# Technical Report

TR-2005-015

**Representations of positive polynomials on non-compact semialgebraic sets via KKT ideals**

by

James Demmel, Jiawang Nie, Victoria Powers

MATHEMATICS AND COMPUTER SCIENCE

EMORY UNIVERSITY

# Representations of positive polynomials on non-compact semialgebraic sets via KKT ideals

James Demmel * Jiawang Nie † and Victoria Powers‡

December 5, 2005

**Abstract**

This paper studies the representation of a positive polynomial $f(x)$ on a noncompact semialgebraic set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_s(x) \geq 0\}$ modulo its KKT (Karush-Kuhn-Tucker) ideal. Under the assumption that the minimum value of $f(x)$ on $S$ is attained at some KKT point, we show that $f(x)$ can be represented as sum of squares (SOS) of polynomials modulo the KKT ideal if $f(x) > 0$ on $S$; furthermore, when the KKT ideal is radical, we have that $f(x)$ can be represented as sum of squares (SOS) of polynomials modulo the KKT ideal if $f(x) \geq 0$ on $S$. This is a generalization of results in [19], which discuss the SOS representations of nonnegative polynomials over gradient ideals.

**Key words:** Polynomials, semialgebraic set, sum of squares (SOS), Karush-Kuhn-Tucker (KKT) system, KKT ideal.

## 1 Introduction

There has been much recent interest in developing algorithms for optimizing polynomial functions on semialgebraic sets using representation theorems from real algebraic geometry for positive polynomials. The idea is to turn a problem of this type into a question about the existence of a representation involving sums of squares (SOS) polynomials and the polynomials defining the semialgebraic set – an SOS representation for short. This can then be implemented as a semidefinite program (SDP), and solved numerically [22, 26]. In the global case, i.e., when the semialgebraic set is the whole space $\mathbb{R}^n$, an SOS representation gives a convex relaxation of the original problem and hence a lower bound for the minimum. In the case of compact semialgebraic sets, using results on SOS representations, Lasserre [15] gave a procedure for finding natural sequences of computationally feasible SDP relaxations of the original problem, whose solutions converge to a solution of the original problem, under a certain constraint qualification condition.

However, these methods do not always work well. In the global case, the resulting SDP might not have a solution even if the polynomial attains a minimum. This can also occur in the case of a semialgebraic set which is not compact. In the compact case, the procedure proposed by Lasserre in [15] can generate a sequence of lower bounds which converge to the

*Department of Mathematics and EECS, University of California, Berkeley, CA 94720. Email: demmel@cs.berkeley.edu. The author was supported in part by the National Science Foundation (ELA-0122599)

†Department of Mathematics, University of California, Berkeley, CA 94720. Email: njw@math.berkeley.edu. The author was supported in part by the National Science Foundation (ELA-0122599).

‡Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322. Email: vicki@mathcs.emory.edu. The author was supported in part by the National Security Agency (H98230-05-1-00).

minimum under a certain constraint qualification condition. Recently, Nie and Schweighofer [20] gave results on the convergence rate of these lower bounds. However, Lasserre's procedure is based on SOS representations of positive polynomials on compact semialgebraic sets and the lower bounds generated usually have only asymptotic convergence, i.e., the finite convergence is usually not guaranteed, as shown in an example due to Stengle [31].

As is well known, most numerical optimization methods targeting local (including global) minimizers are often based on the optimality conditions: the Karush-Kuhn-Tucker (KKT) system. In the unconstrained global case, the KKT system reduces to zero gradient condition. Thus an approach with great potential in global optimization is to look at SOS representations of a polynomial modulo its gradient ideal or an ideal arising from the KKT system. There is some related work in SOS representations of positive polynomials modulo certain ideals, such as Hanzon and Jibetean [10], Laurent [16], Parrilo [26], Jibetean and Laurent [14].

Nie, Demmel and Sturmfels [19] proposed using SOS representations of positive polynomials modulo their gradient ideals, i.e., the ideals generated by all the partial derivatives. This kind of representation works reasonably well in finding the global minimum of a polynomial when the minimum is attained at some point. In this paper, we generalize the results in [19] and give similar representation theorems using a KKT system for polynomials positive on a basic closed semialgebraic set. Note that we do not need to assume that the semialgebraic set is compact, which is necessary in Schmüdgen's or Putinar's Theorem (see below). We will also discuss the application of this representation theorem to finding the minimum of a polynomial on a noncompact basic closed semialgebraic set.

Denote by $\mathbb{R}[X] = \mathbb{R}[x_1, \ldots, x_n]$ the ring of polynomials in $X = (x_1, \cdots, x_n)$ with real coefficients and write $\sum \mathbb{R}[X]^2$ for the cone of polynomials which are sums of squares in $\mathbb{R}[X]$. We say $f(x)$ is SOS if $f \in \sum \mathbb{R}[X]^2$. For a finite set $F = \{g_1, \ldots, g_k\} \subset \mathbb{R}[X]$, let $S(F)$ denote the basic closed semialgebraic set generated by $F$, i.e.,

$$S(F) = \{\alpha \in \mathbb{R}^n \mid g_1(\alpha) \geq 0, \ldots, g_r(\alpha) \geq 0\}.$$

A polynomial $f \in \mathbb{R}[X]$ is *PSD* (resp. *PD*) if $f(\alpha) \geq 0$ (resp. $f(\alpha) > 0$) for all $\alpha \in \mathbb{R}^n$. We define PSD (resp. PD) on a subset $K$ of $\mathbb{R}^n$ similarly and denote these by "$f \geq 0$ on $K$" (resp. "$f > 0$ on $K$").

As is well-known, for $n \geq 2$, there always exist $f \in \mathbb{R}[X]$ that are PSD but not SOS. An SOS decomposition of a polynomial $f$ is an explicit witness to the fact that $f$ is PSD. More generally, one can ask for a witness to the fact that $f > 0$ or $f \geq 0$ on some $S(F)$.

