Exact polynomial optimization strengthened with Fritz John conditions

Ngoc Hoang Anh Mai*

January 24, 2023

Abstract

Let $f, g_1, \ldots, g_m$ be polynomials with real coefficients in a vector of variables $x = (x_1, \ldots, x_n)$. Denote by $\text{diag}(g)$ the diagonal matrix with coefficients $g = (g_1, \ldots, g_m)$ and denote by $\nabla g$ the Jacobian of $g$. Let $C$ be the set of critical points defined by

$$C = \{ x \in \mathbb{R}^n : \text{rank}(\varphi(x)) < m \} \quad \text{with} \quad \varphi := \begin{bmatrix} \nabla g \\ \text{diag}(g) \end{bmatrix}. \quad (1)$$

Assume that the image of $C$ under $f$, denoted by $f(C)$, is empty or finite. (Our assumption holds generically since $C$ is empty in a Zariski open set in the space of the coefficients of $g_1, \ldots, g_m$ with given degrees.) We provide a sequence of values, which returned by semidefinite programs, finitely converges to the minimal value attained by $f$ over the basic semi-algebraic set $S$ defined by

$$S := \{ x \in \mathbb{R}^n : g_j(x) \geq 0, j = 1, \ldots, m \}. \quad (2)$$

Consequently, we can precisely compute the minimal value of any polynomial with real coefficients in $X$ over one of the following sets: the unit ball, the unit hypercube, and the unit simplex. Under a slightly more general assumption, we extend this result to the minimization of any polynomial over a basic convex semi-algebraic set that has a non-empty interior and is defined by the inequalities of concave polynomials.

Keywords: sum-of-squares; Nichtnegativstellensatz; gradient ideal; Fritz John conditions; polynomial optimization

Contents

1 Introduction 2

2 Proof of the main results 6

2.1 Preliminaries ......................................................... 6

2.2 Proof of the representations ...................................... 8

2.3 Proof of the genericity of the assumption ..................... 11

3 Illustrated examples .................................................. 12

*University of Konstanz, Universitätsstraße 10, D-78464 Konstanz, Germany.
1 Introduction

The study of non-negative polynomials is of interest in real algebraic geometry with applications in polynomial optimization. In his seminal paper [6], Hilbert studied the expression of non-negative polynomials as sums of squares of polynomials. We call Positivstellensätze the representations of polynomials positive on a basic semi-algebraic set, a set defined by a system of polynomial inequalities. Putinar showed in [22] a Positivstellensatz that each polynomial positive on a compact basic semi-algebraic set satisfying the so-called Archimedean condition (stated below) can be decomposed as a linear combination of polynomials defining this basic semi-algebraic set with weights which are sums of squares of polynomials. Using Putinar’s Positivstellensatz, Lasserre introduced in [12] a sequence of values returned by semidefinite programs (also known as Lasserre’s hierarchy) to approximate from below as closely as desired the minimal value of a polynomial over a basic semi-algebraic set.

We refer to Nichtnegativstellensätze as the representations of polynomials that are non-negative on a basic semi-algebraic set. They allow us to obtain a sequence of values returned by relaxation programs similar in spirit to Lasserre’s hierarchy that converges finitely to the optimal value for a given polynomial optimization problem. However, not all Nichtnegativstellensätze have clearly practical applications. For instance, the relaxation programs based on Krivine–Stengle’s Nichtnegativstellensätze [10] are not convex, and hence it is hard to obtain the values returned by such programs. It is because Krivine–Stengle’s Nichtnegativstellensätze has non-prescribed denominators. In other words, the corresponding relaxation programs are convex, namely semidefinite programs, if Nichtnegativstellensatz has no denominator or has a prescribed denominator. We refer the readers to recent Nichtnegativstellensätze without denominators by Scheiderer [23,24,26] for some compact basic semi-algebraic sets of low dimensions. His works involve the non-strict extension of Schm"udgen’s Positivstellensatz [27] originally stated that each polynomial positive on a compact basic semi-algebraic set can be written as a linear combination of products of polynomials defining this set with weights which are sums of squares of polynomials.

In this paper we provide some Nichtnegativstellensätze without denominators which have the same forms as Putinar’s and Schm"udgen’s Positivstellensatz. To achieve this, we utilize slack variables and additional polynomial equations for a given basic semi-algebraic set. These polynomial equations, generated
by the input polynomials and their gradients, involve the so-called Fritz John conditions.

Let \( \mathbb{R}[x] \) denote the ring of polynomials with real coefficients in the vector of variables \( x \). Given \( f, g_1, \ldots, g_m \in \mathbb{R}[x] \), consider polynomial optimization problem

\[
f^* := \inf_{x \in S(g)} f(x),
\]

where \( S(g) \) is the basic semi-algebraic set associated with \( g = (g_1, \ldots, g_m) \), i.e.,

\[
S(g) := \{ x \in \mathbb{R}^n : g_j(x) \geq 0, \ j = 1, \ldots, m \}.
\]

Given \( p \in \mathbb{R}[x] \), we denote by \( \nabla p \) the gradient of \( p \), i.e., \( \nabla p = (\frac{\partial p}{\partial x_1}, \ldots, \frac{\partial p}{\partial x_n}) \). We state the Fritz John conditions in the following lemma:

**Lemma 1.** Let \( f, g_1, \ldots, g_m \in \mathbb{R}[x] \). If \( u \) is a local minimizer for problem \((3)\), then the Fritz John conditions hold for problem \((3)\) at \( u \), i.e., the following conditions hold:

\[
\begin{aligned}
\exists (\lambda_0, \ldots, \lambda_m) \in [0, \infty)^{m+1} : \\
\lambda_0 \nabla f(u) & = \sum_{j=1}^m \lambda_j \nabla g_j(u), \\
\lambda_j g_j(u) & = 0, \ j = 1, \ldots, m, \\
\sum_{j=0}^m \lambda_j^2 & = 1
\end{aligned}
\]

\[
\begin{aligned}
\exists (\lambda_0, \ldots, \lambda_m) \in \mathbb{R}^{m+1} : \\
\lambda_0 \nabla f(u) & = \sum_{j=1}^m \lambda_j^2 \nabla g_j(u), \\
\lambda_j^2 g_j(u) & = 0, \ j = 1, \ldots, m, \\
\sum_{j=0}^m \lambda_j^2 & = 1.
\end{aligned}
\]

In the final conditions of the two sides, the sequence of multipliers \( \lambda_j \) is normalized to ensure that the multipliers are not all zeros.

Fritz John derived his conditions in \cite{fritzjohn}. A proof for Lemma 1 can be found in Freund’s lecture note \cite{freund}, Theorem 10.

If \( \lambda_0 > 0 \), the Fritz John conditions are equivalent to the Karush–Kuhn–Tucker conditions (see \cite{karush, kuhn, Tucker}). Note that there are some cases of problem \((3)\) (indicated below) for which the Karush–Kuhn–Tucker conditions do not hold at any global minimizer.

Denote by \( \Sigma^2[x] \) the cone of sum of squares of polynomials in \( \mathbb{R}[x] \). Given \( g_1, \ldots, g_m \in \mathbb{R}[x] \), let \( Q(g)[x] \) be the quadratic module associated with \( g = (g_1, \ldots, g_m) \), i.e.,

\[
Q(g)[x] := \Sigma^2[x] + \sum_{j=1}^m g_j \Sigma^2[x].
\]

We say that \( S(g) \) satisfies the Archimedean condition if there exists \( R > 0 \) such that \( R - x_1^2 - \cdots - x_n^2 \in Q(g)[x] \). Under the Archimedean condition and some standard optimality conditions (containing the Karush–Kuhn–Tucker conditions), Nie utilizes Marshall’s Nichtnegativstellensatz \cite{marshall, Nichtnegativstellensatz} to guarantee in \cite{nie} finite convergence of Lasserre’s hierarchy. In this case, the polynomial optimization problem necessarily has finite global minimizers.

Given \( g_1, \ldots, g_m \in \mathbb{R}[x] \), with \( g = (g_1, \ldots, g_m) \), let \( \Pi g \) be the vector of products of \( g_1, \ldots, g_m \) defined by

\[
\Pi g := (g^\alpha)_{\alpha \in \{0,1\}^m \setminus \{0\}},
\]

where \( \alpha = (\alpha_1, \ldots, \alpha_m) \) and \( g^\alpha := g_1^{\alpha_1} \cdots g_m^{\alpha_m} \). We call \( Q(\Pi g)[x] \) the preorder-
gen-erated by \( g \), denoted by \( P(g)[x] \). Obviously, if \( m = 1 \), it holds that \( P(g)[x] = Q(g)[x] \).
Given \( h_1, \ldots, h_l \in \mathbb{R}[x] \), let \( V(h) \) be the variety defined by \( h = (h_1, \ldots, h_l) \), i.e.,
\[
V(h) := \{ x \in \mathbb{R}^n : h_j(x) = 0, \ j = 1, \ldots, l \}.
\]
and let \( I(h)[x] \) be the ideal generated by \( h \), i.e.,
\[
I(h)[x] := \sum_{j=1}^l h_j \mathbb{R}[x].
\]
(9)

The real radical of an ideal \( I(h)[x] \), denoted by \( \sqrt[\mathbb{R}]{I(h)} \), is defined as
\[
\sqrt[\mathbb{R}]{I(h)} := \{ f \in \mathbb{R}[x] : \exists m \in \mathbb{N} : -f^{2m} \in \Sigma^2[x] + I(h)[x] \}.
\]
(10)

Krivine–Stengle’s Nichtnegativstellensätze [10] imply that
\[
\sqrt[\mathbb{R}]{I(h)} = \{ p \in \mathbb{R}[x] : p = 0 \text{ on } V(h) \}.
\]
(11)

We say that \( I(h)[x] \) is real radical if \( I(h)[x] = \sqrt[\mathbb{R}]{I(h)} \).

Demmel, Nie, and Powers provide in [4] a Nichtnegativstellensatz saying that if \( f \) is non-negative on a subset of \( S(g) \) at which the Karush–Kuhn–Tucker conditions hold for problem (3), then there exists \( q \in P(g)[x, \lambda] \) such that \( f - q \) vanishes on \( V(h_{\text{KKT}}) \), where \( \lambda := (\lambda_1, \ldots, \lambda_m) \) and
\[
h_{\text{KKT}} := (\nabla f - \sum_{j=1}^m \lambda_j \nabla g_j, \lambda_1 g_1, \ldots, \lambda_m g_m).
\]
(12)

Here \( h_{\text{KKT}} \) includes polynomials in the Karush–Kuhn–Tucker conditions. To apply this Nichtnegativstellensatz for exact polynomial optimization, they assume the Karush–Kuhn–Tucker conditions hold at some global minimizer.

Our goal is to deal with the case of problem (3) for which the Karush–Kuhn–Tucker conditions do not hold at any global minimizer or the set of global minimizers has a positive dimension. Given \( g_1, \ldots, g_m \in \mathbb{R}[x] \), let \( \varphi^g : \mathbb{R}^n \to \mathbb{R}^{(n+m) \times m} \) be a function associated with \( g = (g_1, \ldots, g_m) \) defined by
\[
\varphi^g(x) = \begin{bmatrix} \nabla g_1(x) & \cdots & \nabla g_m(x) \\ \text{diag}(g(x)) \end{bmatrix} = \begin{bmatrix} \nabla g_1(x) & \cdots & \nabla g_m(x) \\ g_1(x) & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & g_m(x) \end{bmatrix}.
\]
(13)

Given a real matrix \( A \), we denote by \( \text{rank}(A) \) the dimension of the vector space generated by the columns of \( A \) over \( \mathbb{R} \). We say that a set \( \Omega \) is finite if its cardinal number is a non-negative integer. Let \( C(g) \) be the set of critical points associated with \( g \) defined by
\[
C(g) := \{ x \in \mathbb{R}^n : \text{rank}(\varphi^g(x)) < m \}.
\]

It is easily seen that \( C(g) \) is the set of points at which the Fritz John conditions stated in Lemma [3] hold for problem (3) in the case of \( \lambda_0 = 0 \). In other words, \( C(g) \) is the set of all points at which the Fritz John conditions hold, but the Karush–Kuhn–Tucker conditions do not. From this, the following lemma follows:
Lemma 2. If \( C(g) = \emptyset \), problem \(3\) has no local minimizer or only has local minimizers at which the Karush–Kuhn–Tucker conditions for this problem hold.

We state the first main result in the following theorem:

Theorem 1. Let \( f, g_1, \ldots, g_m \in \mathbb{R}[x] \). Assume that \( f \) is non-negative on \( S(g) \) with \( g := (g_1, \ldots, g_m) \) and \( f(C(g)) \) is finite. Then there exists \( q \in P(g)[x, \lambda] \) such that \( f - q \) vanishes on \( V(h_{FJ}) \), where \( \lambda := (\lambda_0, \ldots, \lambda_m) \) and

\[
h_{FJ} := (\lambda_0 \nabla f - \sum_{j=1}^m \lambda_j \nabla g_j, \lambda_1 g_1, \ldots, \lambda_m g_m, 1 - \sum_{j=0}^m \lambda_j^2). \tag{14}
\]

Moreover, if \( S(g) \) satisfies the Archimedean condition, we can take \( q \in Q(g)[x, \lambda] \).

Here \( h_{FJ} \) includes polynomials in the Fritz John conditions, we state the following theorem, can stated in Lemma 4. By \( \mathbb{R}^d \), it is clear that in Theorem 1 if \( I(h_{FJ})[x, \lambda] \) is real radical, then \( f - q \in I(h_{FJ})[x, \lambda] \). Given \( d \in \mathbb{N} \), let \( \mathbb{R}[x]_d \) be the set of polynomials of degree at most \( d \).

We state the second main result in the following theorem:

Theorem 2. Let \( d_1, \ldots, d_m \) be positive integers. Then there exists a polynomial \( \psi \), which is in the coefficients of polynomials \( g_j \in \mathbb{R}[x]_{d_j} \), for \( j = 1, \ldots, m \) such that if \( \psi \) does not vanish at the input data then \( C(g) \) with \( g := (g_1, \ldots, g_m) \) is empty.

On one hand, Theorem 2 implies that \( C(g) \) is empty, and so is \( f(C(g)) \) on a Zariski open set in the space of the coefficients of \( g_1, \ldots, g_m \) with given degrees. Thus in Theorem 1, our assumption that \( f(C(g)) \) is finite holds generically. However, there exists a case of \( g = (g_1, \ldots, g_m) \) (indicated below) for which both \( C(g) \) and \( f(C(g)) \) are infinite. To overcome this, we provide in Theorem 3 a representation of \( f \) with denominator \( \lambda_0 \) but without assumption on \( f(C(g)) \). On the other hand, combining Theorem 2 and Lemma 2 gives the genericity of the Karush–Kuhn–Tucker conditions.

