

THETA BODIES FOR POLYNOMIAL IDEALS

JOÃO GOUVEIA, PABLO A. PARRILO, AND REKHA R. THOMAS

ABSTRACT. Inspired by a question of Lovász, we introduce a hierarchy of nested semidefinite relaxations of the convex hull of real solutions to an arbitrary polynomial ideal, called theta bodies of the ideal. For the stable set problem in a graph, the first theta body in this hierarchy is exactly Lovász's theta body of the graph. We prove that theta bodies are, up to closure, a version of Lasserre's relaxations for real solutions to ideals, and that they can be computed explicitly using combinatorial moment matrices. Theta bodies provide a new canonical set of semidefinite relaxations for the max cut problem. For vanishing ideals of finite point sets, we give several equivalent characterizations of when the first theta body equals the convex hull of the points. We also determine the structure of the first theta body for all ideals.

1. INTRODUCTION

A central concern in optimization is to understand the convex hull of feasible solutions to a given problem. For example, in the general *integer program* which asks to minimize a linear function $\sum_{i=1}^n c_i x_i$ over the discrete set of points $S = \{\mathbf{x} \in \mathbb{Z}^n : A\mathbf{x} \leq \mathbf{b}\}$ where $A \in \mathbb{Z}^{m \times n}$ and $\mathbf{b} \in \mathbb{Z}^m$, knowing $\text{conv}(S)$, the convex hull of S , allows the integer program to be modeled as the *linear program* $\min\{\sum_{i=1}^n c_i x_i : \mathbf{x} \in \text{conv}(S)\}$, which is conceptually a simpler problem. In many instances, the set of feasible solutions to an optimization problem can be described as the set of real solutions to a system of polynomial equations: $f_1(\mathbf{x}) = f_2(\mathbf{x}) = \dots = f_m(\mathbf{x}) = 0$, where $f_1, \dots, f_m \in \mathbb{R}[\mathbf{x}] := \mathbb{R}[x_1, \dots, x_n]$. This set is the *real variety*, $\mathcal{V}_{\mathbb{R}}(I)$, of the *ideal* I in $\mathbb{R}[\mathbf{x}]$ generated by f_1, \dots, f_m , and our goal is to compute or represent $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ either exactly or at least approximately.

Throughout this paper we will say that f is a *linear* polynomial if it is affine linear of the form $f = a_0 + \sum_{i=1}^n a_i x_i$. The closure of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$, $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$, is the intersection of all half spaces $\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0\}$ as f runs over all linear polynomials that are non-negative on $\mathcal{V}_{\mathbb{R}}(I)$. A classical certificate for the non-negativity of any polynomial f on $\mathcal{V}_{\mathbb{R}}(I)$ is the existence of a *sum of squares (sos)* polynomial $\sum_{j=1}^t h_j^2$ that is congruent to $f \bmod I$ (i.e., $f - \sum_{j=1}^t h_j^2 \in I$). This leads to the natural relaxation of

Date: June 7, 2009.

All authors were partially supported by the NSF Focused Research Group grant DMS-0757371. J. Gouveia was also partially supported by Fundação para a Ciência e Tecnologia and R.R. Thomas by the Robert R. and Elaine K. Phelps Endowed Professorship.

$\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ by the closed convex set:

$$(1) \quad \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \ \forall f \text{ linear and sos mod } I\}.$$

Depending on I , (1) may be strictly larger than $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ since not all polynomials that are non-negative on $\mathcal{V}_{\mathbb{R}}(I)$ may be sos mod I . However, in many interesting cases, (1) will equal $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. A more practical relaxation of $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ is obtained by restricting the degree of the h_j 's in (1) to some positive integer k . As k is increased, we obtain a hierarchy of nested relaxations to $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. In [13], Lovász raises a question that leads to the study of this hierarchy which motivated the work in this paper. To explain his question and motivation we need some definitions.

Definition 1.1. Let f be a polynomial in $\mathbb{R}[\mathbf{x}]$, and I be an ideal in $\mathbb{R}[\mathbf{x}]$ with real variety $\mathcal{V}_{\mathbb{R}}(I) := \{\mathbf{s} \in \mathbb{R}^n : f(\mathbf{s}) = 0 \ \forall f \in I\}$.

- (1) The polynomial f is **non-negative** on $\mathcal{V}_{\mathbb{R}}(I)$ if $f(\mathbf{s}) \geq 0$ for all $\mathbf{s} \in \mathcal{V}_{\mathbb{R}}(I)$.
- (2) Given $g \in \mathbb{R}[\mathbf{x}]$, f is **congruent** to $g \text{ mod } I$, written as $f \equiv g \text{ mod } I$, if $f - g \in I$. (Note that $f \equiv g \text{ mod } I$ if and only if $f(\mathbf{s}) = g(\mathbf{s})$ for all $\mathbf{s} \in \mathcal{V}_{\mathbb{R}}(I)$.)
- (3) The polynomial f is a **sum of squares (sos)** mod I if there exists $h_j \in \mathbb{R}[\mathbf{x}]$ such that $f \equiv \sum_{j=1}^t h_j^2 \text{ mod } I$ for some t . If, in addition, each h_j has degree at most k , then we say that f is k -**sos** mod I .
- (4) The ideal I is k -**sos** if every polynomial that is non-negative on $\mathcal{V}_{\mathbb{R}}(I)$ is k -sos mod I . If every polynomial of degree at most d that is non-negative on $\mathcal{V}_{\mathbb{R}}(I)$ is k -sos mod I , we say that I is (d, k) -**sos**.

Definition 1.1 can be used to define the following hierarchy of nested closed convex sets that all contain $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. The choice of name for these relaxations will be clear shortly.

Definition 1.2. (1) For a positive integer k , the k -th **theta body** of an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ is

$$\text{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every linear } f \text{ that is } k\text{-sos mod } I\}.$$

- (2) An ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ is **TH $_k$ -exact** if the k -th theta body $\text{TH}_k(I)$ coincides with $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$.
- (3) The **theta-rank** of I is the smallest k for which $\text{TH}_k(I)$ equals $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$.

Example 1.3. Consider the principal ideal $I = \langle x_1^2 x_2 - 1 \rangle \subset \mathbb{R}[x_1, x_2]$. Then $\text{conv}(\mathcal{V}_{\mathbb{R}}(I)) = \{(s_1, s_2) \in \mathbb{R}^2 : s_2 > 0\}$, and any linear polynomial that is non-negative over $\mathcal{V}_{\mathbb{R}}(I)$ is of the form $\alpha x_2 + \beta$, where $\alpha, \beta \geq 0$. Since $\alpha x_2 + \beta \equiv (\sqrt{\alpha} x_1 x_2)^2 + (\sqrt{\beta})^2 \text{ mod } I$, I is $(1, 2)$ -sos and TH_2 -exact. Check that x_2 is not 1-sos mod I and so, the theta rank of I is two.

By definition, $\text{TH}_1(I) \supseteq \text{TH}_2(I) \supseteq \dots \supseteq \text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. As seen in Example 1.3, $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ may not always be closed, while $\text{TH}_k(I)$ is closed for

all k . Therefore, the theta-body sequence of I can converge, if at all, only to the closure of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$.

A natural question at this point is whether the algebraic notion of an ideal being $(1, k)$ -sos is equivalent to the geometric notion of being TH_k -exact.

Lemma 1.4. *If an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ is $(1, k)$ -sos then it is TH_k -exact.*

Proof: If I is $(1, k)$ -sos then any linear f non-negative on $\mathcal{V}_{\mathbb{R}}(I)$ and hence on $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ is k -sos mod I which implies that $\text{TH}_k(I) \subseteq \text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. But since the reverse inclusion is always true, we get $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I))) = \text{TH}_k(I)$. \square

The converse of Lemma 1.4 is false in general.

Example 1.5. Consider $I = \langle x^2 \rangle \subset \mathbb{R}[x]$ with $\mathcal{V}_{\mathbb{R}}(I) = \{0\} \subset \mathbb{R}$. All linear polynomials that are non-negative on $\mathcal{V}_{\mathbb{R}}(I)$ are of the form $\pm a^2 x + b^2$ for some $a, b \in \mathbb{R}$. If $b \neq 0$, then $(\pm a^2 x + b^2) \equiv (\frac{a^2}{2b} x \pm b)^2 \pmod{I}$. However, $\pm x$ is not a sum of squares mod I , and hence I is not $(1, k)$ -sos for any k . On the other hand, I is TH_1 -exact since $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ is cut out by the infinitely many linear inequalities of the form $\pm x + b^2 \geq 0$ as b varies over $b \neq 0$.

Remark 1.6. In [5], Lasserre introduces the Putinar-Prestel Bounded Degree Representation (PP-BDR) property for a semialgebraic set. A semialgebraic set $S = \{\mathbf{x} : g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}$ is said to have this property if there exists a positive integer k such that if f is a linear polynomial positive over S , $f(\mathbf{x}) = \sigma_0(\mathbf{x}) + \sum_{i=1}^m \sigma_i(\mathbf{x})g_i$, where the σ_j are sums of squares and the degrees of σ_0 and of all $\sigma_i g_i$ are at most $2k$. We call the smallest such k the PP-BDR rank of S .

If we have an ideal $I = \langle f_1, \dots, f_m \rangle$ then its variety is a semi-algebraic set with generators $\pm f_i$, and so it makes sense to talk of the PP-BDR property for this set. This definition is very strongly linked to both the $(1, k)$ -sos property but its rank depends on the choice of the generators for the ideal, and we demand a representation only for positive polynomials and not for nonnegative ones. However, it is clear that if $\mathcal{V}_{\mathbb{R}}(I)$ has the PP-BDR property with rank k , I has theta-rank at most k .

In Corollary 2.9 we will prove that the converse of Lemma 1.4 is true if I is a *real radical ideal*. Such ideals occur frequently in applications.

Definition 1.7. Let I be an ideal in $\mathbb{R}[\mathbf{x}]$. Then I is

- (1) **radical** if it equals its **radical ideal**

$$\sqrt{I} := \{f \in \mathbb{R}[\mathbf{x}] : f^m \in I, m \in \mathbb{N} \setminus \{0\}\},$$

- (2) **real radical** if it equals its **real radical ideal**

$$\sqrt[\mathbb{R}]{I} := \{f \in \mathbb{R}[\mathbf{x}] : f^{2m} + g_1^2 + \dots + g_t^2 \in I, m \in \mathbb{N} \setminus \{0\}, g_1, \dots, g_t \in \mathbb{R}[\mathbf{x}]\},$$

- (3) and **zero-dimensional** if its **complex variety** $\mathcal{V}_{\mathbb{C}}(I) := \{\mathbf{x} \in \mathbb{C}^n : f(\mathbf{x}) = 0 \forall f \in I\}$ is finite.

Recall that given a set $S \subseteq \mathbb{R}^n$, its *vanishing ideal* in $\mathbb{R}[\mathbf{x}]$ is the ideal $\mathcal{I}(S) := \{f \in \mathbb{R}[\mathbf{x}] : f(\mathbf{s}) = 0 \forall \mathbf{s} \in S\}$. *Hilbert's Nullstellensatz* states that for an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$, $\sqrt{I} = \mathcal{I}(\mathcal{V}_{\mathbb{C}}(I))$ and the *Real Nullstellensatz* states that $\sqrt[\mathbb{R}]{I} = \mathcal{I}(\mathcal{V}_{\mathbb{R}}(I))$. Hence, $I \subseteq \sqrt{I} \subseteq \sqrt[\mathbb{R}]{I}$, and if I is real radical then it is also radical. See [15, Appendix 2].

We can now state Lovász's question using the above definitions and describe the example that inspired it.

Problem 1.8. [13, Problem 8.3] *Which ideals in $\mathbb{R}[\mathbf{x}]$ are $(1, 1)$ -sos? How about $(1, k)$ -sos?*

This question was motivated by a classical problem at the intersection of graph theory and combinatorial optimization that has received a great deal of attention in the optimization literature. Let $G = ([n], E)$ be a graph with vertex set $[n] = \{1, \dots, n\}$ and edge set E . A *stable set* in G is a set $T \subseteq [n]$ such that for all $i, j \in T$, $\{i, j\} \notin E$. The *maximum stable set problem* asks to find the stable set of largest cardinality in G . The largest size of a stable set in G is called the *stability number* of G , denoted as $\alpha(G)$. The maximum stable set problem on graphs can be modeled as follows. For each stable set $T \subseteq [n]$, let $\chi^T \in \{0, 1\}^n$ be its characteristic vector defined as $(\chi^T)_i = 1$ if $i \in T$ and $(\chi^T)_i = 0$ otherwise. Let $S_G \subseteq \{0, 1\}^n$ be the set of characteristic vectors of all stable sets in G . Then $\text{STAB}(G) := \text{conv}(S_G)$ is called the *stable set polytope* of G and the maximum stable set problem is the linear program $\max\{\sum x_i : \mathbf{x} \in \text{STAB}(G)\}$ whose optimal value is $\alpha(G)$. However, this formulation results in an NP-hard problem and so it is reasonable to look for good relaxations of it.