Denote by $M(F)$ the *quadratic module* generated by the $\{g_i\}$, i.e.,

$$M(F) := \left\{ s_0 + s_1 g_1 + \cdots + s_k g_k \,\middle|\, s_i \in \sum \mathbb{R}[X]^2 \right\}.$$

We write $P(F)$ for the *preorder* generated by $\{g_i\}$, i.e.,

$$P(F) = \left\{ \sum_{\epsilon \in \{0,1\}^k} s_\epsilon g_1^{\epsilon_1} \ldots g_k^{\epsilon_k} \,\middle|\, s_\epsilon \in \sum \mathbb{R}[X]^2 \right\}.$$

Note that $P(F)$ is simply the quadratic module generated by the $2^k$ products of the $g_i$'s.

Clearly, if $f \in M(F)$, then $f \geq 0$ on $S(F)$ and an expression $f = s_0 + s_1 g_1 + \cdots + s_k g_k$ is an explicit witness to the fact that $f \geq 0$ on $S(F)$, and similarly for $f \in P(F)$. In general it is not true that $f \geq 0$ on $S(F)$, or $f > 0$ on $S(F)$, implies that $f \in M(F)$. However, we have the following remarkable theorem:

**Theorem 1.1** (Schmüdgen [28])**.** *If $S(F)$ is compact, then $f > 0$ on $S(F)$ implies $f \in P(F)$.*

In general, even with the assumption that $S(F)$ is compact, this does not hold if we replace $P(F)$ by $M(F)$, nor if we assume only that $f \geq 0$ on $S(F)$. See [24] for details.

A quadratic module $M$ is *archimedean* if there exists $p(x) \in M$ such that the set $\{x \in \mathbb{R}^n : p(x) \geq 0\}$ is compact, equivalently, if there exists $N \in \mathbb{N}$ such that $N - \sum_{i=1}^m x_i^2 \in M$, see [5, 5.3.8]. Note that if $M(S)$ or $P(S)$ is archimedean, then $S$ is compact.

**Theorem 1.2** (Putinar [27])**.** . *Suppose $M(F)$ is archimedean, then for any $f \in \mathbb{R}[X]$, $f > 0$ on $S(F)$ implies $f \in M(F)$.*

**Remarks 1.1.** (i) There are examples of compact $S(F)$ for which the corresponding quadratic module $M(F)$ is not archimedean and the conclusion of Putinar's Theorem does not hold, see Example 6.3.1 in [5]. In the case of the preorder $P(F)$, it is a deep theorem of Schmüdgen [28] that if $S(F)$ is compact then $P(F)$ is archimedean.
(ii) The Putinar and Schmüdgen Theorems say that if the conditions are satisfied, then there always exists an SOS representation of $f$ positive on $S(F)$. Thus, in this case, there is trivially an SOS representation modulo the gradient ideal. On the other hand, as is well-known, all of the assumptions of the theorem are necessary.

Fix $F = \{g_1, \dots, g_s\} \subseteq \mathbb{R}[X]$ and let $S = S(F)$, $P = P(F)$, and $M = M(F)$. Given $f \in \mathbb{R}[X]$, let $f^*$ denote the minimum of $f$ on $S$, i.e., the solution to the optimization problem

$$f^* = \min \quad f(x) \tag{1.1}$$
$$s.t. \quad g_i(x) \geq 0, \quad i = 1, \cdots, s, \tag{1.2}$$

The *KKT system* associated to this optimization problem is

$$F \triangleq \nabla f - \sum_{j=1}^{s} \lambda_j \nabla g_j \tag{1.3}$$
$$g_j \geq 0, \;\; \lambda_j g_j = 0, \;\; j = 1, \cdots, s, \tag{1.4}$$

where the variables $\lambda = \begin{bmatrix}\lambda_1 & \cdots \lambda_s\end{bmatrix}^T$ are called Lagrange multipliers and $\nabla f$ denotes the gradient of $f$, i.e., the vector of partial derivatives. Under some regularity conditions (for example, if the gradients of the $g_j$ are linearly independent, see [21]), the local (including global) minimizers of $f(x)$ on $S$ satisfy the KKT system above. Most numerical algorithms targeting local (including global) minimizers often generate a sequence of points $\{(x^{(k)}, \lambda^{(k)})\}$ whose limit or accumulation points satisfy the KKT system (1.3)-(1.4). We refer to [21] and the references therein for general numerical methods in nonlinear programming. Here we ignore the condition that the Lagrange multipliers $\lambda_j$ are nonnegative. As we will see in this paper, we do not need to use the nonnegativeness of $\lambda_j$ in the representation theorems. Actually, taking the sign of $\lambda_j$ into account will make the representation more complicated.

Let $F = (F_1, \dots, F_n)$, i.e., $F_i = \frac{\partial f}{\partial x_i} - \sum_{j=1}^s \lambda_j \frac{\partial g_j}{\partial x_i}$. We now define the *KKT ideal* $I_{KKT}$ and the *varieties* associated with KKT system (1.3)-(1.4) as follows:

$$\begin{aligned}
I_{KKT} &= \langle F_1, \cdots, F_n, \lambda_1 g_1, \cdots, \lambda_s g_s \rangle, \\
V_{KKT} &= \{(x, \lambda) \in \mathbb{C}^n \times \mathbb{C}^s : p(x, \lambda) = 0, \; \forall p \in I_{KKT}\}, \\
V_{KKT}^{\mathbb{R}} &= \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^s : p(x, \lambda) = 0, \; \forall p \in I_{KKT}\}.
\end{aligned}$$

The associated KKT preorder $P_{KKT}$ and KKT quadratic module $M_{KKT}$ are defined as

$$\begin{aligned}
P_{KKT} &= P + I_{KKT} \\
M_{KKT} &= M + I_{KKT}.
\end{aligned}$$

Finally, let $\mathcal{H}$ be the set satisfying constraints (1.2):

$$\mathcal{H} = \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^s : g_j(x) \geq 0, \;\; j = 1, \cdots, t\}.$$

Notice that $I_{KKT}$, $P_{KKT}$, and $M_{KKT}$ are all subsets of $\mathbb{R}[X,\lambda]$ instead of $\mathbb{R}[X]$, where $\lambda = (\lambda_1,\ldots,\lambda_s)$.

