Handelman representation. However, in this case, $P \subseteq Q$ if and only if $e \geq d$.

Interestingly, while the hardness result in Proposition 2.1 indicates the combinatorial complexity of this problem, the order of the Handelman representation is low ($t = 2$) and the number of summands is only linear in the dimension. $\square$
Example 4.8. Let $P$ be the $d$-dimensional unit cube in $\mathcal{H}$-representation as in Example 4.7 and $Q = \text{conv}(\{-1,1\}^d)$ be the $d$-dimensional unit cube in $\mathcal{V}$-representation. Denote by $rP := \{x \in \mathbb{R}^d \mid -r \leq x_i \leq r, i = 1, \ldots, d\}$ the $r$-scaled unit cube with edge length $2r$. Clearly, $rP \subseteq Q$ if and only if $0 \leq r \leq 1$. This containment problem is combinatorially hard since the number of inequalities is equal to $2d + 2d$ and thus exponential in the dimension. Consequently, setting up a Handelman representation of degree $t$ considers $(2d + 2d + t)$ possible summands.

We are interested in the maximal $r$ such that the containment $rP \subseteq Q$ is certified by a certain relaxation degree $t$. On the other hand, we can ask for the minimal relaxation order $t$ such that $P = 1P \subseteq Q$ is certified. Note that such a $t$ only exists in case of finite convergence.

We conjecture that for $t = 2$, $\frac{1}{d}P \subseteq Q$ is certified and for $t = d + 1$, the maximal inclusion $P \subseteq Q$ is certified. Moreover, in our numerical computations, we get that the $t$-th relaxation step certifies containment of $\frac{1}{d}P \subseteq Q$; see Table 1. The bottleneck of computation is extracting the LP from the input. Solving the LP is pretty fast.

We write $1 \circ x_i$, where $\circ \in \{+,-\}$, to denote the constraints of $P$ and $1 \circ z_i * z_{j \neq i}$ to denote the constraints of $Q$ where $\circ \in \{+,-\}$ is fixed and $* \in \{+,-\}^{d-1}$ is arbitrary.

For $r \leq 1/d$, $rP \subseteq Q$ is certified by the Handelman representation

$$dr - x^T z = \frac{1}{2d} \sum_{i=1}^{d} \sum_{(\circ, *) \in \{+,-\}^d} (r \circ x_i) (1 \circ z_i * z_{j \neq i}) \in H_2(P, Q).$$

For $* \in \{+,-\}$, denote by $*^{-1}$ the opposite sign (i.e., if $* = +$, then $*^{-1} = -$, and vice versa). The maximal inclusion $P \subseteq Q$ is certified by the Handelman representation of degree $t = d + 1$

$$(4.5) \quad 1 - x^T z = \frac{1}{2d} \sum_{* \in \{+,-\}^d} (1 *_1 x_1) \cdots (1 *_d x_d) (1 *_{-1}^{-1} z_1 \cdots *_{-d}^{-1} z_d) \in H_{d+1}(P, Q).$$

Table 1 shows the maximal values for $r$ such that containment in dimension $d$ is certified by a given relaxation order $t$.

While the example problem seems to be easier than the cube-in-crosspolytope problem in Example 4.7, the Handelman representation in (4.5) has an exponential number of summands, and we are not aware of a more compact Handelman representation. □
5. Semidefinite relaxations based on Putinar’s Positivstellensatz

In this section, we apply Putinar’s Positivstellensatz to the Polytope containment problem. Our main goal is to show that in generic cases (in a well-defined sense) Putinar’s approach yields a certificate for containment after finitely many steps; see Theorems 5.3 and 5.8.

5.1. Putinar’s Positivstellensatz. Consider a set of polynomials $G = \{g_1, \ldots, g_k\} \subset \mathbb{R}[x]$ in the variables $x = (x_1, \ldots, x_d)$. The quadratic module generated by $g_1, \ldots, g_k$ is defined as

$$QM(G) = \left\{ \sigma_0 + \sum_{i=1}^{k} \sigma_i g_i \mid \sigma_i \in \Sigma[x] \right\},$$

where $\Sigma[x] \subseteq \mathbb{R}[x]$ is the set of sum of squares polynomials. Here, a polynomial $p \in \mathbb{R}[x]$ is called sum of squares (sos) if it can be written in the form $p = \sum_i h_i(x)^2$ for some $h_i \in \mathbb{R}[x]$. Equivalently, $p$ has the form $b(x)^T Q b(x)$, where $b(x)$ is the vector of all monomials in $x$ up to half the degree of $p$ and $Q$ is a semidefinite matrix of appropriate size. Checking whether a polynomial is sos is a semidefinite feasibility problem.

If a polynomial $f(x) \in \mathbb{R}[x]$ lies in $QM(G)$, then we say $f$ has a sum-of-squares decomposition in $g_1, \ldots, g_k$. Obviously, every element in $QM(G)$ is nonnegative on the semialgebraic set $S = \{x \in \mathbb{R}^d \mid g_1(x) \geq 0, \ldots, g_k(x) \geq 0\}$. In [33] Putinar showed that for strictly positive polynomials the converse is true under some regularity assumption.

