JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 29, Number 3, July 2016, Pages 893–913
http://dx.doi.org/10.1090/jams/847
Article electronically published on September 3, 2015

SUMS OF SQUARES AND VARIETIES OF MINIMAL DEGREE

GRIGORIY BLEKHERMAN, GREGORY G. SMITH, AND MAURICIO VELASCO

1. Introduction

The study of nonnegativity and its relation with sums of squares is a basic challenge in real algebraic geometry. The classification of varieties of minimal degree is one of the milestones of classical complex algebraic geometry. The goal of this paper is to establish the deep connection between these apparently separate topics.

To achieve this, let $X \subseteq \mathbb{P}^n$ be an embedded real projective variety with homogeneous coordinate ring R. The variety X has minimal degree if it is nondegenerate (not contained in a hyperplane) and $\deg(X) = 1 + \operatorname{codim}(X)$. A homogeneous element $f \in R$ of even degree is nonnegative if its evaluation at each real point of X is at least zero. Our main theorem is a broad generalization of Hilbert's 1888 classification of nonnegative forms and provides a tight connection between real and complex algebraic geometry.

Theorem 1.1. *Let $X \subseteq \mathbb{P}^n$ be a real irreducible nondegenerate projective subvariety such that the set $X(\mathbb{R})$ of real points is Zariski dense. Every nonnegative real quadratic form on X is a sum of squares of linear forms in R if and only if X is a variety of minimal degree.*

Using the Veronese embedding, this theorem extends to forms of any even degree; see Remark 4.6.

Together with the well-known catalogue for varieties of minimal degree (see Theorem 1 in [EH]), our main theorem produces a complete list of varieties for which nonnegative quadratic forms are sums of squares. There are exactly three families:

- totally real irreducible quadratic hypersurfaces (Example 4.3),
- cones over the Veronese surface (Example 4.4), and
- rational normal scrolls (Example 4.5).

By replacing elements of R with global sections of a line bundle, we also develop an intrinsic version of the main theorem; see Theorem 5.1. Applying this to line bundles on projective space, we recover Hilbert's classification of nonnegative forms in a standard graded polynomial ring—for binary forms, quadratic forms, and ternary quartics, nonnegativity is equivalent to being a sum of squares and, in all other situations, there exist nonnegative forms that are not sums of squares;

Received by the editors January 7, 2014 and, in revised form, May 8, 2015 and July 23, 2015.
2010 *Mathematics Subject Classification.* Primary 14P05; Secondary 12D15, 90C22.

The first author was supported in part by a Sloan Fellowship, NSF Grant DMS-0757212, the Mittag-Leffler Institute, and IPAM.

The second author was supported in part by NSERC, the Mittag-Leffler Institute, and MSRI.
The third author was supported in part by the FAPA grants from Universidad de los Andes.

see Example 5.5. In particular, the exceptional Veronese surface corresponds to the exceptional case of ternary quartics. We obtain the classification of multiforms appearing in [CLR] from line bundles on a product of projective spaces; see Example 5.6. More generally, by working with a projective toric variety or a multigraded polynomial ring, we enumerate the cases in which every nonnegative multihomogeneous polynomial may be expressed as a sum of squares. Specifically, we discover that the ternary quartics belong to an infinite family consisting of cones over the Veronese surface and all other cases come from rational normal scrolls; see Example 5.7, Example 5.8, and Remark 5.9.

Enhancing the intrinsic approach for line bundles on a toric variety yields an analogue of our main theorem for sparse Laurent polynomials. To be more precise, let M be an affine lattice of rank m and let Q be an m-dimensional lattice polytope in $M \otimes_{\mathbb{Z}} \mathbb{R}$. The h^*-polynomial of Q is defined by

$$\mathrm{h}^*_0(Q) + \mathrm{h}^*_1(Q)\, t + \cdots + \mathrm{h}^*_m(Q)\, t^m = (1-t)^{m+1} \sum\nolimits_{k \geqslant 0} |(kQ) \cap M|\, t^k \,.$$

We establish that every nonnegative Laurent polynomial with Newton polytope in $2Q$ is a sum of squares if and only if $\mathrm{h}^*_2(Q) = 0$ and the image of the real points under the associate morphism is dense in the strong topology; see Theorem 6.4. We also describe all of the lattice polytopes Q for which $\mathrm{h}^*_2(Q) = 0$; see Proposition 6.6. This generalizes the main theorem in [BN] classifying degree-one lattice polytopes; see Remark 6.10.

For the proof of Theorem 1.1, convexity provides the bridge between real and complex algebraic geometry. The collections of nonnegative elements and sums of squares both form closed convex cones; see Lemma 2.1. More significantly, the dual of the sums-of-squares cone is a spectrahedron, so its extreme rays have an algebraic characterization; see Observation 2.2. This characterization drives the transition between real and complex algebraic geometry.

Contents of the paper. Section 2 defines the fundamental cones: P_X consists of the nonnegative elements, and Σ_X consists of the sums of squares. The description in Lemma 2.3 of the extreme rays of Σ_X^* is the key. In Section 3, we introduce the quadratic deficiency $\varepsilon(X)$ of the embedded variety $X \subseteq \mathbb{P}^n$. This numerical invariant is an algebraic incarnation of $\mathrm{h}^*_2(Q)$ and forms the pivotal link between quadratic forms and varieties of minimal degree; see Lemma 3.1. As Proposition 3.2 establishes, having $\varepsilon(X) > 0$ is a sufficient condition for the existence of nonnegative real quadratic forms on X that cannot be expressed as sums of squares. Procedure 3.3 constructs nonnegative quadratic forms that are not sums of squares. Proposition 3.5 analyzes the varieties with $\varepsilon(X) = 1$. We prove the main theorem in Section 4. Proposition 4.1 shows that $\varepsilon(X) = 0$ is sufficient. Remark 4.7 connects the main theorem to the truncated moment problem in real analysis. Section 5 translates the main theorem and principal examples into the intrinsic setting of basepoint-free linear series, and Section 6 develops the polyhedral theory.

2. Convexity and spectrahedral properties

In this section, we develop the necessary tools from convex algebraic geometry. We carefully define the fundamental cones and highlight their properties.

Let $X \subseteq \mathbb{P}^n$ be a nondegenerate m-dimensional totally real projective subvariety. In particular, X is a geometrically integral projective scheme over $\mathrm{Spec}(\mathbb{R})$ such that X is not contained in a hyperplane and the set $X(\mathbb{R})$ of real points is Zariski dense. Set $e := n - m = \operatorname{codim}(X)$. If I is the unique saturated homogeneous ideal vanishing on X, then the $\mathbb{Z}$-graded coordinate ring of X is $R := \mathbb{R}[x_0, x_1, \dots, x_n]/I$. For $j \in \mathbb{Z}$, the graded component R_j of degree j is a finite dimensional real vector space. Since X is nondegenerate, we have $\mathbb{R}[x_0, x_1, \dots, x_n]_1 = R_1$. Given $f \in R_{2j}$ and $p \in X(\mathbb{R})$, the ***sign of*** f ***at*** p is

$$\operatorname{sgn}_p(f) := \operatorname{sgn}\bigl(\tilde{f}(\tilde{p})\bigr) \in \{-1, 0, 1\},$$

where the polynomial $\tilde{f} \in \mathbb{R}[x_0, x_1, \dots, x_n]_{2j}$ maps to f and the nonzero real point $\tilde{p} \in \mathbb{A}^{n+1}(\mathbb{R})$ maps to p under the canonical quotient homomorphisms; compare with Section 2.4 in [S1]. Since $p \in X(\mathbb{R})$, the real number $\tilde{f}(\tilde{p})$ is independent of the choice $\tilde{f}$. Similarly, the choice of the affine representative $\tilde{p}$ is determined up to a nonzero real number, so the value of $\tilde{f}(\tilde{p})$ is determined up to the square of a nonzero real number because the degree of f is even. We simply write $f(p) \geqslant 0$ for $\operatorname{sgn}_p(f) \geqslant 0$.

The central objects of study are the following subsets in R_2:

$$\mathrm{P}_X := \{f \in R_2 : f(p) \geqslant 0 \text{ for all } p \in X(\mathbb{R})\}, \qquad \text{and}$$

$$\Sigma_X := \left\{ f \in R_2 : \begin{array}{l} \text{there exists } g_1, g_2, \dots, g_k \in R_1 \\ \text{such that } f = g_1^2 + g_2^2 + \cdots + g_k^2 \end{array} \right\}.$$

We clearly have $\Sigma_X \subseteq \mathrm{P}_X$. To describe the properties of these special subsets, consider the $\mathbb{R}$-linear map $\sigma\colon \operatorname{Sym}^2(R_1) \to R_2$ induced by multiplication in R and let $\sigma^*\colon R_2^* \to \operatorname{Sym}^2(R_1^*) = \bigl(\operatorname{Sym}^2(R_1)\bigr)^*$ be the dual. More explicitly, for a linear functional $\ell \in R_2^*$, $\sigma^*(\ell)$ is the symmetric linear map $R_1 \otimes_{\mathbb{R}} R_1 \to \mathbb{R}$ defined by $g_1 \otimes g_2 \mapsto \ell(g_1 g_2)$. For $p \in X(\mathbb{R})$, evaluation at any affine representative $\tilde{p} \in \mathbb{A}^{n+1}(\mathbb{R})$ determines $\tilde{p}^* \in R_1^*$. Because $p \in X(\mathbb{R})$, the map $\operatorname{Sym}^2(R_1) \to \mathbb{R}$ induced by $\tilde{p}^* \in R_1^*$ annihilates I_2 and defines the element $(\tilde{p}^*)^2 \in R_2^*$. Since evaluations at distinct representatives differ by the square of a nonzero constant, the ray $\operatorname{cone}\bigl((\tilde{p}^*)^2\bigr) := \{\lambda \cdot (\tilde{p}^*)^2 : \lambda \geqslant 0\} \subseteq R_2^*$ is independent of the choice of the affine representative.

The following fundamental lemma is a minor variant of well-known results; compare with Theorem 3.35 in [L2] or Exercise 4.2 in [BPT].

Lemma 2.1. *Both* P_X *and* Σ_X *are pointed full-dimensional closed convex cones in the real vector space* R_2*. We also have*

$$\begin{aligned} \mathrm{P}_X^* &= \operatorname{cone}\bigl((\tilde{p}^*)^2 : p \in X(\mathbb{R})\bigr) \\ &= \{\lambda_1 (\tilde{p}_1^*)^2 + \lambda_2 (\tilde{p}_2^*)^2 + \cdots + \lambda_k (\tilde{p}_k^*)^2 : \tilde{p}_i \in X(\mathbb{R}) \text{ and } \lambda_i \geqslant 0\}, \end{aligned}$$

and $\Sigma_X^* = \{\ell \in R_2^* : \sigma^*(\ell) \text{ is positive-semidefinite}\}$.

Proof. We first consider the nonnegative elements. Set

$$P := \operatorname{cone}\bigl((\tilde{p}^*)^2 : p \in X(\mathbb{R})\bigr).$$

An element $f \in R_2$ belongs to P_X if and only if $f(p) \geqslant 0$, so $P^* = \mathrm{P}_X$. It follows that P_X is a closed convex cone and $(P^*)^* = \mathrm{P}_X^*$. To show that P is closed, fix an inner product on R_2^* and let $\ell \mapsto \|\ell\|$ denote the associated norm. For each $p \in X(\mathbb{R})$, the linear functional $(\tilde{p}^*)^2/\|(\tilde{p}^*)^2\| \in R_2^*$ is independent of the choice of the affine representative. Since $X(\mathbb{R}) \subseteq \mathbb{P}^n(\mathbb{R})$ is compact in the induced metric topology, the spherical section $K := \bigl\{(\tilde{p}^*)^2/\|(\tilde{p}^*)^2\| : p \in X(\mathbb{R})\bigr\}$ of P is compact. Because X is totally real, the convex hull of K does not contain 0. Since P is the conical hull of K, the cone P is closed and $P = \mathrm{P}_X^*$. By hypothesis, the set $X(\mathbb{R})$ of real points is Zariski dense, so P_X cannot contain a nonzero linear subspace.