The main results of this paper are the following: Assume $f^*$ is attained at some KKT point. If $I_{KKT}$ is radical and $f \geq 0$ on $V^{\mathbb{R}}_{KKT} \cap \mathcal{H}$, then $f \in P_{KKT}$; if $I_{KKT}$ is not radical but $f > 0$ on $V^{\mathbb{R}}_{KKT} \cap \mathcal{H}$, then $f \in P_{KKT}$.

This paper is organized as follows. Section 2 introduces some backgrounds in algebraic geometry and real algebra. Section 3 studies the SOS representations of polynomials modulo KKT ideals. Section 4 shows the applications of this kind of SOS representations in optimization on noncompact semialgebraic sets. We draw some conclusions in Section 5.

# 2 Preliminaries

In this section we present some notions and results from algebraic geometry and real algebra needed for our discussion. Readers may consult [2, 3, 4] for more details.

Throughout this section, denote by $\mathbb{R}[Z]$ the ring of polynomials in $Z = (z_1,\cdots,z_m)$ with real coefficients. If we set $Z = (x_1,\cdots,x_n,\lambda_1,\cdots,\lambda_s)$, we get the polynomial ring $\mathbb{R}[x_1,\cdots,x_n,\lambda_1,\cdots,\lambda_s]$ introduced in the preceding section. We present some properties of ideals and varieties in $\mathbb{R}[Z]$, which are also true for $\mathbb{R}[x_1,\cdots,x_n,\lambda_1,\cdots,\lambda_s]$ in particular.

A subset $I \subset \mathbb{R}[Z]$ is called an *ideal* if $p \cdot q \in I$ for any $p \in I$ and $q \in \mathbb{R}[Z]$. Given an ideal $I \subseteq \mathbb{R}[Z]$, define its *variety* to be the set

$$V(I) \quad = \quad \{z \in \mathbb{C}^m : p(z) = 0 \text{ for all } p \in I\},$$

and its *real variety* to be

$$V^{\mathbb{R}}(I) \quad = \quad \{z \in \mathbb{R}^m : p(z) = 0 \text{ for all } p \in I\}.$$

An ideal $I \subseteq \mathbb{R}[X]$ is said to be *zero-dimensional* if its variety $V(I)$ is a finite set. This condition is much stronger than requiring that the real variety $V^{\mathbb{R}}(I)$ be a finite set. For example, $I = \langle Z_1^2 + Z_2^2\rangle$ is not zero-dimensional, however the real variety $V^{\mathbb{R}}(I) = \{(0,0)\}$ consists of one point of the curve $V(I)$.

A nonempty variety $V \subseteq \mathbb{C}^m$ is *irreducible* if there do not exist two proper subvarieties $V_1, V_2 \subsetneq V$ such that $V = V_1 \cup V_2$. The reader should note that in this paper, "irreducible" means that the set of **complex** zeros cannot be written as a proper union of subvarieties defined by **real** polynomials.

Given any ideal $I$ of $\mathbb{R}[Z]$, its *radical ideal* $\sqrt{I}$ is defined to be the following ideal

$$\sqrt{I} \quad = \quad \left\{q \in \mathbb{R}[Z] \ : \ q^{\ell} \in I \text{ for some } \ell \in \mathbb{N}\right\}.$$

Obviously it holds that $I \subseteq \sqrt{I}$. We say that $I$ is a *radical ideal* if $\sqrt{I} = I$.

We need versions of Hilbert's Weak and Strong Nullstellensatz for varieties defined by polynomials in $\mathbb{R}[Z]$. The following are normally stated for ideals in $\mathbb{C}[Z]$. However, keeping in mind that $I(V)$ lies in $\mathbb{C}[Z]$, they hold as stated.

**Theorem 2.1** (Hilbert's Weak Nullstellensatz)**.**
*If $I$ is an ideal in $\mathbb{R}[Z]$ such that $V(I) = \emptyset$ then $1 \in I$.*

Given a variety $V \subset \mathbb{C}^m$, denote by $I(V)$ the ideal consisting of polynomials which vanish on $V$. For any ideal $I$, we obviously have $I \subset I(V(I))$. Actually we have the following strong theorem due to Hilbert:

**Theorem 2.2** (Hilbert's Strong Nullstellensatz)**.**
*If $I$ is an ideal in $\mathbb{R}[Z]$ then $I(V(I)) = \sqrt{I}$.*

Finally, we need the following theorem, which is a real version of Hilbert's Weak Nullstellensatz, see e.g. [5, 4.2.13].

**Theorem 2.3.** *Suppose $S(F)$ and $P(F)$ are defined as above, then $S(F) = \emptyset$ if and only if $-1 \in P(F)$.*

We will also need the following lemma which is the "variety version" of Lagrangian interpolation:

**Lemma 2.4** (Lemma 1 [19])**.** *Let $V_1, \cdots, V_r$ be pairwise disjoint varieties of $\mathbb{C}^m$. Then there exist polynomials $p_1, \cdots, p_r \in \mathbb{C}[X]$ such that $p_i(V_j) = \delta_{ij}$, where $\delta_{ij}$ is the Kronecker delta function.*

**Remark 2.5.** If each $V_\ell$ is conjugate symmetric, i.e., a point $z \in \mathbb{C}^m$ belongs to $V_\ell$ if and only if its complex conjugate $\bar{z} \in V_\ell$, then the polynomials $p_\ell$ can be chosen such that $p_\ell \in \mathbb{R}[Z]$, since we can replace $p_i(Z)$ by $(p_i(Z) + \bar{p}_i(Z))/2$, where $\bar{p}_i(Z)$ is obtained from $p_i(Z)$ by conjugating its coefficients.

# 3 Sums of squares modulo KKT ideals

In this section, we discuss the SOS representation of nonnegative and positive polynomials on a noncompact basic closed semialgebraic set $S$ modulo the corresponding KKT ideals.