A quadratic module $QM(G)$ is called Archimedean if there is a polynomial $p \in QM(G)$ such that the level set $\{x \in \mathbb{R}^d \mid p(x) \geq 0\}$ is compact, or, equivalently, the polynomial $N - (x_1^2 + \cdots + x_d^2) \in QM(G)$ for some positive integer $N$; see Marshall’s book [25] for more equivalent characterizations.

**Proposition 5.1** (Putinar’s Positivstellensatz). [33] (see also [25] Theorem 5.6.1) Let $S = \{x \in \mathbb{R}^d \mid g_1(x) \geq 0, \ldots, g_k(x) \geq 0\}$ for some polynomials $g_1(x), \ldots, g_k(x) \in \mathbb{R}[x]$. If the quadratic module $QM(G)$ is Archimedean, then $QM(G)$ contains every polynomial $f \in \mathbb{R}[x]$ positive on $S$.

The Archimedean condition in the proposition is not very restrictive. Especially, in our case of interest where all polynomials $g_i$ are linear and $S$ is compact, the condition is always fulfilled; see [25] Theorem 7.1.3. We illustrate this fact by an example.

**Example 5.2.** The $d$-dimensional unit cube is given by the $2d$ inequalities $1 \pm x_i \geq 0$. Since

$$2(1 - x_i^2) = (1 - x_i)^2(1 + x_i) + (1 + x_i)^2(1 - x_i) \in QM(1 \pm x_i)$$

for $i \in \{1, \ldots, d\}$, the quadratic module is Archimedean. \hfill \Box

In [29], Nie and Schweighofer stated an exponential upper bound on the degree of a sos-representation for all polynomials $f$ positive on a certain nonempty semialgebraic set $S$ rescaled to fit in the open unit cube. However, in practice, often a small degree suffices; see [29] Theorem 6.

Powers recently showed that if the polynomials $f, g_1, \ldots, g_k$ have only rational coefficients and the polynomial $g_{k+1} = N - x^T x$ is a member of the quadratic module for
some positive integer $N$, then there exists a rational certificate, i.e., $\sigma_i \in \mathbb{Q}[x]$, for $f$ to be positive on $S$ in the terms of $g_1, \ldots, g_k, g_{k+1}$; see [31] Theorem 7.

In order to apply the proposition to polynomial optimization, consider an optimization problem

$$\sup \left\{ f(x) \mid g_i(x) \geq 0, \ i = 1, \ldots, k \right\}$$

with $f, g_1, \ldots, g_k \in \mathbb{R}[x]$. Clearly, this is the same as to infimize a scalar $\mu$ such that $\mu - f(x) \geq 0$ on the set $S$. A natural relaxation of the latter reformulation is to replace the nonnegativity condition by an sos-condition. This is a semi-infinite program since deciding membership can be rephrased as a semi-infinite feasibility problem. In order to get a (finite-dimensional) semidefinite program, we truncate the quadratic module $QM(G)$ by considering only monomials up to a certain degree $2t$,

$$QM_t(G) = \left\{ \sigma_0 + \sum_{i=1}^{k} \sigma_i g_i \mid \sigma_i \in \Sigma[x] \text{ with } \deg(\sigma_0) \leq 2t \text{ and } \deg(\sigma_i g_i) \leq 2t \right\}.$$

The $t$-th sos-relaxation of the polynomial optimization problem (5.1) has the form

$$\mu(t) = \inf \left\{ \mu \mid \mu \geq f(x) \in QM_t(G) \right\}.$$

Clearly, the sequence of truncated quadratic modules is increasing with respect to inclusion as $t$ grows. Thus the sequence of optimal values $\mu(t)$ is monotone increasing and bounded from above by the optimal value of (5.1).

The dual problem to (5.2) can be formulated in terms of moment matrices, again leading to an SDP relaxation of the polynomial optimization problem (5.1). From a computational point of view it is often easier (i.e., faster) to compute the dual side. But extracting a sos certificate out of the dual optimal solution is not an easy task in general. Since we do not use the dual side here, we refer interested readers to Lasserre’s fundamental work [21].

### 5.2. Putinar certificates for Polytope Containment

We study certificates coming from Putinar’s Positivstellensatz. To keep notation simple, we denote the (truncated) quadratic module generated by the linear constraints $a - Ax$ and $1 - BTz$ by $QM_t(P, Q)$. The Putinar (or sos) relaxation of problem (3.1) reads as

$$\mu(t) = \inf \left\{ \mu \mid \mu \geq x^Tz \in QM_t(P, Q) \right\}.$$

Denote the $i$-th constraint defining $P \times Q^2$ by $g_i$. Let $\mu - x^Tz = \sigma_0 + \sum_{i=1}^{k+l} \sigma_i g_i$ be an sos-decomposition. Assume $t = 1$. Then monomials of degree at most 2 appear, i.e., $\deg(\sigma_0) \in \{0, 2\}$ and $\deg(\sigma_i g_i) \leq 2$. Since $\deg(g_i) = 2$ and $\deg(\sigma_i) \leq 2$, the sos condition must be constant (otherwise $\deg(\sigma_i) = 2$ and $\deg(\sigma_i) = 3$, i.e., monomials of degree greater than 2 appear). Thus $\deg(\sum_i \sigma_i g_i) \leq 1$. Moreover, if $\deg(\sigma_0) = 2$, then purely quadratic terms like $x_j^2$ or $z_j^2$ appear for some $j$. Thus $\sigma_0$ is constant as well. As a consequence, the first relaxation order making sense is $t = 2$. We call $t = 2$ the initial relaxation order.