We next examine the sums of squares. For $\ell \in \Sigma_X^*$, we have $\ell(f^2) \geqslant 0$ for all $f \in R_1$, so the bilinear symmetric form $\sigma^*(\ell)$ is positive semidefinite. Conversely, if $\sigma^*(\ell)$ is positive semidefinite, then $\ell(g^2) \geqslant 0$ for all $g \in R_1$. Hence, we have

$$\ell(g_1^2 + g_2^2 + \cdots + g_k^2) = \ell(g_1^2) + \ell(g_2^2) + \cdots + \ell(g_k^2) \geqslant 0$$

for $g_1, g_2, \ldots, g_k \in R_1$, and $\ell \in \Sigma_X^*$. Thus, $\ell \in \Sigma_X^*$ if and only $\sigma^*(\ell)$ is a positive-semidefinite symmetric bilinear form. By duality, the cone Σ_X is a linear projection of the convex cone $\mathbb{S}_+$ of positive-semidefinite symmetric bilinear forms. Since $\mathbb{S}_+$ is full-dimensional and $\sigma\colon \operatorname{Sym}^2(R_1) \to R_2$ is surjective, it follows that Σ_X is also full-dimensional. To complete the proof, fix an inner product on R_1 and let $g \mapsto \|g\|$ denote the associated norm. The spherical section

$$K' := \{g^2 \in R_2 : g \in R_1 \text{ satisfies } \|g\| = 1\}$$

is compact, because it is the continuous image of a compact set. As above, its convex hull does not contain the origin. Therefore, the cone Σ_X is closed. $\square$

The subsequent observation is the key insight from convex geometry needed to prove our main result. Lemma 2.3 is the simple, but crucial, algebraic consequence of this observation.

Observation 2.2. Lemma 2.1 shows that Σ_X^* is a spectrahedron, that is, a section of the convex cone $\mathbb{S}_+$ of positive-semidefinite symmetric bilinear forms. Hence, Theorem 1 in [RG] implies that every face of Σ_X^* is exposed. The unique face containing $\ell \in \Sigma_X^*$ in its relative interior is given by $H_\ell \cap \Sigma_X^*$ where

$$H_\ell := \{\ell' \in R_2^* : \operatorname{Ker}\big(\sigma^*(\ell)\big) \subseteq \operatorname{Ker}\big(\sigma^*(\ell')\big)\}\,.$$

Moreover, Corollary 3 in [RG] characterizes the extreme rays as follows: a point in a spectrahedron is extreme if and only if the kernel of its associated positive semidefinite form is maximal with respect to the inclusion. Hence, if $\ell \in \Sigma_X^*$ is an extreme point and $A \in \operatorname{Im}(\sigma^*)$ such that $\operatorname{Ker}\big(\sigma^*(\ell)\big) \subseteq \operatorname{Ker}(A)$, then we have $\sigma^*(\ell) = \lambda A$ for some $\lambda \in \mathbb{R}$.

Lemma 2.3. *If $\ell \in R_2^*$ generates an extreme ray of Σ_X^*, then either ℓ is given by evaluation at some $p \in X(\mathbb{R})$ or the subspace $\operatorname{Ker}\big(\sigma^*(\ell)\big) \subseteq R_1$ contains a homogeneous system of parameters on R.*

Proof. First, suppose that the linear forms in $\operatorname{Ker}\big(\sigma^*(\ell)\big)$ have a common real zero $p \in X(\mathbb{R})$. Choose an affine representative $\tilde{p} \in \mathbb{A}^{n+1}(\mathbb{R})$. If $\sigma^*\big((\tilde{p}^*)^2\big) \in \operatorname{Sym}^2(R_1^*)$ is the associated symmetric form, then we have $\operatorname{Ker}\big(\sigma^*(\ell)\big) \subseteq \operatorname{Ker}\big(\sigma^*\big((\tilde{p}^*)^2\big)\big)$. Since $\ell \in \Sigma_X^*$ generates an extreme ray, Observation 2.2 implies that $\sigma^*(\ell) = \lambda(\tilde{p}^*)^2$ for some $\lambda \in \mathbb{R}$. As both $\sigma^*(\ell)$ and $(\tilde{p}^*)^2$ are positive semidefinite, it follows that $\lambda > 0$. Hence, by changing the affine representative for $p \in X(\mathbb{R})$ to $\sqrt{\lambda}\tilde{p} \in \mathbb{A}^{n+1}(\mathbb{R})$, we obtain $\ell = (\tilde{p}^*)^2$.

Now, assume that the only common zeroes for the linear forms in $\operatorname{Ker}\big(\sigma^*(\ell)\big)$ have a nonzero complex part. Choose an affine representative $\tilde{\zeta} \in \mathbb{A}^{n+1}(\mathbb{C})$ for one of these complex zeroes. Define $\ell' \in R_2^*$ by $\ell'(f) := \operatorname{Re}\big(f(\tilde{\zeta})\big)$ to be the real part of the evaluation of f at $\tilde{\zeta}$; this is well-defined because $\zeta \in X$. By construction, we have $\operatorname{Ker}\big(\sigma^*(\ell)\big) \subseteq \operatorname{Ker}\big(\sigma^*(\ell')\big)$. Since $\ell \in \Sigma_X^*$ generates an extreme ray, Observation 2.2 implies that $\sigma^*(\ell) = \lambda\sigma^*(\ell')$ for some $\lambda \in \mathbb{R}$. However, there exist $g_1, g_2 \in R_1$ such that $g_1(\tilde{\zeta}) = 1$ and $g_2(\tilde{\zeta}) = \sqrt{-1}$, so $\ell'(g_1^2) = 1$ and $\ell'(g_2^2) = -1$. Hence, $\sigma^*(\ell)$ is not positive semidefinite, which by Lemma 2.1 contradicts the hypothesis that $\ell \in \Sigma_X^*$. In other words, our assumption guarantees that the linear forms in $\operatorname{Ker}\big(\sigma^*(\ell)\big)$ have no common zeroes in X. Therefore, we conclude that $\operatorname{Ker}\big(\sigma^*(\ell)\big) \subseteq R_1$ contains a homogeneous system of parameters via the Nullstellensatz. □

3. Separating the fundamental cones

This section investigates differences between the sums-of-squares cone Σ_X and the nonnegative cone P_X. It relates the positivity of an algebraic invariant associated to an embedded variety $X \subseteq \mathbb{P}^n$ with the proper inclusion of Σ_X in P_X. We construct witnesses that separate Σ_X and P_X. Moreover, we give a general procedure for constructing nonnegative real quadratic forms on X that are not sums of squares.

Emulating Section 5 in [Z], we define the ***quadratic deficiency*** of the subvariety $X \subseteq \mathbb{P}^n$ to be $\varepsilon(X) := \binom{e+1}{2} - \dim(I_2)$ where $e := \operatorname{codim}(X)$ and I is the unique saturated homogeneous ideal vanishing on X. The first lemma provides a couple of elementary reinterpretations for this numerical invariant and recounts the important connection between $\varepsilon(X)$ and varieties of minimal degree.

Lemma 3.1. *The quadratic deficiency $\varepsilon(X)$ equals the coefficient of the quadratic term in the numerator of the Hilbert series for X and*

$$\varepsilon(X) = \dim(R_2) - (m+1)(n+1) + \binom{m+1}{2}.$$

Moreover, $\varepsilon(X)$ is nonnegative and $\varepsilon(X) = 0$ if and only if $\deg(X) = 1 + \operatorname{codim}(X)$.

Proof. Since X is nondegenerate, it follows that $\dim(R_0) = 1$ and $\dim(R_1) = n+1$. Hence, there exists a polynomial $1 + e\,t + \mathrm{h}_2^*(X)\,t^2 + \cdots + \mathrm{h}_n^*(X)\,t^n \in \mathbb{Z}[t]$ such that

$$\sum_{j \geqslant 0} \dim(R_j)\,t^j = \frac{1 + e\,t + \mathrm{h}_2^*(X)\,t^2 + \cdots + \mathrm{h}_n^*(X)\,t^n}{(1-t)^{m+1}}.$$

Using the binomial theorem to compare the coefficients, we obtain

$$\begin{aligned}
\dim(R_2) &= \tbinom{m+2}{2} + e\tbinom{m+1}{1} + \mathrm{h}_2^*(X)\tbinom{m+0}{0} \\
&= \tbinom{m+1}{2} + \tbinom{m+1}{1} + (n-m)(m+1) + \mathrm{h}_2^*(X) \\
&= \tbinom{m+1}{2} - m(m+1) + (n+1)(m+1) + h_2^* \\
&= -\tbinom{m+1}{2} + (n+1)(m+1) + \mathrm{h}_2^*(X).
\end{aligned}$$

Rearranging this equation and using the presentation for R yields

$$\begin{aligned}
\mathrm{h}_2^*(X) &= \dim(R_2) - (m+1)(n+1) + \tbinom{m+1}{2} \\
&= \tbinom{n+2}{2} - \dim(I_2) - (m+1)(n+1) + \tfrac{(m+1)m}{2} \\
&= \tfrac{(n-m+1)(n-m)}{2} - \dim(I_2) = \tbinom{e+1}{2} - \dim(I_2) = \varepsilon(X),
\end{aligned}$$

which establishes the results in the first sentence of the lemma. Both parts of the second sentence are well-known. As Theorem 1.2 in [L3] indicates, they can be deduced from Castelnuovo's Lemma, which states that if $n(n-1)/2$ linearly independent quadrics pass through at least $2n+3$ points in linearly general position in $\mathbb{P}^n$, then these points lie on a rational normal curve. Corollary 5.4 and Corollary 5.8 in [Z] give alternative proofs using properties of secant varieties. □

The subsequent proposition, which extends both Theorem 1.1 and Theorem 1.2 in [B2], provides one of the implications needed for the proof of Theorem 1.1.

Proposition 3.2. *If $\varepsilon(X) > 0$, then Σ_X is a proper subset of P_X.*

Proof. Since $\varepsilon(X) > 0$, Lemma 3.1 gives $\deg(X) > 1 + \operatorname{codim}(X)$. We begin by showing that there exist $h_1, h_2, \dots, h_m \in R_1$ such that $Z := X \cap \mathrm{V}(h_1, h_2, \dots, h_m)$ is a reduced set of points in linearly general position containing at least $e+1$ distinct real points. To achieve this, observe that Bézout's Theorem implies that the intersection of a positive-dimensional irreducible nondegenerate variety with a general hyperplane is nondegenerate; see Proposition 18.10 in [H1]. Next, Bertini's Theorem (see Théorème 6.3 in [J]) establishes that a general hyperplane section of a geometrically integral variety of dimension at least 2 is geometrically integral and that a general hyperplane section of a geometrically reduced variety is geometrically reduced. Third, we see that a geometrically integral real variety is totally real if and only if it contains a nonsingular real point; see Section 1 in [B1]. Finally, we note that the locus of hyperplanes that intersect the nonsingular locus of X transversely contains a nonempty Zariski open set. By combining these four observations, we deduce that the intersection of X with $m-1$ general hyperplanes yields a nondegenerate geometrically integral totally real curve C in $\mathrm{V}(h_1) \cap \mathrm{V}(h_2) \cap \cdots \cap \mathrm{V}(h_{m-1}) \cong \mathbb{P}^{e+1}$. The degree of C, which equals $\deg(X)$, is at least $e+1$; see Corollary 18.12 in [H1]. Any set of $e+1$ distinct real points on C lie in a real hyperplane. Since C is nondegenerate and totally real, the locus of hyperplanes intersecting C in at least $e+1$ distinct real points has dimension at least $e+1$. Hence, there exists a hyperplane $\mathrm{V}(h_m)$ such that intersection with C is a set of points in linearly general position containing at least $e+1$ distinct real points.