The proofs of Theorems 1 and 2 (postponed to Sections 2.2 and 2.3) are inspired by the techniques of Demmel–Nie–Power [4] and Nie [19], respectively. To prove Theorem 1 we claim that the polynomial \( f \) has a finite number of values on the variety \( V(h_{FJ}) \). We prove this by considering \( f \) of each connected component of \( V(h_{FJ}) \) not contained in the hyperplane \( \lambda_0 = 0 \) and then applying the mean value theorem. The remaining case is based on the assumption that \( f(C(g)) \) is finite. The proof of Theorem 2 relies on the existence of a discriminant for a system of polynomial equations generated by \( g_1, \ldots, g_m \) under a simple transformation.

Bucero and Mourrain present in [3] Section 3.3 a variety defined by the Fritz John conditions without giving any representation of polynomials non-negative on semi-algebraic sets in the case of \( C(g) = \emptyset \). Note that the polynomial equation \( 1 - \sum_{j=1}^m \lambda_j^2 = 0 \) does not exist in their variety to ensure the non-zero vector of multipliers as in our variety \( V(h_{FJ}) \) with \( h_{FJ} \) defined as in (14).

Nie provides in [13] a preordering-based representation of polynomial \( f \) non-negative on a basic semi-algebraic set \( S(g) \) with \( g = (g_1, \ldots, g_m) \) by adding to this set a large number of polynomial equations generated by the Jacobian of the polynomial map \( (f, g_1, \ldots, g_m) \). To achieve this, he restricts the number of
polynomials defining $S(g)$ and assumes that the Jacobian of each subset of $g$ has a full rank on their variety. For comparison purposes, to obtain our representations in Theorem 4, we utilize $n + m + 1$ additional polynomial equations (generated by $g_j$ and $\nabla f, \nabla g_j$) for $S(g)$ and $m + 1$ slack variables $\lambda_0, \ldots, \lambda_m$, which are the multipliers in the Fritz John conditions stated in Lemma 1. Here we assume that the matrix $\varphi^\sigma(x)$ (generated by $g_j(x)$ and $\nabla g_j(x)$) does not need to have a full rank for each $x \in \mathbb{R}^n$, but it is required that the image of all real points at which $\varphi^\sigma$ is rank-deficient under $f$ is finite. Under these conditions and the Archimedean condition $S(g)$, we also provide a representation of $f$ involving the quadratic module $Q(g)[x, \lambda]$.

The paper is organized as follows: Section 2 is to prove Theorems 1 and 2. We recall in the following lemma Putinar’s Positivstellensatz [22]:

**Lemma 3.** Let $g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $S(g) = \emptyset$ with $g := (g_1, \ldots, g_m)$. Then it holds that $-1 \in P(g)[x]$.

We recall in the following lemma Putinar’s Positivstellensatz [22]:

**Lemma 4.** Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $f$ is positive on $S(g)$ with $g := (g_1, \ldots, g_m)$, and $S(g) \neq \emptyset$ satisfies the Archimedean condition. Then it holds that $f \in Q(g)[x]$.

Denote by $\deg(p)$ the degree of a given polynomial $p \in \mathbb{R}[x]$. The following lemma is a consequence of Lemmas 3 and 4.

**Lemma 5.** Let $g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $g_m := R - x_1^2 - \cdots - x_n^2$ for some $R > 0$ and $S(g) = \emptyset$ with $g := (g_1, \ldots, g_m)$. Then it holds that $-1 \in Q(g)[x]$.

**Proof.** Since $S(g) = \emptyset$, Lemma 3 yields that there exists $\sigma_\alpha \in \Sigma^2[x]$ such that

\[-1 = \sum_{\alpha \in \{0,1\}^n} \sigma_\alpha g^\alpha. \tag{15}\]

Given $p \in \mathbb{R}[x]$ with $u = \lceil \deg(p)/2 \rceil$, let $\bar{p} := x_0^{2u} p(x/x_0) \in \mathbb{R}[x]$, where $\bar{x} := (x_0, x)$. For instance, $\bar{g}_m = Rx_0^2 - x_1^2 - \cdots - x_n^2$. Let $d$ be an integer number such that $2d \geq \deg(\sigma_\alpha g^\alpha)$. Let $d_j = \lceil \deg(g_j)/2 \rceil$. From (15), we get

\[-x_0^{2d} = \sum_{\alpha \in \{0,1\}^n} \psi_\alpha \bar{g}^\alpha, \tag{16}\]
where $\psi_\alpha = x_0^{2(d-d_j)}\sigma_\alpha(x/x_0) \in \Sigma^2[x]$ and $\bar{g} = (\bar{g}_1, \ldots, \bar{g}_m)$. Denote by $w \in \mathbb{R}[x]$ the polynomial on the right-hand side of \eqref{eq:psi_alpha}. Then $w$ is non-negative on $S(\bar{g}, 1 - x_0^2)$. Since $0 \in S(\bar{g}, 1 - x_0^2)$, we get $S(\bar{g}, 1 - x_0^2) \neq \emptyset$. On the other hand $S(\bar{g}, 1 - x_0^2)$ satisfies the Archimedean condition since

\[(R + 1) - x_0^2 - \cdots - x_n^2 = (R + 1)(1 - x_0^2) + \bar{g}_m \in Q(\bar{g}, 1 - x_0^2)[x]. \tag{17}\]

Applying Lemma 4, we obtain

\[-x_0^{2d} + \frac{1}{2} = w + \frac{1}{2} \in Q(\bar{g}, 1 - x_0^2)[x]. \tag{18}\]

Letting $x_0 = 1$ implies that $-\frac{1}{2} \in Q(\bar{g})[x]$, yielding the result. \hfill \Box

In the following lemma, we obtain the same result as Lemma 5 under a weaker condition:

**Lemma 6.** Let $g_1, \ldots, g_m \in \mathbb{R}[x]$ such that $S(\bar{g})$ with $g := (g_1, \ldots, g_m)$ satisfies the Archimedean condition and $S(\bar{g}) = \emptyset$. Then $-1 \in Q(\bar{g})[x]$.

**Proof.** Since $S(\bar{g})$ satisfies the Archimedean condition, there exists $R > 0$ such that $g_{m+1} := R - x_1^2 - \cdots - x_n^2 \in Q(\bar{g})[x]$. It implies that $S(\bar{g}) \subset S(g_{m+1})$, which gives $S(g, g_{m+1}) = S(g) = \emptyset$. By using Lemma 5, we obtain $-1 \in Q(g, g_{m+1})[x] \subset Q(\bar{g})[x]$, yielding the result. \hfill \Box

Given $h_1, \ldots, h_l \in \mathbb{R}[x]$, let $V_\mathbb{C}(h)$ be the complex variety defined by $h = (h_1, \ldots, h_m)$, i.e.,

\[V_\mathbb{C}(h) := \{ x \in \mathbb{C}^n : h_j(x) = 0, j = 1, \ldots, l \}. \tag{19}\]

Denote by $\delta_{ij}$ the Kronecker delta function at $(i, j) \in \mathbb{N}^2$.

The following lemma is a direct consequence of \cite{4} Lemma 2.4.1 and Remark 2.5:

**Lemma 7.** Let $U_1, \ldots, U_r$ be pairwise disjoint complex varieties defined by finitely many polynomials in $\mathbb{R}[x]$. Then there exist polynomials $p_1, \ldots, p_r \in \mathbb{R}[x]$ such that $p_i(U_j) = \delta_{ij}$.

We generalize the definition of a basic semi-algebraic set as follows: A semi-algebraic subset of $\mathbb{R}$ is a subset of the form

\[\bigcup_{i=1}^{r_1} \bigcap_{j=1}^{r_2} \{ x \in \mathbb{R}^n : f_{ij}(x) *_{ij} 0 \}, \tag{20}\]

where $f_{ij} \in \mathbb{R}[x]$ and $*_{ij}$ is either $>$ or $\leq$. Given two semi-algebraic sets $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^m$, we say that a mapping $f : A \to B$ is semi-algebraic if its graph is a semi-algebraic set in $\mathbb{R}^{n+m}$.

The following lemma can be found in \cite{21} Proposition 1.6.2 (ii):

**Lemma 8.** Compositions of semi-algebraic maps are semi-algebraic.

A semi-algebraic subset $A$ is said to be semi-algebraically path connected if, for every $x, y$ in $A$, there exists a continuous semi-algebraic mapping $\phi : [0, 1] \to A$ such that $\phi(0) = x$ and $\phi(1) = y$.

Combining \cite{4} Theorem 2.4.5 and Proposition 2.5.13, we obtain the following lemma:
Lemma 9. Every semi-algebraic set has a finite number of components, which are semi-algebraically path connected.

The following result is a direct consequence of Lemma 9 since the difference between two real varieties is a semi-algebraic set according to the definition (20):

Lemma 10. The difference between two real varieties has a finite number of components, which are semi-algebraically path connected.

We state in the following lemma a decomposition of the intersection of a given complex variety with a real space:

Lemma 11. Let $f \in \mathbb{R}[x]$ and let $W$ be a complex variety defined by finitely many polynomials in $\mathbb{R}[x]$. Assume that $f(W \cap \mathbb{R}^n)$ is finite. Then there exists a finite sequence of subsets $W_1, \ldots, W_r$ such that the following conditions hold:

1. $W_1, \ldots, W_r$ are pairwise disjoint complex varieties defined by finitely many polynomials in $\mathbb{R}[x]$;

2. for $j = 1, \ldots, r$, $W_j \subset W$, and $f$ is constant on $W_j$;

3. $(W_1 \cup \cdots \cup W_r) \cap \mathbb{R}^n = W \cap \mathbb{R}^n$.

Proof. By assumption, we get $f(W \cap \mathbb{R}^n) = \{t_1, \ldots, t_r\} \subset \mathbb{R}$, where $t_i \neq t_j$ if $i \neq j$. For $j = 1, \ldots, r$, let $W_j := W \cap V_c(f - t_j)$. Then $W_j$ is a complex variety defined by finitely many polynomials in $\mathbb{R}[x]$. It is clear that $f(W_j) = \{t_j\}$. We claim that $W_1, \ldots, W_r$ are pairwise disjoint. Otherwise, let $x \in W_i \cap W_j$ with $i \neq j$. It implies that $t_i = f(x) = t_j$ which is impossible. Let $U = W_1 \cup \cdots \cup W_r$. We show that $W \cap \mathbb{R}^n = U \cap \mathbb{R}^n$. Let $x \in W \cap \mathbb{R}^n$. Then there is $j \in \{1, \ldots, r\}$ such that $f(x) = t_j$ which gives $x \in W_j \subset U$ so we get $x \in U \cap \mathbb{R}^n$. Thus $W \cap \mathbb{R}^n \subset U \cap \mathbb{R}^n$ since $x$ is arbitrary in $W \cap \mathbb{R}^n$. Conversely, suppose that $x \in U \cap \mathbb{R}^n$. By the definition of $U$, there is $j \in \{1, \ldots, r\}$ such that $x \in W_j$. It implies that $x \in W$ by definition of $W_j$. Then $x \in W \cap \mathbb{R}^n$. Thus $U \cap \mathbb{R}^n \subset W \cap \mathbb{R}^n$ since $x$ is arbitrary in $U \cap \mathbb{R}^n$.

We use the technique from the proof of [4] Theorem 3.2 to obtain the following lemma:

Lemma 12. Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $f$ is non-negative on $S(g)$ with $g = (g_1, \ldots, g_m)$. Let $U_1, \ldots, U_r$ be pairwise disjoint complex varieties defined by finitely many polynomials in $\mathbb{R}[x]$. Set $U = U_1 \cup \cdots \cup U_r$. Assume that $f$ is constant on each $U_i$. Then there exists $q \in P(g)[x]$ such that $f - q$ vanishes on $U \cap \mathbb{R}^n$. Moreover, if $S(g)$ satisfies the Archimedean condition, we can take $q \in Q(g)[x]$.

Proof. Let $W_0$ be the union of all $U_j$ whose intersection with $S(g)$ is empty. Then $W_0$ is a complex variety defined by finitely many polynomials in $\mathbb{R}[x]$. Let $W_1, \ldots, W_r$ be the remaining $U_j$. Thus $f$ is constant on $W_j$, for $j = 1, \ldots, r$. Further, since $f$ is non-negative on the non-empty set $S(g) \cap W_j$, there exists $\alpha_j > 0$ such that $f = \alpha_j$ on $W_j$, $j = 1, \ldots, r$. Set $q_j(x) = \alpha_j \in \Sigma^2[x]$ then we get $f = q_j$ on $W_j$. Observe that $U = W_0 \cup W_1 \cup \cdots \cup W_r$, where $W_0, \ldots, W_r$ are pairwise disjoint. By Lemma 7 there exist polynomials $p_0, p_1, \ldots, p_r \in \mathbb{R}[x]$ such that $p_j(W_j) = \delta_j$. By assumption, it holds that $W_0 \cap S(g) = \emptyset$ and hence by Theorem 3 there exists $v_0 \in P(g)[x]$ such that $-1 = v_0$ on $W_0 \cap \mathbb{R}^n$. We
have $f = s_1 - s_2$ for the SOS polynomials $s_1 = (f + \frac{1}{2})^2$ and $s_2 = f^2 + \frac{1}{2}$. It implies that $f = s_1 + v_0 s_2$ on $W_0 \cap \mathbb{R}^n$. Let $q_0 = s_1 + v_0 s_2 \in P(g)[x]$. Now let $q = \sum_{i=0}^n q_i p_i$ then $q \in P(g)[x]$ and we obtain $f - q$ vanishes on $U \cap \mathbb{R}^n$. Assume that $S(g)$ satisfies the Archimedean condition. Following Lemma\textsuperscript{3} we can take $v_0 \in Q(g)[x]$, which implies that $q_0$ is in $Q(g)[x]$ then so is $q$.

\section{Proof of the representations}

Recall the vector of variables $\lambda := (\lambda_0, \ldots, \lambda_m)$. For simplicity of notation, set

$$\{\lambda_0 = 0\} := \{(x, \lambda) \in \mathbb{R}^{n+m+1} : \lambda_0 = 0\}. \hspace{1cm} (21)$$

Let $\pi: \mathbb{R}^{n+m+1} \rightarrow \mathbb{R}^n$ be the projection defined by

$$\pi(x, \lambda) = x, \forall x \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}^{m+1}. \hspace{1cm} (22)$$

We characterize the set of critical points in the following lemma:

**Lemma 13.** Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$. Let $h_{FJ}$ be defined as in \textsuperscript{14}. Set $g := (g_1, \ldots, g_m)$. Then it holds that $C(g) = \pi(V(h_{FJ}) \cap \{\lambda_0 = 0\})$.