When $S = \mathbb{R}^n$, the problem is reduced to the SOS representation of nonnegative or positive polynomials modulo gradient ideals, as discussed in [19]. Nie, Demmel and Sturmfels [19] showed that if a polynomial $f \in \mathbb{R}[X]$ is nonnegative on its real gradient variety and its gradient ideal is radical, then $f$ has a representation as a sum of squares modulo the gradient ideal; if the gradient ideal of $f(x)$ is not radical but $f(x)$ is positive on its real gradient variety, then $f(x)$ also has a representation as a sum of squares modulo its gradient ideal. When $f(x)$ is just nonnegative on its real gradient variety and its gradient ideal is not radical, the polynomial $f(x)$ might not have such an SOS representation modulo its gradient ideal, as shown in Example 1 in [19].

In this section we generalize this result to real polynomials which are nonnegative on a basic closed semialgebraic set. The real gradient variety and real gradient ideal are replaced by a variety and an ideal defined by the KKT system corresponding to the optimization (1.1)-(1.2).

Fix $F = \{g_1, \ldots, g_s\} \subseteq \mathbb{R}[X]$ and let $S = S(F)$, $P = P(F)$, and $M = M(F)$ as in the previous section. Given $f \in \mathbb{R}[X]$, define the ideal, varieties, preorder and quadratic module associated to the KKT system (1.3)-(1.4) as above.

As is well-known, if an ideal $I$ in a polynomial ring is zero-dimensional, then every PSD polynomial $f$ on $V(I)$ is SOS modulo $\sqrt{I}$. This follows easily from the Chinese Remainder Theorem, for a proof see, e.g., [26]. From this fact, we immediately obtain the following representation theorem:

**Theorem 3.1.** *Assume $I_{KKT}$ is zero-dimensional and radical. If $f(x)$ is nonnegative on $V_{KKT}^{\mathbb{R}} \cap \mathcal{H}$, then $f(x)$ belongs to $M_{KKT}$.*

Using a proof similar to that of Theorem 8 in [19], we can remove the restrictive hypothesis that $I_{KKT}$ is zero-dimensional, however to obtain the most general result we must replace the quadratic module $M_{KKT}$ by the preorder $P_{KKT}$.

**Theorem 3.2.** *Assume $I_{KKT}$ is radical. If $f(x)$ is nonnegative on $V_{KKT}^{\mathbb{R}} \cap \mathcal{H}$, then $f(x)$ belongs to $P_{KKT}$.*

To prove the above theorem, we need the following lemma which is a generalization of [19, Lemma 2]:

**Lemma 3.3.** *Let $W$ be an irreducible component of $V_{KKT}$. Then $f(x)$ is constant on $W$.*

*Proof.* We first note that the Lagrangian function

$$\mathcal{L}(x,\lambda) = f(x) + \sum_{i=1}^{s} \lambda_i g_i(x)$$

is equal to $f(x)$ on $V_{KKT}$, which contains $W$. Choose two arbitrary points $(x^{(1)}, \lambda^{(1)})$, $(x^{(2)}, \lambda^{(2)})$ in $W$. Since $W$ is irreducible in $\mathbb{C}^{n+s}$, it is connected in $\mathbb{C}^{n+s}$ (see [30]). Any two points in a connected algebraic variety in $\mathbb{C}^{n+s}$ can be connected by an algebraic curve. This curve may be singular, but it is a projection of some nonsingular curve. Hence there exists a smooth path $\varphi(\tau) = (x(\tau), \lambda(\tau))\,(0 \le \tau \le 1)$ connecting $(x^{(1)}, \lambda^{(1)})$ and $(x^{(2)}, \lambda^{(2)})$. Let $\mu_j(\tau)$ be the principle complex square root of $\lambda_j(\tau),\ 1 \le j \le s$. From the KKT system (1.3)-(1.4), we can see that the function

$$f(x) + \sum_{i=1}^{s} \mu_i^2 g_i(x)$$

has zero gradient on the path $\varphi(\tau)\,(0 \le \tau \le 1)$. By the Mean Value Theorem, it follows that $f(x^{(1)}) = f(x^{(2)})$ and hence that $f$ is constant on $W$. $\square$

*Proof of Theorem 3.2.* Decompose $V_{KKT}$ into its irreducible components, then by Lemma 3.3, $f$ is constant on each of them. Let $W_0$ be the union of all the components whose intersection with $\mathcal{H}$ is empty, and group together the components on which $f$ attains the same value, say $W_1, \ldots, W_r$. Suppose $f = \alpha_i \ge 0$ on $W_i$.

We have $V_{KKT} = W_0 \cup W_1 \cup \cdots \cup W_r$, and the $W_i$ are pairwise disjoint. Note that by our definition of irreducible, each $W_i$ is conjugate symmetric. By Lemma 2.4, there exist polynomials $p_0, p_1, \cdots, p_r \in \mathbb{R}[x,\lambda]$ such that $p_i(W_j) = \delta_{ij}$, where $\delta_{ij}$ is the Kronecker delta function.

By assumption, $W_0 \cap \mathcal{H} = \emptyset$ and so, by Theorem 2.3, there are SOS polynomials $v_\theta$ $(\theta \in \{0,1\}^t)$ such that

$$-1 \equiv \sum_{\theta \in \{0,1\}^s} v_\theta g_1^{\theta_1} \cdots h_s^{\theta_t} \overset{def}{=} v_0 \mod I(W_0).$$

We have $f = (f+\frac{1}{2})^2 - (f^2 + (\frac{1}{2})^2) = f_1 + v_0 \cdot f_2$ for the SOS polynomials $f_1 = (f+\frac{1}{2})^2, f_2 = f^2 + (\frac{1}{2})^2$. Then

$$f \equiv f_1 + v_0 f_2 \equiv \sum_{\theta \in \{0,1\}^s} u_\theta g_1^{\theta_1} \cdots g_t^{\theta_t} \overset{def}{=} q_0 \mod I(W_0)$$

for some SOS polynomials $u_\theta$ $(\theta \in \{0,1\}^s)$. Recall that $f(x) = \alpha_i$, a constant, on each $W_i (1 \le i \le r)$. Set $q_i(x) = \sqrt{\alpha_i}$, then $f(x) = q_i(x)^2$ on $I(W_i)$.