As for the application of Handelman’s Positivstellensatz, asymptotic convergence of the relaxation in the general case and finite convergence in the strict containment case follow easily from the general theory. We have the following analog of Theorem 4.2.

**Theorem 5.3.** Let $P$ be an $H$-polytope and $Q$ be a $V$-polytope with $0 \in \text{int } Q$.
If \( \mu(t) \leq 1 \) for some integer \( t \geq 2 \), then \( P \subseteq Q \).

(2) The relaxation (5.3) converges asymptotically from above to the optimal value \( \mu^* \) of problem (3.1).

(3) If \( P \) is strictly contained in \( Q \), i.e., \( P \subseteq Q \) and \( \partial P \cap \partial Q = \emptyset \), then the relaxation (5.3) converges in finitely many steps to the optimal value of problem (3.1), i.e., there exists an integer \( t \geq 2 \) such that \( \mu(t) = \mu^* \).

Proof. The first statement is clear by construction of the relaxation.

Consider the third statement. Since all constraints are linear in \( x, z \) and the feasible region is bounded, the quadratic module generated by the constraints of problem (3.1) is Archimedean [25, Theorem 7.1.3] and thus contains all polynomials \( f(x, z) \in \mathbb{R}[x, z] \) strictly positive on \( P \times Q^0 \) by Putinar’s Positivstellensatz 5.1.

Let \( P \) be strictly contained in \( Q \). Then the optimal value \( \mu^* \) of problem (3.1) is less than one by part (2) of Proposition 3.1. Thus the polynomial \( 1 - x^T z \) is positive on \( P \times Q^0 \) and, by the above, has a sos-representation of certain degree. This proves part (3) of the statement.

The second statement follows by blowing up \( Q \) such that \( P \subseteq Q \) and \( \partial P \cap \partial Q = \emptyset \). □

A priori it is not clear whether in the non-strict case finite convergence holds. In fact, for general polynomials, there are examples where finite convergence is not possible. We have a deeper look at this in Section 5.3 where we prove an extension of Theorem 5.3.

Similar to Section 4.2, we deduce in Theorem 5.7 a necessary condition for the set of \( H \)-polytopes whose containment in a fixed \( V \)-polytope is certified by the \( t \)-th relaxation step of Putinar’s hierarchy. Here, the criterion will employ the projection of a spectrahedron. To prepare for this criterion, we show several properties of Putinar’s hierarchy for the Polytope containment problem in Statements 5.4–5.6, which are the semidefinite analogs to Lemmas 4.3–4.4.

First we see that in our situation, the moment relaxation is invariant under redundant constraints, i.e., redundant inequalities in the \( H \)-representation of \( P \) or redundant points in the \( V \)-representation of \( Q \). Note that for a general semialgebraic constraint set, this is not always true, even in the case of optimizing a linear function over it; see [13, Section 5.2] for a well-known example (cf. also [8]). Recall that every \( H \)-representation of a certain polytope contains the facet defining halfspaces. Similarly, the vertices are part of each \( V \)-representation.

**Lemma 5.4** (Redundant constraints). Let \( P_A = \{ x \in \mathbb{R}^d \mid a - Ax \geq 0 \} \) and \( Q_B = \text{conv}(B) \) be nonempty polytopes and let \( a_{k+1}, A_{k+1} \in \mathbb{R} \), \( a_{k+1} \in \mathbb{R}^{1 \times d} \), and \( b^{(l+1)} \in \mathbb{R}^d \).

1. If \( a_{k+1} - A_{k+1} x \geq 0 \) is a redundant inequality in the \( H \)-representation of \( P_A \), then it is also redundant in relaxation (5.3), i.e., the inclusion of \( P_A \) in \( Q_B \) is certified by a certain relaxation step if and only if it is certified by the same relaxation step considering \( P_{[A_{k+1}, A_{k+1}]} \) instead.