To complete the proof, we use points in Z to exhibit a linear functional in $\Sigma_X^* \setminus \mathrm{P}_X^*$. We divide the analysis into two cases. In the first case, we assume that the intersection Z contains at least $e+2$ distinct real points. Choose an affine representative $\tilde{p}_j$ where $1 \leqslant j \leqslant e+2$ for each of these points. The points lie in $\mathrm{V}(h_1) \cap \mathrm{V}(h_2) \cap \cdots \cap \mathrm{V}(h_m) \cong \mathbb{P}^e$, so the evaluations $\tilde{p}_j^*$ satisfy a linear equation in R_1^*. The coefficients in this linear equation are nonzero and determine a unique point in $\mathbb{P}^{e+1}$ because $p_1, p_2, \dots, p_{e+2}$ are in linearly general position. Specifically, there are unique nonzero $\lambda_1, \lambda_2, \dots, \lambda_{e+1} \in \mathbb{R}$ such that

$$0 = \lambda_1 \tilde{p}_1^* + \lambda_2 \tilde{p}_2^* + \cdots + \lambda_{e+1} \tilde{p}_{e+1}^* + \tilde{p}_{e+2}^* \,. \tag{1}$$

Fix $\kappa_j > 0$ for $1 \leqslant j \leqslant e+1$, set $\kappa_{e+2} := \left(\frac{\lambda_1^2}{\kappa_1} + \frac{\lambda_2^2}{\kappa_2} + \cdots + \frac{\lambda_{e+1}^2}{\kappa_{e+1}}\right)^{-1}$, and consider

$$\ell := \kappa_1 (\tilde{p}_1^*)^2 + \kappa_2 (\tilde{p}_2^*)^2 + \cdots + \kappa_{e+1} (\tilde{p}_{e+1}^*)^2 - \kappa_{e+2} (\tilde{p}_{e+2}^*)^2 \in R_2^* \,.$$

Since $\kappa_j > 0$ for all $1 \leqslant j \leqslant e+1$, Equation (1) yields

$$\begin{aligned}
\ell &= \sum_{j=1}^{e+1} \big(\sqrt{\kappa_j}\, \tilde{p}_j^*\big)^2 - (\tilde{p}_{e+2}^*)^2 \Big(\sum_{j=1}^{e+1} \big(\tfrac{\lambda_j}{\sqrt{\kappa_j}}\big)^2\Big)^{-1} \\
&= \Big(\sum_{j=1}^{e+1} \big(\tfrac{\lambda_j}{\sqrt{\kappa_j}}\big)^2\Big)^{-1} \left[\Big(\sum_{j=1}^{e+1} \big(\tfrac{\lambda_j}{\sqrt{\kappa_j}}\big)^2\Big) \Big(\sum_{j=1}^{e+1} \big(\sqrt{\kappa_j}\, \tilde{p}_j^*\big)^2\Big) - \Big(\sum_{j=1}^{e+1} \lambda_j \tilde{p}_j^*\Big)^2 \right] .
\end{aligned}$$

Hence, the Cauchy-Schwartz inequality shows that ℓ is nonnegative on squares, whence $\ell \in \Sigma_X^*$ by Lemma 2.1; compare with Theorem 6.1 in [B2]. Nevertheless, there exists $g \in R_1$ such that $\tilde{p}_j^*(g) = g(\tilde{p}_j) = \lambda_j \kappa_j^{-1}$ for all $1 \leqslant j \leqslant e+1$, which implies that $\ell(g^2) = 0$. In addition, choose the κ_j for $1 \leqslant j \leqslant e+1$ so that g does not vanish at any point in Z. Since $g^2 + h_1^2 + h_2^2 + \cdots + h_m^2$ is strictly positive on X and $\ell(g^2 + h_1^2 + \cdots + h_m^2) = 0$, the linear functional ℓ cannot be a nonnegative combination of point evaluations at $X(\mathbb{R})$. Therefore, we have $\ell \in \Sigma_X^* \setminus \mathrm{P}_X^*$.

In the second case, we assume that Z has at most $e+1$ distinct real points. Since $\deg(Z) = \deg(X) > e+1$, the reduced set Z contains at least one pair of complex conjugate points. Let $\tilde{a} \pm \tilde{b}\sqrt{-1} \in \mathbb{A}^{n+1}(\mathbb{C})$, where $\tilde{a}, \tilde{b} \in \mathbb{A}^{n+1}(\mathbb{R})$, be affine representatives for such a pair and choose an affine representative $\tilde{p}_j$ for $1 \leqslant j \leqslant e$ for some real points in Z. As in the other case, the chosen $e+2$ points lie in $\mathrm{V}(h_1) \cap \mathrm{V}(h_2) \cap \cdots \cap \mathrm{V}(h_m) \cong \mathbb{P}^e$, so the evaluations satisfy a linear equation in R_1^*. Again, the coefficients are nonzero and determine a unique point $\mathbb{P}^{e+1}$ because the points in Z are in linearly general position. Adding this linear equation to its conjugate (or simply rescaling by $\sqrt{-1}$), we obtain a linear equation that is invariant under conjugation. Hence, by choosing different affine representatives for $\tilde{a}$ and $\tilde{b}$ if necessary, we may assume that the coefficients of $(\tilde{a} + \tilde{b}\sqrt{-1})^*$ and $(\tilde{a} - \tilde{b}\sqrt{-1})^*$ are equal. Specifically, there are unique nonzero $\lambda_1, \lambda_2, \ldots, \lambda_e \in \mathbb{R}$ such that

$$\begin{aligned} 0 &= \lambda_1 \tilde{p}_1^* + \lambda_2 \tilde{p}_2^* + \cdots + \lambda_e \tilde{p}_e^* + \tfrac{1}{2}(\tilde{a} + \tilde{b}\sqrt{-1})^* + \tfrac{1}{2}(\tilde{a} - \tilde{b}\sqrt{-1})^* \\ &= \lambda_1 \tilde{p}_1^* + \lambda_2 \tilde{p}_2^* + \cdots + \lambda_e \tilde{p}_e^* + \tilde{a}^*. \end{aligned} \tag{2}$$

Taking the real and imaginary parts of $\big((\tilde{a} \pm \tilde{b}\sqrt{-1})^*\big)^2 \in R_2^*$ yields the linear independent real functionals $(\tilde{a}^*)^2 - (\tilde{b}^*)^2 \in R_2^*$ and $2\tilde{a}^*\tilde{b}^* \in R_2^*$. Fix $\kappa_j > 0$ for $1 \leqslant j \leqslant e$, choose κ_{e+1} and κ_{e+2} satisfying

$$(\kappa_{e+1}^2 + \kappa_{e+2}^2)\kappa_{e+1}^{-1} := \left(\tfrac{\lambda_1^2}{\kappa_1} + \tfrac{\lambda_2^2}{\kappa_2} + \cdots + \tfrac{\lambda_e^2}{\kappa_e}\right)^{-1},$$

and consider

$$\ell := \kappa_1(\tilde{p}_1^*)^2 + \kappa_2(\tilde{p}_2^*)^2 + \cdots + \kappa_e(\tilde{p}_e^*)^2 - \kappa_{e+1}\big((\tilde{a}^*)^2 - (\tilde{b}^*)^2\big) + \kappa_{e+2}(2\tilde{a}^*\tilde{b}^*)$$

lying in R_2^*. Completing the square and using Equation (2) yields

$$\begin{aligned} \ell &= \sum_{j=1}^{e}\big(\sqrt{\kappa_j}\,\tilde{p}_j^*\big)^2 - \tfrac{\kappa_{e+1}^2 + \kappa_{e+2}^2}{\kappa_{e+1}}(\tilde{a}^*)^2 + \kappa_{e+1}\Big(\tilde{b}^* + \tfrac{\kappa_{e+2}}{\kappa_{e+1}}\tilde{a}^*\Big)^2 \\ &= \Big(\sum_{j=1}^{e}\big(\tfrac{\lambda_j}{\sqrt{\kappa_j}}\big)^2\Big)^{-1}\left[\Big(\sum_{j=1}^{e}\big(\tfrac{\lambda_j}{\sqrt{\kappa_j}}\big)^2\Big)\Big(\sum_{j=1}^{e}\big(\sqrt{\kappa_j}\,\tilde{p}_j^*\big)^2\Big) - \Big(\sum_{j=1}^{e}\lambda_j\tilde{p}_j^*\Big)^2\right] \\ &\quad + \kappa_{e+1}\Big(\tilde{b}^* + \tfrac{\kappa_{e+2}}{\kappa_{e+1}}\tilde{a}^*\Big)^2. \end{aligned}$$

Since $\kappa_{e+1} > 0$, the Cauchy-Schwartz inequality once more shows that ℓ is nonnegative on squares; compare with Theorem 7.1 in [B2]. By repeating the argument above, we conclude that $\ell \in \Sigma_X^* \setminus \mathrm{P}_X^*$. $\square$

By enhancing the techniques used in the proof of Proposition 3.2, we obtain a way to construct nonnegative polynomials that are not sums of squares. We describe this process below. To make it computationally effective, one needs an explicit bound for the coefficient δ.

Procedure 3.3. *Given an m-dimensional nondegenerate totally real subvariety $X \subseteq \mathbb{P}^n$ such that $\varepsilon(X) > 0$ and $e = \operatorname{codim}(X)$, the following steps yield a polynomial lying in $\mathrm{P}_X \setminus \Sigma_X$.*

Step 1: *Choose general linear forms $h_1, h_2, \dots, h_m \in R_1$ which intersect in $\deg(X)$ distinct points in linearly general position where at least $e+1$ are real and smooth. Fix e smooth real points in the intersection and choose an additional linear form $h_0 \in R_1$ that vanishes only at the selected intersection points. Let L be the ideal in R generated by $h_0, h_1, \dots, h_m$.*

Step 2: *Choose a quadratic form $f \in R \setminus L^2$ that vanishes to order at least two at each of the selected intersection points.*

Step 3: *For every sufficiently small $\delta > 0$, the polynomial*

$$\delta f + h_0^2 + h_1^2 + \cdots + h_m^2$$

is nonnegative on X but not a sum of squares.

Correctness. The existence of the $h_0, h_1, \dots, h_m$ in Step 1 follows from the first paragraph in the proof of Proposition 3.2. The quadratic forms in L^2 have dimension at most $\binom{m+2}{2}$. Since second-order vanishing at e distinct points imposes at most $(m+1)e$ linear conditions, Lemma 3.1 implies that the vector space of suitable f has dimension at least

$$\begin{aligned} \dim(R_2) - (m+1)e - \tbinom{m+2}{2} &= \dim(R_2) - (m+1)\big((n+1) - (m+1)\big) - \tbinom{m+2}{2} \\ &= \dim(R_2) - (m+1)(n+1) + \tbinom{m+1}{2} = \varepsilon(X)\,, \end{aligned}$$

which justifies Step 2. For Step 3, suppose that

$$\delta f + h_0^2 + h_1^2 + \cdots + h_m^2 = g_1^2 + g_2^2 + \cdots + g_k^2$$

for some $g_j \in R_1$. It follows that each g_j vanishes at the selected intersection points. The ideal L contains all linear forms that vanish at the selected intersection points, so $(g_j)^2 \in L^2$. However, this gives a contradiction because $f \notin L^2$.

It remains to show that for a sufficiently small δ, the polynomial

$$\delta f + h_0^2 + h_1^2 + \cdots + h_m^2$$

is nonnegative on X. Let $\tilde{X} \subseteq \mathbb{A}^{n+1}(\mathbb{R})$ denote the affine cone of X, and let $\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_e \in S^n \cap \tilde{X}$ be the affine representatives with unit length for the selected intersection points. Since the selected points are nonsingular on X, the compact set $S^n \cap \tilde{X}$ is a real m-dimensional smooth manifold near each $\tilde{p}_j$ and the differentiable function $h_0^2 + h_1^2 + \cdots + h_m^2$ has a positive definite Hessian at the points $\tilde{p}_j$. Since the $\tilde{p}_j$ are zeroes and critical points for the quadratic form f, it follows that there exists a $\delta_0 > 0$ and a neighborhood U_j of $\tilde{p}_j$ in $S^n \cap \tilde{X}$ for $1 \leqslant j \leqslant e$ such that $\delta_0 f + h_0^2 + h_1^2 + \cdots + h_m^2$ is nonnegative on U_j. On the compact set

$$K'' := (S^n \cap \tilde{X}) \setminus \bigcup\nolimits_j U_j\,,$$

the function $h_0^2 + h_1^2 + \cdots + h_m^2$ is strictly positive, so

$$\delta_1 := (\inf_{K''} h_0^2 + h_1^2 + \cdots + h_m^2)/(\sup_{K''} |f|)$$

is a strictly positive real number. Hence, if $0 < \delta < \min(\delta_0, \delta_1)$, then the polynomial $\delta f + h_0^2 + h_1^2 + \cdots + h_m^2$ is nonnegative on $S^n \cap \tilde{X}$ and X. $\square$

Remark 3.4. In our context, Procedure 3.3 is a generalization of an idea going back to Hilbert. To be more precise, let $\nu_d \colon \mathbb{P}^n \to \mathbb{P}^r$ with $r = \binom{n+d}{n} - 1$ denote the dth Veronese embedding of $\mathbb{P}^n$. For the subvarieties $\nu_3(\mathbb{P}^2) \subset \mathbb{P}^{10}$ and $\nu_2(\mathbb{P}^3) \subset \mathbb{P}^{10}$, Hilbert [H3] uses a similar procedure to prove the existence of nonnegative polynomials that are not sums of squares. By working with concrete forms, Robinson uses this procedure to construct his celebrated form; see Section 4b in [R]. Again for $\nu_3(\mathbb{P}^2) \subset \mathbb{P}^{10}$ and $\nu_2(\mathbb{P}^3) \subset \mathbb{P}^{10}$, [BIK] shows that the form f in Procedure 3.3 is unique up to a constant multiple (that is, the dimension estimates are sharp) and expresses it in terms of the intersection points of the h_j.