Now let $q = q_0(p_0)^2 + \left(\sum_{i=1}^r q_i p_i\right)^2$. Then $f - q$ vanishes on $V_{KKT}$ and hence $f - q \in I_{KKT}$ since $I_{KKT}$ is radical. It follows that $f \in P_{KKT}$. $\square$

**Remark 3.4.** The assumption that $I_{KKT}$ is radical is needed in Theorem 3.2, as shown by Example 3.4 in [19]. However, when $I_{KKT}$ is not radical, the conclusion also holds if $f(x)$ is strictly positive on $V_{KKT}^{\mathbb{R}}$.

**Theorem 3.5.** *If $f > 0$ on $V_{KKT}^{\mathbb{R}} \cap \mathcal{H}$, then $f$ belongs to $P_{KKT}$.*

*Proof.* As in the proof of Theorem 3.2, we decompose $V_{KKT}$ into subvarieties $W_0, W_1, \cdots, W_r$ such that $W_0 \cap \mathcal{H} = \emptyset$, and for $i = 1, \ldots r$, $W_i \cap \mathcal{H} \neq \emptyset$ and $f$ is constant on $W_i$. Since each $W_i$, $i > 0$ contains at least one real point and $f > 0$ on $V_{KKT}^{\mathbb{R}}$, each $\alpha_i > 0$. The $W_i$ were chosen so that each $\alpha_i$ is distinct, hence the $W_i$'s are pairwise disjoint.

Consider the primary decomposition $I_{KKT} = \cap_{i=0}^{r} J_i$ corresponding to our decomposition of $V_{KKT}$, i.e., $V(J_i) = W_i$ for $i = 0, 1, \cdots, r$. Since $W_i \cap W_j = \emptyset$, we have $J_i + J_j = \mathbb{R}[x, \lambda]$ by Theorem 2.1. The Chinese Remainder Theorem, see e.g. [6, 2.13], implies that there is an isomorphism

$$\rho : \mathbb{R}[x, \lambda]/I_{KKT} \to \mathbb{R}[x, \lambda]/J_0 \times \mathbb{R}[x, \lambda]/J_1 \times \cdots \times \mathbb{R}[x, \lambda]/J_r.$$

For any $p \in \mathbb{R}[x, \lambda]$, let $[p]$ and $\rho([p])_i$ denote the equivalence classes of $p$ in $\mathbb{R}[x, \lambda]/I_{KKT}$ and $\mathbb{R}[x, \lambda]/J_i$ respectively.

Recall that that $V(J_0) \cap \mathcal{H} = \emptyset$, hence by Theorem 2.3 there exist SOS polynomials $u_\theta$ $(\theta \in \{0, 1\}^s)$ such that

$$-1 \equiv \sum_{\theta \in \{0,1\}^s} u_\theta \rho([g_1^{\theta_1}])_0 \cdots \rho([g_s^{\theta_s}])_0 \stackrel{def}{=} u_0 \mod J_0 \ .$$

As in the proof of Theorem 3.2, we write $f = f_1 - f_2$ for SOS polynomials $f_1, f_2$ and then we have

$$f \equiv f_1 + u_0 f_2 \equiv \sum_{\theta \in \{0,1\}^s} v_\theta (\rho([g_1^{\theta_1}]))_0 \cdots (\rho([g_s^{\theta_s}]))_0 \stackrel{def}{=} q_0 \mod J_0$$

for some SOS polynomials $v_\theta$ $(\theta \in \{0, 1\}^s)$. Thus the preimage $\rho^{-1}((q_0, 0, \cdots, 0)) \in P_{KKT}$.

Now on each $W_i$, $1 \leq i \leq r$, $f = \alpha_i > 0$, and hence $(f/\alpha_i) - 1$ vanishes on $W_i$. Then by Theorem 2.2 there exists some integer $\ell \in \mathbb{N}$ such that $(f/\alpha_i - 1)^\ell \in J_i$. From the binomial theorem, it follows that

$$\left(1 + \left(\frac{f}{\alpha_i} - 1\right)\right)^{1/2} \equiv \sum_{k=1}^{\ell-1} \binom{1/2}{k} (f/\alpha_i - 1)^k \stackrel{def}{=} q_i/\sqrt{\alpha_i} \mod J_i \ .$$

Thus $(\rho([f]))_i = q_i^2$ is SOS modulo $J_i$, and hence $\rho^{-1}(q_i^2 e_{i+1})$ is SOS modulo $I_{KKT}$, where $e_{i+1}$ is the $(i + 1)$-st standard unit vector in $\mathbb{R}^{r+1}$.

Finally, we see that $\rho([f]) = (q_0, q_1^2, \cdots, q_r^2)$. The preimage of the latter is

$$\rho^{-1}\big((q_0, q_1^2, \cdots, q_r^2)\big) = \rho^{-1}\big(q_0 e_1)\big) + \sum_{i=1}^{r} \rho^{-1}\big(q_i^2 e_{i+1}\big),$$

which implies that $f \in P_{KKT}$. $\square$

*Remark.* The conclusions in Theorem 3.2 and Theorem 3.5 can not be strengthened to show that $f(x) \in M_{KKT}$, as the following example shows.

**Example 3.6.** Let $g_1 = 1 - x_1$, $g_2 = x_2$, and $g_3 = x_3 - x_2 - 1$ and set $F = \{g_1, g_2, g_3\}$. Let $f = (x_3 - x_1^2 x_2)^2 - 1 + \epsilon$, where $0 < \epsilon < 1$. It is easy to see that the minimum of $f^*$ on $S := S(F)$ is $f^* = \epsilon$. In particular, $f > 0$ on $S$. The corresponding KKT ideal

$$I_{KKT} = \Big\langle 2x_1 x_2(x_3 - x_1^2 x_2) - \lambda_1 x_1, 2x_1^2(x_3 - x_1^2 x_2) + \lambda_2 - \lambda_3, \\ 2(x_3 - x_1^2 x_2) - \lambda_3, \lambda_1(1 - x_1^2), \lambda_2 x_2, \lambda_3(x_3 - x_2 - 1)\Big\rangle$$

is radical (verified in Macaulay 2 [7]). However, $f \notin M_{KKT}$. Suppose to the contrary that $f \in M_{KKT}$, then there exist SOS polynomials $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ and general polynomials $\phi_1, \phi_2, \phi_3$ such that

$$f(x) = \sigma_0 + \sigma_1 g_1 + \sigma_2 g_2 + \sigma_3 g_3 + \phi_1(\frac{\partial f}{\partial x_1} - \lambda_1 x_2) + \phi_2(\frac{\partial f}{\partial x_2} - \lambda_2 + \lambda_3) + \phi_3(\frac{\partial f}{\partial x_3} - \lambda_3).$$