2. If \( b^{(l+1)} \) is a redundant point in the \( V \)-representation of \( Q_B \), then it is also redundant in relaxation (5.3), i.e., \( P_A \subseteq Q_B \) is certified by a certain relaxation step if and only if it is certified by the same relaxation step considering \( Q_{[B, b^{(l+1)}]} \) instead.
Proof. We only prove statement (1), the proof of part (2) is analog. Consider a Putinar representation of \( \mu(t) - x^T z \) for some \( t \geq 2 \),
\[
\mu(t) - x^T z = \sigma_0 + \sum_{i=1}^{k+1} \sigma_i (a - Ax)_i + \sum_{i=1}^{l} \sigma_{k+1+i} (1 - B^T x)_i \in \text{QM}(P_{[A,A_{k+1}]}; Q),
\]
where \( \sigma_0, \ldots, \sigma_{k+l+1} \in \Sigma[x, z] \) are sos polynomials with \( \deg \sigma_0 \leq 2t \), \( \deg \sigma_i \leq 2t - 2 \) for \( i \in \{1, \ldots, k + l + 1 \} \). Since \( a_{k+1} - A_{k+1} x \) is redundant in the description of \( P_{[A,A_{k+1}]} \), we can write it as a convex combination of the remaining linear polynomials,
\[
a_{k+1} - A_{k+1} x = \lambda^T (a - Ax), \quad \lambda^T 1_k = 1, \quad \lambda \in \mathbb{R}^k.
\]
Replacing \( \sigma_{k+1}(a_{k+1} - A_{k+1} x) \) in the Putinar representation by
\[
\sigma_{k+1}(a_{k+1} - A_{k+1} x) = \sum_{i=1}^{k} \lambda_i \sigma_{k+1} (a - Ax)_i
\]
yields a Putinar representation of the form
\[
\mu(t) - x^T z = \sigma_0 + \sum_{i=1}^{k} \sigma'_i (a - Ax)_i + \sum_{i=1}^{l} \sigma_{k+i} (1 - B^T x)_i \in \text{QM}(P_A, Q),
\]
where \( \sigma'_i = \sigma_{k+1} + \sigma_i \in \Sigma[x, z] \) with degree \( \deg(\sigma'_i) = \max\{\deg(\sigma_{k+1}), \deg(\sigma_i)\} \leq 2t - 2 \) for \( i \in \{1, \ldots, k\} \). \( \square \)

**Lemma 5.5** (Monotonicity). Let \( P_A = \{x \in \mathbb{R}^d \mid a - Ax \geq 0\} \) be a polytope and let \( a_{k+1} \in \mathbb{R}, A_{k+1} \in \mathbb{R}^{1 \times d} \) such that \( P_{[A,A_{k+1}]} := P_A \cap \{x \in \mathbb{R}^d \mid a_{k+1} - A_{k+1} x \geq 0\} \) is a proper subset of \( P_A \). If for a certain relaxation order \( t \geq 2 \) relaxation (5.3) with respect to \( P_A \) has an optimal value of at most one, then this holds when considering \( P_{[A,A_{k+1}]} \), i.e., if the relaxation certifies containment of \( P_A \) in a \( V \)-polytope \( Q \), then containment of \( P_{[A,A_{k+1}]} \) in \( Q \) is certified as well.

**Proof.** Given an sos decomposition of \( \mu(t) - x^T z \) w.r.t. \( P_A \), by setting the additional sos polynomial \( \sigma_{k+1} \) to the zero-polynomial, i.e. \( \sigma_{k+1} \equiv 0 \), this yields an sos decomposition w.r.t. \( P_{[A,A_{k+1}]} \). \( \square \)

**Lemma 5.6** (Transitivity).

(1) Given a \( V \)-polytope \( Q \) and \( H \)-polytopes \( P \) and \( P' \) such that \( P' \subseteq P \subseteq Q \). If for a certain relaxation order \( t \geq 2 \) relaxation (5.3) certifies containment of \( P \) in \( Q \), then it also certifies containment of \( P' \) in \( Q \).

(2) Given \( V \)-polytopes \( Q \) and \( Q' \), and an \( H \)-polytope \( P \) such that \( P \subseteq Q \subseteq Q' \). If for a certain relaxation order \( t \geq 2 \) relaxation (5.3) certifies containment of \( P \) in \( Q \), then it also certifies containment of \( P \) in \( Q' \).

**Proof.** Starting with \( P \) incorporate the defining inequalities of \( P' \) into the representation of \( P \) step-by-step. By Lemma 5.5 in every step the lower bound of the optimal value in (5.3) can not increase. At the end of this process the defining inequalities of \( P \) are all redundant (since \( P' \subseteq P \)) and thus can be dropped, by Lemma 5.4. This proves part (1) of the statement. The proof of (2) is analog. \( \square \)
Similar to Theorem 4.5 for the Handelman situation, we can now state a necessary characterization for the set of $\mathcal{H}$-polytopes whose containment in a fixed $\mathcal{V}$-polytope is certified by the $t$-th relaxation step of Putinar’s hierarchy.

Recall from (4.3) that $C_d(\bar{x})$ denotes the (formal) natural $\mathcal{H}$-representation of a point $\bar{x}$ considered as $d$-dimensional cube with edge length 0. For $t \geq 2$ and a matrix $B \in \mathbb{R}^{d \times l}$, let $Q = \text{conv}(B)$ and define the set

$$S^t_B = \{ \bar{x} \in \mathbb{R}^d \mid 1 \geq \inf\{ \mu : \mu - x^Tz \in \text{Q.M}_t(C_d(\bar{x}), Q) \} \}.$$ 

Clearly, $S^t_B$ is a subset of $Q$ for each $t \geq 2$. Moreover, $S^t_B$ is the projection of a set defined by semidefinite conditions, i.e., the projection of a spectrahedron.

**Theorem 5.7.** Let $Q = \text{conv} B$ be a fixed $\mathcal{V}$-polytope. If the containment of an $\mathcal{H}$-polytope $P$ in $Q$ is certified by the $t$-th relaxation step of Putinar’s hierarchy (5.3), then $P$ is a subset of $S^t_B$.