In [11], Lovász introduced a convex relaxation of $\text{STAB}(G)$, called the *theta body* of G , denoted as $\text{TH}(G)$. The *theta number* of G which the value of $\max\{\sum x_i : \mathbf{x} \in \text{TH}(G)\}$ is therefore an upper bound on $\alpha(G)$. The historical significance of $\text{TH}(G)$ is that it was the first example of a *semi-definite programming* (SDP) relaxation of a discrete optimization problem. Throughout this paper will always consider SDP's in the form

$$\max \langle \mathbf{x}, \mathbf{c} \rangle \quad \text{s.t.} \quad A_0 + \sum_{i=1}^m x_i A_i \succeq 0,$$

where $\mathbf{c} \in \mathbb{R}^m$ and the A_j are real symmetric matrices. These are generalizations of linear programs to the space of real symmetric matrices and can be solved to arbitrary precision in polynomial time. In particular, the theta number of G can be found in polynomial time in the input size of G . Recall that a graph G is *perfect* if and only if G has no induced odd cycles of length at least five or their complements. Lovász showed that $\text{STAB}(G) = \text{TH}(G)$ if and only if G is perfect. This connection yields the only known polytime algorithm for solving the maximum stable set problem on perfect graphs.

The body $\text{TH}(G)$ has many definitions (see [1, Chapter 9]) but the one relevant for this paper was observed by Lovász and appears without proof in [?]. Let $I_G := \langle x_j^2 - x_j \forall j \in [n], x_i x_j \forall \{i, j\} \in E \rangle \subseteq \mathbb{R}[\mathbf{x}]$. Then check

add references,
Bruce Shepherd
etc

that $\mathcal{V}_{\mathbb{R}}(I_G) = S_G$ and that I_G is both zero-dimensional and real radical. Lovász observed that

$$\text{TH}(G) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \ \forall \ \text{linear } f \text{ that is 1-sos mod } I_G\}.$$

By Definition 1.2 (1), $\text{TH}(G)$ is exactly the first theta body, $\text{TH}_1(I_G)$, of the ideal I_G . Hence I_G is TH_1 -exact (i.e., $\text{TH}_1(I_G) = \text{STAB}(G)$) if and only if G is perfect. Lovász observed that in fact, I_G is $(1, 1)$ -sos if and only if G is perfect. This motivated him to ask Problem 1.8 where he refers to a $(1, 1)$ -sos ideal as a *perfect ideal*. A $(1, 1)$ -sos ideal I would have the property that its first and simplest theta body, $\text{TH}_1(I)$, coincides with $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ which is a valuable property for linear optimization over $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. See Section 3 for more details on $\text{TH}(G)$ and the above assertions.

Our contributions in this paper. In Section 2 we prove that the theta body sequence of an ideal I is a version of a hierarchy of relaxations for the convex hull of a basic semialgebraic set, due to Lasserre, arising from the theory of moments [3, 4]. In particular, each theta body is the closure of the projection of a *spectrahedron* (feasible region of a semidefinite program), and an explicit representation of this type is possible using the *combinatorial moment matrices* introduced by Laurent [9]. The latter allows linear optimization over $\text{TH}_k(I)$ to be modeled as a standard semidefinite program which in turn admits polynomial time algorithms in the size of the input. When I is a real radical ideal, we prove that I is $(1, k)$ -sos if and only if I is TH_k -exact which impacts the later sections. Several situations in which the theta body sequence of an ideal I is guaranteed to converge to $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ are indicated.

In Section 3 we apply the results from Section 2 to well-known problems in combinatorial optimization to obtain a hierarchy of optimization problems that converge to the original one. In each case, we describe the theta body sequence of the ideals that arise, and establish when these ideals are $(1, 1)$ -sos.

The typical problem in combinatorial optimization requires the convexification of a finite set $S \subset \mathbb{R}^n$, say a subset of $\{0, 1\}^n$. The vanishing ideal of S is then real radical and zero-dimensional. In Section 4 we study Problem 1.8 for this class of ideals. Theorem 4.2 characterizes finite sets with TH_1 -exact (equivalently, $(1, 1)$ -sos) vanishing ideals via four equivalent conditions, answering Lovász’s question for such ideals. Several corollaries follow from this structure theorem. If S is finite and $\mathcal{I}(S)$ is $(1, 1)$ -sos then S is affinely equivalent to a subset of $\{0, 1\}^n$ (Corollary 4.5) and its convex hull can have at most 2^n facets (Theorem 4.7). If S is the vertex set of a down-closed 0/1-polytope in \mathbb{R}^n , then $\mathcal{I}(S)$ is $(1, 1)$ -sos if and only if $\text{conv}(S)$ is the stable set polytope of a perfect graph (Theorem 4.10 and Corollary 4.11). Families of finite sets in growing dimension with $(1, 1)$ -sos vanishing ideals are exhibited.

Finally, in Section 5, we give an intrinsic description of the first theta body, $\text{TH}_1(I)$, of an arbitrary polynomial ideal I in terms of the convex quadrics in I . This result leads to non-trivial examples of TH_1 -exact ideals with arbitrarily high-dimensional real varieties and reveals the algebraic-geometric structure of $\text{TH}_1(I)$. Analogous descriptions for the higher theta bodies remain open.

Acknowledgments. We thank Monique Laurent for several valuable comments and corrections on the first version of this paper, and for hosting João Gouveia at CWI for two weeks while this work was in progress. We also thank Ting Kei Pong for useful inputs. The organization of the paper benefitted considerably from the referee reports. We thank the referees for their thorough reading and useful comments.

2. THETA BODIES

In Definition 1.2 we introduced the k -th theta body of a polynomial ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ and observed that these bodies create a nested sequence of closed convex relaxations of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ with $\text{TH}_k(I) \supseteq \text{TH}_{k+1}(I) \supseteq \text{conv}(\mathcal{V}_{\mathbb{R}}(I))$.

Lasserre [3] and Parrilo [16, 18] have independently introduced hierarchies of semidefinite relaxations for polynomial optimization over basic semialgebraic sets in \mathbb{R}^n using results from real algebraic geometry and the theory of moments. In Corollary 2.7 we prove that $\text{TH}_k(I)$ is, up to closure, a version of Lasserre's relaxation of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ and hence is a semidefinite relaxation of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. We adopt the point of view in [15] of Lasserre's method to prove our results. Proposition 2.8 shows that if I is real radical then being $(1, k)$ -sos is equivalent to being TH_k -exact. Using results in the literature we point out several situations in which the theta body sequence of an ideal I is guaranteed to converge to $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. We also show how theta bodies can be handled computationally, by directly computing SOS decompositions in the quotient ring $\mathbb{R}[\mathbf{x}]/I$ [19], or dually, by using the combinatorial moment matrices from [9].

2.1. Lasserre's hierarchy and theta bodies.

Definition 2.1. Let I be an ideal in $\mathbb{R}[\mathbf{x}]$. The **quadratic module** of I is

$$\mathcal{M}(I) := \{s + I : s \text{ is a sum of squares in } \mathbb{R}[\mathbf{x}]\}.$$

The k -th **truncation** of $\mathcal{M}(I)$ is

$$\mathcal{M}_k(I) := \{s + I : s \text{ is } k\text{-sos}\}.$$

Both $\mathcal{M}(I)$ and $\mathcal{M}_k(I)$ are cones in the \mathbb{R} -vector space $\mathbb{R}[\mathbf{x}]/I$. Let $(\mathbb{R}[\mathbf{x}]/I)^*$ denote the set of linear forms on $\mathbb{R}[\mathbf{x}]/I$ and π_I be the projection map from $(\mathbb{R}[\mathbf{x}]/I)^*$ to \mathbb{R}^n defined as $\pi_I(y) = (y(x_1 + I), \dots, y(x_n + I))$. Also let $\mathcal{M}_k(I)^* \subseteq (\mathbb{R}[\mathbf{x}]/I)^*$ denote the *dual cone* to $\mathcal{M}_k(I)$, the set of all linear forms on $\mathbb{R}[\mathbf{x}]/I$ that are non-negative on $\mathcal{M}_k(I)$, and $\mathbb{R}[\mathbf{x}]_t$ the set of all polynomials of degree at most t .

Definition 2.2. For each $y \in (\mathbb{R}[\mathbf{x}]/I)^*$, let H_y be the symmetric bilinear form

$$\begin{aligned} H_y : \mathbb{R}[\mathbf{x}]/I \times \mathbb{R}[\mathbf{x}]/I &\longrightarrow \mathbb{R} \\ (f + I, g + I) &\longmapsto y(fg + I) \end{aligned}$$

and $H_{y,t}$ be the restriction of H_y to the subspace $(\mathbb{R}[\mathbf{x}]_t + I)$.

Recall that a symmetric bilinear form $H : V \times V \rightarrow \mathbb{R}$, where V is a \mathbb{R} -vector space, is *positive semidefinite* (written as $H \succeq 0$) if $H(v, v) \geq 0$ for all non-zero elements $v \in V$. Given a basis B of V , the matrix indexed by the elements of B with (b_i, b_j) -entry equal to $H(b_i, b_j)$ is called the *matrix representation* of H in the basis B . The form H is positive semidefinite if and only if its matrix representation in any basis is positive semidefinite.

Lemma 2.3. *Let I be an ideal and k a positive integer. Then*

$$\mathcal{M}_k(I)^* = \{y \in (\mathbb{R}[\mathbf{x}]/I)^* : H_{y,k} \succeq 0\}.$$

Proof: Note that $y \in \mathcal{M}_k(I)^*$ if and only if $y(s + I) \geq 0$ for all k -sos polynomials s . By linearity this is equivalent to $y(h^2 + I) \geq 0$ for all $h \in \mathbb{R}[\mathbf{x}]_k$ which is the definition of $H_{y,k}$ being positive semidefinite. \square

The original Lasserre relaxation introduced in [3], approximate the convex hull of a semialgebraic set $S = \{\mathbf{x} : g_i(\mathbf{x}) \geq 0, i = 1, \dots, m\}$ by considering all the forms $y \in \mathbb{R}[\mathbf{x}]^*$ that are nonnegative on all sos polynomials and products of sos polynomials by the generators g_i , up to some fixed degree, and then evaluating them at the monomials x_1, \dots, x_n . Both Lasserre [4] (for 0/1 sets), and [9] (more generally for finite varieties), propose to make these computations modulo an ideal to deal more economically with the presence of equalities in the definition of S . In this case, since we only are interested in the restrictions imposed by the ideal, we'll take this approach to define an approximation hierarchy.

Definition 2.4. Let I be an ideal and k a positive integer, and \mathcal{Y}_1 the hyperplane of all forms $y \in \mathbb{R}[\mathbf{x}]/I)^*$ verifying $y(1 + I) = 1$. The *k -th modified Lasserre relaxation* $Q_k(I)$ of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ is

$$Q_k(I) := \pi_I(\mathcal{M}_k(I)^* \cap \mathcal{Y}_1).$$

It's worth noticing that while $\mathcal{M}_k(I)^* \cap \mathcal{Y}_1$ is always closed, $Q_k(I)$ might not be, as it will be shown in example 2.13.

Lemma 2.5. *Let I be an ideal and k a positive integer then*

$$\text{conv}(\mathcal{V}_{\mathbb{R}}(I)) \subseteq Q_k(I).$$

Proof: For $\mathbf{p} \in \mathcal{V}_{\mathbb{R}}(I)$, consider $y^{\mathbf{p}} \in (\mathbb{R}[\mathbf{x}]/I)^*$ defined as $y^{\mathbf{p}}(f + I) := f(\mathbf{p})$. Then $y^{\mathbf{p}} \in \mathcal{M}_k(I)^*$ and $y^{\mathbf{p}}(1 + I) = 1$. Therefore, $\pi_I(y^{\mathbf{p}}) = \mathbf{p} \in Q_k(I)$ for all k , and $\text{conv}(\mathcal{V}_{\mathbb{R}}(I)) \subseteq Q_k(I)$ since $Q_k(I)$ is convex. \square

For each positive integer k we have $Q_{k+1}(I) \subseteq Q_k(I)$, so these bodies create a nested sequence of relaxations of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ as intended. We

now establish the relationship between $Q_k(I)$ and the k -th theta body of I (Definition 1.2).