When $\varepsilon(X) = 1$, we can clarify the difference between Σ_X and P_X. Proposition 5.10 in [Z] shows that $\varepsilon(X) = 1$ if and only if X is either a hypersurface of degree $d \geqslant 3$ or a linearly normal variety such that $\deg(X) = 2 + \operatorname{codim}(X)$ (also known as a variety of almost minimal degree). Given $\ell \in R_2^*$, $\sigma^*(\ell)$ denotes the corresponding symmetric linear map; see Section 2. Let $I(\ell)$ be the Gorenstein ideal in R generated by all homogeneous $g \in R$ such that either $\ell(fg) = 0$ for all $f \in R_{2-\deg(g)}$ or $\deg(g) > 2$.

Proposition 3.5. *Assume that X is arithmetically Cohen-Macaulay and $\varepsilon(X) = 1$. If $\ell \in \Sigma_X^*$ is an extreme ray not contained in P_X^*, then $\sigma^*(\ell)$ is positive semidefinite with $\dim \operatorname{Ker}\bigl(\sigma^*(\ell)\bigr) = m + 1$. Dually, if f lies in the boundary of Σ_X and not in the boundary of P_X, then the element f can be expressed as a sum of $m+1$ squares, but not as a sum of fewer squares.*

Proof. Lemma 2.3 asserts that the subspace $\operatorname{Ker}\bigl(\sigma^*(\ell)\bigr) \subseteq R_1$ contains a homogeneous system of parameters $h_0, h_1, \ldots, h_m$ on R. Since R is Cohen-Macaulay, this system of parameters is a regular sequence. On the other hand, Remark 4.5 in [BS] establishes that a projective variety of almost minimal degree is arithmetically Cohen-Macaulay if and only if it is arithmetically Gorenstein. Hence, the quotient ring $R' := R/(h_0, h_2, \ldots, h_m)$ is Gorenstein. Lemma 3.1 implies that the Hilbert function of R' is $(1, e, 1)$. The ideal generated by the image of $I(\ell)$ in R' under the canonical map either is trivial or contains the socle. By definition, the elements in $I(\ell)_2$ are annihilated by ℓ, so the second possibility cannot occur. Hence, we have $I(\ell) = (h_0, h_2, \ldots, h_m)$ and $\dim \operatorname{Ker}\bigl(\sigma^*(\ell)\bigr) = m + 1$.

If $f = g_1^2 + g_2^2 + \cdots + g_k^2$ lies in the boundary of Σ_X, then there exists an extreme ray $\ell \in \Sigma_X^*$ such that $\ell(f) = 0$, so $g_1, g_2, \ldots, g_k$ lie in $\operatorname{Ker}\bigl(\sigma^*(\ell)\bigr)$. Since f is not in the boundary of P_X, the element f is strictly positive on $X(\mathbb{R})$ and ℓ is not defined by evaluation at a point. The previous paragraph proves that $\dim \operatorname{Ker}\bigl(\sigma^*(\ell)\bigr) = m+1$ and this ensures that f is a sum of at most $m+1$ squares. To finish the proof, suppose that $f = g_1^2 + g_2^2 + \cdots + g_k^2$ where $k \leqslant m$ and $g_1, g_2, \ldots, g_k$ are linearly independent. If $k < m$, then choose general linear forms $g_{k+1}, g_{k+2}, \ldots, g_m$ in $\operatorname{Ker}\bigl(\sigma^*(\ell)\bigr)$. Since f is strictly positive on $X(\mathbb{R})$, the ideal J generated by $g_1, g_2, \ldots, g_m$ defines a subscheme of X that has no real zeroes. By perturbing J if necessary, we obtain a subvariety Z of X that consists of $\deg(X)$ reduced points none of which are real. Every element of R_2 vanishing at all the points in Z lies in $\operatorname{Ker}\bigl(\sigma^*(\ell)\bigr)$, so it follows that ℓ can be expressed as a linear combination of the evaluations at points in Z. As in the proof of Corollary 4.3 in [B2], we deduce that the set Z contains at most one pair of complex zeroes. Because $\deg(X) \geqslant 3$, we conclude the set Z must contain at least one real zero that produces the required contradiction. $\square$

4. Equality of the fundamental cones

This section focuses on sufficient conditions for the equality of the sums-of-squares cone Σ_X and the nonnegative cone P_X. We complete the proof of our main theorem by showing that Σ_X equals P_X whenever the quadratic deficiency vanishes. Combining our main theorem with the celebrated classification for varieties of minimal degree (see Theorem 1 in [EH]), we describe in detail the varieties for which equality holds. Using the Veronese map, we also generalize the main theorem to nonnegative forms of higher degree.

Our first proposition provides the second implication needed for the proof of Theorem 1.1.

Proposition 4.1. *If $\varepsilon(X) = 0$, then we have $\Sigma_X = \mathrm{P}_X$.*

Proof. It suffices to prove that $\mathrm{P}_X^* = \Sigma_X^*$. Given the descriptions for P_X^* and Σ_X^* in Lemma 2.1, this reduces to showing that every extreme ray of Σ_X^* is generated by evaluation at some point $p \in X(\mathbb{R})$. Suppose otherwise and consider an $\ell \in \Sigma_X^*$ that generates an extreme ray but is not determined by evaluation at a point $p \in X(\mathbb{R})$. Lemma 2.3 establishes that there exists a homogeneous system of parameters $g_0, g_1, \dots, g_m \in \mathrm{Ker}\big(\sigma^*(\ell)\big)$. Since $\varepsilon(X) = 0$, Lemma 3.1 establishes that X is a variety of minimal degree; varieties of minimal degree are arithmetically Cohen-Macaulay (see Section 4 in [EG]), so $g_0, g_1, \dots, g_m$ are also a regular sequence. Let J denote the homogeneous ideal in R generated $g_0, g_1, \dots, g_m$. Since we have $\ell(f g_j) = 0$ for all $f \in R_1$ and all $0 \leqslant j \leqslant m$, the linear functional $\ell \in R_2^*$ annihilates the subspace J_2. By taking the degree-two graded components of the associated Koszul complex and using Lemma 3.1, we obtain

$$\dim\big(\tfrac{R}{J}\big)_2 = \dim(R_2) - (m+1)\dim(R_1) + \tbinom{m+1}{2}\dim(R_0) = \varepsilon(X) = 0\,,$$

where $R_2 = J_2$. However, this yields a contradiction because the linear functional $\ell \in R_2^*$ is nonzero and does not annihilate all of R_2. Therefore, every extreme ray of Σ_X^* is generated by evaluation at some point $p \in X(\mathbb{R})$ as required. $\square$

Remark 4.2. In the proof of Proposition 4.1, the hypothesis that X is totally real is not required to establish that $\mathrm{P}_X^* = \Sigma_X^*$. As an illustration, consider the quadratic hypersurface $X' \subset \mathbb{P}^2$ defined by $x_0^2 + x_1^2 \in \mathbb{R}[x_0, x_1, x_2]$. The unique point $[0 : 0 : 1]$ in $X'(\mathbb{R})$ is singular, so X' is not totally real. Hence, $\mathrm{P}_{X'}$ is the closed half-space in R_2 for which the coefficient of x_2^2 is nonnegative. If the coefficient of x_2^2 in $f \in \mathbb{R}[x_0, x_1, x_2]_2$ is positive, then the quadratic form $f + \lambda(x_0^2 + x_1^2)$ is positive semidefinite for all sufficiently large $\lambda \in \mathbb{R}$, so $\Sigma_{X'}$ contains the interior of $\mathrm{P}_{X'}$ and $\mathrm{P}_{X'}^* = \Sigma_{X'}^*$. Amusingly, the quadratic form $\lambda(x_0^2 + x_1^2) - x_0 x_2$ for $\lambda \in \mathbb{R}$ is never positive semidefinite, so $\Sigma_{X'} \neq \mathrm{P}_{X'}$.

Proof of Theorem 1.1. If X is not a variety of minimal degree, then we have the inequality $\varepsilon(X) > 0$ and Proposition 3.2 establishes that Σ_X is a proper subset of P_X. Conversely, if X is a variety of minimal degree, then Lemma 3.1 establishes that $\varepsilon(X) = 0$ and Proposition 4.1 states that $\Sigma_X = \mathrm{P}_X$. $\square$

Beyond the conceptual explanation for the equality $\mathrm{P}_X = \Sigma_X$, Theorem 1.1 allows us to explicitly exhibit all the varieties that satisfy this condition. The classical characterization for varieties of minimal degree (see Theorem 1 in [EH]) states that a variety of minimal degree is a cone over a smooth variety of minimal degree, and a smooth variety of minimal degree is a quadratic hypersurface, the Veronese surface $\nu_2(\mathbb{P}^2) \subset \mathbb{P}^5$, or a rational normal scroll. Together with Theorem 1.1, this yields precisely the following three families in which nonnegativity is equivalent to being a sum of squares.

Example 4.3. Let $X \subset \mathbb{P}^n$ be a cone over a totally real irreducible quadric hypersurface. In other words, $R = \mathbb{R}[x_0, x_1, \dots, x_n]/I$ where I is the principal ideal generated by an indefinite quadratic form. It follows that $\deg(X) = 2 = 1 + \operatorname{codim}(X)$, so Theorem 1.1 implies that every nonnegative element of R_2 is a sum of squares.

Example 4.4. For $n \geqslant 5$, let $X \subset \mathbb{P}^n$ be the cone over the Veronese surface $\nu_2(\mathbb{P}^2) \subset \mathbb{P}^5$. Given suitable coordinates $x_0, x_1, \dots, x_n$ on $\mathbb{P}^n$, the homogeneous ideal I for X is defined by the (2×2)-minors of the generic symmetric matrix,

$$\begin{bmatrix} x_0 & x_1 & x_2 \\ x_1 & x_3 & x_4 \\ x_2 & x_4 & x_5 \end{bmatrix} .$$

In this case, we have $\deg(X) = 4 = 1 + \operatorname{codim}(X)$, so Theorem 1.1 implies that every nonnegative element of $R_2 = (\mathbb{R}[x_0, x_1, \dots, x_n]/I)_2$ is a sum of squares.

Example 4.5. For $k \geqslant 0$ and $d_k \geqslant d_{k-1} \geqslant \cdots \geqslant d_0 \geqslant 0$ with $d_k > 0$, consider the integer $n := k + d_0 + d_1 + \cdots + d_k$ and let $X \subset \mathbb{P}^n$ be the associated rational normal scroll; X is the image of the projectivized vector bundle $\mathscr{O}_{\mathbb{P}^1}(d_0) \oplus \mathscr{O}_{\mathbb{P}^1}(d_1) \oplus \cdots \oplus \mathscr{O}_{\mathbb{P}^1}(d_k)$ under the complete linear series of the tautological line bundle. In particular, X is the rational normal curve of degree n in $\mathbb{P}^n$ when $k = 0$, and X is $\mathbb{P}^n$ when $d_{k-1} = 0$ and $d_k = 1$. For suitable coordinates $x_{0,0}, x_{0,1}, \dots, x_{0,d_0}, x_{1,0}, x_{1,1}, \dots, x_{1,d_1}, \dots, x_{k,0}, x_{k,1}, \dots, x_{k,d_k}$ on $\mathbb{P}^n$, the homogeneous ideal I for X is defined by the (2×2)-minors of the block Hankel matrix,

$$\begin{bmatrix} x_{0,0} & \cdots & x_{0,d_0-1} & x_{1,0} & \cdots & x_{1,d_1-1} & \cdots & x_{k,0} & \cdots & x_{k,d_k-1} \\ x_{0,1} & \cdots & x_{0,d_0} & x_{1,1} & \cdots & x_{1,d_1} & \cdots & x_{k,1} & \cdots & x_{k,d_k} \end{bmatrix} .$$

Since $\deg(X) = d_0 + d_1 + \cdots + d_k = n - k = 1 + \operatorname{codim}(X)$, Theorem 1.1 implies that every nonnegative element of $R_2 = (\mathbb{R}[x_{0,0}, x_{0,1}, \dots, x_{k,d_k}]/I)_2$ is a sum of squares.