**Proof.** Assume that the inclusion $P \subseteq Q$ is certified by the $t$-th relaxation step, i.e.,

$$1 \geq \inf\{ \mu : \mu - x^Tz \in \text{Q.M}_t(P, Q) \},$$

and assume $P \not\subseteq S^t_B$. Then there exists $\bar{x} \in P \setminus S^t_B$. Considering $\bar{x}$ as fixed, we have

$$1 < \alpha := \inf\{ \mu : \mu - x^Tz \in \text{Q.M}_t(C_d(\bar{x}), Q) \}.$$ 

But, by Lemma 5.6, this implies a contradiction to (5.4). \hfill \Box

### 5.3. Finite convergence

By Theorem 5.3, there exists a Putinar representation of the polynomial $1 - x^Tz$ over $P \times Q^c$ whenever $P$ is strictly contained in $Q$. This is a severely limited case. It does not take into account that $P$ and $Q$ may have common boundary points or $P$ is not contained in $Q$. In this section, we prove a partial extension of Theorem 5.3 to the case where the bilinear optimization problem (3.1) has only finitely many optimal solutions (as characterized in Lemma 3.5).

**Theorem 5.8.** Let $P_A = \{ x \in \mathbb{R}^d \mid a - Ax \geq 0 \}$ be an $\mathcal{H}$-polytope with nonempty interior and let $Q_B = \text{conv}(B)$ be a $\mathcal{V}$-polytope containing the origin in its interior. Assume that one of the equivalent statements in Lemma 3.2 holds (e.g., there are only finitely many optimal solutions to problem (3.1)). Then $\mu^* - x^Tz$ lies in the quadratic module generated by the linear inequalities defining $P_A$ and $Q_B^c$, and thus relaxation (5.3) converges in finitely many steps to the optimal value of (3.1).

To prove the theorem, we introduce a sufficient convergence condition by Marshall (see [25]), which is based on a boundary Hessian condition.

Given $g_1, \ldots, g_k \in \mathbb{R}[x]$ and a boundary point $\bar{x}$ of $S = \{ x \in \mathbb{R}^d \mid g_1(x) \geq 0, \ldots, g_k(x) \geq 0 \}$. We assume that (say, by an application of the inverse function theorem), there exists a local parameterization for $\bar{x}$ in the following sense: There exist open sets $U, V \subseteq \mathbb{R}^d$ such that $\bar{x} \in U$, $\phi : U \rightarrow V$, $x \mapsto t := (t_1, \ldots, t_d)$ is bijective, the inverse $\phi^{-1} : V \rightarrow U$ is a continuously differentiable function on $V$, and the region $R$ defined by $t_1 \geq 0, \ldots, t_r \geq 0$ (for some $r \in \{1, \ldots, d\}$) equals the set $S \cap U$. Given a polynomial $f \in \mathbb{R}[x]$, denote by $f_1$ and $f_2$ the linear and quadratic part of $f$ in the localizing parameters $t_1, \ldots, t_d$, respectively.
Condition 5.9 (Boundary Hessian condition, BHC). If the linear form \( f_1 = c_1 t_1 + \cdots + c_r t_r \) has only positive coefficients and the quadratic form \( f_2(0, \ldots, 0, t_{r+1}, \ldots, t_d) \) is positive definite, then the restriction \( f|_R \) has a local minimum in \( p \).

Using this condition, the following generalization of Putinar’s Theorem can be stated.

Proposition 5.10. \([25, \text{Theorem 9.5.3}]\) Let \( f, g_1, \ldots, g_k \in \mathbb{R}[x] \), and suppose that the quadratic module \( \text{QM} \) generated by \( g_1, \ldots, g_k \) is Archimedean. Further assume that for each global maximizer \( \bar{x} \) of \( f \) over \( S = \{ x \in \mathbb{R}^d \mid g_1(x) \geq 0, \ldots, g_k(x) \geq 0 \} \) there exists an index set \( I \subseteq \{1, \ldots, k\} \) such that (after renaming the variables w.r.t. the indices in \( I \) and w.r.t. the indices not in \( I \)) \( f \) satisfies BHC at \( \bar{x} \). In this situation, if \( f \geq 0 \) on \( S \) then \( f \in \text{QM} \).

Our goal is to show that under the assumptions of Theorem 5.8 the boundary Hessian condition holds. We will use the following version of the Karush-Kuhn-Tucker conditions adapted to the bilinear situation.

Lemma 5.11. Let \( f(x, z) \in \mathbb{R}[x, z] \) be a continuously differentiable function and let \( P := \mathbb{P} := P_A \times P_B = \{(x, z) \in \mathbb{R}^{2d} \mid a - Ax \geq 0, b - Bz \geq 0\} \) be the product of two nonempty polytopes. If \( f \) attains a local maximum in \((\bar{x}, \bar{z})\) on \( \mathbb{P} \), i.e., there exists \( \varepsilon > 0 \) such that for all \((x, z) \in \mathbb{P} \cap U_\varepsilon(\bar{x}, \bar{z})\) the relation \( f(\bar{x}, \bar{z}) \leq f(x, z) \) holds, then there exists \((\alpha, \beta)\) such that

\[
\nabla f(\bar{x}, \bar{z}) = \begin{bmatrix} A^T & 0 \\ 0 & B^T \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}
\]

\[(5.5)\]

\[0 = \alpha_i(a - A\bar{x})_i = \beta_j(b - B\bar{z})_j, \quad i = 1, \ldots, k, \quad j = 1, \ldots, l\]

\[\alpha \geq 0, \quad \beta \geq 0,\]

and the set of positive multipliers corresponds to linearly independent columns in \( A \) and \( B \), respectively.