Theorem 2.6. *Let I be an ideal and k a positive integer. Then*

$$\text{cl}(Q_k(I)) = \text{TH}_k(I).$$

Proof: We first prove that $\text{cl}(Q_k(I)) \subseteq \text{TH}_k(I)$. Since theta bodies are closed, it is enough to show $Q_k(I) \subseteq \text{TH}_k(I)$. Pick $\mathbf{p} \in Q_k(I)$ and $y \in \mathcal{M}_k(I)^* \cap \mathcal{Y}_1$ such that $\pi_I(y) = \mathbf{p}$. Let $f = a_0 + \sum_{i=1}^n a_i x_i$ and $f + I \in \mathcal{M}_k(I)$. Then

$$f(\mathbf{p}) = f(\pi_I(y)) = a_0 y(1 + I) + \sum_{i=1}^n a_i y(x_i + I) = y(f + I) \geq 0,$$

so $p \in \text{TH}_k(I)$.

To prove the reverse inclusion suppose that we have $\mathbf{p} \in \text{TH}_k(I)$ that is, such that $f(\mathbf{p}) \geq 0$ for all linear f such that $f + I \in \mathcal{M}_k(I)$. Choose a basis f_1, f_2, \dots of the \mathbb{R} -vector space I such that f_1, \dots, f_t is a basis for $I \cap \mathbb{R}[\mathbf{x}]_1$. Complete it to a basis of $\mathbb{R}[\mathbf{x}]$ by adding polynomials h_1, h_2, \dots such that $f_1, \dots, f_t, h_1, \dots, h_s$ is a basis for $\mathbb{R}[\mathbf{x}]_1$. We can then define a linear form y' in $\mathbb{R}[\mathbf{x}]^*$ by setting

$$y' \left(\sum_j \alpha_j f_j + \sum_i \beta_i h_i \right) := \sum_{i=1}^s \beta_i h_i(\mathbf{p}).$$

Since $y'(I) = \{0\}$ by definition, we can descend y' to a linear operator $y \in (\mathbb{R}[\mathbf{x}]/I)^*$ by simply setting $y(f + I) = y'(f)$ for all f . Since for $i = 1, \dots, t$, $f_i + I = -f_i + I = I \subseteq \mathcal{M}_k(I)$, we must have $f_i(\mathbf{p}) = 0$ for $i = 1, \dots, t$, since the f_i are linear. Let g be any other linear polynomial. Then $g(\mathbf{x}) = \sum_{j=1}^t \alpha_j f_j(\mathbf{x}) + \sum_{i=1}^s \beta_i h_i(\mathbf{x})$ for some constants α_j, β_i , and

$$y(g + I) = \sum_{i=1}^s \beta_i h_i(\mathbf{p}) = g(\mathbf{p}).$$

In particular, $\pi_I(y) = \mathbf{p}$, $y(1 + I) = 1$, and if $g + I \in \mathcal{M}_k(I)$, we have $y(g + I) = g(\mathbf{p}) \geq 0$. To conclude the proof it suffices to show that that $\pi_I(y)$ belongs to the closure of $Q_k(I)$. Note that

$$(2) \quad y \in (R_1 \cap \mathcal{M}_k(I))^* \cap \mathcal{Y}_1 = \text{cl}(\mathcal{M}_k(I)^* + R_1^\perp) \cap \mathcal{Y}_1$$

where $R_1 := (\mathbb{R}[\mathbf{x}]_1 + I)$ is the set of all cosets of I with a representative of degree at most one, and $R_1^\perp \subset (\mathbb{R}[\mathbf{x}]/I)^*$ the set of all forms that are zero there. This equality follows from the dual of the intersection of two cones being the closure of the sum of the duals where the closure is taken in the weak* topology of $(\mathbb{R}[\mathbf{x}]/I)^*$. We then have $y = \lim_\lambda (y_\lambda + z_\lambda)$ for

some nets $y_\lambda \in \mathcal{M}_k(I)^*$ and $z_\lambda \in R_1^\perp$. Since $z_\lambda(1+I) = 0$ we must have $\lim y_\lambda(1+I) = y(1+I) = 1$ and so

$$y = \lim_\lambda \frac{y_\lambda}{y_\lambda(1+I)} + z_\lambda \in \text{cl}(\mathcal{M}_k(I)^* \cap \mathcal{Y}_1 + R_1^\perp).$$

This means that the projection $\pi_I(y)$ must be in the set

$$\pi_I(\text{cl}(\mathcal{M}_k(I)^* \cap \mathcal{Y}_1 + R_1^\perp)) \subseteq \text{cl}(\pi_I(\mathcal{M}_k(I)^* \cap \mathcal{Y}_1 + R_1^\perp)),$$

and because R_1^\perp lies in the kernel of π_I , we conclude that $\pi_I(y)$ lies in

$$\text{cl}(\pi_I(\mathcal{M}_k(I)^* \cap \mathcal{Y}_1)) = \text{cl}(Q_k(I)).$$

□

Corollary 2.7. *Let I be any ideal in $\mathbb{R}[\mathbf{x}]$ such that $\mathbb{R}[\mathbf{x}]/I$ is finite dimensional and k be a positive integer. If there exists a linear polynomial g such that $g + I$ is in the relative interior of $\mathcal{M}_k(I)$ then $\text{TH}_k(I) = Q_k(I)$.*

Proof: Just note that if there exists such g , the closure in (2) can be dropped [21, Corollary 16.4.2]. Then we just have to observe that $(\mathcal{M}_k(I)^* + R_1^\perp) \cap \mathcal{Y}_1$ is the same as $\mathcal{M}_k(I)^* \cap \mathcal{Y}_1 + R_1^\perp$ and the argument used in the proof of the Theorem 2.6 will yield the intended result. □

2.2. Basic properties. Let I be an ideal and k a positive integer. Recall that if I is $(1, k)$ -sos (every linear polynomial which is nonnegative on $\mathcal{V}_\mathbb{R}(I)$ is also k -sos modulo the ideal), then I is TH_k -exact ($\text{TH}_k(I) = \text{cl}(\text{conv}(\mathcal{V}_\mathbb{R}(I)))$). The reverse implication can be false. Indeed, it was shown in Example 1.5 that the ideal $I = \langle x^2 \rangle$ is TH_1 -exact but not $(1, 1)$ -sos. However, as we show below, the reverse implication does hold when I is a real radical ideal.

Proposition 2.8. *Let I be an ideal and k a positive integer. If I is TH_k -exact and f is a linear polynomial non-negative on $\mathcal{V}_\mathbb{R}(I)$ then $f + I$ belongs to the closure of the set $\mathcal{M}_k(I)$.*

Proof:

Let f be a linear polynomial that is non-negative on $\mathcal{V}_\mathbb{R}(I)$ and suppose that $f + I$ it is not in $\text{cl}(\mathcal{M}_k(I))$. By the separation theorem, there exists $y \in (\mathbb{R}[\mathbf{x}]/I)^*$ such that $y(f) < 0$ and $y \in \mathcal{M}_k(I)^*$. Given any real number r we then have

$$0 \leq y((f + r + I)^2) = y(f^2 + I) + 2ry(f + I) + r^2y(1 + I).$$

Since $y(f + I) < 0$, the inequality implies $y(1 + I) > 0$, and we can scale y so that $y(1 + I) = 1$, thus $0 > y(f + I) = f(\pi_I(y))$. However, $y \in \mathcal{M}_k(I)^* \cap \mathcal{Y}_1$ so $\pi_I(y) \in \text{TH}_k(I) = \text{cl}(\text{conv}(\mathcal{V}_\mathbb{R}(I)))$, which implies by hypothesis that $f(\pi_I(y)) \geq 0$, so we reached a contradiction and $f + I$ belongs to $\text{cl}(\mathcal{M}_{I,k})$ as intended. □

Corollary 2.9. *If I is a real radical ideal then I is $(1, k)$ -sos if and only if I is TH_k -exact.*

Proof: Directly from Proposition 2.8 since $\mathcal{M}_k(I)$ is a closed set when I is a real radical ideal by [20, Prop 2.6]. \square

We close this subsection with a brief discussion of several instances in which the sequence of theta bodies of an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ is guaranteed to converge (finitely or asymptotically) to the closure of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$.

- (1) If $\mathcal{V}_{\mathbb{R}}(I)$ is finite the results in [9, 6] imply that I is TH_k -exact for some finite k . It will follow from Section 2.3 that in this case k can be bounded above by the maximum degree of a linear basis of $\mathbb{R}[\mathbf{x}]/I$. However, as in $I = \langle x^2 \rangle$, we cannot guarantee that I is $(1, k)$ -sos for any k even when I is zero-dimensional. If in addition to $\mathcal{V}_{\mathbb{R}}(I)$ being finite I is radical then in fact, I is $(1, k)$ -sos for finite k . Parrilo proved that in this case a polynomial f is nonnegative in $\mathcal{V}_{\mathbb{R}}(I)$ if and only if it is sos mod I [17], [10, Theorem 2.4]. The proof uses a set of interpolating polynomials of $\mathcal{V}_{\mathbb{R}}(I)$ to write the sos representations. Since interpolators can be constructed to have degree at most $|\mathcal{V}_{\mathbb{C}}(I)| - 1$, I is $(1, k)$ -sos for $k \leq |\mathcal{V}_{\mathbb{C}}(I)| - 1$. Better bounds follow from Proposition ?? in some instances.
- (2) When $\mathcal{V}_{\mathbb{R}}(I)$ is not finite but is still compact, Schmüdgen's Positivstellensatz [15, Chapter 3] implies that the theta body sequence of I converges (perhaps not finitely) to $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ i.e., the intersection of all theta bodies TH_k will exactly match this closure. Under strong additional requirements on smoothness and curvature, results of Helton and Nie [2] can guarantee that I is TH_k -exact for finite k .
- (3) If $\mathcal{V}_{\mathbb{R}}(I)$ is not even compact the study of the hierarchy becomes much harder. Scheiderer [15, Chapter 2] has identified several instances of ideals I with $\mathcal{V}_{\mathbb{R}}(I)$ not necessarily compact but of dimension at most two for which every $f \geq 0 \text{ mod } I$ is sos mod I . In all these cases, the theta body sequence of I converges to $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$.

The results of Schmüdgen and Scheiderer mentioned above fit within a general framework in real algebraic geometry that is concerned with when an arbitrary polynomial f that is positive or non-negative over a basic semi-algebraic set can be written as a sum of squares modulo certain algebraic objects defined by the set. We only care about real varieties and whether linear polynomials that are non-negative over them are sos mod their ideals. Therefore, it is natural that there are ideals I that are TH_k -exact or $(1, k)$ -sos for which there are non-linear polynomials f such that $f \geq 0 \text{ mod } I$ but f is not sos mod I . For instance, the proof of Theorem 5.6 implies that the ideal $J_n := \langle \sum_{i=1}^n x_i^2 - 1 \rangle$ is $(1, 1)$ -sos for all n , but a result of Scheiderer [15,

Theorem 2.6.3] implies that when $n \geq 4$, there is always some non-linear f nonnegative on the sphere that is not sos mod J_n .

2.3. Combinatorial moment matrices. For computations with theta bodies we must work with the truncated quadratic module $\mathcal{M}_k(I)$. This requires computing sum of squares decompositions over the quotient $\mathbb{R}[\mathbf{x}]/I$, as described in [19], or dually, using the combinatorial moment matrices introduced by Laurent in [9]. We describe the latter viewpoint here, as it more directly connects with the theta bodies. Let $\mathcal{B} = \{f_0 + I, f_1 + I, \dots\}$ be a basis for $\mathbb{R}[\mathbf{x}]/I$. Define $\deg(f_i + I) := \min_{f \in f_i + I} \deg f$. For a positive integer k , let $\mathcal{B}_k := \{f_l + I \in \mathcal{B} : \deg(f_l + I) \leq k\}$, and $\mathbf{f}_k := (f_l + I : f_l + I \in \mathcal{B}_k)$ denote the vector of elements in \mathcal{B}_k . We may assume that the elements of \mathcal{B} are indexed in order of increasing degree.

Let $\lambda^{(g+I)} := (\lambda_l^{(g+I)})$ be the vector of coordinates of $g + I$ with respect to \mathcal{B} . Note that $\lambda^{(g+I)}$ has only finitely many non-zero coordinates.

Definition 2.10. Let $\mathbf{y} \in \mathbb{R}^{\mathcal{B}}$. Then the **combinatorial moment matrix** $M_{\mathcal{B}}(\mathbf{y})$ is the (possibly infinite) matrix indexed by \mathcal{B} whose (i, j) entry is

$$\lambda^{(f_i f_j + I)} \cdot \mathbf{y} = \sum \lambda_l^{(f_i f_j + I)} y_l.$$

The **k -th-truncated combinatorial moment matrix** $M_{\mathcal{B}_k}(\mathbf{y})$ is the finite (upper left principal) submatrix of $M_{\mathcal{B}}(\mathbf{y})$ indexed by \mathcal{B}_k .