**Proof.** The result follows thanks to the following equivalences:

$$x \in C(g) \iff \text{rank}(\varphi^g(x)) < m$$

$$\iff \exists \lambda \in \mathbb{R}^n : \sum_{j=m}^n \lambda_j^2 = 1, \sum_{j=1}^m \lambda_j \nabla g_j(x) = 0, \lambda_j g_j(x) = 0$$

$$\iff \exists \lambda \in \mathbb{R}^{m+1} : \sum_{j=0}^m \lambda_j^2 = 1, \lambda_0 = 0, \lambda_0 \nabla f(x) = \sum_{j=1}^m \lambda_j \nabla g_j(x), \lambda_j g_j(x) = 0$$

$$\iff \exists \lambda \in \mathbb{R}^{m+1} : (x, \lambda) \in V(h_{FJ}) \cap \{\lambda_0 = 0\}$$

$$\iff x \in \pi(V(h_{FJ}) \cap \{\lambda_0 = 0\}) \hspace{1cm} (23)$$

The following lemma is given in \textsuperscript{21} Theorem 1.8.1:

**Lemma 14.** Let $f : (a, b) \rightarrow \mathbb{R}$ be a semi-algebraic function. Then there are $a = a_0 < a_1 < \cdots < a_s < a_{s+1} = b$ such that, for each $i = 0, \ldots, s$, the restriction $f|_{(a_i, a_{i+1})}$ is analytic.

The following lemma follows from the mean value theorem:

**Lemma 15.** Let $f : [0, 1] \rightarrow \mathbb{R}$ be a continuous piecewise-differentiable function, i.e., there exist $0 = a_0 < a_1 < \cdots < a_r = 1$ such that $f$ is continuous, and $f$ is differentiable on each open interval $(a_i, a_{i+1})$. Assume that $f$ has zero subgradient. Then $f(0) = f(1)$.

**Proof.** By using the mean value theorem on each open interval $(a_i, a_{i+1})$, we get $f(a_i) = f(a_{i+1})$. Hence $f(0) = f(a_1) = \cdots = f(a_r) = f(1)$ yields the result.

The following lemma extends \textsuperscript{4} Lemma 3.3] to the case of varieties defined by the Fritz John conditions:

**Lemma 16.** Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $f(C(g))$ with $g := (g_1, \ldots, g_m)$ is finite. Let $h_{FJ}$ be defined as in \textsuperscript{14}. Let $W$ be a semi-algebraically path connected component of $V(h_{FJ})$. Then $f$ is constant on $W$.  

9
Proof. Let \((x^{(0)}, \lambda^{(0)})\) and \((x^{(1)}, \lambda^{(1)})\) in \(W\). We claim that \(f(x^{(0)}) = f(x^{(1)})\). By assumption, there exists a continuous piecewise-differentiable path \(\phi(\tau) = (x(\tau), \lambda(\tau))\), for \(\tau \in [0,1]\), lying inside \(W\) such that \(\phi(0) = (x^{(0)}, \lambda^{(0)})\) and \(\phi(1) = (x^{(1)}, \lambda^{(1)})\) (see, e.g., [21, Theorem 1.8.1]). Since \(\tau \mapsto \lambda_0(\tau)\) is continuous on \([0,1]\), the set \(\lambda_0^{-1}(0) \subset [0,1]\) is closed. Then there exists a sequence of intervals \([a_j, b_j] \subset [0,1]\) for \(j = 1, \ldots, r\) such that

- \(0 \leq a_1 < b_1 < a_2 < b_2 < \cdots < a_r \leq b_r \leq 1;\)
- \(\forall \tau \in [0,1], \lambda_0(\tau) = 0 \Leftrightarrow \tau \in \bigcup_{j=1}^r [a_j, b_j].\)

We claim that \(\tau \mapsto f(x(\tau))\) is constant on \([a_j, b_j]\) for \(j = 1, \ldots, r\). Let \(j \in \{1, \ldots, r\}\) be fixed. Assume by contradiction that there are \(\tau_1, \tau_2 \in [a_j, b_j]\) such that \(\tau_1 < \tau_2\) and \(f(x(\tau_1)) \neq f(x(\tau_2))\). Since \(\tau \mapsto f(x(\tau))\) is continuous on \([a_j, b_j]\), the set \(f(x([\tau_1, \tau_2]))\) is infinite. This contradicts the assumption \(\pi(V(h_{FJ}) \cap \{\lambda_0 = 0\})\) (according to Lemma 14) is finite since \(\pi(V(h_{FJ}) \cap \{\lambda_0 = 0\}) \supset f(x([\tau_1, \tau_2]))\) which is due to the fact that

\[
\pi(V(h_{FJ}) \cap \{\lambda_0 = 0\}) \supset \pi(W \cap \{\lambda_0 = 0\}) \supset x([\tau_1, \tau_2]).
\]

On the other hand \(\tau \mapsto \lambda_0(\tau)\) has no zero value on each of the following open intervals:

\[
(0, a_1), (b_1, a_2), \ldots, (b_{r-1}, a_r), (b_r, 1)
\]

(25)

We claim that \(\tau \mapsto f(x(\tau))\) is constant on each of open intervals in (25).

Let us prove this for the interval \((0, a_1)\). The proof for the other intervals is similar. Let \(\tau_1, \tau_2\) be arbitrary in \((0, a_1)\). The Lagrangian function

\[
L(x, \lambda) = f(x) - \sum_{j=1}^m \frac{\lambda_j}{\lambda_0} g_j(x).
\]

(26)

is equal to \(f(x)\) on \(V(h_{FJ}) \setminus \{\lambda_0 = 0\}\), which contains \(\phi([\tau_1, \tau_2])\). By Lemma 8, the function \(L \circ \phi\) is semi-algebraic. Moreover, the function \(L \circ \phi\) is continuous since \(L\) and \(\phi\) are continuous. It implies that \(L \circ \phi\) is a continuous piecewise-differentiable function thanks to Lemma 14. Note that the function \(L \circ \phi\) has zero subgradient on \([\tau_1, \tau_2]\). From Lemma 13 it follows that \(f(x(\tau_1)) = (L \circ \phi)(\tau_1) = (L \circ \phi)(\tau_2) = f(x(\tau_2))\). We now obtain \(f(x(\tau_1)) = f(x(\tau_2))\).

By its continuity, \(\tau \mapsto f(x(\tau))\) is constant on the following closed intervals:

\[
[0, a_1], [b_1, a_2], \ldots, [b_{r-1}, a_r], [b_r, 1]
\]

(27)

It implies that \(\tau \mapsto f(x(\tau))\) is constant on \([0,1]\) since it is constant on each \([a_j, b_j]\). We now obtain \(f(x^{(0)}) = f(x^{(1)})\), and hence \(f\) is constant on \(W\).

Proof of Theorem 1

Proof. Using Lemma 9 we decompose \(V(h_{FJ})\) into semi-algebraically path connected components: \(Z_1, \ldots, Z_s\). Lemma 10 yields that \(f\) is constant on \(Z_i\). Thus \(f(V(h_{FJ}))\) is finite. Observe that \(V(h_{FJ}) = V_C(h_{FJ}) \cap \mathbb{R}^{n+m+1}\). By using Lemma 11 we obtain a finite sequence of subsets \(W_1, \ldots, W_r\) such that the following conditions hold:

- \(W_1, \ldots, W_r\) are pairwise disjoint complex varieties defined by finitely many polynomials in \(\mathbb{R}[x, \lambda];\)
• for \( j = 1, \ldots, r \), \( W_j \subset V_{\mathcal{C}(h_{FJ})} \), and \( f \) is constant on \( W_j \);
• \((W_1 \cup \cdots \cup W_r) \cap \mathbb{R}^{n+m+1} = V_{\mathcal{C}(h_{FJ})} \cap \mathbb{R}^{n+m+1} \).

Let \( D \) be the union of \( W_1, \ldots, W_r \). Let \( b = 1 - \lambda_0^2 - \cdots - \lambda_m^2 \). From this, Lemma \[12\] yields that there exists \( p \in P(g, b)[x, \bar{\lambda}] \) such that \( f - p \) vanishes on \( D \cap \mathbb{R}^{n+m+1} = V(h_{FJ}) \). We write
\[
p = \sum_{\alpha \in \{0, 1\}^m} \sigma_{\alpha} g^\alpha + b \sum_{\beta \in \{0, 1\}^m} \psi_{\beta} g^\beta, \tag{28}
\]
for some \( \sigma_{\alpha}, \psi_{\beta} \in \Sigma^2[x, \bar{\lambda}] \). Let \( q = \sum_{\alpha \in \{0, 1\}^m} \sigma_{\alpha} g^\alpha \in P(g)[x, \bar{\lambda}] \). Since \( b = 0 \) on \( V(h_{FJ}) \), it holds that \( f = p = q \) on \( V(h_{FJ}) \).

Assume that \( S(g) \) satisfies the Archimedean condition. Then there exists \( R > 0 \) such that \( g_{m+1} = R - x_1^2 - \cdots - x_n^2 \in Q(g)[x] \). It implies that \( S(g, b) \) with \( b = 1 - \lambda_0^2 - \cdots - \lambda_m^2 \) satisfies the Archimedean condition. It is because
\[
(R + 1) - x_1^2 - \cdots - x_n^2 - \lambda_0^2 - \cdots - \lambda_m^2 = b + g_{m+1} \in Q(g, b)[x, \bar{\lambda}] \tag{29}.
\]
From this, Lemma \[12\] shows that there exists \( p \in Q(g, b)[x, \bar{\lambda}] \) such that \( f - p \) vanishes on \( D \cap \mathbb{R}^{n+m+1} = V(h_{FJ}) \). We write
\[
p = \sigma_0 + \sum_{j=1}^m \sigma_j g_j + b \sigma_{m+1}, \tag{30}
\]
for some \( \sigma_j \in \Sigma^2[x, \bar{\lambda}] \). Let \( q = \sigma_0 + \sum_{j=1}^m \sigma_j g_j \in Q(g)[x, \bar{\lambda}] \). Since \( b = 0 \) on \( V(h_{FJ}) \), \( f = p = q \) on \( V(h_{FJ}) \). This completes the proof. \( \square \)

2.3 Proof of the genericity of the assumption

Given \( p = (p_1, \ldots, p_m) \) with \( p_1, \ldots, p_m \in \mathbb{R}[x] \), we denote the Jacobian of \( p \) by \( \nabla p = [\nabla p_1 \ldots \nabla p_m] \). Given a complex matrix \( A \), we denote by \( \text{rank}_{\mathbb{C}}(A) \) the dimension of the vector space generated by the columns of \( A \) over \( \mathbb{C} \). Note that \( \text{rank}(A) \) is the dimension of the vector space generated by the columns of \( A \) over \( \mathbb{R} \).

In the following lemma, we recall Nie’s result in \[17\] involving discriminants:

Lemma 17. Let \( p_1, \ldots, p_m \in \mathbb{R}[x] \) be of degrees \( d_1, \ldots, d_m \), respectively, and \( m \leq n \). Suppose at least one \( d_i > 1 \). Then there exists a polynomial \( \psi \) in the coefficients of \( p_1, \ldots, p_m \) having the property that with \( p := (p_1, \ldots, p_m) \), \( \psi(p) = 0 \) if and only if there exists \( u \in \mathbb{C}^n \) satisfying
\[
p_1(u) = \cdots = p_m(u) = 0 \quad \text{and} \quad \text{rank}_{\mathbb{C}}(\nabla p(u)) < m. \tag{31}
\]
Polynomial \( \psi \) in Lemma \[17\] is called a discriminant of \( p \), denoted by \( \Delta(p) \).

Proof of Theorem \[2\]

Proof. Let \( y = (y_1, \ldots, y_m) \). Let \( p_j(x, y) = g_j(x) - y_j^2 \in \mathbb{R}[x, y] \). Then \( \text{deg}(p_j) > 1 \). With \( p = (p_1, \ldots, p_m) \), we get
\[
\nabla p(x, y) = \begin{bmatrix} \nabla g_1(x) & \ldots & \nabla g_m(x) \\ -2y_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & -2y_m \end{bmatrix}. \tag{32}
\]
Let \( \psi(g) := \Delta(p) \). Then \( \psi \) is a polynomial in the coefficients of \( g_1, \ldots, g_m \). We make the following implications:

\[
\begin{align*}
C(g) \neq \emptyset & \implies \exists x \in C(g) \\
& \implies \exists x \in \mathbb{R}^n : \text{rank}(\varphi^p(x)) < m \\
& \implies \exists x \in \mathbb{C}^n : \exists \lambda \in \mathbb{C}^m \setminus \{0\} : \sum_{j=1}^m \lambda_j \nabla g_j(x) = 0, \lambda_j g_j(x) = 0 \\
& \implies \exists (x,y) \in \mathbb{C}^{n+m} : g_j(x) = y_j, \exists \lambda \in \mathbb{C}^m \setminus \{0\} : \sum_{j=1}^m \lambda_j \nabla g_j(x) = 0, -2\lambda_j y_j = 0 \\
& \implies \Delta(p) = 0 \quad \text{(according to Lemma 17)} \\
& \implies \psi(g) = 0.
\end{align*}
\]

Thus it holds that if \( \psi(g) \neq 0 \), then \( C(g) = \emptyset \), yielding the result. \( \square \)

### 3 Illustrated examples

In this section, we illustrate our Nichtnegativstellensätze stated in Theorem 1 with several explicit examples that belong to the following cases of problem (3):

- unconstrained case (see Example 1) and constrained case (see Examples 2, 3, 4, 5, 6, 7, 8, 9, 10);
- the Karush–Kuhn–Tucker conditions do not hold at any global minimizer (see Examples 2, 3, 4, 5);
- the basic semi-algebraic set \( S(g) \) is non-compact (see Example 1, 3, 4, 5) and the set of global minimizers is infinite (see Example 6);
- both the set of critical points \( C(g) \) and its image under \( f \) are empty (see Examples 6, 7, 8, 9);
- both the set of critical points \( C(g) \) and its image under \( f \) are finite (see Examples 2, 3);
- neither the set of critical points \( C(g) \) nor its image under \( f \) is finite (see Example 4);
- the set of critical points \( C(g) \) is infinite but its image under \( f \) is finite (see Example 5);
- the minimal value \( f^* \) is not attained (see Example 9);
- the set \( f(C) \) is infinite and there does not exist \( q \in P(g)[x, \bar{\lambda}] \) such that \( f - q \) vanishes on \( V(h_{FJ}) \) with \( h_{FJ} \) defined as in (14) (see Example 10);
- the set \( f(C) \) is infinite but there exists \( q \in P(g)[x, \bar{\lambda}] \) such that \( f - q \) vanishes on \( V(h_{FJ}) \) (see Example 5).

We begin with the following example in the case where the basic semi-algebraic set is the whole space:
Example 1. Consider the unconstrained case of problem (3). We can assume that \( f^* = 0, m = 1, \) and \( g = (g_1) = (1) \). It is obvious that \( \text{rank}(\varphi^0(x)) = 1 = m \), which implies \( C(g) \) is empty and so is \( f(C(g)) \). Using Theorem 4, we get \( q \in \Sigma^2[x, \lambda_0, \lambda_1] \) such that \( f - q \) vanishes on \( V(h_{FJ}) \) with \( h_{FJ} \) defined as in (14), i.e.,

\[
h_{FJ} := (\lambda_0 \nabla f, \lambda_1, 1 - \lambda_0^2 - \lambda_1^2).
\]

With \( \lambda_0 = 1 \) and \( \lambda_1 = 0 \), \( g(x, 1, 0) \in \Sigma^2[x] \), and \( f - g(x, 1, 0) \) vanishes on \( V(\nabla f) \). It is exactly the same as the result of Nie, Demmel and Siefels in [20].