In the lemma, only multipliers corresponding to active constraints can be positive, since otherwise one of the equations \((5.5)\) is violated.

Proof. Denote by \( I \) and \( J \) the index sets of the active constraints in \( \bar{x} \) and \( \bar{z} \), respectively. We only have to show the linear independence statement, since all other parts of the lemma are the Karush-Kuhn-Tucker conditions together with the well-known fact that in case of linear constraints a constraint qualification in the KKT Theorem is not required (see, e.g., \([3, \text{Section 5.1}]\)).

By Carathéodory’s Theorem \([34, \text{Corollary 7.1}]\), there exists an index set \( K \subseteq (I, J) \) such that the corresponding columns are linearly independent and \( \nabla f(\bar{x}, \bar{z}) \) is a strictly positive combination of these columns. Hence, \((\alpha, \beta)\) can be chosen in this way. \(\square\)

We are now able to prove Theorem 5.8. In a more general setting, Nie used the Karush-Kuhn-Tucker optimality conditions to certify the BHC; see \([28]\). Because of the special structure of problem \((3.1)\), we do not need the whole machinery used by Nie. In particular, the local parameterization needed for the BHC (see the paragraph before Condition 5.9) comes from an affine variable transformation. As a consequence, for Polytope Containment, our direct approach allows to prove a stronger result than we would obtain.
just by applying Nie’s Theorem. Specifically, we obtain a geometric characterization of the degenerate situations as given in Theorem 5.8

Proof (of Theorem 5.8). Let \((\bar{x}, \bar{z}) \in P_A \times Q^+_P\) be an arbitrary but fixed optimal solution. By Lemma 5.11 there exists \((\alpha, \beta) \in \mathbb{R}^{k+l}\) such that

\[
\begin{align*}
(\bar{z}, \bar{x}) &= (A^T \alpha, B \beta) \\
0 &= \alpha_i(a - A \bar{x})_i = \beta_j(1 - B^T \bar{z})_j, \quad i = 1, \ldots, k, \ j = 1, \ldots, l \\
\alpha &\geq 0, \ \beta \geq 0,
\end{align*}
\]

and the set of positive multipliers corresponds to linearly independent rows in \(A\) and \(B^T\), respectively. As mentioned before, only multipliers corresponding to active constraints can be positive. Denote by \(I\) and \(J\) the index sets of linearly independent, active constraints in \(\bar{x}\) and \(\bar{z}\), respectively. Then \(|I| \leq d\) and \(|J| \leq d\).

Assume \(|I| < d\). Since \(\bar{z} \in \text{cone}\{A^T \alpha \mid \forall i \in I\}\), \(\bar{z}\) lies in the outer normal cone of an at least one-dimensional face \(F\) of \(P_A\) containing \(\bar{x}\). Thus, \(x^T \bar{z} = \bar{x}^T \bar{z}\) for all \(x \in F\), in contradiction to the assumption of the theorem and Lemma 3.5. By a symmetric argument, \(|J| < d\) is not possible either.

We apply the affine variable transformation \(\phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}\) defined by

\[
\phi(x, z) = \left[ \begin{array}{c}
(a - Ax)_I \\
(1 - B^T z)_J
\end{array} \right]
\]

and denote the new variables by \((s, t) := (s_1, \ldots, s_d, t_1, \ldots, t_d) = (\phi_1(x, z), \ldots, \phi_2d(x, z))\).

Clearly, \(\phi\) is a local parameterization at \((\bar{x}, \bar{z})\) in the sense of Condition 5.9. The inverse of \(\phi\) is given by

\[
(s, t) \mapsto \left[ \begin{array}{c}
A_I^{-1} (s - a_I) \\
(B_J^{-1})^{-1} (t - 1_J)
\end{array} \right].
\]

Setting \(M := B_J^{-1} A_I^{-1}\), the objective \(x^Tz\) has the form

\[
f(s, t) := (A_I^{-1} (s - a_I))^T ((B_J^{-1})^{-1} (t - 1_J)) = s^T M^T I_J - s^T M^T 1_J - a_I^T M^T t + a_I^T M^T 1_J
\]

in the local parameterization space. Denote by \(f_1\) the homogeneous part of degree 1. Then \((\bar{x}, \bar{z}) = \phi^{-1}(0) = (-A_I^{-1}a_I, -(B_J^{-1})^{-1}1_J)\) implies

\[
\nabla_{s,t} f_1(0) = ( -1_J^T B_J^{-1} A_I^{-1}, -B_J^{-1} A_I^{-1} a_I) = (\bar{z}^T A_I^{-1}, B_J^{-1} \bar{x}) = (\alpha_I, \beta_J),
\]

where the last equation follows from the first identity in (5.6). Thus the first part of Condition 5.9 is satisfied. Since \(|I| + |J| = r = 2d\) (where \(r\) is from Condition 5.9), the second assumption in Condition 5.9 is obsolete. Therefore, by Proposition 5.10 \(\mu^* - x^T z \in \text{QM}(P, Q)\). \(\square\)