Plugging $\lambda = (0, 0, 0)$ into the above identity yields

$$0 = 1 - \epsilon + \sigma_0 + \sigma_1(1 - x_1^2) + \sigma_2 x_2 + \sigma_3(x_3 - x_2 - 1) + \phi(x_3 - x_1^2 x_2)$$

where $\phi = -4x_1\phi_1 - x_1^2\phi_2 + 2\phi_3 - (x_3 - x_1^2 x_2)$. Now substitute $x_3 = x_1^2 x_2$ in the above, yielding

$$\sigma_3((1 - x_1^2)x_2 + 1) = 1 - \epsilon + \sigma_0 + \sigma_1(1 - x_1^2) + \sigma_2 x_2.$$

Here $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ are now considered SOS polynomials in $(x_1, x_2)$. Since $1 - \epsilon > 0$, $\sigma_3$ cannot be the zero polynomial. If $\sigma_3 = \sigma_3(x_1)$ is independent of $x_2$, we can derive a contradiction using an argument identical to the argument in the proof of of [24, Thm. 2]. Thus $2m = \deg_{x_2}\sigma_3(x_1, x_2) \geq 2$ and $2d = \deg_{x_1}\sigma_3(x_1, x_2) \geq 0$. On the left hand side, the leading term is of the form $A \cdot x_1^{2d+2} x_2^{2m+1}$ with coefficient $A < 0$. Since the degree in $x_2$ on the left hand side is odd, the leading term on the right hand side must come from $\sigma_2(x_1, x_2)x_2$, and is of the form $B \cdot x_1^{2d} x_2^{2m+1}$ with $B > 0$. This is a contradiction. Therefore we can conclude that $f \notin M_{KKT}$.

# 4 Applications in Optimization

Given $f, g_1, \ldots, g_s \in \mathbb{R}[X]$, recall the optimization problem from the Introduction

$$f^* = \min \quad f(x) \tag{4.1}$$

$$s.t. \quad g_i(x) \geq 0, \quad i = 1, \cdots, s, \tag{4.2}$$

and suppose we are interested in computing numerically the optimal value $f^*$. In other words, we wish to compute the minimum of $f$ on the basic closed semialgebraic set $S(F)$, where $F = \{g_1, \ldots, g_s\}$.

Finding the global optimal solutions to (4.1) – (4.2) is an NP-hard problem, even if $f$ is quadratic and the $g_i$ are linear. For instance, the Maximum-Cut problem for graphs is of this form, and it is NP-hard [9]. Recently, the techniques of sum of squares (SOS) relaxations and moment matrix methods have made it possible to approximate the global optimal solutions to (4.1)-(4.2) by approximating nonnegative polynomials with SOS polynomials, which allows the problem to be implemented as a semidefinite program which can then be solved numerically. For details about these methods and their applications, see [14, 15, 16, 18, 19, 22, 23, 29].

In the case where $S$ is compact, the SOS methods are based on representation theorems for positive polynomials on compact semialgebraic sets, i.e., the theorems of Schmüdgen and Putinar. However, these theorems do not hold in the case where $S$ is not compact. As discussed in the Introduction, a more traditional approach in numerical optimization methods is to use the first order optimality conditions (the Karush-Kuhn-Tucker (KKT) system in the constrained case). Using Theorem 3.2 and Theorem 3.5, we combine these two methods to give a procedure for approximating $f^*$ in the case where the semialgebraic set is not necessarily compact.

Recall the KKT system corresponding to (4.1)-(4.2):

$$F \triangleq \nabla f(x) - \sum_{j=1}^{s} \lambda_j \nabla g_j(x) \tag{4.3}$$

$$g_j(x) \geq 0, \ \lambda_j g_j(x) = 0, \ j = 1, \cdots, s, \tag{4.4}$$

Let $f^*_{KKT}$ be the global minimum of $f(x)$ over the KKT system defined by (4.3)-(4.4). Assume the KKT system holds at at least one global optima. Then we claim that $f^* = f^*_{KKT}$. First, $f^* \le f^*_{KKT}$ follows immediately from the fact that all solutions to the KKT system are feasible. Now let $x^*$ be a global minimizer such that $f(x^*) = f^*$, then by assumption, there exist Lagrange multipliers $\lambda^*$ such that $(x^*, \lambda^*)$ satisfies the above KKT system. Thus $f^* \ge f^*_{KKT}$ and hence they are equal.

In order to implement membership in $P_{KKT}$ as a semidefinite programming problem, we need a bound on the degrees of the sums of squares involved. Thus we define the truncated KKT ideal

$$I_{N,KKT} = \Big\{ \sum_{k=1}^{n} \phi_k F_k + \sum_{j=1}^{s} \psi_j \lambda_j g_j \Big| deg(\phi_k F_k), deg(\psi_j \lambda_j g_j) \le 2N \Big\}.$$

and the truncated preorder

$$P_{N,KKT} = \left\{ \sum_{\theta \in \{0,1\}^s} \sigma_\theta g_1^{\theta_1} g_2^{\theta_2} \cdots g_s^{\theta_t} \middle| deg(\sigma_\theta g_1^{\theta_1} \cdots g_s^{\theta_s}) \le 2N \right\} + I_{N,KKT}.$$

Then we define a sequence $\{f_N^*\}$ of *SOS relaxations* of the optimization problem (4.1)-(4.2) as follows:

$$f_N^* = \max_{\gamma \in \mathbb{R}} \quad \gamma \tag{4.5}$$

$$s.t. \quad f(x) - \gamma \in P_{N,KKT}. \tag{4.6}$$

Obviously each $\gamma$ feasible in (4.6) is a lower bound of $f^*$. So $f_N^* \le f^*$. When we increase $N$, the feasible region defined by (4.6) is increasing, and hence the sequence of lower bounds $\{f_N^*\}$ is also monotonically increasing. Thus we have

$$f_1^* \le f_2^* \le f_3^* \le \cdots \le f^*.$$

It can be shown that the sequence of lower bounds $\{f_N^*\}$ obtained from (4.5)-(4.6) converges to $f^*$ in (1.1)-(1.2), provided that $f^*$ is attained at one KKT point, which is summarized in the following theorem.