In the following four examples, \( f^* = 0 \) is attained, but the Karush–Kuhn–Tucker conditions do not hold at any global minimizer for problem (3):

Example 2. Consider the problem (3) with \( n = m = 1, f = x, \) and \( g = (g_1) = (-x^2) \). Then 0 is the unique global minimizer for this problem. The condition \( \text{rank}(\varphi^0(x)) < m \) can be expressed as

\[
\exists \lambda_1 \in \mathbb{R} \setminus \{0\} \ : \ \lambda_1 \nabla g_1(x) = -2\lambda_1 x = 0, \ \lambda_1 g_1(x) = -\lambda_1 x^2 = 0.
\]

It is equivalent to \( x = 0 \). Thus \( C(g) = \{0\} \) is singleton, and so is \( f(C(g)) \). Since \( 1 - x^2 = 1 + g_1 \in Q(g[x]) \), \( S(g) \) satisfies the Archimedean condition. From this, Theorem 2 yields that there exists \( q \in Q(g[x], \lambda) \) with \( \lambda = (\lambda_0, \lambda_1) \) such that \( f - q \) vanishes on \( V(h_{FJ}) \) with \( h_{FJ} \) defined as in (14), namely \( h_{FJ} = (\lambda_0 + 2\lambda_1 x, -\lambda_1 x^2, 1 - \lambda_0^2 - \lambda_1^2) \). It is easy to check that \( V(h_{FJ}) = \{(0, 0, \pm 1)\} \). It implies the selection \( q = 0 \) in this case.

The following example involves the case where the basic semi-algebraic set is non-compact:

Example 3. Consider the problem (3) with \( n = 2, m = 1, f = (x_1 + 1)^2 + x_2^2 - 1, \) and \( g = (g_1) = (x_1^3 - x_2^3) \). Then \((0, 0)\) is the unique global minimizer for this problem. The condition \( \text{rank}(\varphi^0(x)) < m \) can be expressed as

\[
\exists \lambda_1 \in \mathbb{R} \setminus \{0\} \ : \ \lambda_1 \nabla g_1(x) = \lambda_1 \begin{bmatrix} 3x_1^2 \\ -2x_2 \end{bmatrix} = 0, \ \lambda_1 g_1(x) = \lambda_1 (x_1^3 - x_2^2) = 0.
\]

It is equivalent to \( x = (x_1, x_2) = 0 \). Thus \( C(g) = \{0\} \) is singleton then so is \( f(C(g)) \). From this, Theorem 2 yields that there exists \( q \in P(g[x], \lambda) \) with \( \lambda = (\lambda_0, \lambda_1) \) such that \( f - q \) vanishes on \( V(h_{FJ}) \) with \( h_{FJ} \) defined as in (14), i.e.,

\[
h_{FJ} = (\lambda_0 \begin{bmatrix} 2(x_1 + 1) \\ 2x_2 \end{bmatrix} - \lambda_1 \begin{bmatrix} 3x_1^2 \\ -2x_2 \end{bmatrix}, \lambda_1 (x_1^3 - x_2^2), 1 - \lambda_0^2 - \lambda_1^2).
\]

Let \((x, \lambda) \in V(h_{FJ}) \). We get \( 2(\lambda_0 + \lambda_1)x_2 = 0 \), so either \( x_2 = 0 \) or \( \lambda_0 + \lambda_1 = 0 \). If \( \lambda_0 + \lambda_1 = 0 \), then we obtain \( \lambda_0 = -\lambda_1 \in \{\pm \frac{1}{\sqrt{2}}\} \) and \( 2(x_1 + 1) + 3x_1^2 = 0 \), which has no real solution, and hence this is impossible. Consequently \( x_2 = 0 \) from which we obtain \( \lambda_1 x_1^3 = 0 \) which gives \( \lambda_1 = 0 \) or \( x_1 = 0 \). If \( \lambda_1 = 0 \), we obtain \( \lambda_0 \in \{\pm 1\} \) which implies \( x_1 = -1 \). If \( x_1 = 0 \), we get \( \lambda_0 = 0 \) which implies \( \lambda_1 \in \{\pm 1\} \). These give \( V(h_{FJ}) = \{(-1, 0, \pm 1, 0), (0, 0, 0, \pm 1)\} \). It is not hard to check that \( q = g_1 \) satisfies \( q \in P(g[x], \lambda) \) and \( f - q \) vanishes on \( V(h_{FJ}) \).

In the last two examples, the sets of critical points are finite. We indicate in the following two examples the case where the sets of critical points are infinite:
Example 4. Consider the problem \([3]\) with \(n = 2\), \(m = 1\), \(f = x_1 + x_2^2\), and \(g = (g_1) = (-x_1^2)\). Then \((0, 0)\) is the unique global minimizer. The condition rank\(\varphi^q(x)\) < \(m\) can be expressed as

\[
\exists \lambda_1 \in \mathbb{R} \setminus \{0\} : \lambda_1 \nabla g_1(x) = \lambda_1 \begin{bmatrix} -2x_1 \\ 0 \end{bmatrix} = 0, \quad \lambda_1 g_1(x) = -\lambda_1 x_1^2 = 0. \tag{38}
\]

It is equivalent to \(x_1 = 0\). Thus \(C(g) = \{(0, t) : t \in \mathbb{R}\}\) is infinite, which implies that \(f(C(g)) = \{t^2 : t \in \mathbb{R}\} = [0, \infty)\) is infinite. Let \(h_{FJ}\) be as in \([14]\). Then we obtain

\[
h_{FJ} = (\lambda_0 \begin{bmatrix} 1 \\ 2x_2 \end{bmatrix} - \lambda_1 \begin{bmatrix} -2x_1 \\ 0 \end{bmatrix}, -\lambda_1 x_1^2, 1 - \lambda_0^2 - \lambda_1^2). \tag{39}
\]

Let \((x, \bar{x}) \in V(h_{FJ})\). We get \(-\lambda_1 x_1^2 = 0\), so either \(\lambda_1 = 0\) or \(x_1 = 0\). If \(\lambda_1 = 0\), we get \(\lambda_0 \in \{\pm 1\}\) (since \(1 = \lambda_0^2 + \lambda_1^2\)) and \(\lambda_0 = 0\). Hence this is impossible. Consequently, \(x_1 = 0\) implies \(\lambda_0 = 0\), which gives \(\lambda_1 \in \{\pm 1\}\). It implies that \(V(h_{FJ}) = \{(0, t, 0, \pm 1) : t \in \mathbb{R}\}\). Since \(f(C(g))\) is infinite, the assumption of Theorem \([7]\) does not hold in this case. However, it is not hard to check that \(q = x_2^2\) satisfies \(q \in P(g)[x, \bar{x}]\) and \(f - q\) vanishes on \(V(h_{FJ})\). It is interesting to know clearly the extension of Theorem \([14]\) to the case where \(f(C(g))\) is infinite.

In the last three examples, the sets of global minimizers are finite. The following example shows the case where the set of global minimizers is infinite.

Example 5. Consider the problem \([3]\) with \(n = 2\), \(m = 1\), \(f = x_1 - x_2\) and \(g = (g_1) = -(x_1 - x_2)^2\). Then the set of global minimizers \(\{(t, t) : t \in \mathbb{R}\}\) is infinite. The condition rank\(\varphi^q(x)\) < \(m\) can be expressed as

\[
\exists \lambda_1 \in \mathbb{R} \setminus \{0\} : \lambda_1 \nabla g_1(x) = \lambda_1 \begin{bmatrix} -2(x_1 - x_2) \\ 2(x_2 - x_1) \end{bmatrix} = 0, \quad \lambda_1 g_1(x) = -\lambda_1(x_1 - x_2)^2 = 0. \tag{40}
\]

It is equivalent to \(x_2 = x_1\). Thus \(C(g) = \{(t, t) : t \in \mathbb{R}\}\) is infinite. However, \(f(C(g)) = \{0\}\) is singleton. Thus Theorem \([7]\) yields that there exists \(q \in P(g)[x, \bar{x}]\) with \(\bar{x} = (\lambda_0, \lambda_1)\) such that \(f - q\) vanishes on \(V(h_{FJ})\) with \(h_{FJ}\) defined as in \([14]\), i.e.,

\[
h_{FJ} = (\lambda_0 \begin{bmatrix} 1 \\ -1 \end{bmatrix} - \lambda_1 \begin{bmatrix} -2(x_1 - x_2) \\ 2(x_2 - x_1) \end{bmatrix}, -\lambda_1(x_1 - x_2)^2, 1 - \lambda_0^2 - \lambda_1^2). \tag{41}
\]

It is easy to check that \(V(h_{FJ}) = \{(t, t, 0, \pm 1) : t \in \mathbb{R}\}\). It yields the selection \(q = 0\) in this case.

We show in the following three examples of problem \([3]\) that the set of critical points \(C(g)\) is empty, and so is its image under \(f\):

Example 6. Consider the problem \([3]\) with \(m = 1\) and \(g = (1 - x_1^2 - \cdots - x_n^2)\). Then \(S(g)\) is the unit ball. The condition rank\(\varphi^q(x)\) < \(m\) can be expressed as

\[
\exists \lambda_1 \in \mathbb{R} \setminus \{0\} : \lambda_1 \nabla g_1(x) = -2\lambda_1 x = 0, \quad \lambda_1 g_1(x) = \lambda_1(1 - x_1^2 - \cdots - x_n^2) = 0. \tag{42}
\]

It is equivalent to \(x = 0\) and \(1 - x_1^2 - \cdots - x_n^2 = 0\), and hence it is impossible. Thus we get \(C(g) = \emptyset\), which implies \(f(C(g)) = \emptyset\). Note that \(S(g)\) satisfies the
Archimedean condition in this case. From this, Theorem 2 yields that if \( f \) is non-negative on \( S(g) \), there exists \( q \in Q(g)[x, \bar{x}] \) with \( \bar{x} = (\lambda_0, \lambda_1) \) such that \( f - q \) vanishes on \( V(h_{FJ}) \) with \( h_{FJ} \) defined as in (43), i.e.,
\[
\begin{align*}
  h_{FJ} &= (\lambda_0 \nabla f - 2\lambda_1 x, \lambda_1 (1 - x_1^2 - \cdots - x_n^2), 1 - \lambda_0^2 - \lambda_1^2). 
\end{align*}
\]

Let \( e_1, \ldots, e_n \) be the canonical basis of \( \mathbb{R}^n \).

**Example 7.** Consider the problem (3) with \( m = n \) and \( g = (1 - x_1^2, \ldots, 1 - x_n^2) \). Then \( S(g) \) is the unit hypercube. The condition \( \text{rank}(\varphi^q(x)) < m \) can be expressed as
\[
\exists \lambda \in \mathbb{R}^n \setminus \{0\} : \sum_{j=1}^n \lambda_j \nabla g_j(x) = -2 \sum_{j=1}^n \lambda_j x_j e_j = 0, \lambda_j g_j(x) = \lambda_j (1 - x_j^2) = 0. 
\]
\[
\text{It implies that } \lambda_j = \lambda_j (1 - x_j^2) - \frac{1}{2} x_j (-2 \lambda_j x_j) = 0, j = 1, \ldots, m \text{ which contradicts } \lambda \neq 0. \text{ Thus we get } C(g) = \emptyset, \text{ which implies } f(C(g)) = \emptyset. \text{ Note that } S(g) \text{ satisfies the Archimedean in this case. From this, Theorem 1 yields that if } f \text{ is non-negative on } S(g), \text{ there exists } q \in Q(g)[x, \bar{x}] \text{ with } \bar{x} = (\lambda_0, \lambda_1) \text{ such that } f - q \text{ vanishes on } V(h_{FJ}) \text{ with } h_{FJ} \text{ defined as in (43), i.e.,}
\]
\[
\begin{align*}
  h_{FJ} &= (\lambda_0 \nabla f - 2 \sum_{j=1}^n \lambda_j x_j e_j, \lambda_1 (1 - x_1^2), \ldots, \lambda_n (1 - x_n^2), 1 - \lambda_0^2 - \lambda_1^2). 
\end{align*}
\]

**Example 8.** Consider the problem (3) with \( m = n + 1 \) and \( g = (x_1, \ldots, x_n, 1 - x_1 - \cdots - x_n) \). Then \( S(g) \) is the unit simplex. The condition \( \text{rank}(\varphi^q(x)) < m \) can be expressed as
\[
\exists \lambda \in \mathbb{R}^{n+1} \setminus \{0\} : \left\{ \begin{array}{l}
\sum_{j=1}^n \lambda_j \nabla g_j(x) + \lambda_{n+1} \nabla g_{n+1}(x) = \sum_{j=1}^n (\lambda_j - \lambda_{n+1}) e_j = 0, \\
\lambda_j g_j(x) = \lambda_j x_j, j = 1, \ldots, n, \\
\lambda_{n+1} g_{n+1}(x) = \lambda_{n+1} (1 - x_1 - \cdots - x_n) = 0.
\end{array} \right. 
\]
\[
\text{It implies that } \lambda_j = \lambda_{n+1}, j = 1, \ldots, n, \text{ and } \lambda_{n+1} = \lambda_{n+1} \sum_{j=1}^n x_j = \sum_{j=1}^n \lambda_j x_j = 0. \text{ Thus we get } \lambda = 0 \text{ which contradicts } \lambda \neq 0. \text{ Thus we obtain } C(g) = \emptyset \text{ which implies } f(C(g)) = \emptyset. \text{ Note that } S(g) \text{ satisfies the Archimedean in this case, as shown in (7). From this, Theorem 1 shows that if } f \text{ is non-negative on } S(g), \text{ there exists } q \in Q(g)[x, \bar{x}] \text{ with } \bar{x} = (\lambda_0, \lambda_1) \text{ such that } f - q \text{ vanishes on } V(h_{FJ}) \text{ with } h_{FJ} \text{ defined as in (43), i.e.,}
\]
\[
\begin{align*}
  h_{FJ} &= (\lambda_0 \nabla f - \sum_{j=1}^n (\lambda_j - \lambda_{n+1}) e_j, \lambda_1 x_1, \ldots, \lambda_n x_n, \lambda_{n+1} (1 - \sum_{j=1}^n x_j), 1 - \sum_{j=0}^{n+1} \lambda_j^2). 
\end{align*}
\]

Next, we consider an example of problem (5) whose optimal value \( f^* \) is not attained:

**Example 9.** Consider the problem (3) with \( m = 2, n = 1, f = x_1 \) and \( g = (g_1) = (x_1 x_2^2 - 1) \). It is not hard to prove that \( f^* = 0 \). Note that problem (3) does not have any global minimizer. The condition \( \text{rank}(\varphi^q(x)) < m \) can be expressed as
\[
\exists \lambda_1 \in \mathbb{R} \setminus \{0\} : \lambda_1 \nabla g_1(x) = \lambda_1 \left[ \begin{array}{c}
  x_2^2 \\
  2x_1 x_2
\end{array} \right] = 0, \lambda_1 g_1(x) = \lambda_1 (x_1 x_2^2 - 1) = 0. 
\]
It implies that $\lambda_1 x_2^2 = 0$ which gives $\lambda_1 = \lambda_2 x_1 x_2^2 = 0$. This contradicts $\lambda_1 \neq 0$. Thus we get $C(g) = \emptyset$, in consequence, $f(C(g)) = \emptyset$. From this, Theorem 7 yields that there exists $q \in P(g)[x, \lambda]$ with $\lambda = (\lambda_0, \lambda_1)$ such that $f - q$ vanishes on $V(h_{FJ})$ with $h_{FJ}$ defined as in (43), namely

$$h_{FJ} = (\lambda_0 \begin{array}{c} 1 \\ 0 \end{array}, \lambda_1 \begin{array}{c} x_2^2 \\ 2x_1 x_2^2 \end{array}, \lambda_1 (x_1^2 - 1), 1 - \lambda_0^2 - \lambda_2^2).$$

(49)

Let $(x, \lambda) \in V(h_{FJ})$. We get $\lambda_1 x_1 x_2 = 0$ which gives $\lambda_1 = \lambda_2 x_1 x_2 = 0$. It follows that $\lambda_0 = 0$ which is impossible since $1 = \lambda_1^2 + \lambda_2^2$. Thus $V(h_{FJ}) = \emptyset$ yields the selection $q = 0$.