Although only a finite number of the components in $\lambda^{(f_i f_j + I)}$ are non-zero, for practical purposes we need to control exactly which indices can be non-zero. One way to do this is by choosing \mathcal{B} such that if $f + I$ has degree k then $f + I \in \text{span}(\mathcal{B}_k)$. This is true for instance if \mathcal{B} is the set of standard monomials of a term order that respects degree. If \mathcal{B} has this property then the matrix $M_{\mathcal{B}_k}(\mathbf{y})$ only depends on the entries of y indexed by \mathcal{B}_{2k} . These definitions allow a practical characterization of theta bodies, up to closure.

Theorem 2.11. For each positive integer k ,

$$\text{proj}_{\mathbb{R}^{\mathcal{B}_1}} \{\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} : M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0, y_0 = 1\} = \mathbf{f}_1(Q_k(I)),$$

where $\text{proj}_{\mathbb{R}^{\mathcal{B}_1}}$ is the projection onto the coordinates indexed by \mathcal{B}_1 , and \mathbf{f}_1 is the function that sends a point \mathbf{p} in \mathbb{R}^n to the vector $(f_i(\mathbf{p}))_{f_i + I \in \mathcal{B}_1}$.

Proof: Note that we can see any $\mathbf{y} = (y_i) \in \mathbb{R}^{\mathcal{B}_{2k}}$ as an operator $\bar{y} \in (\mathbb{R}[\mathbf{x}]/I)^*$ by setting $\bar{y}(f_i + I) = y_i$ if $f_i + I \in \mathcal{B}_{2k}$ and zero otherwise. But then $M_{\mathcal{B}_k}(\mathbf{y})$ is simply the matrix representation of $H_{\bar{y}, k}$ in the basis \mathcal{B} , and we get that $\text{proj}_{\mathbb{R}^{\mathcal{B}_1}} \{\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} : M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0\}$ equals

$$\{(\bar{y}(f_i + I))_{\mathcal{B}_1} : \bar{y} \in (\mathbb{R}[\mathbf{x}]/I)^*, H_{\bar{y}, k} \succeq 0\},$$

since because of the assumptions we made, if $\deg f_i + I, \deg f_j + I \leq k$ then $\bar{y}(f_i f_j + I)$ depends only on the value of \bar{y} on \mathcal{B}_{2k} . Furthermore,

$$(\bar{y}(f_i + I))_{\mathcal{B}_1} = (f_i(\pi_I(\bar{y})))_{\mathcal{B}_1} =: \mathbf{f}_1(\pi_I(\bar{y}))$$

so by Theorem 2.3, $\text{proj}_{\mathbb{R}^{\mathcal{B}_1}} \{\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} : M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0, y_0 = 1\} = \mathbf{f}_1(Q_k(I))$.
 \square

Corollary 2.12. *Suppose $\mathcal{B}_1 = \{1 + I, x_1 + I, \dots, x_n + I\}$ and denote by y_0, y_1, \dots, y_n the first $n + 1$ coordinates of $\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}}$, then*

$$Q_k(I) = \{(y_1, \dots, y_n) : \mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} \text{ with } M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0 \text{ and } y_0 = 1\}.$$

Note that Corollary 2.12 implies that optimizing a linear objective function over $Q_k(I)$, hence over TH_k , is an SDP and so can be solved to arbitrary precision in polynomial time, provided we have a way to explicitly write the combinatorial moment matrix.

Example 2.13. Consider again the ideal $I = \langle x_1^2 x_2 - 1 \rangle \subset \mathbb{R}[x, y]$ from Example 1.3. We saw that $\text{conv}(\mathcal{V}_{\mathbb{R}}(I)) = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\}$ was not closed but $\text{TH}_2(I) = \text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I))) = \{(x_1, x_2) \in \mathbb{R}^2 : y \geq 0\}$ and I was $(1, 2)$ -sos. Note that $\mathcal{B} = \bigcup_{k \in \mathbb{N}} \{x_1^k + I, x_2^k + I, x_1 x_2^k + I\}$ is a degree-compatible monomial basis for $\mathbb{R}[x, y]/I$ for which

$$\mathcal{B}_4 = \{1, x_1, x_2, x_1^2, x_1 x_2, x_2^2, x_1^3, x_1 x_2^2, x_2^3, x_1^4, x_1 x_2^3, x_2^4\} + I.$$

The combinatorial moment matrix $M_{\mathcal{B}_2}(\mathbf{y})$ for $\mathbf{y} = (1, y_1, \dots, y_{11}) \in \mathbb{R}^{\mathcal{B}_4}$ is

$$\begin{array}{c} 1 \quad x_1 \quad x_2 \quad x_1^2 \quad x_1 x_2 \quad x_2^2 \\ \begin{pmatrix} 1 & y_1 & y_2 & y_3 & y_4 & y_5 \\ y_1 & y_3 & y_4 & y_6 & 1 & y_7 \\ y_2 & y_4 & y_5 & 1 & y_7 & y_8 \\ y_3 & y_6 & 1 & y_9 & y_1 & y_2 \\ y_4 & 1 & y_7 & y_1 & y_2 & y_{10} \\ y_5 & y_7 & y_8 & y_2 & y_{10} & y_{11} \end{pmatrix} \end{array}$$

If $M_{\mathcal{B}_2}(\mathbf{y}) \succeq 0$, then the principal minor indexed by x_1 and $x_1 x_2$ is non-negative which implies that $y_2 y_3 \geq 1$ and so in particular, $y_2 \neq 0$ for all $\mathbf{y} \in Q_2(I)$. However, since $Q_2(I) \supseteq \text{conv}(\mathcal{V}_{\mathbb{R}}(I)) = \{(x_1, x_2) \in \mathbb{R}^2 : y > 0\}$, it must be that $Q_2(I) = \text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ which shows that $Q_2(I) \subset \mathbb{R}^2$ is not closed.

Remark 2.14. Example 2.13 can be modified to show that these relaxations may not be closed even if $\mathcal{V}_{\mathbb{R}}(I)$ is finite. To see this, choose sufficiently many pairs of points $(\pm t, 1/t^2)$ on the curve $x^2 y = 1$ to form a set S such that the ideal $\mathcal{I}(S)$ has a monomial basis \mathcal{B}' in which \mathcal{B}'_4 equals the \mathcal{B}_4 from above. For instance, $S = \{(\pm t, 1/t^2) : t = 1, \dots, 7\}$ will work. Then $Q_2(\mathcal{I}(S))$ coincides with $Q_2(I)$ computed above and so is not a closed set.

In the particular case of vanishing ideals of 0/1 vectors, the most common setting for combinatorial optimization, the following proposition shows that the closure in Proposition 2.6 is not needed. Remark 2.14 implies that this result does not extend to arbitrary finite sets.

Proposition 2.15. *If S is a set of 0/1 vectors in \mathbb{R}^n and $I = \mathcal{I}(S)$ then for all integers $k > 0$, $\text{TH}_k(I) = Q_k(I)$.*

Proof: By Corollary 2.7 it is enough to show that there is a linear polynomial $g \in \mathbb{R}[\mathbf{x}]$ such that $g \equiv \mathbf{f}_k^t A \mathbf{f}_k \pmod I$ for a *positive definite* matrix A and some basis of $\mathbb{R}[\mathbf{x}]/I$ with respect to which \mathbf{f}_k was determined. Let \mathcal{B} be a monomial basis for $\mathbb{R}[\mathbf{x}]/I$ and $\mathcal{B}_k = \{1, p_1, \dots, p_l\} + I$. Let $\mathbf{c} \in \mathbb{R}^{m+l}$ be the vector with all entries -2 , and $D \in \mathbb{R}^{(m+l) \times (m+l)}$ be the diagonal matrix with all entries equal to four. Since $x_i^2 \equiv x_i \pmod I$ for $i = 1, \dots, m$ and \mathcal{B} is a monomial basis, for any $f + I \in \mathcal{B}$, $f \equiv f^2 \pmod I$. Therefore,

$$l + 1 \equiv \mathbf{f}_k^t \begin{bmatrix} l + 1 & \mathbf{c}^t \\ \mathbf{c} & D \end{bmatrix} \mathbf{f}_k \pmod I,$$

and it is enough to prove that the square matrix in this representation is positive definite. But since D is clearly positive definite it is enough to note that its Schur complement $(l + 1) - \mathbf{c}^t D^{-1} \mathbf{c} = 1$ is positive. \square

3. COMBINATORIAL EXAMPLES

In this section, we apply the methodology developed in Section 2 to some well known examples from combinatorial optimization. The typical problem in this area involves finding an object of maximum *weight* from a finite collection of combinatorial objects. In the usual *NP*-hard instance from combinatorial optimization, we have the integer program $\max \{\mathbf{c}\mathbf{x} : \mathbf{x} \in S\}$ where $S \subset \mathbb{Z}^n$ is the collection of characteristic vectors of the finitely many objects being optimized over and \mathbf{c} is the weight vector. This discrete optimization problem is implicitly the linear program $\max \{\mathbf{c}\mathbf{x} : \mathbf{x} \in \text{conv}(S)\}$ and it is useful to know relaxations of $\text{conv}(S)$ that yield approximations of the optimal value of the original problem in polynomial time in the size of the input data. A good example is the maximum stable set problem in a graph $G = ([n], E)$, described in the introduction, which was modeled as the linear program

$$\max \left\{ \sum_{i=1}^n x_i : \mathbf{x} \in S_G \right\} = \max \left\{ \sum_{i=1}^n x_i : \mathbf{x} \in \text{conv}(S_G) = \text{STAB}(G) \right\}.$$

If $I = \mathcal{I}(S)$ is the vanishing ideal of S , then the theta bodies of I developed in Section 2 provide a hierarchy of convex relaxations of $\text{conv}(S)$. Recall that the theta body, $\text{TH}(G)$, of a graph G is an instance of a “good” relaxation in the sense that one can optimize a linear function over $\text{TH}(G)$ in time polynomial in the size of G . By Lovász’s observation, $\text{TH}(G) = \text{TH}_1(I_G)$ where I_G is the vanishing ideal of S_G . In this section we justify this observation and revisit the stable set problem in a more general context and provide further results about its theta body hierarchy. We then compute the theta body relaxations for two different formulations of the *maximum cut problem* in a graph which results in new semidefinite relaxations for this well studied problem in combinatorial optimization.

3.1. Examples from simplicial complexes. Let Δ be an *abstract simplicial complex* (or *independence system*) with vertex set $[n]$ recorded as a collection of subsets of $[n]$, called the *faces* of Δ . The *Stanley-Reisner ideal* of Δ is the ideal J_Δ generated by the squarefree monomials $x_{i_1}x_{i_2}\cdots x_{i_k}$ such that $\{i_1, i_2, \dots, i_k\} \subseteq [n]$ is not a face of Δ . If $I_\Delta := J_\Delta + C$ where $C = \langle x_i^2 - x_i : i \in [n] \rangle$, then $\mathcal{V}_\mathbb{R}(I_\Delta) = \{\mathbf{s} \in \{0, 1\}^n : \text{support}(\mathbf{s}) \in \Delta\}$. For $T \subseteq [n]$, let $\mathbf{x}^T := \prod_{i \in T} x_i$. Then $\mathcal{B} := \{\mathbf{x}^T : T \in \Delta\} + I_\Delta$ is a basis for $\mathbb{R}[\mathbf{x}]/I_\Delta$ containing $1 + I_\Delta, x_1 + I_\Delta, \dots, x_n + I_\Delta$. Therefore, by Corollary 2.12 and Proposition 2.15, the k -th theta body of I_Δ is

$$\text{TH}_k(I_\Delta) = \text{proj}_{y_1, \dots, y_n} \{\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} : M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0, y_0 = 1\}.$$

Since \mathcal{B} is in bijection with the faces of Δ , and $x_i^2 - x_i \in I_\Delta$ for all $i \in [n]$, the theta body can be written explicitly as follows:

$$\text{TH}_k(I_\Delta) = \left\{ \mathbf{y} \in \mathbb{R}^n : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{|\mathcal{B}_k| \times |\mathcal{B}_k|} \text{ such that} \\ M_{\emptyset\emptyset} = 1, \\ M_{\emptyset\{i\}} = M_{\{i\}\emptyset} = M_{\{i\}\{i\}} = y_i \\ M_{UU'} = 0 \text{ if } U \cup U' \notin \Delta \\ M_{UU'} = M_{WW'} \text{ if } U \cup U' = W \cup W' \end{array} \right\}.$$

This yields a hierarchy of semidefinite programming relaxations of linear programs over $\mathcal{V}_\mathbb{R}(I_\Delta)$. If the dimension of Δ is $d - 1$ (i.e., the largest faces in Δ have size d), then I_Δ is $(1, d)$ -sos and therefore, TH_d -exact since all elements of \mathcal{B} have degree at most d . However, the theta-rank of I_Δ could be much less than d . We examine two specific cases of the above set up.