**Theorem 4.1.** *Assume $f(x)$ has a minimum $f^* := f(x^*)$ at one KKT point $x^*$ of (1.1)-(1.2). Then $\lim_{N\to\infty} f_N^* = f^*$. Furthermore, if $I_{KKT}$ is radical, then there exists some $N \in \mathbb{N}$ such that $f_N^* = f^*$, i.e., the SOS relaxations (4.5)-(4.6) converge in finitely many steps.*

*Proof.* The sequence $\{f_N^*\}$ is monotonically increasing, and $f_N^* \le f^*$ for all $N \in \mathbb{N}$, since $f^*$ is attained by $f(x)$ in the KKT system (1.3)-(1.4) by assumption and the constraint (4.6) implies that $\gamma \le f^*$. Now for arbitrary $\epsilon > 0$, let $\gamma_\epsilon = f^* - \epsilon$ and replace $f(x)$ by $f(x) - \gamma_\epsilon$ in (1.1)-(1.2). The KKT system remains unchanged, and $f(x) - \gamma_\epsilon$ is strictly positive on $V^{\mathbb{R}}_{KKT}$. By Theorem 3.5, $f(x) - \gamma_\epsilon \in P_{KKT}$. Since $f(x) - \gamma_\epsilon$ is fixed, there must exist some integer $N_1$ such that $f(x) - \gamma_\epsilon \in P_{N_1,KKT}$. Hence $f^* - \epsilon \le f^*_{N_1} \le f^*$. Therefore we have that $\lim_{N\to\infty} f_N^* = f^*$.

Now assume that $I_{KKT}$ is radical. Replace $f(x)$ by $f(x) - f^*$ in (1.1)-(1.2). The KKT system still remains the same, and $f(x) - f^*$ is now nonnegative on $V^{\mathbb{R}}_{KKT}$. By Theorem 3.2, $f(x) - f^* \in P_{KKT}$. So there exists some integer $N_2$ such that $f(x) - f^* \in P_{N_2,KKT}$, and hence $f^*_{N_2} \ge f^*$. Then $f_N^* \le f^*$ for all $N$ implies that $f^*_{N_2} = f^*$. $\square$

*Remarks:* The assumption in Theorem 4.1 that $f$ has a minimum at a KKT point is non-trivial and cannot be removed, as the following example shows.

**Example 4.2.** Consider the optimization: $\min\ x$ s.t. $x^3 \geq 0$. Obviously $f^* = 0$ and the global minimizer $x^* = 0$. However, the KKT system

$$1 - \lambda \cdot 3x^2 = 0, \quad \lambda \cdot x^3 = 0, \quad x^3 \geq 0, \quad \lambda \geq 0$$

is not satisfied, since $V_{KKT} = \emptyset$. Actually we can see that the lower bounds $\{f_N^*\}$ given by (4.5)-(4.6) tend to infinity. By Theorem 2.3, $V_{KKT} = \emptyset$ implies that $1 \in P_{KKT}$, i.e.,

$$(1 + 3\nu x^2)(1 - 3\nu x^2) + 9\nu^2 x \cdot \nu x^3 = 1.$$

In the SOS relaxation (4.5)-(4.6), for arbitrarily large $\gamma$, $x - \gamma \in P_{KKT}$, since

$$x - \gamma = (x - \gamma)(1 + 3\nu x^2)(1 - 3\nu x^2) + 9\nu^2 x(x - \gamma) \cdot \nu x^3 \in P_{KKT}.$$

Thus $p_8^* = \infty$.

The SOS relaxation (4.5)-(4.6) is essentially a semidefinite program [22, 23, 33] and can be solved numerically. The dual problem of (4.5)-(4.6) is to minimize a linear functional over some linear moment matrix inequalities. It can also be obtained by applying moment matrix methods [15] to minimize $f$ over the semialgebraic set defined by KKT system (1.3)-(1.4). Using software like Gloptipoly [11] and SOSTOOLS [25], the SOS program (4.5)-(4.6) or its dual problem can be solved, and in many cases, the global minimizer $x^*$ and the Lagrange multiplier $\lambda^*$ can be extracted. For more details about extracting minimizers from SOS relaxations or moment matrix methods, we refer to [12].

**Example 4.3** (Exercise 2.18, [13])**.** Consider the global optimization problem:

$$\begin{array}{ll} \min & (-4x_1^2 + x_2^2)(3x_1 + 4x_2 - 12) \\ s.t. & 3x_1 - 4x_2 \leq 12, \quad 2x_1 - x_2 \leq 0, \quad -2x_1 - x_2 \geq 0. \end{array}$$

The semialgebraic set $S$ defined by the constraints is non-compact. The global minimum $f^* = -\frac{1024}{55} \approx -18.6182$ and the minimizer $x^* = (-24/55, 128/55) \approx (-0.4364, 2.3273)$. The lower bound obtained from (4.5)-(4.6) is $f_4^* \approx -18.6182$. The extracted minimizer $\hat{x} = (-0.4364, 2.3273)$.

**Example 4.4.** Consider the Quadratically Constrained Quadratic Program (QCQP):

$$\begin{array}{ll} \min & -\dfrac{4}{3}x_1^2 + \dfrac{2}{3}x_2^2 - 2x_1x_2 \\ s.t. & x_2^2 - x_1^2 \geq 0, \quad -x_1x_2 \geq 0. \end{array}$$

The global minimum $f^* = 0$ and minimizer $x^* = (0, 0)$. The semialgebraic set $S$ defined by the constraints is non-compact. The lower bound returned by (4.5)-(4.6) is $f_4^* = -2.6 \times 10^{-15}$ (Note: this computation was done in double precision floating point, with round off error bounded by $2^{-53} \sim 10^{-16}$). The extracted minimizer is $\hat{x} = (6.1 \times 10^{-16}, -9.0 \times 10^{-17})$.