We indicate the following counterexample, and consequently, the assumption that $f(C(g))$ is finite in Theorem 7 cannot be removed:

**Example 10.** Consider the problem (43) with $n = 2, m = 2, f = x_1 + x_2$ and $g = (-x_1^2, -x_2^2)$. Then we get $f^* = 0$ and $(0, 0)$ is the unique global minimizer for this problem. It is not hard to prove the Karush–Kuhn–Tucker conditions do not hold for this problem at any global minimizer. The condition rank($\varphi^g(x)$) < $m$ can be expressed as

$$\exists \lambda \in \mathbb{R}^2 \setminus \{0\} : \begin{cases} \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x) = \lambda_1 \begin{bmatrix} -2x_1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ -2x_2 \end{bmatrix} = 0, \\ \lambda_j g_j(x) = -\lambda_j x_j^2 = 0, j = 1, 2. \end{cases}$$

(50)

Since $\lambda \neq 0$, either $\lambda_1$ or $\lambda_2$ is non-zero. If $\lambda_1 \neq 0$, we get $x_1 = 0$. If $\lambda_2 \neq 0$, we get $x_2 = 0$. It implies that $x_1 = 0$ or $x_2 = 0$. Thus we get $C(g) = \{(t, 0), (0, t) : t \in \mathbb{R}\}$ which gives that $f(C(g)) = \mathbb{R}$ is infinite. Let $h_{FJ}$ be as in (47). Then we obtain

$$h_{FJ} = (\lambda_0 \begin{array}{c} 1 \\ 0 \end{array} - \lambda_1 \begin{bmatrix} -2x_1 \\ 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0 \\ -2x_2 \end{bmatrix}, -\lambda_1 x_1^2, -\lambda_2 x_2^2, 1 - \lambda_0^2 - \lambda_1^2 - \lambda_2^2).$$

(51)

Let $(x, \lambda) \in V(h_{FJ})$. We get $\lambda_1 x_1 = \lambda_2 x_2 = 0$ which gives $\lambda_0 = 0$. Since $\lambda_1 x_1 = 0$, either $\lambda_1 = 0$ or $x_1 = 0$ holds. If $\lambda_1 = 0$, we get $\lambda_2 \in \{\pm 1\}$ which gives $x_2 = 0$. If $x_1 = 0$, we consider $\lambda_2 x_2 = 0$ which gives $\lambda_2 = 0$ or $x_2 = 0$. If $x_1 = 0$ and $\lambda_2 = 0$, it implies that $\lambda_1 \in \{\pm 1\}$. If $x_1 = 0$ and $x_2 = 0$, it follows that $\lambda_1^2 + \lambda_2^2 = 1$. Thus we see that

$$V(h_{FJ}) = \{(t, 0, 0, \pm 1), (0, t, 0, \pm 1), (0, 0, 0, \text{cost}, \sin t) : t \in \mathbb{R}\}. \quad (52)$$

Since $f(C(g))$ is infinite, the assumption of Theorem 7 does not hold in this case. Assume that there exists $q \in P(g)[x, \lambda]$ such that $f - q$ vanishes on $V(h_{FJ})$. We write $q = \sigma_0 + \sigma_1 g_1 + \sigma_2 g_2 + \sigma_3 g_1 g_2$ for some $\sigma_j \in \Sigma^2[x]$. For all $t \in \mathbb{R}$, since $(t, 0, 0, 0, 1) \in V(h_{FJ})$, we have $f(t, 0) = q(t, 0, 0, 0, 1)$. Setting $\psi_j = \sigma_j (t, 0, 0, 0, 1) \in \Sigma^2[t]$, we obtain $q = \psi_0 = \psi_1 t^2$. It implies that $\psi_0 = t(1 + t^2) \psi_1$ which leads to $\psi_0 = t^2 \xi$ for some $\xi \in \Sigma^2[t]$ since $\psi_0 \in \Sigma^2[t]$. We now get $t = t^2 (\xi - \psi_1)$ which is impossible. Hence there does not exist $q \in P(g)[x, \lambda]$ such that $f - q$ vanishes on $V(h_{FJ})$. 

16
4 Exact polynomial optimization

4.1 Moment-SOS relaxations

In this subsection we recall some preliminaries of the Moment-SOS relaxations originally developed by Lasserre in [12]. Given $d \in \mathbb{N}$, let $\mathbb{N}^0_n := \{ \alpha \in \mathbb{N}^n : \sum_{j=1}^n \alpha_j \leq d \}$. Given $d \in \mathbb{N}$, we denote by $v_d$ the vector of monomials in $x$ of degree at most $d$, i.e., $v_d = (x^\alpha)_{\alpha \in \mathbb{N}^0_n}$ with $x^\alpha := x_1^{\alpha_1} \ldots x_n^{\alpha_n}$. For each $p \in \mathbb{R}[x]_d$, we write $p = c(p)^TV_d$, where $c(p)$ is denoted by the vector of coefficient of $p$, i.e., $c(p) = (p_\alpha)_{\alpha \in \mathbb{N}^0_n}$ with $p_\alpha \in \mathbb{R}$. Given $A \in \mathbb{R}^{r \times r}$ being symmetric, we say that $A$ is positive semidefinite, denoted by $A \succeq 0$, if every eigenvalue of $A$ is non-negative.

The following lemma shows the connection between sums of squares and semidefinite programs (see, e.g., [2]):

**Lemma 18.** Let $\sigma \in \mathbb{R}[x]$ and $d \in \mathbb{N}$ such that $2d \geq \deg(\sigma)$. Then $\sigma \in \Sigma^2[x]$ iff there exists $G \succeq 0$ such that $\sigma = v_j^TGV_d$.

Given $y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subset \mathbb{R}$, let $L_y : \mathbb{R}[x] \to \mathbb{R}$ be the Riesz linear functional defined by $L_y(p) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha y_\alpha$ for every $p \in \mathbb{R}[x]$. Given $d \in \mathbb{N}$, $p \in \mathbb{R}[x]$, $y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subset \mathbb{R}$, let $M_d(y)$ be the moment matrix of order $d$ defined by $(y_{\alpha + \beta})_{\alpha, \beta \in \mathbb{N}^0_n}$ and let $M_d(py)$ be the localizing matrix of order $d$ associated with $p$ defined by $(\sum_{\gamma \in \mathbb{N}^n} p_\gamma y_{\alpha + \beta + \gamma})_{\alpha, \beta \in \mathbb{N}^0_n}$.

Given $g_1, \ldots, g_m \in \mathbb{R}[x]$, let $Q_d(g)[x]$ be the truncated quadratic module of order $d$ associated with $g = (g_1, \ldots, g_m)$ defined by

$$Q_d(g)[x] = \{ \sigma_0 + \sum_{j=1}^m \sigma_j g_j : \sigma_j \in \Sigma^2[x], \deg(\sigma_0) \leq 2d, \deg(\sigma_j g_j) \leq 2d \}. \quad (53)$$

Given $h_1, \ldots, h_l \in \mathbb{R}[x]$, let $I_d(h)$ be the truncated ideal of order $d$ associated with $h = (h_1, \ldots, h_l)$ defined by

$$I_d(h)[x] = \{ \sum_{j=1}^l \psi_j h_j : \psi_j \in \mathbb{R}[x], \deg(\psi_j h_j) \leq 2d \}. \quad (54)$$

Given $k \in \mathbb{N}$ and $f, g_1, \ldots, g_m, h_1, \ldots, h_l \in \mathbb{R}[x]$, consider the following primal-dual semidefinite programs associated with $f$, $g = (g_1, \ldots, g_m)$ and $h = (h_1, \ldots, h_l)$:

$$\tau_k(f, g, h) := \inf_{y} L_y(f)$$
$$\text{s.t. } M_k(y) \succeq 0, \quad M_{k-d_j}(g_j y) \succeq 0, \quad j = 1, \ldots, m,$$
$$M_{k-r_i}(h_i y) = 0, \quad t = 1, \ldots, l, \quad y_0 = 1, \quad (55)$$

$$\rho_k(f, g, h) := \sup_{\xi, G, u} \xi$$
$$\text{s.t. } G_j \succeq 0,$$
$$f - \xi = v_k^T G_0 v_k + \sum_{j=1}^m g_j v_{k-d_j}^T G_j v_{k-d_j} + \sum_{t=1}^l h_t u_t^T v_{r_t}, \quad (56)$$

where $d_j = \lceil \deg(g_j)/2 \rceil$ and $r_t = \lceil \deg(h_t)/2 \rceil$. Using Lemma 18 we obtain

$$\rho_k(f, g, h) := \sup_{\xi \in \mathbb{R}} \{ \xi : f - \xi \in Q_k(g)[x] + I_k(h)[x] \}. \quad (57)$$
Primal-dual semidefinite programs (55)-(56) are known as the Moment-SOS relaxations of order \( k \) for problem

\[
\tilde{f}^* := \inf_{x \in S(g) \cap V(h)} f(x).
\]  

We state in the following lemma some recent results involving the Moment-SOS relaxations:

**Lemma 19.** Let \( f, g_1, \ldots, g_m, h_1, \ldots, h_l \in \mathbb{R}[x] \). Let \( \tilde{f}^* \) be as in (58) with \( g = (g_1, \ldots, g_m) \) and \( h = (h_1, \ldots, h_l) \). Then the following statements hold:

1. For every \( k \in \mathbb{N} \), \( \tau_k(f, g, h) \leq \tau_{k+1}(f, g, h) \) and \( \rho_k(f, g, h) \leq \rho_{k+1}(f, g, h) \).
2. For every \( k \in \mathbb{N} \), \( \rho_k(f, g, h) \leq \tau_k(f, g, h) \leq \tilde{f}^* \).
3. If \( S(g) \cap V(h) \) has a non-empty interior, for \( k \in \mathbb{N} \) sufficiently large, the Slater condition holds for the Moment relaxation (55) of order \( k \).
4. If \( S(g) \cap V(h) \) satisfies the Archimedean condition, \( \rho_k(f, g, h) \to \tilde{f}^* \) as \( k \to \infty \).
5. If there exists \( R > 0 \) such that \( g_m + h_l = R - x_1^2 - \cdots - x_n^2 \), for \( k \in \mathbb{N} \) sufficiently large, the Slater condition holds for the SOS relaxation (56) of order \( k \).
6. If there exists \( q \in Q(g)[x] \) such that \( f - \tilde{f}^* - q \) vanishes on \( V(h) \), then there exists \( k \in \mathbb{N} \) such that \( \rho_k(f, g, h) = \tilde{f}^* \).

**Proof.** The first four statements are proved by Lasserre in [12]. The proof of the fifth statement can be found in [13]. The final statement is based on Nie’s technique in [19], which is sketched as follows: Let \( u = f - \tilde{f}^* - q \). By assumption, we get \( u = 0 \) on \( V(h) \). From this, Krivine–Stengle’s Nichtnega-tivstellensätze [10] yields that there exists a positive integer \( r \) and \( \sigma \in \Sigma^r[x] \) such that \( u^{2r} + \sigma \in I(h)[x] \). Let \( c = \frac{1}{2^r} \). Then it holds that \( 1 + t + ct^{2r} \in \Sigma^r[t] \).

Thus for all \( \varepsilon > 0 \), we have

\[
f - \tilde{f}^* + \varepsilon = q + \varepsilon (1 + \frac{u}{\varepsilon} + c \left( \frac{u}{\varepsilon} \right)^{2r} - \varepsilon^1 - 2r (u^{2r} + \sigma) + \varepsilon 1 - 2r \sigma \in Q(g)[x] + I(h)[x].
\]

Moreover, the degree of the right-hand side has an upper bound independent from \( \varepsilon \). This implies that there exists \( k \in \mathbb{N} \) such that for all \( \varepsilon > 0 \), \( f - \tilde{f}^* + \varepsilon \in Q_k(g)[x] + I_k(h)[x] \). Then we for all \( \varepsilon > 0 \), \( \tilde{f}^* - \varepsilon \) is a feasible solution of (57) of the value \( \rho_k(f, g, h) \). It gives \( \rho_k(f, g, h) \geq \tilde{f}^* - \varepsilon \), for all \( \varepsilon > 0 \), and, in consequence, we get \( \rho_k(f, g, h) \geq \tilde{f}^* \). Using the second statement, we obtain that \( \rho_k(f, g, h) = \tilde{f}^* \), yielding the final statement.

**Remark 1.** In the final statement of Lemma 19, if we assume further that \( I(h) \) is real radical, and \( \rho_k(f, g, h) \) is attained. However, there is a case where \( \rho_k(f, g, h) \) is not attained, although \( \rho_k(f, g, h) = \tilde{f}^* \). Example is an instance for this (see at the end of [12] Section 3). Fortunately, if there exists \( R > 0 \) such that \( g_m + h_l = R - x_1^2 - \cdots - x_n^2 \), for \( k \in \mathbb{N} \) sufficiently large, the fifth statement of Theorem shows that the Slater condition holds for the SOS relaxation (56) of order \( k \), and hence the Moment relaxation (55) has at least one global minimizer.
As a consequence of Lemma 1, the following lemma is obtained:

Lemma 20. Let \(f, g_1, \ldots, g_m \in \mathbb{R}[x]\). Let \(f^*\) be as in problem (3) with \(g = (g_1, \ldots, g_m)\). Let \(h_{FJ}\) be as in (14). If problem (3) has a global minimizer, then it holds that

\[
f^* = \min_{x, \lambda} \ f(x) \\
\text{s.t.} \quad x \in S(g), \ (x, \lambda) \in V(h_{FJ}).
\]

We present in the following theorem the main application of Theorem 1 to polynomial optimization:

Theorem 3. Let \(f, g_1, \ldots, g_m \in \mathbb{R}[x]\). Let \(f^*\) be as in problem (3) with \(g = (g_1, \ldots, g_m)\). Let \(h_{FJ}\) be as in (14). Assume that problem (3) has at least one global minimizer, and \(f(C(g))\) is finite. Then there exists \(k \in \mathbb{N}\) such that \(\rho_k(f, \Pi g, h_{FJ}) = f^*\), where \(\Pi g\) is defined as in (7). Moreover, if \(S(g)\) satisfies the Archimedean condition, there exists \(k \in \mathbb{N}\) such that \(\rho_k(f, g, h_{FJ}) = f^*\). Furthermore, if these exists \(R > 0\) such that \(g_m = R - x_1^2 - \cdots - x_n^2\), for \(k \in \mathbb{N}\) sufficiently large, the Slater condition holds for the SOS relaxation (56) of order \(k\) with \(h = h_{FJ}\).