5.4. Examples. We apply the semidefinite hierarchy to the examples in Section 4.4.

Example 5.12. Let \(P = \{-1 \leq x_i \leq 1, \ i = 1, \ldots, d\}\) be the \(d\)-dimensional unit cube and let \(Q^e = \{-1 \leq ez_i \leq 1, \ i = 1, \ldots, d\}\), i.e., \(Q\) is the \(d\)-dimensional unit cross polytope scaled by a positive integer \(e\). Clearly, \(P \subseteq Q\) if and only if \(e \geq d\).
Consider the Putinar representation of order $t = 2$

$$\frac{d}{e} - x^Tz = \frac{1}{8e} \sum_{i=1}^{d} \left[ (1 - x_i)(1 + x_i)^2 + (1 + ez_i)^2 \right] + \frac{1}{8e} \sum_{i=1}^{d} \left[ (1 - ez_i)(1 + x_i)^2 + (1 + x_i)(1 - ez_i)^2 \right].$$

If $e \geq d$, then $1 - x^Tz \geq \frac{d}{e} - x^Tz \geq 0$, certifying the containment $P \subseteq Q$ (with strictness if $e > d$). If $e < d$, then $1 - x^Tz < \frac{d}{e} - x^Tz$. This is not a certificate for non-containment, since there might be a different sos-representation. However, in this case this is not possible since $e \geq d$ is a necessary condition for containment.

Note that like in the application of Handelman’s Positivstellensatz, Example 4.7, the necessary relaxation order is low and the number of terms is linear in the dimension. □

**Example 5.13.** Consider the $d$-dimensional $r$-scaled $H$-unit cube in $V$-unit cube as described in Example 4.8. Again, we are interested in the maximal $r$ such that the containment $rP \subseteq Q$ is certified by a certain relaxation degree $t$. On the other hand, we could ask for the minimal $t$ such that $P_A = 1P_A \subseteq Q_B$ is certified. Note that a priori the existence of such a $t$ is not clear since neither Theorem 5.13 nor Theorem 5.18 applies.

Modifying the Handelman representation, we get the same bound for Putinar’s hierarchy. Recall the notation from Example 4.8. Then, for $r \leq 1/d$, $rP \subseteq Q$ is certified by the Putinar representation

$$dr - x^Tz = \frac{1}{2d+1} \sum_{i=1}^{d} \sum_{(\circ, \ast) \in \{+, -\}^d} \left( r \circ x_i \right) \left( 1 \circ^{-1} z_i \ast z_{j \neq i} \right)^2 + \frac{dr}{2d+1} \sum_{\ast \in \{+, -\}^d} \left( 1 \ast_1 \cdots \ast_d z_d \right) \left( 1 \ast_1^{-1} \cdots \ast_d^{-1} z_d \right)^2 \in \text{QM}_2(P, Q).$$

We are not aware of a more compact Putinar representation. Numerically, for $t = 2$ and $d \leq 5$, we get $r(d) = \sqrt{d}/d$; see Table 2. Comparing the table with Table 1 we
see that in this situation, the initial Putinar relaxation is strictly better than the initial Handelman relaxation. On the other hand, while the \((d + 1)\)-st Handelman relaxation is exact (see (4.5)), it is not clear whether this is true for Putinar’s relaxation.

\[\square\]

6. Open questions

In this paper, we studied Handelman certificates and Putinar certificates for Polytope containment. We close with a short discussion of open questions. We believe that these questions will be very relevant in improving the understanding of relaxations for low-degree geometric problems, such as the one studied here.

In Theorem 5.8 we saw that the Putinar relaxation always finitely converges (under mild preconditions). Does the Handelman relaxation always finitely converge for Polytope containment? Note that by Theorem 3.3 the optima of the bilinear programming formulation are always attained at the boundary, which would allow for a positive answer of the question. (Recall that the Handelman hierarchy cannot finitely converge whenever there exists an optimizer in the interior of the feasible region).

For the Polytope containment problem, can the structure of the certificates be better characterized? Such as, what are improved degree bounds with regard to Polytope containment or, somewhat more general, with regard to general bilinear programming problems? How is Fourier-Motzkin-elimination (as an \(H\)-in-\(V\) conversion algorithm) related to the Handelman certificates of Polytope containment?