# 5 Conclusions

This paper studies representations of positive polynomials on non-compact semialgebraic sets via the KKT ideal. We give a representation theorem for polynomials positive on a basic closed semialgebraic set, even in the case where the semialgebraic set is not compact. This theorem can be used to numerically solve an optimization problem of the form (1.1)-(1.2) in the case where the feasible region is not compact. However, we must make the assumption that one of the global minimizers satisfies the KKT system. As discussed in [19], this assumption is sometimes very restrictive. Also, in general, the SOS relaxations (4.5)-(4.6) are very hard to solve when there are many constraints, since this introduces many Lagrange multipliers. The structure of (4.5)-(4.6) should be exploited to improve the efficiency of the method.

**Acknowledgment.** The authors thank Prof. Sturmfels for illuminating discussions and many helpful comments on the paper.

# References

[1] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry*. Springer-Verlag, 2003.

[2] J. Bochnak, M. Coste and M-F. Roy. *Real Algebraic Geometry*, Springer, 1998.

[3] D.A. Cox, J.B. Little, and D.O'Shea. *Ideals, Varieties and Algorithms: an Introduction to Computational Algebraic Geometry and Commutative Algebra*, Second Edition. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1997.

[4] D.A. Cox, J.B. Little, and D.O'Shea. *Using Algebraic Geometry*, Graduate Texts in Mathematics, Vol. 185. Springer-Verlag, New York, 1998.

[5] C. Delzell and A. Prestel, *Positive Polynomials*, Monographs in Mathematics, Springer, Berlin, 2001.

[6] D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, Graduate Texts in Mathematics, Vol. 150. Springer-Verlag, New York, 1995.

[7] D. Eisenbud, with Daniel R. Grayson, Michael Stillman, and Bernd Sturmfels (Eds.)). *Computations in Algebraic Geometry with Macaulay 2*, Algorithms and Computation in Mathematics. Springer-Verlag, 2002.

[8] C.A. Floudas and P.M. Pardalos. *A collection of test problems for constrained global optimization algorithms*, Lecture Notes in Computer Science, **455**, Springer-Verlag, Berlin, 1990.

[9] M. R. Garey and D. S. Johnson. *Computers and Intractability: A guide to the theory of NP-completeness.* W. H. Freeman and Company, 1979.

[10] B. Hanzon and D. Jibetean: Global minimization of a multivariate polynomial using matrix methods. *Journal of Global Optimization*, 27:1-23, 2003.

[11] D. Henrion and J. Lasserre. GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi. *ACM Trans. Math. Soft.*, 29:165-194, 2003.

[12] D. Henrion and J. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. In *Positive Polynomials in Control*, D. Henrion and A. Garulli, eds., Lecture Notes on Control and Information Sciences, Springer Verlag, 2005.

[13] R. Horst, P. Pardalos and N. Thoai. *Introduction to global optimization*, second edition, Kluwer Academic Publishers, 2000.

[14] D. Jibetean and M. Laurent. Semidefinite approximations for global unconstrained polynomial optimization. To appear in *SIAM Journal on Optimization*.

[15] J. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization* **11** (2001), No. 3, 796–817.

[16] M. Laurent. Semidefinite representations for finite varieties. To appear in *Mathematical Programming*.

[17] M. Marshall. Optimization of polynomial functions, *Canad. Math. Bull.*, **46** (2003) 575–587.

[18] J. Nie and J. Demmel. Minimum ellipsoid bounds for solutions of polynomial systems via sum of squares, to appear in *Journal of Global Optimization*.

[19] J. Nie, J. Demmel and B. Sturmfels. Minimizing Polynomials via Sum of Squares over the Gradient Ideal. Submitted. To appear in *Mathematical Programming*.

[20] J. Nie and M. Schweighofer. On the complexity of Putinar's Positivstellensatz. *Preprint*, 2005. `ArXiv: math.AG/0510309`.

[21] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.

[22] P. Parrilo and B. Sturmfels. Minimizing polynomial functions, Proceedings of the DIMACS Workshop on Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science (March 2001), (eds. S. Basu and L. Gonzalez-Vega), American Mathematical Society, 2003, pp. 83–100.

[23] P. Parrilo. Semidefinite Programming relaxations for semialgebraic problems. *Mathematical Programming, Ser. B* **96** (2003), No. 2, 293–320.

[24] V. Powers and B. Reznick. Polynomials positive on unbounded rectangles. In *Positive Polynomials in Control*, Springer Lecture Notes in Control and Information Sciences, Vol. 312, 2005.

[25] S. Prajna, A. Papachristodoulou and P. Parrilo. *SOSTOOLS User's Guide*. `http://www.mit.edu/~parrilo/SOSTOOLS/`.

[26] P. Parrilo. An explicit construction of distinguished representations of polynomials nonnegative over finite sets. IfA Technical Report AUT02-02, March 2002.

[27] M. Putinar. Positive polynomials on compact semi-algebraic sets, *Ind. Univ. Math. J.* **42** (1993) 203–206.

[28] K. Schmüdgen. The K-moment problem for compact semialgebraic sets. *Math. Ann.* **289** (1991), 203–206.

[29] M. Schweighofer. Optimization of polynomials on compact semialgebraic sets. *SIAM Journal on Optimization* **15** (2005), No. 3, 805-825.

[30] Shafarevich. Basic algebraic geometry. Die Grundlehren der mathematischen Wissenschaften. Band 213. Springer-Verlag, 1974.

[31] G. Stengle, Complexity estimates for the Schmüdgen Positivstellensatz. *Journal of Complexity* 12 (1996), 167-174. MR 97d:14080

[32] B. Sturmfels. Solving Systems of Polynomial Equations, Amer.Math.Soc., CBMS Regional Conferences Series, No 97, Providence, Rhode Island, 2002.

[33] L. Vandenberghe and S. Boyd. Semidefinite Programming, *SIAM Review* **38** (1996) 49-95.