Proof. We first prove \(\rho_k(f, \Pi g, h_{FJ}) = f^*\) for some \(k \in \mathbb{N}\). Since \(S(g) = S(\Pi g)\), Lemma 20 implies that

\[
f^* = \min_{x, \lambda} \ f(x) \\
\text{s.t.} \quad x \in S(\Pi g), \ (x, \lambda) \in V(h_{FJ}),
\]

By assumption, Theorem 1 yields that there exists \(g \in P(g)[x, \bar{\lambda}] = Q(\Pi g)[x, \bar{\lambda}]\) such that \(f - f^* - g\) vanishes on \(V(h_{FJ})\). Applying the final statement of Lemma 19 (by replacing \(g\) with \(\Pi g\), we obtain the first statement. Assume that \(S(g)\) satisfies the Archimedean condition. The proof of the equality \(\rho_k(f, g, h_{FJ}) = f^*\) for some \(k \in \mathbb{N}\) is similar. Now assume that there exists \(R > 0\) such that \(g_m = R - x_1^2 - \cdots - x_n^2\). Let us prove the final statement. By definition of \(h_{FJ}\), the final entry of \(h_{FJ}\) is \(b = 1 - \lambda_0^2 - \cdots - \lambda_m^2\) which implies that

\[
g_m + b = (R + 1) - x_1^2 - \cdots - x_n^2 - \lambda_0^2 - \cdots - \lambda_m^2.
\]

From this, the fifth statement of Lemma 19 shows that for \(k \in \mathbb{N}\) sufficiently large, the Slater condition holds for the SOS relaxation (56) of order \(k\) with \(h = h_{FJ}\), yielding the final statement.

Combining Theorem 3, Examples 6, 7, and 8, we obtain the following corollary:

Corollary 1. Let \(f, g_1, \ldots, g_m \in \mathbb{R}[x]\). Let \(f^*\) be as in problem (3) with \(g = (g_1, \ldots, g_m)\). Let \(h_{FJ}\) be as in (14). Assume that problem (3) has at least one global minimizer, and one of the following conditions holds:

1. \(g = (1 - x_1^2 - \cdots - x_n^2)\);
2. \(g = (1 - x_n^2, \ldots, 1 - x_n^2)\);
3. \(g = (x_1, \ldots, x_n, 1 - x_1 - \cdots - x_n)\).

Then there exists \(k \in \mathbb{N}\) such that \(\rho_k(f, g, h_{FJ}) = f^*\).
The following two examples are given in [19, Example 3.3]:

**Example 11.** Let \( \varepsilon > 0 \). Consider the problem \((3)\) with \( n = 3, m = 1, f = x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^2 - 3 x_1^2 x_2^2 x_3^2 + \varepsilon (x_1^2 + x_2^2 + x_3^2) \) and \( g = (1 - x_1^2 - x_2^2 - x_3^2) \).

For \( \varepsilon > 0 \) sufficiently small, Lasserre’s hierarchy for this problem does not have finite convergence, as shown by Marshall [14, Example 2.4]. However, Corollary 1 yields that there exists \( k \in \mathbb{N} \) such that \( \rho_k(f,g,h_{FJ}) = f^* \).

**Example 12.** Consider the problem \((3)\) with \( n = 2, m = 3, f = x_1 x_2 + x_1^2 + x_2^2 \) and \( g = (x_1, x_2, 1 - x_1 - x_2) \). As shown by Scheiderer [25, Remark 3.9], Lasserre’s hierarchy for this problem does not have finite convergence. However, Corollary 1 yields that there exists \( k \in \mathbb{N} \) such that \( \rho_k(f,g,h_{FJ}) = f^* \).

The following two examples are given in [18, Example 5.6]:

**Example 13.** Consider the problem \((3)\) with \( n = m = 3, f = x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^2 x_4^2 - 3 x_1^2 x_2^2 x_3 x_4^2 \) and \( g = (1 - x_1^2, 1 - x_2^2, 1 - x_3^2) \). As shown by Nie [18, Example 5.6], Lasserre’s hierarchy for this problem does not have finite convergence. However, Corollary 1 yields that there exists \( k \in \mathbb{N} \) such that \( \rho_k(f,g,h_{FJ}) = f^* \).

### 4.2 Numerical experiments

In this subsection, we report numerical results produced by the SOS relaxations of the values \( \rho_k(f,g,0) \) and \( \rho_k(f,g,h_{FJ}) \) for problem \((3)\), where \( h_{FJ} \) is defined as in [14]. The former is the standard semidefinite program in [12] while the latter is the semidefinite program modeled by our method in this paper. We indicate the data of each semidefinite program, namely “value”, “time”, and “size” correspond to the numerical value of the optimal value of the semidefinite program, the running time in seconds to obtain this numerical value, and the size of the semidefinite program, respectively. Here the size of the semidefinite program includes the largest matrix size, the number of affine constraints, the number of scalar variables, and the number of matrix variables.

The experiments are performed in Julia 1.3.1 with softwares TSSOS [28] and Mosek 9.1 [16]. We use a desktop computer with an Intel(R) Core(TM) i7-865U CPU @ 1.9GHz × 8 and 31.2 GB of RAM.

Our test problem is taken from Example 3 namely \( n = 2, m = 1, f = (x_1 + 1)^2 + x_2^2 - 1 \) and \( g = (g_1) = (x_1^2 - x_2^2) \). It is clear that \( f^* = 0 \) which is attained at the unique global minimizer \((0,0)\) for problem \((3)\). Moreover, the Karush–Kuhn–Tucker conditions do not hold at this minimizer. By using [13, Proposition 3.4], we get \( \rho_k(f,g,0) < f^* \) for all \( k \in \mathbb{N} \). As shown in Example 3 \( f(C(g)) \) is singleton. From this, Theorem 3 yields that \( \rho_d(f,g,h_{FJ}) = f^* \) for some \( d \in \mathbb{N} \). Thus we obtain \( \rho_k(f,g,0) < f^* = \rho_d(f,g,h_{FJ}) \), for all \( k \in \mathbb{N} \).

We display the numerical results in Table 1. It shows that the numerical value \(-0.00831\) of \( \rho_6(f,g,h_{FJ}) \) is the best lower bound on \( f^* \). It takes around 7 seconds to obtain this numerical value. It is worth pointing out that the standard SOS relaxations of the values \( \rho_k(f,g,0) \) and \( \rho_k(f,g,b,0) \) cannot reach the bound \(-0.00831\) in less than 10 seconds despite using the additional ball constraint \( b = 1 - x_1^2 - x_2^2 \). Another observation is that the size of the SOS relaxations of the value \( \rho_k(f,g,0) \) grows faster than the ones of the values \( \rho_k(f,g,0) \) and \( \rho_k(f,g,b,0) \) when \( k \) increases.
Table 1: Lower bounds on $f^* = 0$

| $k$ | value     | time     | size             |
|-----|-----------|----------|------------------|
| 2   | -0.99984  | 0.03     | (15, 70, 34, 1)  |
| 3   | -0.48463  | 0.08     | (35, 210, 271, 2)|
| 4   | -0.00920  | 0.33     | (70, 495, 1501, 2)|
| 5   | -0.00860  | 1.63     | (126, 1001, 6231, 2)|
| 6   | -0.00831  | 7.64     | (210, 1820, 20973, 2)|

| $k$ | value     | time     | size             |
|-----|-----------|----------|------------------|
| 16  | -0.04760  | 2.34     | (153, 561, 1, 2) |
| 17  | -0.04672  | 3.30     | (171, 630, 1, 2) |
| 18  | -0.04528  | 5.26     | (190, 703, 1, 2) |
| 19  | -0.04476  | 6.73     | (210, 780, 1, 2) |
| 20  | -0.04337  | 11.4     | (231, 861, 1, 2) |

5 Variations

5.1 Representations

We state in the following theorem the first variation of Theorem 1, where we assume that the image of the intersection of a semi-algebraic set with the set of its singularities is finite:

**Theorem 4.** Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $f$ is non-negative on $S(g)$ with $g := (g_1, \ldots, g_m)$ and $f(C(g) \cap S(g))$ is finite. Then there exists $q \in P(g)[x, \bar{\lambda}]$ such that $f - q$ vanishes on $(S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ})$, where $\bar{\lambda} := (\lambda_0, \ldots, \lambda_m)$ and $h_{FJ}$ is defined as in (14). Moreover, if $S(g)$ satisfies the Archimedean condition, we can take $q \in Q(g)[x, \bar{\lambda}]$.

The proof of Theorem 4 is postponed to Section 5.2.

Given a real matrix $A$, we denote by $\text{rank}^+(A)$ the largest number of columns of $A$ whose convex hull over $\mathbb{R}$ has no zero. Let $C^+(g)$ be the set of critical points associated with $g$ defined by

$$C^+(g) := \{x \in \mathbb{R}^n : \text{rank}^+(\varphi^g(x)) < m\}.$$ 

We state the second variation of Theorem 1 in the following theorem:

**Theorem 5.** Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $f$ is non-negative on $S(g)$ with $g := (g_1, \ldots, g_m)$, and $f(C^+(g))$ is finite. Then there exists $q \in P(g)[x, \bar{\lambda}]$ such that $f - q$ vanishes on $V(h_{FJ}^+)$, where $\bar{\lambda} := (\lambda_0, \ldots, \lambda_m)$ and

$$h_{FJ}^+ := (\lambda_0 \nabla f - \sum_{j=1}^{m} \lambda_j^2 \nabla g_j, \lambda_1^2 g_1, \ldots, \lambda_m^2 g_m, 1 - \sum_{j=0}^{m} \lambda_j^2).$$ (63)

Moreover, if $S(g)$ satisfies the Archimedean condition, we can take $q \in Q(g)[x, \bar{\lambda}]$. 


Proof. To prove Theorem 6 we do similarly to the proof of Theorem 1 by replacing $C(g)$ and $h_{FJ}^r$ in Section 2.2 with $C^+(g)$ and $h_{FJ}^+$, respectively. Note that the equality $C^+(g) = \pi(V(h_{FJ}^+) \cap \{\lambda_0 = 0\})$ follows from the following equivalences:

$$x \in C^+(g)$$
$$\Leftrightarrow \text{rank}^+ (p^2(x)) < m$$
$$\Leftrightarrow \exists \lambda \in \mathbb{R}^m : \sum_{j=1}^m \lambda_j^2 = 1, \sum_{j=1}^m \lambda_j^2 \nabla g_j(x) = 0, \lambda_j^2 g_j(x) = 0$$
$$\Leftrightarrow \exists \lambda \in \mathbb{R}^{m+1} : \sum_{j=0}^m \lambda_j^2 = 1, \lambda_0 = 0, \lambda_0^2 \nabla f(x) = \sum_{j=1}^m \lambda_j^2 \nabla g_j(x), \lambda_j^2 g_j(x) = 0$$
$$\Leftrightarrow \exists \lambda \in \mathbb{R}^{m+1} : (x, \lambda) \in V(h_{FJ}^+) \cap \{\lambda_0 = 0\}$$
$$\Leftrightarrow x \in \pi(V(h_{FJ}^+) \cap \{\lambda_0 = 0\})$$.

(64)

Moreover, the Lagrangian function becomes $L(x, \bar{\lambda}) = f(x) + \sum_{j=1}^m \left( \frac{\lambda_j}{\lambda_0} \right)^2 g_j(x)$. □

We state the third variation of Theorem 1 in the following theorem:

**Theorem 6.** Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $f$ is non-negative on $S(g)$ with $g := (g_1, \ldots, g_m)$ and $f(C^+(g) \cap S(g))$ is finite. Then there exists $q \in P(g)[x, \bar{\lambda}]$ such that $f - q$ vanishes on $S(g) \cap V(h_{FJ}^+)$, where $\bar{\lambda} := (\lambda_0, \ldots, \lambda_m)$ and $h_{FJ}^+$ is defined as in (30). Moreover, if $S(g)$ satisfies the Archimedean condition, we can take $q \in Q(g)[x, \bar{\lambda}]$.

**Proof.** The proof is processed similarly to the one of Theorem 1 by replacing $C(g)$ and $h_{FJ}^r$ in Section 2.2 with $C^+(g)$ and $h_{FJ}^+$, respectively. □

**Remark 2.** Since $C(g) \cap S(g)$, $C^+(g)$ and $C^+(g) \cap S(g)$ are subsets of $C(g)$, the assumptions of Theorems 2, 5 and 6 that $f(C(g) \cap S(g))$, $f(C^+(g))$ and $f(C^+(g) \cap S(g))$ are finite hold generically thanks to Theorem 3 respectively.

### 5.2 Proofs

We use the same notation as in Section 2. We generalize Lemma 11 in the following lemma:

**Lemma 21.** Let $f \in \mathbb{R}[x]$, let $W$ be a complex variety defined by finitely many polynomials in $\mathbb{R}[x]$, and let $A$ be a semi-algebraic subset of $\mathbb{R}^n$. Assume that $f(A \cap W)$ is finite. Then there exists a finite sequence of subsets $W_1, \ldots, W_r$ such that the following conditions hold:

1. $W_1, \ldots, W_r$ are pairwise disjoint complex varieties defined by finitely many polynomials in $\mathbb{R}[x]$;
2. for $j = 1, \ldots, r$, $W_j \subset W$, and $f$ is constant on $W_j$;
3. $(W_1 \cup \cdots \cup W_r) \cap A = W \cap A$.

The proof of Lemma 21 is similar to the one of Lemma 11 since we only need to replace $\mathbb{R}^n$ in the proof of Lemma 11 with $A$.

The following lemma extends [1] Lemma 3.3] to the case of varieties defined by the Fritz John conditions:
Lemma 22. Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$. Assume that $f(C(g))$ with $g := (g_1, \ldots, g_m)$ is finite. Let $h_{FJ}$ be defined as in (13). Let $W$ be a semi-algebraically path connected component of $(S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ})$. Then $f$ is constant on $W$.