3.1.1. Stable sets in graphs. Recall that the maximum stable set problem on a graph $G = ([n], E)$ is the linear program $\max \{\sum_{i=1}^n x_i : \mathbf{x} \in \mathcal{V}_\mathbb{R}(I_G)\}$ where

$$I_G := \langle x_i^2 - x_i (\forall i \in [n]), x_i x_j (\forall \{i, j\} \in E) \rangle$$

is the zero-dimensional, real radical vanishing ideal of S_G , the set of characteristic vectors of the stable sets in G . The ideal I_G is I_Δ for the simplicial complex Δ on $[n]$, of stable sets in G . Therefore, $\mathcal{B} := \{\mathbf{x}^U + I_G : U \text{ stable set in } G\}$ is a basis of $\mathbb{R}[\mathbf{x}]/I_G$ and

$$\text{TH}_k(I_G) = \left\{ \mathbf{y} \in \mathbb{R}^n : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{|\mathcal{B}_k| \times |\mathcal{B}_k|} \text{ such that} \\ M_{\emptyset\emptyset} = 1, \\ M_{\emptyset\{i\}} = M_{\{i\}\emptyset} = M_{\{i\}\{i\}} = y_i \\ M_{UU'} = 0 \text{ if } U \cup U' \text{ is not stable in } G \\ M_{UU'} = M_{WW'} \text{ if } U \cup U' = W \cup W' \end{array} \right\}.$$

In particular,

$$\text{TH}_1(I_G) = \left\{ \mathbf{y} \in \mathbb{R}^n : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{(n+1) \times (n+1)} \text{ such that} \\ M_{00} = 1, \\ M_{0i} = M_{i0} = M_{ii} = y_i \forall i \in [n] \\ M_{ij} = 0 \forall \{i, j\} \in E \end{array} \right\}.$$

In [11] Lovász introduced the theta body, $\text{TH}(G)$, of G [1, Chapter 9]. There are multiple descriptions of $\text{TH}(G)$, but the one in [14, Lemma 2.17], for instance, is exactly the description of $\text{TH}_1(I_G)$ given above and hence $\text{TH}(G) = \text{TH}_1(I_G)$. The theta bodies of I_G extend $\text{TH}(G)$ to a hierarchy of semidefinite relaxations for the stable set problem. Using Corollary 2.9 and the fact that $\text{TH}(G) = \text{TH}_1(I_G)$, we get an extension of the following well-known set of equivalences for a perfect graph.

Theorem 3.1. [1, Chapter 9] *The following are equivalent for a graph G .*

- (1) G is perfect.
- (2) $\text{STAB}(G) = \text{TH}(G)$.
- (3) $\text{TH}(G)$ is a polytope.
- (4) The complement \overline{G} of G is perfect.

Corollary 3.2. *The following are equivalent for a graph G .*

- (1) G is perfect.
- (2) I_G is TH_1 -exact and $(1, 1)$ -sos.
- (3) $I_{\overline{G}}$ is TH_1 -exact and $(1, 1)$ -sos.

Since no monomial in the basis \mathcal{B} of $\mathbb{R}[\mathbf{x}]/I_G$ has degree larger than $\alpha(G)$, for any G , I_G is $(1, \alpha(G))$ -sos and $\text{STAB}(G) = \text{TH}_{\alpha(G)}(I_G)$. However, for many non-perfect graphs the theta-rank of I_G can be a lot smaller than $\alpha(G)$. For instance if G is a $(2k + 1)$ -cycle, then $\alpha(G) = k$ while Proposition 3.4 below shows that the theta-rank of I_G is two.

Theorem 3.3. [24, Corollary 65.12a] *If $G = ([n], E)$ is an odd cycle with $n \geq 5$, then $\text{STAB}(G)$ is determined by the following inequalities:*

$$x_i \geq 0 \quad \forall i \in [n], \quad 1 - \sum_{i \in K} x_i \geq 0 \quad \forall \text{cliques } K \text{ in } G, \quad \alpha(G) - \sum_{i \in [n]} x_i \geq 0.$$

Proposition 3.4. *If G is an odd cycle with at least five vertices, then I_G is $(1, 2)$ -sos and therefore, TH_2 -exact.*

Proof: Let $n = 2k + 1$ and G be an n -cycle. Then $I_G = \langle x_i^2 - x_i, x_i x_{i+1} \mid \forall i \in [n] \rangle$ where $x_{n+1} = x_1$. Therefore, $(1 - x_i)^2 \equiv 1 - x_i$ and $(1 - x_i - x_{i+1})^2 \equiv 1 - x_i - x_{i+1} \pmod{I_G}$. This implies that, mod I_G ,

$$p_i^2 := ((1 - x_1)(1 - x_{2i} - x_{2i+1}))^2 \equiv p_i = 1 - x_1 - x_{2i} - x_{2i+1} + x_1 x_{2i} + x_1 x_{2i+1}.$$

Summing over $i = 1, \dots, k$, we get

$$\sum_{i=1}^k p_i^2 \equiv k - kx_1 - \sum_{i=2}^{2k+1} x_i + \sum_{i=3}^{2k} x_1 x_i \pmod{I_G}$$

since $x_1 x_2$ and $x_1 x_{2k+1}$ lie in I_G . Define $g_i := x_1(1 - x_{2i+1} - x_{2i+2})$. Then $g_i^2 - g_i \in I_G$ and mod I_G we get that

$$\sum_{i=1}^{k-1} g_i^2 \equiv (k-1)x_1 - \sum_{i=3}^{2k} x_1 x_i, \quad \text{which implies} \quad \sum_{i=1}^k p_i^2 + \sum_{i=1}^{k-1} g_i^2 \equiv k - \sum_{i=1}^{2k+1} x_i.$$

To prove that I_G is $(1, 2)$ -sos it suffices to show that the left hand sides of the inequalities in the description of $\text{STAB}(G)$ in Theorem 3.3 are 2-sos mod I_G since by Farkas Lemma [23], all other linear inequalities that are non-negative over S_G are non-negative real combinations of a set of inequalities defining $\text{STAB}(G)$. Clearly, $x_i \equiv x_i^2 \pmod{I_G}$ for all $i \in [n]$ and one can check that for each clique K , $(1 - \sum_{i \in K} x_i) \equiv (1 - \sum_{i \in K} x_i)^2 \pmod{I_G}$. The previous paragraph shows that $k - \sum_{i=1}^{2k+1} x_i$ is also 2-sos mod I_G . \square

An induced odd cycle C_{2k+1} in G , yields the well-known *odd cycle inequality* $\sum_{i \in C_{2k+1}} x_i \leq \alpha(C_{2k+1}) = k$ that is satisfied by S_G [1, Chapter 9]. Proposition 3.4 implies that for any graph G , $\text{TH}_2(I_G)$ satisfies all odd cycle inequalities from G since every stable set U in G restricts to a stable set in an induced odd cycle in G . This general result can also be proved indirectly using results from [14] and [7]. The direct arguments used in the proof of Proposition 3.4 are examples of the algebraic *inference rules* outlined by Lovász in [12]. Similarly, one can also show that other well-known classes of inequalities such as the *odd antihole* and *odd wheel* inequalities [1, Chapter 9] are also valid for $\text{TH}_2(I_G)$. Schoenebeck [22] has recently shown that there is no constant k such that $\text{STAB}(G) = \text{TH}_k(I_G)$ for all graphs G (as expected, unless $\text{P}=\text{NP}$). A construction remains open.

The Lasserre relaxations of $\text{STAB}(G)$ have been studied in the literature. These relaxations are usually set up from the following initial linear programming relaxation of $\text{STAB}(G)$:

$$\text{FRAC}(G) := \{\mathbf{x} \in \mathbb{R}^n : x_i \geq 0 \forall i \in [n], 1 - x_i - x_j \geq 0 \forall \{i, j\} \in E\}.$$

Note that $S_G = \text{FRAC}(G) \cap \{0, 1\}^n$. The k -th Lasserre relaxation of $\text{STAB}(G)$ (see [4], [7]) uses both the ideal $C = \langle x_i^2 - x_i : i \in [n] \rangle$ and the inequality system describing $\text{FRAC}(G)$, whereas in the theta body formulation, $\text{TH}_k(I_G)$, there is only the ideal I_G and no inequalities. Despite this difference, [7, Lemma 20] proves that the usual Lasserre hierarchy is exactly our theta body hierarchy for the stable set problem.

3.1.2. Cuts in graphs. Given an undirected connected graph $G = ([n], E)$ and a partition of its vertex set $[n]$ into two parts V_1 and V_2 , the set of edges $\{i, j\} \in E$ such that exactly one of i or j is in V_1 and the other in V_2 is the *cut* in G induced by the partition (V_1, V_2) . The cuts in G are in bijection with the 2^{n-1} distinct partitions of $[n]$ into two sets. The (weighted) maximum cut problem in G is the problem of finding the largest cut in G (with respect to the weight vector). This problem is NP-hard and has received a great deal of attention in the literature. One of the celebrated results is an approximation algorithm for the maximum cardinality cut problem, due to Goemans and Williamson [?], that guarantees a cut of size at least 0.878 of the optimal cut. It relies on a simple semidefinite programming relaxation of the problem.

In the rest of this section, we give two different formulations of the weighted max cut problem and exhibit the theta body relaxations in each

case. Let

$$SG := \{\chi^F : F \subseteq E \text{ is contained in a cut of } G\} \subseteq \{0, 1\}^E.$$

Then the max cut problem with non-negative weights w_e on the edges $e \in E$ is the integer program $\max \{\sum_{e \in E} w_e x_e : \mathbf{x} \in SG\}$, and the vanishing ideal

$$\mathcal{I}(SG) = \langle x_e^2 - x_e (\forall e \in E), \mathbf{x}^T (\forall \text{ odd cycles } T \text{ in } G) \rangle.$$

This again fits the simplicial complex setting with Δ equal to the complex of edge sets of G without odd cycles. A basis of $\mathbb{R}[\mathbf{x}]/\mathcal{I}(SG)$ is

$$\mathcal{B} = \{\mathbf{x}^U + \mathcal{I}(SG) : U \subseteq E \text{ does not contain an odd cycle in } G\}$$

and hence its elements correspond to certain subsets of E . Therefore,

$$\text{TH}_k(\mathcal{I}(SG)) = \left\{ \mathbf{y} \in \mathbb{R}^E : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{|\mathcal{B}_k| \times |\mathcal{B}_k|} \text{ such that} \\ M_{\emptyset\emptyset} = 1, \\ M_{\emptyset\{i\}} = M_{\{i\}\emptyset} = M_{\{i\}\{i\}} = y_i \\ M_{UU'} = 0 \text{ if } U \cup U' \text{ has an odd cycle} \\ M_{UU'} = M_{WW'} \text{ if } U \cup U' = W \cup W' \end{array} \right\}.$$

In particular,

$$\text{TH}_1(\mathcal{I}(SG)) = \left\{ \mathbf{y} \in \mathbb{R}^E : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{(|E|+1) \times (|E|+1)} \text{ such that} \\ M_{00} = 1, \\ M_{0e} = M_{e0} = M_{ee} = y_e \forall e \in E \end{array} \right\}.$$

Proposition 3.5. *The ideal $\mathcal{I}(SG)$ is TH_1 -exact if and only if G is a bipartite graph.*

Since the maximum degree of a monomial in \mathcal{B} is the size of the max cut in G , the theta-rank of $\mathcal{I}(SG)$ is bounded from above by the size of the max cut in G . We state some related observations.

Proposition 3.6. *There is no constant k such that $\mathcal{I}(SG)$ is TH_k -exact for all graphs G .*

Proof: Let G be a $(2k+1)$ -cycle. Then $\text{TH}_k(\mathcal{I}(SG)) \neq \text{conv}(SG)$ since the linear constraint imposed by the cycle in the definition of $\text{TH}_k(\mathcal{I}(SG))$ will not appear in theta bodies of index k or less. \square

Note that for any graph G , $\text{TH}_1(\mathcal{I}(SG))$ is the unit cube in \mathbb{R}^E which may not be equal to $\text{conv}(SG)$. This stands in contrast to the case of stable sets for which $\text{TH}_1(I_G)$ is a polytope if and only if $\text{TH}_1(I_G) = \text{STAB}(G)$. If G is a $(2m+1)$ -cycle then in fact, $\text{TH}_1(\mathcal{I}(SG)) = \text{TH}_2(\mathcal{I}(SG)) = \dots = \text{TH}_m(\mathcal{I}(SG))$ is the unit cube in \mathbb{R}^E even though $\text{TH}_m(\mathcal{I}(SG)) \neq \text{conv}(SG)$. Therefore, the k -th theta body need not strictly contain the $(k+1)$ -th body in the theta body hierarchy.