The following remark explains why it is sufficient to consider quadratic forms.

Remark 4.6. The union of Theorem 1.1 with the classification for varieties of minimal degree also allows us to identify when every nonnegative form on X of degree $2d$ for $d > 1$ is a sum of squares. Geometrically, this is equivalent to recognizing when the dth Veronese embedding of $X \subseteq \mathbb{P}^n$ is a variety of minimal degree. The degree of every curve on the image $\nu_d(X)$ is a multiple of d, so $\nu_d(X)$ does not contain any lines. Assume that $\nu_d(X)$ is a variety of minimal degree. It cannot be a cone over a smooth variety of minimal degree or a rational normal scroll with $k > 0$ because these varieties contain lines. It follows that $\nu_d(X)$ is either a rational normal curve or the Veronese surface $\nu_2(\mathbb{P}^2) \subset \mathbb{P}^5$. Therefore, every nonnegative form on X of degree $2d$ for $d > 1$ is a sum of squares if and only if $\nu_d(X)$ is a nondegenerate curve of degree n in $\mathbb{P}^n$, or $X = \mathbb{P}^2$ and $d = 2$.

As an example, the rational quartic curve in $C \subset \mathbb{P}^3$ defined by

$$[y_0 : y_1] \mapsto [y_0^4 : y_0^3 y_1 : y_0 y_1^3 : y_1^4]$$

is not a variety of minimal degree. However, its image under the second Veronese map $\nu_2(C) \subset \mathbb{P}^8$ is the rational normal curve of degree eight which is a variety of minimal degree. Hence, every nonnegative quartic form on C is a sum of squares.

We conclude this section by viewing our main theorem through the lens of measure theory.

Remark 4.7. Fix a positive integer d and let X be a real projective variety with homogeneous coordinate ring R. Let $W := S^n \cap \tilde{X}$ be the intersection of the affine cone $\tilde{X} \subseteq \mathbb{A}^{n+1}(\mathbb{R})$ of X with the unit sphere S^n. A measure on $X(\mathbb{R})$ corresponds to a measure on W that is invariant under the antipodal map. Any such measure μ defines a linear functional $\ell \in R_{2d}^*$ by sending $f \in R_{2d}$ to $\int_W f \, d\mu$. The ***truncated moment problem*** asks for a characterization of the $\ell \in R_{2d}^*$ that come from integration with respect to a measure on X; see Definition 3.1 in [L1]. Such functionals are nonnegative and belong to $\mathrm{P}_{\nu_d(X)}^*$. Moreover, every element of $\mathrm{P}_{\nu_d(X)}^*$ has this form. As a result, the truncated moment problem on X can be reinterpreted as asking for a characterization of the cone $\mathrm{P}_{\nu_d(X)}^*$. If B_ℓ is the moment matrix of ℓ (that is, the matrix associated to the quadratic form of ℓ with respect to a monomial basis for R_d), then it is necessary that B_ℓ be positive semidefinite or equivalently that $\ell \in \Sigma_{\nu_d(X)}^*$. From this viewpoint, Theorem 1.1 classifies the varieties X for which the truncated moment problem in degree two is equivalent to deciding positive semidefiniteness of the moment matrix. As in Remark 4.6, this equivalence holds for the truncated moment problem in degree $2d$ where $d > 1$ if and only if $\nu_d(X)$ is either a rational normal curve or the Veronese surface $\nu_2(\mathbb{P}^2) \subset \mathbb{P}^5$.

5. The intrinsic perspective

In this section, we shift our perspective from an embedded variety to linear series on an abstract variety. This approach gives us greater flexibility which will be used in applications. For example, by working with positively multigraded polynomial rings, we list the cases in which every nonnegative multihomogeneous polynomial is a sum of squares.

Let Y be an m-dimensional totally real projective variety; it is a geometrically integral projective scheme over $\operatorname{Spec}(\mathbb{R})$ such that the set $Y(\mathbb{R})$ of real points is Zariski dense. Consider a Cartier divisor D on Y that is locally defined by rational functions with real coefficients, and fix a nondegenerate basepoint-free linear series $V \subseteq H^0\big(Y, \mathcal{O}_Y(D)\big)$. Since D is defined over $\mathbb{R}$, we may regard V as a real vector space. Let $\sigma \colon \operatorname{Sym}^2\big(H^0\big(Y, \mathcal{O}_Y(D)\big)\big) \to H^0\big(Y, \mathcal{O}_Y(2D)\big)$ denote the canonical multiplication map and let $2V := \sigma\big(\operatorname{Sym}^2(V)\big) \subseteq H^0\big(Y, \mathcal{O}_Y(2D)\big)$. Given a real point $p \in Y(\mathbb{R})$ and a section $s \in H^0\big(Y, \mathcal{O}_Y(2D)\big)$, the ***sign of*** s ***at*** p is $\operatorname{sgn}_p(s) := \operatorname{sgn}(\lambda) \in \{-1, 0, 1\}$ where $U \subseteq Y$ is a neighborhood of the point $p \in Y$ over which the line bundle $\mathcal{O}_Y(D)$ is trivial, $\varsigma \in H^0\big(U, \mathcal{O}_Y(D)\big)$ is a generator of $\mathcal{O}_Y(D)|_U$, and the section $\lambda \in H^0(U, \mathcal{O}_Y)$ is defined by $s|_U = \lambda \varsigma^2$. The sign of s at p is independent of the choice of U and ς; see Section 2.4 in [S1]. The section s is ***nonnegative*** if $s(p) = \operatorname{sgn}_p(s) \geqslant 0$ for all $p \in Y(\mathbb{R})$.

The central objects of study, in this intrinsic setting, become

$$\mathrm{P}_{Y,V} := \{ s \in 2V : s(p) \geqslant 0 \text{ for all } p \in Y(\mathbb{R}) \}, \qquad \text{and}$$

$$\Sigma_{Y,V} := \left\{ s \in 2V : \begin{array}{l} \text{there exist } t_1, t_2, \ldots, t_k \in V \text{ such} \\ \text{that } s = \sigma(t_1^2) + \sigma(t_2^2) + \cdots + \sigma(t_k^2) \end{array} \right\}.$$

We again have $\Sigma_{Y,V} \subseteq \mathrm{P}_{Y,V}$. To describe the properties of these subsets, let n be the projective dimension of $|V|$, let $\varphi\colon Y \to \mathbb{P}^n$ be the associated morphism, and let $X := \varphi(Y)$. The linear series $|V|$ is nondegenerate if and only if $X \subseteq \mathbb{P}^n$ is nondegenerate. The kernel of the composition of the canonical homomorphisms of graded rings $\mathbb{R}[x_0, x_1, \dots, x_n] \cong \mathrm{Sym}(V) \to \mathrm{Sym}\big(H^0(Y, \mathcal{O}_Y(D))\big)$ and $\mathrm{Sym}\big(H^0(Y, \mathcal{O}_Y(D))\big) \to \bigoplus_{j\in\mathbb{N}} H^0\big(Y, \mathcal{O}_Y(jD)\big)$ is the unique saturated ideal I vanishing on X. It follows that the homogeneous coordinate ring of X is the quotient $R = \mathrm{Sym}(V)/I$, and the induced inclusion of graded rings is the homomorphism $\varphi^\sharp\colon R \to \bigoplus_{j\in\mathbb{N}} H^0\big(Y, \mathcal{O}_Y(jD)\big)$.

The next proposition shows that these collections of $\mathrm{P}_{Y,V}$ and $\Sigma_{Y,V}$ are closely related to the cones P_X and Σ_X, and it provides an alternative version of Theorem 1.1.

Theorem 5.1. *We have $\varphi^\sharp(\Sigma_X) = \Sigma_{Y,V}$. If $\varphi\big(Y(\mathbb{R})\big)$ is dense in the strong topology on $X(\mathbb{R})$, then we also have $\varphi^\sharp(\mathrm{P}_X) = \mathrm{P}_{Y,V}$, and $\mathrm{P}_{Y,V} = \Sigma_{Y,V}$ if and only if X is a variety of minimal degree.*

Proof. By construction, we have $\varphi^\sharp(R_1) = V$ and $\varphi^\sharp(R_2) = 2V$, which establishes the first assertion. Since φ sends a real point to a real point, we have the inclusion $\mathrm{P}_{Y,V} \subseteq \varphi^\sharp(\mathrm{P}_X)$. Conversely, each real point in X lies in the closure of the image of a real point in Y by assumption, so we have $\varphi^\sharp(\mathrm{P}_X) \subseteq \mathrm{P}_{Y,V}$. Combining the first two parts with Theorem 1.1 yields the third part. ☐

Remark 5.2. When the map φ has finite fibers of odd lengths, the condition on φ in Theorem 5.1 is automatically satisfied. In particular, the hypothesis holds when φ is an embedding. Indeed, complex conjugation fixes the fiber over a real point. Since the fibers have odd lengths, conjugation must fix at least one point in each fiber over a real point, so φ maps $Y(\mathbb{R})$ surjectively onto $X(\mathbb{R})$.

Without placing some restrictions on the map φ, the theorem is false.

Example 5.3. Consider the linear series

$$V = \langle x_0^2, x_1^2, \dots, x_n^2 \rangle \subseteq H^0\big(\mathbb{P}^n, \mathcal{O}_{\mathbb{P}^n}(2)\big)\,.$$

The corresponding morphism $\varphi\colon \mathbb{P}^n \to \mathbb{P}^n$ is not surjective on real points. In this case, $(\varphi^\sharp)^{-1}(\mathrm{P}_{Y,V})$ consists of all quadratic forms that are nonnegative on the closed nonnegative orthant in $\mathbb{R}^{n+1}$ (that is, the copositive forms), and this collection is strictly larger than the cone of all nonnegative quadratic forms; see Section 3.6.1 in [BPT].

The following explains our restriction on the linear series.

Observation 5.4. If $2V \neq H^0\big(Y, \mathcal{O}_Y(2D)\big)$, then we claim that there is a nonnegative section in $H^0\big(Y, \mathcal{O}_Y(2D)\big)$ that is not a sum of squares. Since the linear series V is basepoint-free, there exists $t_0, t_1, \dots, t_n \in V$ with no common zeroes, so $\sigma(t_0^2) + \sigma(t_1^2) + \cdots + \sigma(t_n^2) \in H^0\big(Y, \mathcal{O}_Y(2D)\big)$ is strictly positive on $Y(\mathbb{R})$. Our assumption on $2V$ implies that there is a section $s \in H^0\big(Y, \mathcal{O}_Y(2D)\big) \setminus 2V$. It follows that the section

$$\sigma(t_0^2) + \sigma(t_1^2) + \cdots + \sigma(t_n^2) - \delta s \in H^0\big(Y, \mathcal{O}_Y(2D)\big)$$

cannot be a sum of squares for all $\delta \in \mathbb{R}$. On the other hand, this section is nonnegative for all sufficiently small $\delta > 0$, because $Y(\mathbb{R})$ is a compact set and, for any section s, we have

$$\overline{\{p \in X(\mathbb{R}) : \operatorname{sgn}_p(s) < 0\}} \subseteq \{p \in X(\mathbb{R}) : \operatorname{sgn}_p(s) \leqslant 0\} .$$

To illustrate the power of Theorem 5.1, we capture all of the previously known situations in which nonnegativity is equivalent to being a sum of squares.