References

[1] G. Averkov. Constructive proofs of some Positivstellensätze for compact semialgebraic subsets of \(\mathbb{R}^d\). J. Optim. Theory Appl., 158(2):410–418, 2013.
[2] D. Avis, D. Brenner, and R. Seidel. How good are convex hull algorithms? Comput. Geom., 7(5-6):265–301, 1997.
[3] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. Nonlinear Programming. Wiley-Interscience, Hoboken, NJ, third edition, 2006.
[4] C. Böhm and H.-P. Kriegel. Determining the convex hull in large multidimensional databases. In Data Warehousing and Knowledge Discovery, pages 294–306. Springer, 2001.
[5] V. Chandru and J.-L. Lassez. Qualitative theorem proving in linear constraints. In Verification: Theory and Practice, pages 395–406. Springer, 2003.
[6] R.M. Freund and J.B. Orlin. On the complexity of four polyhedral set containment problems. Math. Program., 33(2):139–145, 1985.
[7] G. Gallo and A. Ülküçü. Bilinear programming: An exact algorithm. Math. Program., 12(1):173–194, 1977.
[8] B. Ghaddar, J.C. Vera, and M.F. Anjos. An iterative scheme for valid polynomial inequality generation in binary polynomial programming. In Integer Programming and Combinatorial Optimization, volume 6655 of Lecture Notes in Comput. Sci., pages 207–222. Springer, Heidelberg, 2011.
[9] J. Gouveia, R. Grappe, V. Kaibel, K. Pashkovich, R. Robinson, and R. Thomas. Which nonnegative matrices are slack matrices? Linear Algebra Appl., 439(10):2921–2933, 2013.
[10] P. Gritzmann and V. Klee. Computational complexity of inner and outer \(j\)-radii of polytopes in finite-dimensional normed spaces. Math. Program., 59(2, Ser. A):163–213, 1993.
[11] P. Gritzmann, V. Klee, and D. Larman. Largest \(j\)-simplices in \(n\)-polytopes. Discrete Comput. Geom., 13(3-4):477–515, 1995.
[12] D. Handelman. Representing polynomials by positive linear functions on compact convex polyhedra. Pacific J. Math., 132(1):35–62, 1988.
[13] D. Henrion. On semidefinite representations of plane quartics. Technical Report 08444, LAAS-CNRS, 2008.
[14] M. Joswig and G.M. Ziegler. Convex hulls, oracles, and homology. J. Symb. Comp., 38(4):1247–1259, 2004.
[15] V. Kaibel and M.E. Pfetsch. Some algorithmic problems in polytope theory. In M. Joswig and N. Takayama, editors, Algebra, Geometry, and Software Systems, pages 23–47. Springer, Berlin, 2003.
[16] K. Kellner, T. Theobald, and C. Trabandt. A semidefinite hierarchy for containment of spectrahedra. Preprint, arXiv:1308.5076.
[17] K. Kellner, T. Theobald, and C. Trabandt. Containment problems for polytopes and spectrahedra. SIAM J. Optim., 23(2):1000–1020, 2013.
[18] L. Khachiyan, E. Boros, K. Borys, K. Elbassioni, and V. Gurvich. Generating all vertices of a polyhedron is hard. Discrete Comput. Geom., 39(1-3):174–190, 2008.
[19] L.G. Khachiyan. Polynomial algorithms in linear programming. USSR Computational Mathematics and Mathematical Physics, 20(1):53–72, 1980.
[20] H. Konno. A cutting plane algorithm for solving bilinear programs. Math. Program., 11(1):14–27, 1976.
[21] J.B. Lasserre. Global optimization with polynomials and the problem of moments. SIAM J. Optim., 11(3):796, 2001.
[22] J.B. Lasserre. Semidefinite programming vs. LP relaxations for polynomial programming. Math. Oper. Res., 27(2):347–360, 2002.
[23] J.B. Lasserre. Moments, Positive Polynomials and their Applications, volume 1. Imperial College Press, London, 2010.
[24] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In M. Putinar and S. Sullivant, editors, Emerging Applications of Algebraic Geometry, volume 149 of IMA Vol. Math. Appl., pages 157–270. Springer, New York, 2009.
[25] M. Marshall. Positive Polynomials and Sums of Squares. American Mathematical Society, Providence, Rhode Island, 2008.
[26] P. McMullen. The maximum numbers of faces of a convex polytope. Mathematika, 17:179–184, 1970.
[27] D. Monniaux. A quantifier elimination algorithm for linear real arithmetic. In Logic for Programming, Artificial Intelligence, and Reasoning, pages 243–257. Springer, 2008.
[28] J. Nie. Optimality conditions and finite convergence of Lasserre’s hierarchy. Math. Program., 146(1, Ser. A):97–121, 2014.
[29] J. Nie and M. Schweighofer. On the complexity of Putinar’s Positivstellensatz. J. Complexity, 23(1):135–150, 2007.
[30] A. Packer. Polynomial-time approximation of largest simplices in V-polytopes. Discrete Appl. Math., 134(1-3):213–237, 2004.
[31] V. Powers. Rational certificates of positivity on compact semialgebraic sets. Pacific J. Math., 251(2):385–391, 2011.
[32] V. Powers and B. Reznick. A new bound for Pólya’s theorem with applications to polynomials positive on polyhedra. J. Pure Appl. Algebra, 164(1):221–229, 2001.
[33] M. Putinar. Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J., 42(3):969–984, 1993.
[34] A. Schrijver. Theory of Linear and Integer Programming. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons Ltd., Chichester, 1986.
[35] G.M. Ziegler. Lectures on Polytopes. Graduate Texts in Mathematics. Springer, New York, 1995.

Goethe-Universität, FB 12 – Institut für Mathematik, Postfach 11 19 32, D–60054 Frankfurt am Main, Germany
E-mail address: {kellner, theobald}@math.uni-frankfurt.de