3.2. Cuts revisited: a new formulation. The integer programming model of the weighted max cut problem in Section 3.1.2 is not standard. In [?], we study the theta bodies of the standard formulation of the max cut problem which yields a new canonical series of semidefinite relaxations for max cut that directly exploits the structure of the graph. In this section we summarize the main results in [?] with references but without proofs or details.

Each cut C in an undirected graph $G = ([n], E)$ can be recorded by its *cut vector* $\chi^C \in \{\pm 1\}^E$ with $\chi_{\{i,j\}}^C = 1$ if $\{i,j\} \notin C$ and $\chi_{\{i,j\}}^C = -1$ if $\{i,j\} \in C$. Let E_n denote the edge set of the complete graph K_n , and π_E be the natural projection map from \mathbb{R}^{E_n} to \mathbb{R}^E . The *cut polytope* of G is

$$\text{CUT}(G) := \text{conv}\{\chi^C : C \text{ is a cut in } G\} \subseteq \mathbb{R}^E = \pi_E(\text{CUT}(K_n)),$$

and the weighted max cut problem can be formulated as

$$\max \left\{ \frac{1}{2} \sum_{e \in E} w_e (1 - x_e) : \mathbf{x} \in \text{CUT}(G) \right\},$$

for some weights $w_e \in \mathbb{R}$.

Definition 3.7. Let $G = ([n], E)$ be an undirected graph, $T \subseteq [n]$ have even cardinality and $F \subseteq E$. Then F is called a **T -join** if T is exactly the set of vertices of odd-degree in the subgraph of G induced by F .

Let $\mathbb{R}E := \mathbb{R}[x_e : e \in E]$ and $IG \subseteq \mathbb{R}E$ be the vanishing ideal of the cut vectors in G . Identify a subgraph F of G with its set of edges $E(F) \subseteq E$. For a subgraph F in G , let \mathbf{x}^F denote the squarefree monomial

$$\prod_{e \in E(F)} x_e \in \mathbb{R}E.$$

Theorem 3.8. [?, Theorem 3.4] *The ideal IG is generated by the binomials $x_e^2 - 1$, $\forall e \in E$ and $1 - \mathbf{x}^D$, \forall chordless circuits D in G . For an even subset $T \subseteq [n]$, let F_T be a T -join in G and let $\mathcal{B} := \{\mathbf{x}^{F_T} + IG : T \subseteq [n], |T| \text{ even}\}$. Then \mathcal{B} is a basis of $\mathbb{R}E/IG$.*

Set \mathcal{B} as in Theorem 3.8 with the additional stipulation that for each $T \subseteq [n]$ of even cardinality, the T -join F_T chosen to represent a coset in \mathcal{B} is one with smallest number of edges among all T -joins in G . Index a coset $\mathbf{x}^F + IG \in \mathcal{B}$ by F and let $F_1 \Delta F_2$ denote the symmetric difference of the sets F_1 and F_2 . For each positive integer k , let \mathcal{B}_k be the elements of \mathcal{B} of degree at most k which are exactly the cosets in \mathcal{B} given by T -joins of size at most k . When G is a simple graph (i.e., G has no loops or parallel edges) we show in Section 3 of [?] that the k -th theta body, $\text{TH}_k(IG)$, of IG is

$$\left\{ \mathbf{y} \in \mathbb{R}^E : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{|\mathcal{B}_k| \times |\mathcal{B}_k|} \text{ such that} \\ M_{\emptyset, \emptyset} = 1 \\ M_{F_1, F_2} = M_{F_3, F_4} \text{ if } F_1 \Delta F_2 \Delta F_3 \Delta F_4 \text{ is a cycle in } G \end{array} \right\}.$$

In particular,

$$\text{TH}_1(IG) = \left\{ \mathbf{y} \in \mathbb{R}^E : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{(|E|+1) \times (|E|+1)} \text{ such that} \\ M_{\emptyset, \emptyset} = M_{e, e} = 1 \ \forall e \in E \\ M_{e, f} = M_{\emptyset, g} \text{ if } \{e, f, g\} \text{ is a triangle in } G, \\ M_{e, f} = M_{g, h} \text{ if } \{e, f, g, h\} \text{ is a circuit in } G \end{array} \right\}.$$

In Section 3.3 of [?] we compare the theta bodies shown above to other known relaxations of $\text{CUT}(G)$ such as those in [8]. In particular, we show that if G^* is the suspension of G from a new vertex then $\pi_E(\text{TH}_1(IG^*))$ is contained in the Goemans-Williamson relaxation of $\text{CUT}(G)$. In [?, Corollary 4.12] we prove that IG is TH_1 -exact if and only if G has no K_5 -minors and no induced cycles of length at least five which answers Problem 8.4 posed by Lovász in [13]. See [?] for additional results and the proofs of the above statements.

A recent trend in theoretical computer science has been to study the computational complexity of approximating problems in combinatorial optimization via the standard hierarchies of convex relaxations to these problems such as those in [14] and [3, 4]. Our sums of squares approach via theta bodies provides a new mechanism to establish such complexity results.

4. VANISHING IDEALS OF FINITE SETS OF POINTS

Recall that when $S \subset \mathbb{R}^n$ is finite, its vanishing ideal $\mathcal{I}(S)$ is zero-dimensional and real radical.

Definition 4.1. We say that a finite set $S \subset \mathbb{R}^n$ is *exact* if its vanishing ideal $\mathcal{I}(S) \subseteq \mathbb{R}[\mathbf{x}]$ is TH_1 -exact.

Theorem 4.2 answers Problem 1.8 for vanishing ideals of finite sets.

Theorem 4.2. *For a finite set $S \subset \mathbb{R}^n$, the following are equivalent.*

- (1) S is exact.
- (2) $\mathcal{I}(S)$ is $(1, 1)$ -sos.
- (3) There is a linear inequality description of $\text{conv}(S)$, of the form $g_i(x) \geq 0$ ($i = 1, \dots, m$), where each g_i is 1-sos modulo $\mathcal{I}(S)$.
- (4) There is a linear inequality description of $\text{conv}(S)$, of the form $g_i(x) \geq 0$ ($i = 1, \dots, m$), where each g_i is idempotent modulo $\mathcal{I}(S)$ i.e., $g_i^2 - g_i \in \mathcal{I}(S)$.
- (5) There is a linear inequality description of $\text{conv}(S)$, of the form $g_i(x) \geq 0$ ($i = 1, \dots, m$), where each g_i takes at most two different values in S i.e., S is contained in the union of the hyperplane $g_i(\mathbf{x}) = 0$ and one unique parallel translate of it.

Proof: Since $\mathcal{I}(S)$ is real radical, by Proposition 2.9, (1) \Leftrightarrow (2).

The implication (2) \Rightarrow (3) follows from the fact that $\text{conv}(S)$ has a finite linear inequality description, since S is finite. The implication (3) \Rightarrow (2) follows from Farkas' lemma, which implies that any valid inequality on S is a conic combination of the linear inequalities $g_i(x) \geq 0$.

Suppose (3) holds and $\text{conv}(S)$ is a full-dimensional polytope. Let F be a facet of $\text{conv}(S)$, and $g(\mathbf{x}) \geq 0$ its defining inequality in the given description of $\text{conv}(S)$. Then $g(\mathbf{x})$ is 1-sos mod $\mathcal{I}(S)$ if and only if there are linear polynomials $h_1, \dots, h_l \in \mathbb{R}[\mathbf{x}]$ such that $g \equiv h_1^2 + \dots + h_l^2 \pmod{\mathcal{I}(S)}$. In particular, since $g(\mathbf{x}) = 0$ on the vertices of F , and all the h_i^2 are non-negative, each h_i must be zero on all the vertices of F . Hence, since the h_i 's are linear, they must vanish on the affine span of F which is the hyperplane defined by $g(\mathbf{x}) = 0$. Thus each h_i must be a multiple of g and $g \equiv \alpha g^2 \pmod{\mathcal{I}(S)}$ for some $\alpha > 0$. We may assume that $\alpha = 1$ by replacing $g(\mathbf{x})$ by $g'(\mathbf{x}) := \alpha g(\mathbf{x})$. If $\text{conv}(S)$ is not full-dimensional, then since mod $\mathcal{I}(S)$, all linear polynomials can be assumed to define hyperplanes whose normal vectors are parallel to the affine span of S , the proof still holds. Therefore, (3) implies (4). Conversely, since if $g \equiv g^2 \pmod{\mathcal{I}(S)}$ then g is 1-sos mod $\mathcal{I}(S)$, (4) implies (3).

The equivalence (4) \Leftrightarrow (5) follows since $g \equiv g^2 \pmod{\mathcal{I}(S)}$ if and only if $g(\mathbf{s})(1 - g(\mathbf{s})) = 0 \forall \mathbf{s} \in S$. \square

Recall from the end of Section 2.1 that by results of Parrilo, if $I \subseteq \mathbb{R}[\mathbf{x}]$ is a zero-dimensional radical ideal, then the theta-rank of I is at most $|\mathcal{V}_{\mathbb{C}}(I)| - 1$. Better upper bounds can be derived using the following extension of Parrilo's theorem whose proof is similar.

Remark 4.3. Suppose $S \subseteq \mathbb{R}^n$ is a finite point set such that for each facet F of $\text{conv}(S)$ there is a hyperplane H_F such that $H_F \cap \text{conv}(S) = F$ and S is contained in at most $t+1$ parallel translates of H_F . Then $\mathcal{I}(S)$ is TH_t -exact. To see this just note that if F is such a facet, there is a facet defining linear inequality $h_F(\mathbf{x}) \geq 0$ such that h_F only takes $t+1$ distinct values in S . It is then easy to construct a degree t interpolator g for the values of $\sqrt{h_F}$ in S , and we have $h_T \equiv g^2 \pmod{I}$. The result then follows from Farkas' Lemma.

Remark 4.4. The theta-rank of $\mathcal{I}(S)$ could be much smaller than the upper bound in Proposition ???. Consider a $(2t+1)$ -cycle G and the set S_G of characteristic vectors of its stable sets. Using Theorem 3.3 check that for each facet F of $\text{STAB}(G)$, S_G is contained in at most $t+1$ parallel translates of the hyperplane spanned by F and that exactly $t+1$ translates are needed for the facet cut out by $\sum_{i \in [n]} x_i = t$. However, Proposition 3.4 shows that $\mathcal{I}(S_G)$ is TH_2 -exact.

In the rest of this section we derive various consequences of Theorem 4.2. Finite point sets with property (5) in Theorem 4.2 have been studied in various contexts. In particular, Corollaries 4.5, 4.9 and 4.11 below were also observed independently by Greg Kuperberg, Raman Sanyal, Axel Werner and Günter Ziegler (personal communication). In their work, $\text{conv}(S)$ is called a **2-level polytope** when property (5) in Theorem 4.2 holds.

If S is a finite subset of \mathbb{Z}^n and \mathcal{L} is the smallest lattice in \mathbb{Z}^n containing S , then the lattice polytope $\text{conv}(S)$ is said to be **compressed** if every

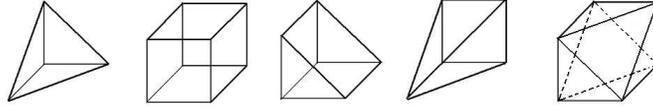


FIGURE 1. Convex hulls of exact 0/1 point sets in \mathbb{R}^3 .

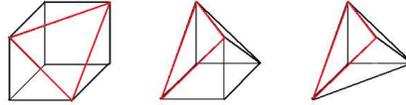


FIGURE 2. Convex hulls of non-exact 0/1 point sets in \mathbb{R}^3 .

reverse lexicographic triangulation of the lattice points in $\text{conv}(S)$ is *unimodular* with respect to \mathcal{L} . Compressed polytopes were introduced by Stanley [25]. Corollary 4.5 (4) and Theorem 2.4 in [26] (see also the references after Theorem 2.4 in [26] for earlier citations of part or unpublished versions of this result), imply that a finite set $S \subset \mathbb{R}^n$ is exact if and only if $\text{conv}(S)$ is affinely equivalent to a compressed polytope.