Example 5.5. For $n \geqslant 0$ and $d \geqslant 1$, consider $Y = \mathbb{P}^n$ and the linear series $V = H^0\big(\mathbb{P}^n, \mathcal{O}_{\mathbb{P}^n}(d)\big)$. The corresponding map φ is the Veronese embedding, so Theorem 5.1 implies that every nonnegative homogeneous polynomial of degree $2d$ is a sum of squares (that is, $\mathrm{P}_{Y,V} = \Sigma_{Y,V}$) if and only if $X = \varphi(\mathbb{P}^n)$ is a variety of minimal degree. Moreover, we have $\deg(X) = d^n = \binom{n+d}{n} - n = 1 + \operatorname{codim}(X)$ in only three cases:

- $n = 1$: all nonnegative binary forms are sums of squares, and X is a rational normal curve;
- $d = 1$: all nonnegative quadratic forms are sums of squares, and $X = \mathbb{P}^n$;
- $d = 2$ and $n = 2$: all nonnegative ternary quartics are sums of squares, and X is the Veronese surface.

In particular, we recover Hilbert's famous characterization of when every nonnegative homogeneous polynomial is a sum of squares; see [H3] or Section 3.1.2 in [BPT]. Even better, we provide a new geometric interpretation for the exceptional case of ternary quartics.

Example 5.6. For $k \geqslant 2$, $n_i \geqslant 1$, and $d_i \geqslant 1$ where $1 \leqslant i \leqslant k$, consider the product $Y = \mathbb{P}^{n_1} \times \mathbb{P}^{n_2} \times \cdots \times \mathbb{P}^{n_k}$ and the linear series

$$V = H^0\big(Y, \mathcal{O}_{\mathbb{P}^{n_1}}(d_1) \boxtimes \mathcal{O}_{\mathbb{P}^{n_2}}(d_2) \boxtimes \cdots \boxtimes \mathcal{O}_{\mathbb{P}^{n_k}}(d_k)\big) .$$

The corresponding map φ is the Segre-Veronese embedding, so Theorem 5.1 implies that every nonnegative multihomogeneous polynomial of degree $(2d_1, 2d_2, \ldots, 2d_k)$ is a sum of squares (that is, $\mathrm{P}_{Y,V} = \Sigma_{Y,V}$) if and only if $X = \varphi(Y)$ is a variety of minimal degree. Moreover, we have

$$\begin{aligned} \deg(X) &= d_1^{n_1} d_2^{n_2} \cdots d_k^{n_k} \tbinom{n_1+n_2+\cdots+n_k}{n_1,n_2,\ldots,n_k} = d_1^{n_1} d_2^{n_2} \cdots d_k^{n_k} \tfrac{(n_1+n_2+\cdots+n_k)!}{n_1!n_2!\cdots n_k!} \\ &= \tbinom{n_1+d_1}{n_1}\tbinom{n_2+d_2}{n_2} \cdots \tbinom{n_k+d_k}{n_k} - n_1 - n_2 - \cdots - n_k \\ &= 1 + \operatorname{codim}(X) \end{aligned}$$

in precisely two cases:

- $k = 2$, $n_1 = 1$, and $d_2 = 1$,
- $k = 2$, $n_2 = 1$, and $d_1 = 1$.

By symmetry, both cases assert that all nonnegative biforms that are quadratic in one set of variables and binary in the other set of variables are sums of squares, and X is a rational normal scroll associated to a vector bundle of the form $\bigoplus_j \mathcal{O}_{\mathbb{P}^1}(1)$. In particular, we recover and provide a new interpretation for Theorem 8.4 in [CLR].

Since two of the three families of varieties of minimal degree are toric varieties, the intrinsic descriptions can be expressed in terms of a polynomial ring with an appropriate grading.

Example 5.7. For $n \geqslant 5$, consider the cone $Y \subset \mathbb{P}^n$ over the Veronese surface $\mathbb{P}^2 \hookrightarrow \mathbb{P}^5$ and the complete linear series $V = H^0\big(Y, \mathcal{O}_Y(H)\big)$ where H is a hyperplane divisor (compare with Example 4.4). Hence, Y is a simplicial normal toric variety with class group $\mathbb{Z}^1$, and the Cox homogeneous coordinate ring is $S := \mathbb{R}[y_0, y_1, \dots, y_{n-3}]$ where $\deg(y_i) = 1$ for $0 \leqslant i \leqslant 2$ and $\deg(y_j) = 2$ for $3 \leqslant j \leqslant n-3$. Since $\operatorname{Pic}(X)$ has index two within the class group, it follows that $V = H^0\big(Y, \mathcal{O}_Y(H)\big) \cong S_2$. The image of Y is a variety of minimal degree, so Theorem 5.1 implies that every nonnegative element in S_4 is a sum of squares. An element of S_4 is a linear combination of the 15 monomials $y_0^4, y_0^3 y_1, \dots, y_2^4$, the $6n-30$ monomials $y_0^2 y_j, y_0 y_1 y_j, \dots, y_2^2 y_j$ where $3 \leqslant j \leqslant n-3$, and the $\binom{n-4}{2}$ monomials $y_3^2, y_3^2 y_4, \dots, y_{n-3}^2$; the vector space S_4 has dimension $\frac{1}{2}n^2 + \frac{3}{2}n - 5$. Contrary to the sentence preceding Theorem 8.4 in [CLR], this inserts the exceptional case of ternary quartics from Example 5.5 into an infinite family.

Example 5.8. For integers $k > 0$ and $d_k \geqslant d_{k-1} \geqslant \cdots \geqslant d_0 > 0$, consider the projectivized vector bundle $Y = \mathbb{P}\big(\mathcal{O}_{\mathbb{P}^1}(d_0) \oplus \mathcal{O}_{\mathbb{P}^1}(d_1) \oplus \cdots \oplus \mathcal{O}_{\mathbb{P}^1}(d_k)\big)$ and the complete linear series $V = H^0\big(Y, \mathcal{O}_Y(1)\big)$; compare with Example 4.5. Hence, Y is a $(k+1)$-dimensional smooth toric variety with class group $\mathbb{Z}^2 = \operatorname{Pic}(X)$; see pages 6–7 in [EH]. By choosing a suitable basis for the class group, the Cox homogeneous coordinate ring is $S := \mathbb{R}[y_0, y_1, \dots, y_{k+2}]$ where the degree of y_j in $\mathbb{Z}^2$ is given by the jth column of the matrix

$$\begin{bmatrix} 1 & 1 & 0 & d_0 - d_1 & d_0 - d_2 & \cdots & d_0 - d_k \\ 0 & 0 & 1 & 1 & 1 & \cdots & 1 \end{bmatrix} .$$

It follows that $V = H^0\big(Y, \mathcal{O}_Y(1)\big) \cong S_{(1,1)}$. Since the image of Y is a variety of minimal degree, Theorem 5.1 implies that every nonnegative element in $S_{(2,2)}$ is a sum of squares. An element of $S_{(2,2)}$ is a linear combination of monomials that are quadratic in the variables $y_2, y_3, \dots, y_{k+2}$; the vector space $S_{(2,2)}$ has dimension

$$(3 - 2d_0)\binom{k+2}{2} + (k+2)(d_0 + d_1 + \cdots + d_k) .$$

The special case $d_0 = d_1 = \cdots = d_k = 1$ retrieves Example 5.6.

Remark 5.9. Example 5.8 excludes two types of scrolls: a cone over a rational normal curve (that is, $k = 0$ or $d_{k-1} = 0$), which has class group isomorphic to $\mathbb{Z}^1$, and a cone over a smooth rational normal scroll (that is, $d_{k-1} \neq 0$ and $d_0 = 0$), which has class group isomorphic to $\mathbb{Z}^2$. The minor modifications to Example 5.8 required for both types are left to the interested reader.

Remark 5.10. The multihomogeneous forms in Example 5.8 also have a useful interpretation in terms of matrix polynomials. By viewing $f \in S_{(2,2)}$ as a quadratic form in the variables $y_2, y_3, \dots, y_{k+2}$, we obtain a symmetric matrix F with homogeneous entries in $\mathbb{R}[y_0, y_1]$. Lemma 3.78 in [BPT] basically shows that F is pointwise positive semidefinite if and only if f is nonnegative and

$$F = G_1^{\mathsf{T}} G_1 + G_2^{\mathsf{T}} G_2 + \cdots + G_k^{\mathsf{T}} G_k$$

for some matrices $G_1, G_2, \dots, G_k$ with entries in $\mathbb{R}[y_0, y_1]$ if and only if f is a sum of squares. Hence, the fact that every nonnegative element in $S_{(2,2)}$ is a sum of squares becomes a slight strengthening of Theorem 3.80 in [BPT] in which each entry is homogeneous (although not necessarily of the same degree).

6. Nonnegative sparse polynomials

This section examines certain sparse Laurent polynomials—those Laurent polynomials in which the exponent vector of each monomial appearing with a nonzero coefficient lies in a fixed lattice polytope. We characterize the Newton polytopes Q such that every nonnegative polynomial with support contained in $2Q$ is a sum of squares.

Let M be an m-dimensional affine lattice, let $M_{\mathbb{R}} := M \otimes_{\mathbb{Z}} \mathbb{R}$ be the associated real vector space, and let $T := \operatorname{Spec}(\mathbb{R}[M])$ be the corresponding split real torus. By choosing an isomorphism $M \cong \mathbb{Z}^m$, we identify the group ring $\mathbb{R}[M]$ with the Laurent polynomial ring $\mathbb{R}[z_1^{\pm 1}, z_2^{\pm 1}, \dots, z_m^{\pm 1}]$. Given $f = \sum_{\mathbf{u} \in M} c_{\mathbf{u}} \, z^{\mathbf{u}} \in \mathbb{R}[M]$, its ***Newton polytope*** is $\operatorname{New}(f) := \operatorname{conv}\{\mathbf{u} \in M : c_{\mathbf{u}} \neq 0\} \subset M_{\mathbb{R}}$. The Laurent polynomial f is ***nonnegative***, denoted by $f \geqslant 0$, if the evaluation of f at every point in $T(\mathbb{R})$ is nonnegative. Equivalently, for all $\mathbf{w} \in \mathbb{N}^m$ with sufficiently large entries, the polynomial $z^{2\mathbf{w}} f \in \mathbb{R}[z_1, z_2, \dots, z_m]$ is nonnegative on $\mathbb{R}^m$. Fix an m-dimensional lattice polytope Q in $M_{\mathbb{R}}$. For $k \in \mathbb{N}$, we write kQ is the k-fold Minkowski sum of Q. The lattice polytope Q is ***k-normal*** if, for each $\mathbf{u} \in (kQ) \cap M$, there exist $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in Q \cap M$ such that $\mathbf{u} = \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k$; compare with Definition 2.2.9 in [CLS]. Following Section 3 in [S2], the h^****-polynomial*** of Q defined by the equation

$$\mathrm{h}_0^*(Q) + \mathrm{h}_1^*(Q)\, t + \cdots + \mathrm{h}_m^*(Q)\, t^m := (1-t)^{m+1} \sum_{k \geqslant 0} \big|(kQ) \cap M\big| \, t^k \, .$$

The central objects of study, in this polyhedral setting, are

$$\mathrm{P}_Q := \{ f \in \mathbb{R}[M] : \operatorname{New}(f) \subseteq 2Q \text{ and } f \geqslant 0 \} \qquad \text{and}$$

$$\Sigma_Q := \left\{ f \in \mathbb{R}[M] : \begin{array}{l} \text{there exists } g_1, g_2, \dots, g_k \in \mathbb{R}[M] \text{ such} \\ \text{that } \operatorname{New}(g_j) \subseteq Q \text{ for all } 1 \leqslant j \leqslant k \\ \text{and } f = g_1^2 + g_2^2 + \cdots + g_k^2 \end{array} \right\} .$$

Once again, we have $\Sigma_Q \subseteq \mathrm{P}_Q$. To describe the properties of these subsets, let $X \subseteq \mathbb{P}^n$ be the embedded projective toric variety determined by the lattice polytope Q. More explicitly, the number of lattice points in Q is $n+1 = |Q \cap M|$, the polyhedral affine monoid associated to Q is $C(Q) := \mathbb{N}\cdot\{(q,1) : q \in Q \cap M\} \subset M \oplus \mathbb{Z}$, and the toric variety is $X = \operatorname{Proj}\big(\mathbb{R}[C(Q)]\big) \subseteq \mathbb{P}^n$; compare with Section 2.3 in [CLS]. The lattice points in Q also yield the canonical inclusion map $\eta \colon T \to X$.