Proof. The proof is similar to the one of Lemma 10. When proving that $\tau \mapsto f(x(\tau))$ is constant on $[a_j, b_j]$, we note that
\[
f(S(g) \cap C(g)) = f(S(g) \cap \pi(V(h_{FJ}) \cap \{\lambda_0 = 0\}))
\]
is finite since $f(S(g) \cap \pi(V(h_{FJ}) \cap \{\lambda_0 = 0\})) \supset f(x([\tau_1, \tau_2]))$, which is because
\[
S(g) \cap \pi(V(h_{FJ}) \cap \{\lambda_0 = 0\}) = \pi((S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ}) \cap \{\lambda_0 = 0\}) \\
\supset \pi(W \cap \{\lambda_0 = 0\}) \supset x([\tau_1, \tau_2]).
\]

Proof of Theorem 1

Proof. Using Lemma 9, we decompose $(S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ})$ into semi-algebraically path connected components: $Z_1, \ldots, Z_s$. Lemma 22 yields that $f$ is constant on $Z_i$. Thus $f((S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ}))$ is finite. Observe that $(S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ}) = (S(g) \times \mathbb{R}^{m+1}) \cap V_C(h_{FJ})$. By using Lemma 21 we obtain a finite sequence of subsets $W_1, \ldots, W_r$ such that the following conditions hold:

- $W_1, \ldots, W_r$ are pairwise disjoint complex varieties defined by finitely many polynomials in $\mathbb{R}[x, \lambda]$;
- for $j = 1, \ldots, r$, $W_j \subset V_C(h_{FJ})$, and $f$ is constant on $W_j$;
- $(W_1 \cup \cdots \cup W_r) \cap (S(g) \times \mathbb{R}^{m+1}) = V_C(h_{FJ}) \cap (S(g) \times \mathbb{R}^{m+1}) = (S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ})$.

Let $D$ be the union of $W_1, \ldots, W_r$. Let $b = 1 - \lambda_2^2 - \cdots - \lambda_m^2$. From this, Lemma 12 yields that there exists $p \in P(g, b)[x, \lambda]$ such that $f - p$ vanishes on $D \cap \mathbb{R}^{n+m+1} \supset (S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ})$. We write $p$ as in (28) for some $\sigma_0, \psi_0 \in \Sigma^2[x, \lambda]$. Let $q = \sum_{\alpha \in \{0, 1\}^m} \sigma_0 g^\alpha \in P(g)[x, \lambda]$. Since $b = 0$ on $V(h_{FJ})$, it holds that $f = p = q$ on $(S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ})$.

Assume that $S(g)$ satisfies the Archimedean condition. Then there exists $R > 0$ such that $g_{m+1} = R - x_1^2 - \cdots - x_n^2 \in Q(g)[x]$. It implies that $S(g, b)$ with $b = 1 - \lambda_2^2 - \cdots - \lambda_m^2$ satisfies the Archimedean condition due to (29).

From this, Lemma 12 yields that there exists $p \in Q(g, b)[x, \lambda]$ such that $f - p$ vanishes on $D \cap \mathbb{R}^{n+m+1} \supset (S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ})$. We write $p$ as in (29) for some $\sigma_j \in \Sigma^2[x, \lambda]$. Let $q = \sigma_0 + \sum_{j=1}^m \sigma_j g_j \in Q(g)[x, \lambda]$. Since $b = 0$ on $V(h_{FJ})$, $f = p = q$ on $(S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ})$. This completes the proof.

5.3 Examples

This section illustrates our Nichtnegativstellensätze stated in Theorems 1, 5, and 6 with several explicit examples. The following lemma shows a case where $S(g)$ is convex and $C^+(g)$ is empty:
Lemma 23. Let \( g = (g_1, \ldots, g_m) \) with \( g_j \in \mathbb{R}[x] \). Assume that each \( g_j \) is concave, and \( S(g) \) has a non-empty interior. Then \( S(g) \) is convex and \( C^+(g) = \emptyset \).

Proof. It is a simple matter to prove that \( S(g) \) is convex. Assume by contradiction that there is \( y \in C^+(g) \). Then by (64), there exists \( \lambda \in \mathbb{R}^m \setminus \{0\} \) such that \( \sum_{j=1}^{m} \lambda_j^2 \nabla g_j(y) = 0 \) and \( \lambda_j^2 g_j(y) = 0 \) for all \( j \). Let \( G = \sum_{j=1}^{m} \lambda_j^2 g_j(x) \). Then \( G \) is concave since all \( g_j \) are concave. In addition, \( \nabla G(y) = 0 \) yields that \( G(y) = 0 \) is the maximal value of \( G \). Let \( a \) be in the interior of \( S(g) \). Then \( 0 \geq G(a) = \sum_{j=1}^{m} \lambda_j^2 g_j(a) \) implies that \( \lambda_1 = \cdots = \lambda_m = 0 \) since all \( g_j(a) \) are positive. This contradicts \( \lambda \neq 0 \), and hence it holds that \( C^+(g) = \emptyset \).

The following lemma is a consequence of Lemma 23 and Theorem 5:

Lemma 24. Let \( f, g_1, \ldots, g_m \in \mathbb{R}[x] \). Assume that \( f \) is non-negative on \( S(g) \) with \( g := (g_1, \ldots, g_m) \), each \( g_j \) is concave, and \( S(g) \) has a non-empty interior. Then \( S(g) \) is convex, and there exists \( q \in P(g)[x, \lambda] \) such that \( f - q \) vanishes on \( V(h^+_{FJ}) \), where \( \lambda := (\lambda_1, \ldots, \lambda_m) \), and \( h^+_{FJ} \) is defined as in (63). Moreover, if \( S(g) \) satisfies the Archimedean condition, we can take \( q \in Q(g)[x, \lambda] \).

To illustrate the representations in Lemma 24 see Examples 6, 7 and 8. Note that if \( C(g) = \emptyset \), then so is \( C^+(g) \) since \( C^+(g) \subset C(g) \).

Contrary to Lemma 24 the following example shows the representation of any polynomial non-negative on a non-convex semi-algebraic set \( S(g) \):

Example 14. Consider the problem (60) with \( n = 2, m = 3 \) and \( g = (g_1, g_2, g_3) = (x_1 + 1, 1 - x_3^2, 1 - (x_1 - 1)^2 - x_2^2) \). Then \( S(g) \) is non-convex since it contains all points in the hypercube \([-1, 1]^2 \) but not in the open ball of center \((1, 0)\) with unit radius. The condition rank \(^+(\varphi^q(x)) < m \) can be expressed as

\[
\begin{align*}
\exists \lambda &\in \mathbb{R}^3 \setminus \{0\} : \sum_{j=1}^{3} \lambda_j \nabla g_j(x) = \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ -2x_2 \end{bmatrix} + \lambda_3 \begin{bmatrix} 2(x_1 - 1) \\ -2x_2 \end{bmatrix} = 0, \\
\lambda_1 g_1(x) &= \lambda_1 (x_1 + 1) = 0, \quad \lambda_2 g_2(x) = \lambda_2 (1 - x_3^2) = 0, \\
\lambda_3 g_3(x) &= \lambda_3 (1 - (x_1 - 1)^2 - x_2^2) = 0.
\end{align*}
\]

(67)

It implies that \( \lambda_1 = 0 \) or \( x_1 = -1 \). If \( \lambda_1 = 0 \), then \( 2\lambda_3 (x_1 - 1) = 0 \) which gives \( \lambda_3 = 0 \) or \( x_1 = 1 \). If \( \lambda_1 = \lambda_3 = 0 \), then \( \lambda_2 \neq 0 \) which implies \( -2x_2 = 0 \) and \( 1 - x_3^2 = 0 \), hence this is impossible. If \( \lambda_1 = 0 \) and \( x_1 = 1 \), then \(-2\lambda_3 + 1)x_2 = 0 \) and \( \lambda_2 (1 - x_3^2) = \lambda_3 (1 - x_3^2) = 0 \). If \( \lambda_1 = 0 \), \( x_1 = 1 \) and \( x_2 = 0 \), then \( \lambda_2 = \lambda_3 = 0 \) which contradicts \( \lambda \neq 0 \). If \( \lambda_1 = 0 \), \( x_1 = 1 \) and \( x_2 = 0 \), then \( \lambda_2 = \lambda_3 = 0 \) (since \( \lambda_j \geq 0 \) which also contradicts \( \lambda \neq 0 \)).

In the following example, we reconsider Example 10 where Theorem 1 is inapplicable, but Theorem 4 is applicable:

Example 15. Consider the problem (69) with \( n = 2, m = 2 \), \( f = x_1 + x_2 \) and \( g = (-x_1^2, -x_2^2) \). Then we get \( f^* = 0 \), \( S(g) = \{(0, 0)\} \) and \((0, 0)\) is the unique
global minimizer for this problem. We get \( C(g) = \{(t,0),(0,t) : t \in \mathbb{R}\} \) which gives that \( f(C(g)) = \mathbb{R} \) is infinite. However \( C(g) \cap S(g) = \{(0,0)\} \) implies \( f(C(g) \cap S(g)) = \{0\} \) is finite. Let \( h_{FJ} \) be as in \([13]\). Then

\[
V(h_{FJ}) = \{(t,0,0,0,\pm 1),(0,t,0,\pm 1,0),(0,0,0,\text{cost},\sin t) : t \in \mathbb{R}\}. \quad (68)
\]

By Theorem \([7]\) there exists \( q \in P(g)[x,\lambda] \) such that \( f - q \) vanishes on

\[
(S(g) \times \mathbb{R}^{m+1}) \cap V(h_{FJ}) = \{(0,0,0,\text{cost}, \sin t) : t \in \mathbb{R}\}. \quad (69)
\]

It is not hard to take \( q = 0 \).

The following lemma shows a case where \( C^+(g) \) is a singleton:

**Example 16.** Consider the problem \([3]\) with \( n = 2, m = 3, f = x_1 - 1 \) and \( g = (g_1, g_2, g_3) = (x_1, x_2,(x_1 - 1)^3 - x_2) \). It is easy to check that \( S(g) \) is non-convex and \( f^* = 0 \). Moreover, \((1,0)\) is the unique global minimizer for this problem, and Karush–Kuhn–Tucker conditions do not hold for \((3)\) at this point. The condition \( \text{rank}^+(\varphi^g(x)) < m \) can be expressed as

\[
\left\{ \begin{array}{l}
\exists \lambda \in \mathbb{R}^3 \setminus \{0\} : \sum_{j=1}^3 \lambda_j \nabla g_j(x) = \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_3 \begin{bmatrix} 3(x_1 - 1)^2 \\ -1 \end{bmatrix} = 0, \\
\lambda_1 g_1(x) = \lambda_1 x_1 = 0, \lambda_2 g_2(x) = \lambda_2 x_2 = 0, \lambda_3 g_3(x) = \lambda_3 ((x_1 - 1)^3 - x_2) = 0.
\end{array} \right. \quad (70)
\]

It implies that \( \lambda_3 = \lambda_2 \) so \( \lambda_1 + 3\lambda_2(x_1 - 1)^2 = 0 \). It follows that \( \lambda_1 = 0 \) and \( \lambda_2(x_1 - 1) = 0 \). If \( \lambda_2 = 0 \), then \( \lambda = 0 \), and hence this is impossible. Thus we get \( \lambda_2 > 0 \) which gives \( x_1 = 1 \) and \( x_2 = 0 \). Thus \( C^+(g) = \{(1,0)\} \) is a singleton then so is \( f(C^+(g)) \). From this, Theorem \([5]\) shows that there exists \( q \in P(g)[x,\lambda] \) with \( \lambda = (\lambda_0,\lambda_1,\lambda_2,\lambda_3) \) such that \( f - q \) vanishes on \( V(h_{FJ}^+) \) with \( h_{FJ}^+ \) defined as in \([69]\), i.e.,

\[
h_{FJ}^+ = (\lambda_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \lambda_3 \begin{bmatrix} 3(x_1 - 1)^2 \\ -1 \end{bmatrix}, 1 - \sum_{j=0}^3 \lambda_j^2). \quad (71)
\]

Let \( (x,\lambda) \in V(h_{FJ}^+) \). We get \( \lambda_3^2 = \lambda_2^2 \), so \( \lambda_2^2 - \lambda_1^2 - 3\lambda_2((x_1 - 1)^3 - x_2) = 0 \) and \( \lambda_2^2 x_1 = \lambda_3^2 x_2 = \lambda_3^2((x_1 - 1)^3 - x_2) = 0 \). If \( \lambda_2 = 0 \), then \( \lambda_0 = \lambda_1 = \lambda_3 = 0 \). If \( \lambda_2 \neq 0 \), then \( x_2 = 0, x_1 = 1 \) and \( \lambda_1 = \lambda_0 = 0 \) which gives \( \lambda_2^2 = \lambda_3^2 = \frac{1}{2} \). Thus we obtain

\[
V(h_{FJ}^+) = \{(0, t, \frac{1}{\sqrt{2}}r_1, \frac{1}{\sqrt{2}}r_2, 0, 0), (1, 0, 0, 0, \frac{1}{\sqrt{2}}r_1, \frac{1}{\sqrt{2}}r_2) : r_j = \pm 1, t \in \mathbb{R}\}. \quad (72)
\]

It is not hard to check that \( q = (x_1 - 1)^3 = g_2 + g_3 \in Q(g)[x,\lambda] \subset P(g)[x,\lambda] \) and \( f - q \) vanishes on \( V(h_{FJ}^+) \).

**Remark 3.** From Lemma \([20]\) and Examples \([4, 66]\) if \( C^+(g) \) is finite, then there are cases where \( S(g) \) is convex, and there are cases where \( S(g) \) is non-convex. This implies that the finiteness of \( C^+(g) \) does not depend on the convexity of \( S(g) \). Thus it is still open to find all explicit cases of \( g \) where \( C^+(g) \) is finite.
Next, we show in the following example that Theorems 5 and 6 are inapplicable to Example 4.

Example 17. Consider the problem with $n = 2$, $m = 1$, $f = x_1 + x_2^2$ and $g = (g_1) = (-x_1^2)$. Then the point $(0, 0)$ is the unique global minimizer. The condition $\text{rank}^+(\varphi^g(x)) < m$ can be expressed as

$$\exists \lambda \in \mathbb{R}^+ \setminus \{0\} : \lambda_1 \nabla g_1(x) = \lambda_1 \begin{bmatrix} -2x_1 \\ 0 \end{bmatrix} = 0, \lambda_1 g_1(x) = -\lambda_1 x_2^2 = 0. \quad (73)$$

It is equivalent to $x_1 = 0$. Thus $C(g) = C^+(g) = S(g) = \{(0, t) : t \in \mathbb{R}\}$ is infinite, which implies that $f(C(g)) = f(C^+(g)) = f(C(g) \cap S(g)) = f(C^+(g) \cap S(g)) = \{t^2 : t \in \mathbb{R}\} = [0, \infty)$ is infinite.