Corollary 4.5. *Let $S, S' \subset \mathbb{R}^n$ be exact sets. Then*

- (1) *all points of S are vertices of $\text{conv}(S)$,*
- (2) *the set of vertices of any face of $\text{conv}(S)$ is again exact,*
- (3) *the product $S \times S'$ is exact, and*
- (4) *$\text{conv}(S)$ is affinely equivalent to a 0/1 polytope.*

Proof: The first three properties follow from Theorem 4.2 (5). If the dimension of $\text{conv}(S)$ is $d (\leq n)$, then $\text{conv}(S)$ has at least d non-parallel facets. If $\mathbf{a} \cdot \mathbf{x} \geq b$ cuts out a facet in this collection, then $\text{conv}(S)$ is supported by both $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a} \cdot \mathbf{x} = b\}$ and a parallel translate of it. Taking these two parallel hyperplanes from each of the d facets gives a parallelepiped. By Theorem 4.2, S is contained in the vertices of this parallelepiped intersected with the affine hull of S . This proves (4). \square

By Corollary 4.5 (4), it essentially suffices to look at subsets of $\{0, 1\}^n$ to obtain all exact finite varieties. In \mathbb{R}^2 , the set of vertices of any 0/1-polytope verify this property. In \mathbb{R}^3 there are eight full-dimensional 0/1-polytopes up to affine equivalence. In Figures 1 & 2 the convex hulls of the exact and non-exact 0/1 configurations in \mathbb{R}^3 are shown.

Example 4.6. The vertices of the following 0/1-polytopes in \mathbb{R}^n are exact for every n : (1) hypercubes, (2) (regular) cross polytopes, (3) hypersimplices (includes simplices), (4) joins of 2-level polytopes, and (5) stable set polytopes of perfect graphs on n vertices.

Theorem 4.7. *If S is a finite exact point set then $\text{conv}(S)$ has at most 2^d facets and vertices, where $d = \dim \text{conv}(S)$. Both bounds are sharp.*

Proof: The bound on the number of vertices is immediate by Corollary 4.5 and is achieved by $[0, 1]^d$.

For a polytope P with an exact vertex set S , define a **face pair** to be an unordered pair (F_1, F_2) of proper faces of P such that $S \subseteq F_1 \cup F_2$ and F_1 and F_2 lie in parallel hyperplanes, or equivalently, there exists a linear form $h_{F_1, F_2}(\mathbf{x})$ such that $h_{F_1, F_2}(F_1) = 0$ and $h_{F_1, F_2}(F_2) = 1$. We will show that if $\dim P = d$ then P has at most $2^d - 1$ face pairs and 2^d facets.

If $d = 1$, then an exact S consists of two distinct points and P has two facets and one face pair as desired. Assume the result holds for $(d - 1)$ -polytopes with exact vertex sets and consider a d -polytope P with exact vertex set S . Let F be a facet of P which by Theorem 4.2, is in a face pair (F, F') of P . Since exactness does not depend on the affine embedding, we may assume that P is full-dimensional and that F spans the hyperplane $\{\mathbf{x} : x_d = 0\}$, while F' lies in $\{\mathbf{x} : x_d = 1\}$. By Corollary 4.5, F satisfies the induction hypothesis and so has at most $(2^{d-1} - 1)$ face pairs. Any face pair of P besides (F, F') induces a face pair of F by intersection with F , and every facet of P is in a face pair of P since S is exact. The plan is to count how many face pairs of P induce the same face pair of F and the number of facets they contain.

Fix a face pair (F_1, F_2) of F , with h_{F_1, F_2} the associated linear form depending only on x_1, \dots, x_{d-1} . Suppose (F_1, F_2) is induced by a face pair of P with associated linear form $H(\mathbf{x})$. Since H and h_{F_1, F_2} agree on every vertex of F , a facet of P , $H(\mathbf{x}) = h_{F_1, F_2}(x_1, \dots, x_{d-1}) + cx_d$ for some constant c .

If $h_{F_1, F_2}(x_1, \dots, x_{d-1})$ takes the same value v on all of F' , then $H(F') = v + c = 0$ or 1 which implies that $c = -v$ or $c = 1 - v$. The two possibilities lead to the face pairs $(\text{conv}(F_1 \cup F'), F_2)$ and $(\text{conv}(F_2 \cup F'), F_1)$ of P . Each such pair contains at most one facet of P .

If $h_{F_1, F_2}(x_1, \dots, x_{d-1})$ takes more than one value on the vertices of F' , then these values must be v and $v + 1$ for some v since H takes values 0 and 1 on the vertices of F' . In that case, $c = -v$, so H is unique and we get at most one face pair of P inducing (F_1, F_2) . This pair will contain at most two facets of P .

Since there are at most $2^{d-1} - 1$ face pairs in F , they give us at most $2(2^{d-1} - 1)$ face pairs and facets of P . Since we have not counted (F, F') as a face pair of P , and F and F' as possible facets of P , we get the desired result. The bound on the number of facets is attained by cross-polytopes. \square

Remark 4.8. Günter Ziegler has pointed out that our proof of Theorem 4.7 can be refined to yield that P (as used above) has $2^d - 1$ face pairs if and only if it is a simplex and 2^d facets if and only if it is a regular cross-polytope.

Recall that Problem 1.8 was inspired by perfect graphs. Theorem 4.2 adds to the characterizations of perfect graphs in Corollary 3.2 as follows.

Corollary 4.9. *For a graph G , let S_G denote the set of characteristic vectors of stable sets in G . Then the following are equivalent.*

- (1) *The graph G is perfect.*
- (2) *The stable set polytope, $\text{STAB}(G)$, is a 2-level polytope.*

A polytope P in $\mathbb{R}_{\geq 0}^n$ is said to be **down-closed** if for all $\mathbf{v} \in P$ and $\mathbf{v}' \in \mathbb{R}_{\geq 0}^n$ such that $\mathbf{v}'_i \leq \mathbf{v}_i$ for $i = 1, \dots, n$, $\mathbf{v}' \in P$. For a graph G , $\text{STAB}(G)$ is a down-closed 0/1-polytope, and G is perfect if and only if the vertex set of $\text{STAB}(G)$ is exact. We now prove that all down-closed 0/1-polytopes with exact vertex sets are stable set polytopes of perfect graphs.

Theorem 4.10. *Let $P \subseteq \mathbb{R}^n$ be a down-closed 0/1-polytope and S be its set of vertices. Then S is exact if and only if all facets of P are either defined by non-negativity constraints on the variables or by an inequality of the form $\sum_{i \in I} x_i \leq 1$ for some $I \subseteq [n]$.*

Proof: If P is not full-dimensional then since it is down-closed, it must be contained in a coordinate hyperplane $x_i = 0$ and the arguments below can be repeated in this lower-dimensional space. So we may assume that P is n -dimensional. Then since P is down-closed, S contains $\{\mathbf{0}, \mathbf{e}_1, \dots, \mathbf{e}_n\}$.

If all facets of P are of the stated form, using that $S \subseteq \{0, 1\}^n$, it is straight forward to check that S is exact.

Now assume that S is exact and $g(\mathbf{x}) \geq 0$ is a facet inequality of P that is not a non-negativity constraint. Then $g(\mathbf{x}) := c - \sum_{i=1}^n a_i x_i \geq 0$ for some integers c, a_1, \dots, a_n with $c \neq 0$. Since $\mathbf{0} \in S$ and S is exact, we get that $g(\mathbf{s})$ equals 0 or c for all $\mathbf{s} \in S$. Therefore, for all i , $g(\mathbf{e}_i) = c - a_i$ equals 0 or c , so a_i is either 0 or c . Dividing through by c , we get that the facet inequality $g(\mathbf{x}) \geq 0$ is of the form $\sum_{i \in I} x_i \leq 1$ for some $I \subseteq [n]$. \square

Corollary 4.11. *Let $P \subseteq \mathbb{R}^n$ be a full-dimensional down-closed 0/1-polytope and S be its vertex set. Then S is exact if and only if P is the stable set polytope of a perfect graph.*

Proof: By Corollary 4.9 we only need to prove the “only-if” direction. Suppose S is exact. Then by Theorem 4.10, all facet inequalities of P are either of the form $x_i \geq 0$ for some $i \in [n]$ or $\sum_{i \in I} x_i \leq 1$ for some $I \subseteq [n]$. Define the graph $G = ([n], E)$ where $\{i, j\} \in E$ if and only if $\{i, j\} \subseteq I$ for some I that indexes a facet inequality of P .

We prove that $P = \text{STAB}(G)$ and that G is perfect. Let $K \subseteq [n]$ such that its characteristic vector $\chi^K \in S$. If there exists $i, j \in K$ such that $i, j \in I$ for some I that indexes a facet inequality of P , then $1 - \sum_{i \in I} x_i$ takes three different values when evaluated at the points $\mathbf{0}, \mathbf{e}_i, \chi^K$ in S which contradicts that S is exact. Therefore, K is a stable set of G and $P \subseteq \text{STAB}(G)$. If $K \subseteq [n]$ is a stable set of G then, by construction, for every I indexing a facet inequality of P , χ^K lies on either $\sum_{i \in I} x_i = 1$ or $\sum_{i \in I} x_i = 0$. Therefore $\chi^K \in P$ and $\text{STAB}(G) \subseteq P$. Since all facet inequalities of

$\text{STAB}(G)$ are either non-negativities or clique inequalities, G is perfect by [1, Theorem 9.2.4 iii.]. \square

5. ARBITRARY TH_1 -EXACT IDEALS

In this last section we describe $\text{TH}_1(I)$ for an arbitrary (not necessarily radical or zero-dimensional) ideal $I \subseteq \mathbb{R}[\mathbf{x}]$. The main structural result is Theorem 5.6 which allows the construction of non-trivial high-dimensional TH_1 -exact ideals as in Example 5.7.

In this study, the *convex quadrics* in $\mathbb{R}[\mathbf{x}]$ play a particularly important role. These are precisely the polynomials of degree two that can be written as $F(\mathbf{x}) = \mathbf{x}^t A \mathbf{x} + \mathbf{b}^t \mathbf{x} + c$, where $A \neq 0$ is an $n \times n$ positive semidefinite matrix, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Note that every sum of squares of linear polynomials in $\mathbb{R}[\mathbf{x}]$ is a convex quadric.

Lemma 5.1. *For $I \subseteq \mathbb{R}[\mathbf{x}]$, $\text{TH}_1(I) \neq \mathbb{R}^n$ if and only if there exists some convex quadric $F \in I$.*

Proof: If $\text{TH}_1(I) \neq \mathbb{R}^n$, there exists a degree one polynomial f that is strictly positive on $\text{TH}_1(I)$, hence 1-sos modulo I . Then $f(\mathbf{x}) \equiv g(\mathbf{x}) \pmod{I}$ for some 1-sos $g(\mathbf{x}) \neq 0$ and $g(\mathbf{x}) - f(\mathbf{x}) \in I$ is a convex quadric.

Conversely, suppose $\mathbf{x}^t A \mathbf{x} + \mathbf{b}^t \mathbf{x} + c \in I$ with $A \succeq 0$. Then for any $\mathbf{d} \in \mathbb{R}^n$, $(\mathbf{x} + \mathbf{d})^t A (\mathbf{x} + \mathbf{d}) = \mathbf{x}^t A \mathbf{x} + 2\mathbf{d}^t A \mathbf{x} + \mathbf{d}^t A \mathbf{d} \equiv (2\mathbf{d}^t A - \mathbf{b}^t) \mathbf{x} + \mathbf{d}^t A \mathbf{d} - c \pmod{I}$. Therefore, since $(\mathbf{x} + \mathbf{d})^t A (\mathbf{x} + \mathbf{d})$ is a sum of squares of linear polynomials, the linear polynomial $(2\mathbf{d}^t A - \mathbf{b}^t) \mathbf{x} + \mathbf{d}^t A \mathbf{d} - c$ is 1-sos mod I and $\text{TH}_1(I)$ must satisfy it. Since \mathbf{d} can be chosen so that $(2\mathbf{d}^t A - \mathbf{b}^t) \neq 0$, $\text{TH}_1(I)$ is not trivial. \square

Example 5.2. For $S \subseteq \mathbb{R}^2$, $\text{TH}_1(\mathcal{I}(S)) \neq \mathbb{R}^2$ if and only if S is contained in a pair of parallel lines, a parabola, or an ellipse.

Corollary 5.3. *If $\mathbf{x}^t A \mathbf{x} + \mathbf{b}^t \mathbf{x} + c \in I$ with $A \succeq 0$ and of full rank then $\text{TH}_1(I)$ is bounded.*

Proof: In this case $\{2\mathbf{d}^t A - \mathbf{b}^t : \mathbf{d} \in \mathbb{R}^n\} = \mathbb{R}^n$ and so there are valid inequalities for $\text{TH}_1(I)$ with all possible elements of \mathbb{R}^n as normals. \square

Corollary 5.4. *For an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$, $\text{TH}_1(I) = \bigcap \text{TH}_1(\langle F \rangle)$, where F varies over all convex quadrics in I .*

Proof: If $F \in I$ then $\langle F \rangle \subseteq I$. Also, if f is linear and 1-sos mod $\langle F \rangle$ then it is also 1-sos mod I . Therefore, $\text{TH}_1(I) \subseteq \text{TH}_1(\langle F \rangle)$.