We elucidate this framework with a couple of examples.

Example 6.1. The arithmetic–geometric inequality establishes that the celebrated Motzkin polynomial $f := z_1^4 z_2^2 + z_1^2 z_2^4 + 1 - 3 z_1^2 z_2^2$ is nonnegative on $\mathbb{R}^2$. Since $\operatorname{New}(f) = \operatorname{conv}\{(4,2), (2,4), (0,0)\}$, it follows that $f \in \mathrm{P}_Q$ where

$$Q := \operatorname{conv}\{(2,1), (1,2), (0,0)\} \, .$$

The polytope Q is 2-normal because every lattice point in $\operatorname{New}(f)$ is a sum of lattice points in Q; for instance, $(2,2) = (1,1) + (1,1)$. Moreover, the h^*-polynomial is

$$1 + t + t^2 = (1-t)^3 \sum_{k \geqslant 0} \big|(kQ) \cap M\big| \, t^k = (1-t)^3 \sum_{k \geqslant 0} \big(\tfrac{3}{2} k^2 + \tfrac{3}{2} k + 1\big) t^k \, ,$$

so $\mathrm{h}_2^*(Q) = 1$.

Example 6.2. If Q is an $(m-2)$-fold pyramid over the simplex

$$\operatorname{conv}\{(0,0),(2,0),(0,2)\} \subset \mathbb{R}^2\,,$$

then the embedded projective toric variety X is the cone over the Veronese surface defined in Example 4.4. Likewise, if Q is the Cayley polytope of the line segments $[0,d_0],[0,d_1],\dots,[0,d_k]$ (see Definition 2.1 in [BN]), then the embedded projective toric variety X is the rational normal scroll defined in Example 4.5.

To establish that 2-normality is a necessary condition for $\mathrm{P}_Q = \Sigma_Q$, we have a better version of Observation 5.4, that provides an explicit bound on the coefficient δ.

Lemma 6.3. *If $Q \subset M_{\mathbb{R}}$ is a lattice polytope that is not 2-normal, then Σ_Q is a proper subset of P_Q.*

Proof. Since Q is not 2-normal, there is a lattice point $\mathbf{u} \in 2Q \cap M$ that cannot be written as a sum of lattice points in $Q \cap M$. If $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ denote the vertices of Q, then $\mathbf{u}$ is a convex rational linear combination of $2\mathbf{v}_1, 2\mathbf{v}_2, \dots, 2\mathbf{v}_k$, which are the vertices of $2Q$. By clearing the denominators, we obtain

$$(r_1 + r_2 + \cdots + r_k)\mathbf{u} = 2r_1\mathbf{v}_1 + 2r_2\mathbf{v}_2 + \cdots + 2r_k\mathbf{v}_k,$$

where $r_1, r_2, \dots, r_k \in \mathbb{N}$ and $r_1 + r_2 + \cdots + r_k > 0$. Consider the Laurent polynomial

$$f := r_1 z^{2\mathbf{v}_1} + r_2 z^{2\mathbf{v}_2} + \cdots + r_k z^{2\mathbf{v}_k} - (r_1 + r_2 + \cdots + r_k) z^{\mathbf{u}}\,.$$

Clearly, $\operatorname{New}(f) \subseteq 2Q$, and our choice of $\mathbf{u}$ guarantees that f is not a sum of squares. On the other hand, the inequality of weighted arithmetic and geometric means shows that f is nonnegative. Therefore, we have $f \in \mathrm{P}_Q \setminus \Sigma_Q$. $\square$

The following result is a strengthening of Theorem 5.1 for projective toric varieties, because the condition on real points is now both necessary and sufficient.

Theorem 6.4. *We have $\mathrm{P}_Q = \Sigma_Q$ if and only if $\mathrm{h}_2^*(Q) = 0$ and $\eta\big(T(\mathbb{R})\big)$ is dense in the strong topology on $X(\mathbb{R})$.*

Proof. We first verify that Q is 2-normal. If $\mathrm{P}_Q = \Sigma_Q$, then Lemma 6.3 shows that Q is 2-normal. Assuming that $\mathrm{h}_2^*(Q) = 0$, we confirm that Q is 2-normal by induction on the dimension m. Since every lattice polytope of dimension at most 2 is normal (that is, k-normal for all k), the base case for the induction holds; see Corollary 2.2.13 in [CLS]. If $m \geqslant 3$, then our assumption combined together with inequality (4) in [S3] proves that $\mathrm{h}_m^*(Q) = 0$. Similarly, inequality (6) in [S3] (with $i = 1$) shows that $\mathrm{h}_{m-1}^*(Q) = 0$. Hence, Ehrhart-Macdonald reciprocity (see Theorem 4.4 in [BR]) establishes that neither Q nor $2Q$ have any interior lattice points. It follows that every lattice point $\mathbf{u} \in (2Q) \cap M$ is contained in a face of $2Q$. Since every facet of $2Q$ equals $2F$ for some face F of Q and the monotonicity of h^*-polynomials (see Theorem 3.3 in [S2]) ensures that $\mathrm{h}_2^*(F) \leqslant \mathrm{h}_2^*(Q) = 0$, the induction hypothesis shows that F is 2-normal. In particular, we have $\mathbf{u} = \mathbf{v}_1 + \mathbf{v}_2$ for some $\mathbf{v}_1, \mathbf{v}_2 \in F \cap M \subset Q \cap M$, and we conclude that Q is also 2-normal.

The 2-normality of Q ensures that $R_2 = \mathbb{R}[C(Q)]_2 \cong \mathbb{R} \cdot \{(2Q) \cap M\}$ and, by definition, we have $\mathrm{h}_1^*(Q) = n + 1 = |Q \cap M| = \dim \mathbb{R}[C(Q)]_1$, which together imply that $\mathrm{P}_Q = \mathrm{P}_X$ and $\Sigma_Q = \Sigma_X$. Since we have $\mathrm{h}_0^*(Q) = 1 = \dim \mathbb{R}[C(Q)]_0$, Lemma 3.1 establishes that $\mathrm{h}_2^*(Q) = \varepsilon(X)$ and we have $\mathrm{h}_2^*(Q) = 0$ if and only if X is a variety of minimal degree. If $\eta\big(T(\mathbb{R})\big)$ is dense in the strong topology on $X(\mathbb{R})$, then Theorem 1.1 proves that $\mathrm{P}_Q = \Sigma_Q$ if and only if $\mathrm{h}_2^*(Q) = \varepsilon(X) = 0$.

Thus, it remains to show that $\mathrm{P}_Q = \Sigma_Q$ implies that $\eta\big(T(\mathbb{R})\big)$ is dense in the strong topology on $X(\mathbb{R})$.

Assume that $\mathrm{P}_Q = \Sigma_Q$ and suppose that $\eta\big(T(\mathbb{R})\big)$ is not dense in the strong topology on $X(\mathbb{R})$. By translating Q in $M_{\mathbb{R}}$ if necessary, we may assume that Q contains the origin and this lattice point corresponds to the 0th coordinate of the map $\eta\colon T \to X \subseteq \mathbb{P}^n$. Let $U_0 \cong \mathbb{A}^n$ denote the distinguished open subset of $\mathbb{P}^n$ determined by the vanishing of the 0th coordinate and set $W := X \cap U_0 \subset \mathbb{A}^n$. Since $\eta\big(T(\mathbb{R})\big) \subseteq W$, our supposition implies that $\eta\big(T(\mathbb{R})\big)$ is not dense in the strong topology on $W(\mathbb{R})$. As a consequence, there exists a point $p \in W(\mathbb{R})$ and a real number $\delta > 0$ such that the open ball $B_\delta(p)$ of radius δ centered at p is completely contained in $W(\mathbb{R}) \setminus \overline{\eta\big(T(\mathbb{R})\big)}$. Choose coordinates $x_0, x_1, \dots, x_n$ on $\mathbb{P}^n$ with $p = [1 : p_1 : p_2 : \cdots : p_n] \in \mathbb{P}^n(\mathbb{R})$. Consider the polynomial in $\mathbb{R}[x_0, x_1, \dots, x_n]$

$$\hat{f} := (x_1 - p_1 x_0)^2 + (x_2 - p_2 x_0)^2 + \cdots + (x_n - p_n x_0)^2 - \delta x_0^2$$

and the corresponding Laurent polynomial $f = \eta^\sharp(\hat{f}) \in \mathbb{R}[M]$ where the canonical ring homomorphism associated to η is $\eta^\sharp\colon \mathbb{R}[x_0, x_1, \dots, x_n] \to \mathbb{R}[M]$. By construction, we have $\operatorname{New}(f) \subseteq 2Q$ and f is nonnegative on $T(\mathbb{R})$, so $f \in \mathrm{P}_Q$. The assumption $\mathrm{P}_Q = \Sigma_Q$ guarantees that there exists $g_1, g_2, \dots, g_k \in \mathbb{R}[M]$ such that $f = g_1^2 + g_2^2 + \cdots + g_k^2$. It follows that $\operatorname{New}(g_j) \subseteq \frac{1}{2}\operatorname{New}(f) = Q$, so there are linear forms $\hat{g}_j \in \mathbb{R}[x_0, x_1, \dots, x_n]$ satisfying $g_j = \eta^\sharp(\hat{g})$ for $1 \leqslant j \leqslant k$. Since $\eta^\sharp$ is injective, we obtain $\hat{f} = \hat{g}_1^2 + \hat{g}_2^2 + \cdots + \hat{g}_k^2$. However, this is impossible because $\hat{f}(p) = -\delta < 0$. Therefore, we conclude that $\eta\big(T(\mathbb{R})\big)$ is dense in the strong topology on $X(\mathbb{R})$. $\square$

Remark 6.5. Combining Example 6.1 and Theorem 6.4, we only establish that there exists a nonnegative polynomial with the same Newton polytope as the Motzkin polynomial, which is not a sum of squares. Nevertheless, one can easily prove that the Motzkin polynomial is not a sum of squares; see Exercise 3.97 in [BPT].

The ensuing propositions, which practically classify the lattice polytopes Q with $\mathrm{h}_2^*(Q) = 0$, increase the utility of Theorem 6.4. They also advance the general program of classifying polytopes based on their h^*-polynomials.

Proposition 6.6. *Let $Q \subset M_{\mathbb{R}}$ be an m-dimensional lattice polytope. We have $\mathrm{h}_2^*(Q) = 0$ if and only if Q is 2-normal and Q is the affine $\mathbb{Z}$-linear image, surjective on integral points, of a polytope Q' where $Q' \subset M'_{\mathbb{R}}$ is either the $(m-2)$-fold pyramid over* $\operatorname{conv}\{(0,0),(2,0),(0,2)\} \subset \mathbb{R}^2$ *or the Cayley polytope of m line segments.*

Proof. The first paragraph in the proof of Theorem 6.4 shows that Q is 2-normal whenever $\mathrm{h}_2^*(Q) = 0$, and the second paragraph shows that the 2-normality of Q implies that $0 = h_2^* = \varepsilon(X)$ and X is a variety of minimal degree. Since X is a toric variety, the classification for varieties of minimal degree (see Theorem 1 in [EH]) establishes that X is either a cone over the Veronese surface or a rational normal scroll. It follows from Example 6.2 that X is projectively equivalent to the embedded toric variety X' determined by a polytope Q' where Q' is either an $(m-2)$-fold pyramid over $\operatorname{conv}\{(0,0),(2,0),(0,2)\} \subset \mathbb{R}^2$ or the Cayley polytope of m line segments. The $\mathbb{R}$-algebras $\mathbb{R}[C(Q)]$ and $\mathbb{R}[C(Q')]$ are isomorphic, so Theorem 2.1 in [G] implies that the affine monoids $C(Q)$ and $C(Q')$ are also isomorphic. This isomorphism extends to a $\mathbb{Z}$-linear homomorphism $\beta\colon M' \oplus \mathbb{Z} \to M \oplus \mathbb{Z}$, because $C(Q')$ contains a lattice basis, and β is injective, because Q is full-dimensional.