To prove the reverse inclusion, we need to show that if f is a linear polynomial that is nonnegative on $\text{TH}_1(I)$, it is also nonnegative on $\bigcap_{F \in I} \text{TH}_1(\langle F \rangle)$, where F is a convex quadric. It suffices to show that whenever f is linear and 1-sos mod I , then there is a convex quadric $F \in I$ such that $f(\mathbf{x}) \geq 0$ is valid for $\text{TH}_1(\langle F \rangle)$, or equivalently that f is 1-sos mod $\langle F \rangle$. Since f is

1-sos mod I , there is a sum of squares of linear polynomials $g(\mathbf{x})$ such that $f(\mathbf{x}) \equiv g(\mathbf{x}) \pmod{I}$. But g is a convex quadric, hence so is $g(\mathbf{x}) - f(\mathbf{x})$. Thus f is 1-sos mod the ideal $\langle g(\mathbf{x}) - f(\mathbf{x}) \rangle$ and we can take $F(\mathbf{x}) = g(\mathbf{x}) - f(\mathbf{x})$. \square

Lemma 5.5. *If $F(\mathbf{x}) = \mathbf{x}^t A \mathbf{x} + \mathbf{b}^t \mathbf{x} + c$ with $A \succeq 0$, then $\text{TH}_1(\langle F \rangle) = \text{conv}(\mathcal{V}_{\mathbb{R}}(F))$.*

Proof: We know that $\text{conv}(\mathcal{V}_{\mathbb{R}}(F)) \subseteq \text{TH}_1(\langle F \rangle)$ and, since F is convex, $\text{conv}(\mathcal{V}_{\mathbb{R}}(F)) = \{\mathbf{x} \in \mathbb{R}^n : F(\mathbf{x}) \leq 0\}$. Thus, if for every $\mathbf{x} \in \mathcal{V}_{\mathbb{R}}(F)$ $\text{grad}F(\mathbf{x}) \neq \mathbf{0}$, then $\text{conv}(\mathcal{V}_{\mathbb{R}}(F))$ is supported by the tangent hyperplanes to $\mathcal{V}_{\mathbb{R}}(F)$. In this case, to show that $\text{TH}_1(\langle F \rangle) \subseteq \text{conv}(\mathcal{V}_{\mathbb{R}}(F))$, it suffices to prove that all tangent hyperplanes to $\mathcal{V}_{\mathbb{R}}(F)$ are 1-sos mod $\langle F \rangle$. The proof of the “if” direction of Lemma 5.1 shows that it would suffice to prove that a tangent hyperplane to $\mathcal{V}_{\mathbb{R}}(F)$ has the form $(2\mathbf{d}^t A - \mathbf{b}^t)\mathbf{x} + \mathbf{d}^t A \mathbf{d} - c = 0$, for some $\mathbf{d} \in \mathbb{R}^n$. The tangent at $\mathbf{x}_0 \in \mathcal{V}_{\mathbb{R}}(F)$ has equation $0 = (2A\mathbf{x}_0 + \mathbf{b})^t(\mathbf{x} - \mathbf{x}_0)$ which can be rewritten as

$$0 = (2\mathbf{x}_0^t A + \mathbf{b}^t)\mathbf{x} - 2\mathbf{x}_0^t A \mathbf{x}_0 - \mathbf{b}^t \mathbf{x}_0 = (2\mathbf{x}_0^t A + \mathbf{b}^t)\mathbf{x} - \mathbf{x}_0^t A \mathbf{x}_0 + c,$$

and so setting $\mathbf{d} = -\mathbf{x}_0$ gives the result.

Suppose there is an \mathbf{x}_0 such that $F(\mathbf{x}_0) = 0$ and $\text{grad}F(\mathbf{x}_0) = \mathbf{0}$. By translation we may assume that $\mathbf{x}_0 = \mathbf{0}$, hence, $c = 0$ and $\mathbf{b} = \mathbf{0}$. Therefore $F = \mathbf{x}^t A \mathbf{x} = \sum h_i^2$ where the h_i are linear. Since $\mathcal{V}_{\mathbb{R}}(\langle F \rangle) = \mathcal{V}_{\mathbb{R}}(\langle h_1, \dots, h_m \rangle)$ it is enough to prove that all inequalities $\pm h_i \geq 0$ are valid for $\text{TH}_1(\langle F \rangle)$. For any $\epsilon > 0$ we have

$$(\pm h_l + \epsilon)^2 + \sum_{i \neq l} h_i^2 = F \pm 2\epsilon h_l + \epsilon^2 \equiv 2\epsilon(\pm h_l + \epsilon/2) \pmod{\langle F \rangle},$$

so $\pm h_l + \epsilon/2$ is 1-sos mod $\langle F \rangle$ for all l and all $\epsilon > 0$. This implies that all the inequalities $\pm h_l + \epsilon/2 \geq 0$ are valid for $\text{TH}_1(\langle F \rangle)$, therefore so are the inequalities $\pm h_l \geq 0$. \square

Theorem 5.6. *Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be any ideal, then*

$$\text{TH}_1(I) = \bigcap_{\substack{F \in I \\ F \text{ convex quadric}}} \text{conv}(\mathcal{V}_{\mathbb{R}}(F)) = \bigcap_{\substack{F \in I \\ F \text{ convex quadric}}} \{\mathbf{x} \in \mathbb{R}^n : F(\mathbf{x}) \leq 0\}.$$

Proof: Immediate from corollary 5.4 and lemma 5.5. \square

Example 5.7. Theorem 5.6 shows that some non-principal ideals such as $I = \langle x^2 - z, y^2 - z \rangle \subseteq \mathbb{R}[x, y, z]$ are TH_1 -exact. Since $\mathcal{V}_{\mathbb{R}}(I) = \{(t, \pm t, t^2) : t \geq 0\}$, fixing the third coordinate we get the four points $(|x|, |y|, t^2)$ where $|x| = |y| = |t|$. This implies

$$\text{conv}(\mathcal{V}_{\mathbb{R}}(I)) \subseteq \{(x, y, t^2) : |x| \leq t, |y| \leq t, t \geq 0\},$$

but it is easy to see that the RHS is equal to $\{(x, y, z) : x^2 \leq z, y^2 \leq z\}$ which is exactly $\text{conv}(\mathcal{V}_{\mathbb{R}}(x^2 - z)) \cap \text{conv}(\mathcal{V}_{\mathbb{R}}(y^2 - z))$ and so contains the theta body

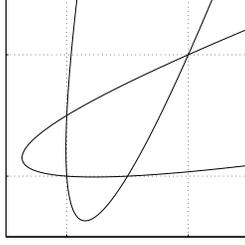


FIGURE 3. Example 5.8

hence also $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$, and all inclusions must be equalities. This kind of reasoning allows us to construct non-trivial examples of TH_1 -exact ideals with high-dimensional varieties.

Example 5.8. Consider the set $S = \{(0, 0), (1, 0), (0, 1), (2, 2)\}$. Then the family of all quadratic curves in $\mathcal{I}(S)$ is

$$a(x^2 - x) + b(y^2 - y) - \left(\frac{a+b}{2}\right)xy = (x, y) \begin{bmatrix} a & -\left(\frac{a+b}{4}\right) \\ -\left(\frac{a+b}{4}\right) & b \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - ax - by.$$

Since the case where both a and b are zero is trivial, we may normalize by setting $a + b = 1$ and get the matrix in the quadratic to be

$$\begin{bmatrix} \lambda & -1/4 \\ -1/4 & 1 - \lambda \end{bmatrix}$$

with $\lambda \geq 0$. This matrix is positive semidefinite if and only if $\lambda(1 - \lambda) - 1/16 \geq 0$, or equivalently, if and only if $\lambda \in [1/2 - \sqrt{3}/4, 1/2 + \sqrt{3}/4]$.

This means that $(x, y) \in \text{TH}_1(\mathcal{I}(S))$ if and only if, for all such λ ,

$$\lambda(x^2 - x) + (1 - \lambda)(y^2 - y) - \frac{1}{2}xy \leq 0.$$

Since the right-hand-side does not depend on λ , and the left-hand-side is a convex combination of $x^2 - x$ and $y^2 - y$, the inequality holds for every $\lambda \in [1/2 - \sqrt{3}/4, 1/2 + \sqrt{3}/4]$ if and only if it holds at the end points of the interval. Equivalently, if and only if

$$\left(\frac{1}{2} - \frac{\sqrt{3}}{4}\right)(x^2 - x) + \left(\frac{1}{2} + \frac{\sqrt{3}}{4}\right)(y^2 - y) - \frac{1}{2}xy \leq 0,$$

and

$$\left(\frac{1}{2} + \frac{\sqrt{3}}{4}\right)(x^2 - x) + \left(\frac{1}{2} - \frac{\sqrt{3}}{4}\right)(y^2 - y) - \frac{1}{2}xy \leq 0.$$

But this is just the intersection of the convex hull of the two curves obtained by turning the inequalities into equalities. Figure 3 shows this intersection.

REFERENCES

- [1] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, second edition, 1993.
- [2] J. William Helton and Jiawang Nie. Sufficient and necessary conditions for semidefinite representability of convex hulls and sets. arXiv:0709.4017v3 [math.OC], 2007.
- [3] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2001.
- [4] Jean B. Lasserre. An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM J. Optim.*, 12(3):756–769, 2002.
- [5] Jean B. Lasserre. Convex sets with semidefinite representation. *Math. Program.*, 120:457–477, 2009.
- [6] Jean Bernard Lasserre, Monique Laurent, and Philipp Rostalski. Semidefinite characterization and computation of zero- dimensional real radical ideals. *Found. Comput. Math.*, 8(5):607–647, 2008.
- [7] Monique Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. *Math. Oper. Res.*, 28(3):470–496, 2003.
- [8] Monique Laurent. Semidefinite relaxations for max-cut. In *The sharpest cut*, MPS/SIAM Ser. Optim., pages 257–290. SIAM, Philadelphia, PA, 2004.
- [9] Monique Laurent. Semidefinite representations for finite varieties. *Math. Program.*, 109(1, Ser. A):1–26, 2007.
- [10] Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, volume 149 of *IMA Volumes in Mathematics and its Applications*. Springer, 2009.
- [11] László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979.
- [12] László Lovász. Stable sets and polynomials. *Discrete Math.*, 124(1-3):137–153, 1994. Graphs and combinatorics (Qawra, 1990).
- [13] László Lovász. Semidefinite programs and combinatorial optimization. In *Recent advances in algorithms and combinatorics*, volume 11 of *CMS Books Math./Ouvrages Math. SMC*, pages 137–194. Springer, New York, 2003.
- [14] László Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.*, 1(2):166–190, 1991.
- [15] Murray Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Math Society, Providence, RI, 2008.
- [16] Pablo A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, May 2000. Available at <http://resolver.caltech.edu/CaltechETD:etd-05062004-055516>.
- [17] Pablo A. Parrilo. An explicit construction of distinguished representations of polynomials nonnegative over finite sets. IfA Tech. Report AUT02-02, ETH Zurich, 2002.
- [18] Pablo A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Prog.*, 96(2, Ser. B):293–320, 2003.
- [19] Pablo A. Parrilo. Exploiting algebraic structure in sum of squares programs. In D. Henrion and A. Garulli, editors, *Positive Polynomials in Control*, volume 312 of *Lecture Notes in Control and Information Sciences*. Springer, 2005.
- [20] Victoria Powers and Claus Scheiderer. The moment problem for non-compact semi-algebraic sets. *Adv. Geom.*, 1(1):71–88, 2001.
- [21] Ralph T. Rockafellar. *Convex Analysis (Princeton Landmarks in Mathematics and Physics)*. Princeton University Press, December 1996.
- [22] Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *FOCS*, pages 593–602. IEEE Computer Society, 2008.

- [23] A. Schrijver. *Theory of Linear and Integer Programming*. Wiley-Interscience Series in Discrete Mathematics and Optimization, New York, 1986.
- [24] Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency. Vol. B*, volume 24 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003. Matroids, trees, stable sets, Chapters 39–69.
- [25] Richard P. Stanley. Decompositions of rational convex polytopes. *Ann. Discrete Math.*, 6:333–342, 1980. Combinatorial mathematics, optimal designs and their applications (Proc. Sympos. Combin. Math. and Optimal Design, Colorado State Univ., Fort Collins, Colo., 1978).
- [26] Seth Sullivant. Compressed polytopes and statistical disclosure limitation. *Tohoku Math. J. (2)*, 58(3):433–445, 2006.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, BOX 354350, SEATTLE, WA 98195, USA, AND CMUC, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COIMBRA, 3001-454 COIMBRA, PORTUGAL

E-mail address: `jgouveia@math.washington.edu`

DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE, LABORATORY FOR INFORMATION AND DECISION SYSTEMS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MA 02139-4307, USA

E-mail address: `parrilo@mit.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, BOX 354350, SEATTLE, WA 98195, USA

E-mail address: `thomas@math.washington.edu`