Restricting to the affine slice at height 1, we obtain the affine map $\alpha\colon M' \to M$ such that $\alpha(Q') = Q$. Since β, and hence α, sends the generators of $C(Q')$ to the generators of $C(Q)$, every lattice point in Q is the image of a lattice point in Q'. $\square$

Corollary 6.7. *Let $Q' \subset M'_{\mathbb{R}}$ be either the $(m-2)$-fold pyramid over*

$$\operatorname{conv}\{(0,0),(2,0),(0,2)\} \subset \mathbb{R}^2$$

or the Cayley polytope of m line segments, and let $\alpha\colon M' \to M$ be an affine map. If $Q := \alpha(Q')$, every integer point in Q comes from an integer point in Q', and the determinant of the linear component of α is a nonzero odd integer, then we have $\mathrm{P}_Q = \Sigma_Q$.

Proof. Propositon 6.6 implies that $\mathrm{h}_2^*(Q) = 0$, so it is enough to prove, by Theorem 6.4, that $\eta\big(T(\mathbb{R})\big)$ is dense in the strong topology on $X(\mathbb{R})$. The embedded projective toric variety $X \subseteq \mathbb{P}^n$ determined Q is a compactification of the dense algebraic torus $T'' := X \cap \{x_0x_1\cdots x_n \neq 0\}$, so it suffices to show that the induced map $\eta''\colon T(\mathbb{R}) \to T''(\mathbb{R})$ obtained from η is surjective. If M'' denotes the sublattice generated by the lattice points in Q, then induced map η'' corresponds to an injective ring homomorphism from $\mathbb{R}[M''] \to \mathbb{R}[M]$. Since $\mathbb{R}[M'']$ is the image of map $\mathbb{R}[M'] \to \mathbb{R}[M]$ defined by the linear component of α, it follows that η'' is a finite morphism with degree equal to the determinant of the linear component. As in Remark 5.2, η'' is surjective when the degree is odd. $\square$

To refine our classification, we need an auxiliary invariant. The ***degree*** of Q is the smallest $j \in \mathbb{N}$ such that, for $1 \leqslant k \leqslant m-j$, k-fold Minkowski sum kQ contains no interior lattice point.

Remark 6.8. One can directly verify that a pyramid over

$$\operatorname{conv}\{(0,0),(2,0),(0,2)\} \subset \mathbb{R}^2$$

or a Cayley polytope of line segments has degree one.

With a few small adjustments to the proof of Proposition 6.6, we obtain the following.

Proposition 6.9. *For an m-dimensional lattice polytope $Q \subset M_{\mathbb{R}}$, the following are equivalent:*

(a) *Q is normal and $\mathrm{h}_2^*(Q) = 0$,*
(b) *Q is a polytope of degree one,*
(c) *we have $\mathrm{h}_2^*(Q) = \mathrm{h}_3^*(Q) = \cdots = \mathrm{h}_m^*(Q) = 0$.*

Proof. (a) $\implies$ (b): Since $\mathrm{h}_2^*(Q) = 0$, the proof of Proposition 6.6 provides $\mathbb{Z}$-linear homomorphism $\beta\colon M' \oplus \mathbb{Z} \to M \oplus \mathbb{Z}$. By changing bases on the source and target, we can assume that β is represented by a diagonal matrix (for example, its Smith normal form), which sends a lattice basis in $C(Q')$ to certain multiplies in $C(Q)$. Since Q is normal, the monoid $C(Q)$ also contains a lattice basis. It follows that β is a lattice isomorphism. By restricting to the affine slice at height 1, we conclude that Q and Q' are affinely isomorphic.

(b) $\implies$ (c): As in the proof of Proposition 6.6, this follows immediately from Ehrhart-Macdonald reciprocity (see Theorem 4.4 in [BR]).

(c) $\implies$ (a): We need to show that Q is k-normal for all $k > 1$. Since every m-dimensional polytope is k-normal for all $k \geqslant m-1$ (see Theorem 2.2.12 in

[CLS]), we may assume $k < m-1$. When $2 \leqslant k \leqslant m-j$, one can adapt the arguments from the first paragraph in the proof of Proposition 6.6 to show Q is k-normal. □

Remark 6.10. By combining Proposition 6.6 and the proof of Propositon 6.9, we obtain a new interpretation and a new proof for the main theorem in [BN]. Specifically, Theorem 2.5 in [BN] characterizes the m-dimensional lattice polytopes of degree one as either an $(m-2)$-fold pyramid over $\operatorname{conv}\{(0,0),(2,0),(0,2)\} \subset \mathbb{R}^2$ or the Cayley polytope of m line segments.

We end with a family of nonnormal polytopes Q for which we have $\mathrm{h}_2^*(Q) = 0$. By examining the proof of Proposition 6.6, we see that smallest such example must have dimension at least 5.

Example 6.11. Let $m \geqslant 5$ be an odd integer and fix $k \in \mathbb{N}$. If $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m$ denotes the standard basis for $\mathbb{Z}^m$, then consider the simplex

$$Q := \operatorname{conv}\left\{ \begin{array}{l} \mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \\ \mathbf{e}_1 + \cdots + \mathbf{e}_{\frac{m-1}{2}} + k\,\mathbf{e}_{\frac{m+1}{2}} + \cdots + k\,\mathbf{e}_{m-1} + (k+1)\,\mathbf{e}_m \end{array} \right\}.$$

The h^*-polynomial for Q is $1 + k\,t^{(m+1)/2}$, so $\mathrm{h}_2^*(Q) = 0$; see Section 1 in [H2]. When k is even, Corollary 6.7 implies that $\mathrm{P}_Q = \Sigma_Q$. When k is odd, $\eta\big(T(\mathbb{R})\big)$ is not dense in the strong topology on $X(\mathbb{R})$, so Theorem 6.4 implies that $\mathrm{P}_Q \neq \Sigma_Q$.

Acknowledgements

The authors thank Bernd Sturmfels for stimulating their interest in convex algebraic geometry. They also thank Matthias Beck, Mircea Mustaţă, Maurice Rojas, and Mike Roth for helpful conversations.

References

[BN] Victor Batyrev and Benjamin Nill, *Multiples of lattice polytopes without interior lattice points* (English, with English and Russian summaries), Mosc. Math. J. **7** (2007), no. 2, 195–207, 349. MR2337878 (2008g:52018)

[BR] Matthias Beck and Sinai Robins, *Computing the continuous discretely*, Undergraduate Texts in Mathematics, Springer, New York, 2007. Integer-point enumeration in polyhedra. MR2271992 (2007h:11119)

[B1] Eberhard Becker, *Valuations and real places in the theory of formally real fields*, Real algebraic geometry and quadratic forms (Rennes, 1981), Lecture Notes in Math., vol. 959, Springer, Berlin-New York, 1982, pp. 1–40. MR683127 (84f:12011)

[B2] Grigoriy Blekherman, *Nonnegative polynomials and sums of squares*, J. Amer. Math. Soc. **25** (2012), no. 3, 617–635, DOI 10.1090/S0894-0347-2012-00733-4. MR2904568

[BPT] *Semidefinite optimization and convex algebraic geometry*, MOS-SIAM Series on Optimization, vol. 13, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013. Edited by Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas. MR3075433

[BIK] Grigoriy Blekherman, Sadik Iliman, and Martina Kubitzker, *Dimensional differences between faces of the cones of nonnegative polynomials and sums of squares*, Int. Math. Res. Not. IMRN, DOI 10.1093/imrn/rnu202.

[BS] Markus Brodmann and Peter Schenzel, *Arithmetic properties of projective varieties of almost minimal degree*, J. Algebraic Geom. **16** (2007), no. 2, 347–400, DOI 10.1090/S1056-3911-06-00442-5. MR2274517 (2008b:14085)

[CLS] David A. Cox, John B. Little, and Henry K. Schenck, *Toric varieties*, Graduate Studies in Mathematics, vol. 124, American Mathematical Society, Providence, RI, 2011. MR2810322 (2012g:14094)

[CLR] Man Duen Choi, Tsit Yuen Lam, and Bruce Reznick, *Real zeros of positive semidefinite forms. I*, Math. Z. **171** (1980), no. 1, 1–26, DOI 10.1007/BF01215051. MR566480 (81d:10012)

[EG] David Eisenbud and Shiro Goto, *Linear free resolutions and minimal multiplicity*, J. Algebra **88** (1984), no. 1, 89–133, DOI 10.1016/0021-8693(84)90092-9. MR741934 (85f:13023)

[EH] David Eisenbud and Joe Harris, *On varieties of minimal degree (a centennial account)*, Algebraic geometry, Bowdoin, 1985 (Brunswick, Maine, 1985), Proc. Sympos. Pure Math., vol. 46, Amer. Math. Soc., Providence, RI, 1987, pp. 3–13, DOI 10.1090/pspum/046.1/927946. MR927946 (89f:14042)

[G] Joseph Gubeladze, *The isomorphism problem for commutative monoid rings*, J. Pure Appl. Algebra **129** (1998), no. 1, 35–65, DOI 10.1016/S0022-4049(97)00063-7. MR1626643 (99h:20094)

[H1] Joe Harris, *Algebraic geometry*, Graduate Texts in Mathematics, vol. 133, Springer-Verlag, New York, 1992. A first course. MR1182558 (93j:14001)

[H2] Akihiro Higashitani, *Counterexamples of the conjecture on roots of Ehrhart polynomials*, Discrete Comput. Geom. **47** (2012), no. 3, 618–623, DOI 10.1007/s00454-011-9390-4. MR2891252

[H3] David Hilbert, *Ueber die Darstellung definiter Formen als Summe von Formenquadraten* (German), Math. Ann. **32** (1888), no. 3, 342–350, DOI 10.1007/BF01443605. MR1510517

[J] Jean-Pierre Jouanolou, *Théorèmes de Bertini et applications* (French), Progress in Mathematics, vol. 42, Birkhäuser Boston, Inc., Boston, MA, 1983. MR725671 (86b:13007)

[L1] Jean Bernard Lasserre, *Moments, positive polynomials and their applications*, Imperial College Press Optimization Series, vol. 1, Imperial College Press, London, 2010. MR2589247 (2011c:90001)

[L2] Monique Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging applications of algebraic geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 157–270, DOI 10.1007/978-0-387-09686-5_7. MR2500468 (2010j:13054)

[L3] S. L'vovsky, *On inflection points, monomial curves, and hypersurfaces containing projective curves*, Math. Ann. **306** (1996), no. 4, 719–735, DOI 10.1007/BF01445273. MR1418349 (99e:14033)

[RG] Motakuri Ramana and A. J. Goldman, *Some geometric results in semidefinite programming*, J. Global Optim. **7** (1995), no. 1, 33–50, DOI 10.1007/BF01100204. MR1342934 (96i:90059)

[R] Bruce Reznick, *Some concrete aspects of Hilbert's 17th Problem*, Real algebraic geometry and ordered structures (Baton Rouge, LA, 1996), Contemp. Math., vol. 253, Amer. Math. Soc., Providence, RI, 2000, pp. 251–272, DOI 10.1090/conm/253/03936. MR1747589 (2001i:11042)

[S1] Claus Scheiderer, *A Positivstellensatz for projective real varieties*, Manuscripta Math. **138** (2012), no. 1-2, 73–88, DOI 10.1007/s00229-011-0484-3. MR2898748

[S2] Richard P. Stanley, *A monotonicity property of h-vectors and h^*-vectors*, European J. Combin. **14** (1993), no. 3, 251–258, DOI 10.1006/eujc.1993.1028. MR1215335 (94f:52016)

[S3] A. Stapledon, *Inequalities and Ehrhart δ-vectors*, Trans. Amer. Math. Soc. **361** (2009), no. 10, 5615–5626, DOI 10.1090/S0002-9947-09-04776-X. MR2515826 (2010e:52026)

[Z] F. L. Zak, *Projective invariants of quadratic embeddings*, Math. Ann. **313** (1999), no. 3, 507–545, DOI 10.1007/s002080050271. MR1678545 (2000b:14071)

School of Mathematics, Georgia Tech, 686 Cherry Street, Atlanta, Georgia, 30332
E-mail address: greg@math.gatech.edu

Department of Mathematics & Statistics, Queen's University, Kingston, Ontario, K7L 3N6, Canada
E-mail address: ggsmith@mast.queensu.ca

Departamento de Matemáticas, Universidad de los Andes, Carrera 1 No. 18a 10, Edificio H, Primer Piso, 111711 Bogotá, Colombia
E-mail address: mvelasco@uniandes.edu.co