We prove in the following example that the finiteness assumptions of $f(C^+(g))$ and $f(S(g) \cap C^+(g))$ in Theorems 5 and 6 cannot be removed, respectively:

Example 18. Consider problem with $n = 2$, $m = 3$, $f = x_1 + x_2$, and $g = (g_1, g_2, g_3) = (x_1^2, x_2^2, -x_1 x_2)$. It is easy to check that $S(g) = \{(0, t), (t, 0) : t \geq 0\}$ and $f^* = 0$. Moreover, $(0, 0)$ is the unique global minimizer for this problem, and Karush–Kahn–Tucker conditions do not hold for at this point. The condition $\text{rank}^+(\varphi^g(x)) < m$ can be expressed as

$$\exists \lambda \in \mathbb{R}^3 \setminus \{0\} : \sum_{j=1}^3 \lambda_j \nabla g_j(x) = \lambda_1 \begin{bmatrix} 2x_1^2 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 2x_2^2 \end{bmatrix} + \lambda_3 \begin{bmatrix} -x_2 \\ -x_1 \end{bmatrix} = 0, \lambda_1 g_1(x) = \lambda_1 x_1^2 = 0, \lambda_2 g_2(x) = \lambda_2 x_2^3 = 0, \lambda_3 g_3(x) = -\lambda_3 x_1 x_2 = 0. \quad (75)$$

It implies that $\lambda_3 x_1 = \lambda_3 x_2 = \lambda_1 x_1 = \lambda_2 x_2 = 0$. If $\lambda_3 \neq 0$, then $x = 0$. If $\lambda_3 = 0$, we have either $x_1 = 0$ or $x_2 = 0$. If $\lambda_1 = \lambda_3 = 0$, then $x_2 = 0$ so $x_2 = 0$. Thus $C^+(g) = \{(0, t), (t, 0) : t \in \mathbb{R}\}$ is infinite then so are $f(C^+(g))$ and $f(S(g) \cap C^+(g))$. Let $h_{P,J}^+$ be as in (63). Then we get

$$h_{P,J}^+ = (\lambda_0^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda_2^2 \begin{bmatrix} 2x_1^2 \\ 0 \end{bmatrix} - \lambda_3^2 \begin{bmatrix} 0 \\ 2x_2^2 \end{bmatrix} - \lambda_3 \begin{bmatrix} -x_2 \\ -x_1 \end{bmatrix}, \lambda_1^2 x_1^2, \lambda_2^2 x_2^3, -\lambda_3^2 x_1 x_2, 1 - \sum_{j=0}^3 \lambda_j^2). \quad (76)$$

Let $(x, \lambda) \in V(h_{P,J}^+)$. Then we have $\lambda_1 x_1 = \lambda_2 x_2 = \lambda_3^2 + \lambda_2^2 x_2 = \lambda_3 x_2 + \lambda_2 x_2 = 0$. If $\lambda_3 = 0$, then $\lambda_0 = \lambda_1 x_1 = \lambda_2 x_2 = 0$, which implies $\lambda_1 = 0$ or $x_1 = 0$. If $\lambda_3 = \lambda_0 = \lambda_1 = 0$, then $x_2 = \pm 1$, which yields $x_2 = 0$. If $\lambda_3 = \lambda_0 = x_1 = 0, then \lambda_2 = 0 or x_2 = 0$. If $\lambda_3 \neq 0, then x_1 = x_2 = -\lambda_0^2/\lambda_3$ and $\lambda_0 x_1 = \lambda_0 x_2 = 0$.

Thus we obtain

$$V(h_{P,J}^+) = \{(t_1, 0, 0, 0, \pm 1, 0), (0, t_2, 0, \pm 1, 0, 0), (0, 0, \cos t_3, \sin t_3, 0, (0, 0, \cos t_3, \sin t_3, 0), (-\mu_1^2, -\mu_2^2, 0, \mu_1, \mu_2, \mu_3) : t_1 \in \mathbb{R}, \mu_j \in \mathbb{R}, \mu_3 \neq 0, \sum_{i=0}^3 \mu_i^2 = 1, \mu_0 \mu_1 = \mu_0 \mu_2 = 0\}. \quad (77)$$

Let us prove that there does not exist $q \in P(g)[x, \lambda]$ such that $f - q$ vanishes on $V(h_{P,J}^+)$. Assume by contradiction that there is $\sigma_\alpha \in \Sigma^2[x, \lambda]$ such that $f - q$...
\[
\sum_{e \in \{0,1\}^t} \sigma_e g^e \ \text{vanishes on } V(h_{FJ}) . \] Since \((t,0,0,0,1,0) \in V(h_{FJ})\) for all \(t \in \mathbb{R}\), we get
\[
t - \sigma_{(0,0,0)}(t,0,0,0,1,0) - \sigma_{(1,0,0)}(t,0,0,0,1,0) \epsilon^3 = 0 .
\] (78)

Set \(\psi_0(t) = \sigma_{(0,0,0)}(t,0,0,0,1,0) \in \Sigma^2[t]\) and \(\psi_1(t) = \sigma_{(1,0,0)}(t,0,0,0,1,0) \in \Sigma^2[t]\). Then it implies that \(\psi_0 = t(1 - t^2 \psi_1), \) yielding \(\psi_0 = t^2 \hat{\psi}_0\) for some \(\hat{\psi}_0 \in \Sigma^2[t]\) (since \(\psi_0 \in \Sigma^2[t]\)). Thus, \(t = \psi_0 + t^3 \psi_1 = t^2 (\hat{\psi}_0 + \psi_1)\), and hence this is impossible.

### 5.4 Application to exact polynomial optimization

In this subsection, we apply the representations stated in Section 5.1 for computing precisely the optimal value of a polynomial optimization problem using semidefinite programming.

We state in the following lemma an extension of the sixth statement of Lemma 19.

**Lemma 25.** Let \(f,g_1,\ldots,g_m,h_1,\ldots,h_l \in \mathbb{R}[x]\). Let \(\hat{f}^*\) be as in (58) with \(g = (g_1,\ldots,g_m)\) and \(h = (h_1,\ldots,h_l)\). Assume that there exists \(q \in P(g)[x]\) such that \(f - \hat{f}^* - q\) vanishes on \(V(h) \cap S(g)\). Then there exists \(k \in \mathbb{N}\) such that \(\rho_k(f,\Pi g, h) = \hat{f}^*\).

**Proof.** The proof is proved similar to the sixth statement of Lemma 19. It is sketched as follows: Let \(u = f - \hat{f}^* - q\). By assumption, we get \(u = 0\) on \(S(g) \cap V(h)\). From this, Krivine-Stengle’s Nichtnegativstellensätze [10] say that there exists a positive integer \(r\) and \(w \in P(g)[x]\) such that \(u^{2r} + w \in I(h)[x]\). Let \(c = \frac{1}{2r}\). Then it holds that \(1 + t + ct^2r \in \Sigma^2[t]\). Thus for all \(\varepsilon > 0\), we have
\[
f - \hat{f}^* + \varepsilon = q + \varepsilon (1 + \frac{u}{\varepsilon} + \varepsilon \left(\frac{u}{\varepsilon}\right)^{2r}) - \varepsilon c^{-1-2r}(u^{2r} + w) + \varepsilon^{1-2r} w \in P(g)[x] + I(h)[x].
\] (79)

Moreover, the degree of the right-hand side has an upper bound independent from \(\varepsilon\). This implies that there exists \(k \in \mathbb{N}\) such that for all \(\varepsilon > 0\), \(f - \hat{f}^* + \varepsilon \in P_k(g)[x] + I_k(h)[x]\). Then we for all \(\varepsilon > 0\), \(f^* - \varepsilon\) is a feasible solution of (57) of the value \(\rho_k(f,\Pi g, h)\). It gives \(\rho_k(f,\Pi g, h) \geq f^* - \varepsilon\), for all \(\varepsilon > 0\), and, in consequence, we get \(\rho_k(f,\Pi g, h) \geq \hat{f}^*\). Using the second statement of Lemma 19, we obtain that \(\rho_k(f,\Pi g, h) = \hat{f}^*\), yielding the result. \(\square\)

**Remark 4.** In Lemma 19, if \(m \geq 2\), and there exists \(q \in Q(g)[x]\) such that \(f - \hat{f}^* - q\) vanishes on \(V(h) \cap S(g)\), we are not sure that \(\rho_k(f,g,h) = \hat{f}^*\) for some \(k \in \mathbb{N}\).

As a consequence of Lemma 19, the following lemma is obtained:

**Lemma 26.** Let \(f,g_1,\ldots,g_m \in \mathbb{R}[x]\). Let \(\hat{f}^*\) be as in problem (53) with \(g = (g_1,\ldots,g_m)\). Let \(h_{FJ}^+\) be as in (63). If problem (53) has a global minimizer, then it holds that
\[
f^* = \min_{x,\lambda} f(x)
\]
\[s.t. \quad x \in S(g), (x,\lambda) \in V(h_{FJ}^+).
\] (80)

We present in the following two theorems the main application of Theorems 5 and 6 to polynomial optimization:
Theorem 7. Let \( f, g_1, \ldots, g_m \in \mathbb{R}[x] \). Let \( f^* \) be as in problem (3) with \( g = (g_1, \ldots, g_m) \). Let \( h_{F^+} \) be as in (14) and \( h_{F^+}^+ \) be as in (63). Assume that problem (3) has at least one global minimizer. Let \( \Pi g \) be as in (7). Then the following statements hold:

1. If \( f(C(g) \cap S(g)) \) is finite, there exists \( k \in \mathbb{N} \) such that \( \rho_k(f, \Pi g, h_{F^+}) = f^* \).

2. If \( f(C^+(g) \cap S(g)) \) is finite, there exists \( k \in \mathbb{N} \) such that \( \rho_k(f, \Pi g, h_{F^+}) = f^* \).

Remark 5. Since \( C^+(g) \cap S(g) \subset C^+(g) \), the second statement of Theorem 7 holds if \( f(C^+(g)) \) is finite.

The proof of Theorem 7, which relies on Theorems 4 and Lemma 25, is similar to the one of Theorem 3.

Theorem 8. Let \( f, g_1, \ldots, g_m \in \mathbb{R}[x] \). Let \( f^* \) be as in problem (3) with \( g = (g_1, \ldots, g_m) \). Let \( h_{F^+} \) be as in (63). Assume that problem (3) has at least one global minimizer and \( S(g) \) satisfies the Archimedean condition. If \( f(C^+(g)) \) is finite, there exists \( k \in \mathbb{N} \) such that \( \rho_k(f, g, h_{F^+}) = f^* \). Furthermore, if these exist \( R > 0 \) such that \( g_m = R - x_1^2 - \cdots - x_n^2 \), for \( k \in \mathbb{N} \) sufficiently large, the Slater condition holds for the SOS relaxation (58) of order \( k \) with \( h = h_{F^+} \) or \( h = h_{F^+}^+ \).

The proof of Theorem 8, which relies on Theorem 5 together with the fifth and sixth statements of Lemma 19, is similar to the one of Theorem 3.

Combining Lemmas 24 and 25, we obtain the following corollary:

Corollary 2. Let \( f, g_1, \ldots, g_m \in \mathbb{R}[x] \). Let \( f^* \) be as in problem (3) with \( g = (g_1, \ldots, g_m) \). Let \( h_{F^+} \) be as in (63). Assume that problem (3) has at least one global minimizer, each \( g_j \) is concave, and \( S(g) \) has a non-empty interior. Then \( S(g) \) is convex and there exists \( k \in \mathbb{N} \) such that \( \rho_k(f, \Pi g, h_{F^+}) = f^* \) with \( \Pi g \) being defined as in (7). Moreover, if \( S(g) \) satisfies the Archimedean condition, there exists \( k \in \mathbb{N} \) such that \( \rho_k(f, g, h_{F^+}) = f^* \).

### 5.5 Numerical examples

In this subsection we report numerical results produced by the SOS relaxations of the values \( \rho_k(f, g, 0) \) and \( \rho_k(f, g, h_{F^+}) \) for problem (3), where \( h_{F^+} \) is defined as in (63). We use the same notation as in Section 4.2.

Our test problem is taken from Example 16, namely \( n = 2, m = 3, f = x_1 - 1 \) and \( g = (g_1, g_2, g_3) = (x_1, x_2, (x_1 - 1)^3 - x_2) \). It is clear that \( f^* = 0 \) which is attained at the unique global minimizer \((1, 0)\) for problem (3). Moreover, the Karush–Kuhn–Tucker conditions do not hold at this minimizer. By using [19] Proposition 3.4, we get \( \rho_k(f, g, 0) < f^* \) for all \( k \in \mathbb{N} \). As shown in Example 16 there exists \( q \in Q(g)[x, \lambda] \) and \( f - q \) vanishes on \( V(h_{F^+}) \). From this, the sixth statement of Theorem 15 yields that \( \rho_d(f, g, h_{F^+}) = f^* \) for some \( d \in \mathbb{N} \). Thus we obtain \( \rho_k(f, g, 0) < f^* = \rho_d(f, g, h_{F^+}) \), for all \( k \in \mathbb{N} \). Since \( h_{F^+} \) has sign symmetry at \( \lambda \), we use TSSOS [28] to exploit this structure when computing \( \rho_k(f, g, h_{F^+}) \).

We display the numerical results in Table 2. It shows that the numerical value \(-0.00822\) of \( \rho_0(f, g, h_{F^+}) \) is the best lower bound on \( f^* \). It takes around 2 seconds to obtain this numerical value. It is worth pointing out that the
Table 2: Lower bounds on $f^* = 0$

| $k$ | $\rho_k(f, g, h_F)$ | value | time  | size                        |
|-----|---------------------|-------|-------|-----------------------------|
| 3   | $-0.38154$          | 0.09  | (21,186,216,19) |
| 4   | $-0.01199$          | 0.38  | (42,476,1056,54) |
| 5   | $-0.00822$          | 2.07  | (84,1110,4393,100) |

| $k$ | $\rho_k(f, g, 0)$ | value | time  | size                        |
|-----|-------------------|-------|-------|-----------------------------|
| 13  | $-0.02419$        | 1.27  | (105,378,1,4) |
| 14  | $-0.02254$        | 2.08  | (120,435,1,4) |
| 15  | $-0.02254$        | 3.13  | (136,496,1,4) |

| $k$ | $\rho_k(f, (g, b), 0)$ | value | time  | size                        |
|-----|------------------------|-------|-------|-----------------------------|
| 13  | $-0.01490$            | 1.68  | (105,378,1,5) |
| 14  | $-0.01305$            | 2.43  | (120,435,1,5) |
| 15  | $-0.01267$            | 3.03  | (136,496,1,5) |

standard SOS relaxations of the values $\rho_k(f, g, 0)$ and $\rho_k(f, (g, b), 0)$ cannot reach the bound $-0.00822$ in less than 3 seconds in spite of using the additional ball constraint $b = 2 - x^2 - x_2$. Another observation is that the size of the SOS relaxations of the value $\rho_k(f, g, h_F)$ grows faster than the ones of the values $\rho_k(f, g, 0)$ and $\rho_k(f, (g, b), 0)$ when $k$ increases.

6 Representations with denominators

We close the paper with the following Nichtnegativstellensätze:

**Theorem 9.** Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$ such that $f$ is non-negative on $S(g)$ with $g := (g_1, \ldots, g_m)$. Then there exists $q \in P(g)[x, \lambda]$ such that $\lambda \rho_k(f - q)$ vanishes on $V(h_F)$, where $\lambda := (\lambda_0, \ldots, \lambda_m)$ and $h_F$ is defined as in (14). Moreover, if $S(g)$ satisfies the Archimedean condition, we can take $q \in Q(g)[x, \lambda]$. The proof of Theorem 9 which has the same idea as Theorem 1 is postponed to below. Although the representations in Theorem 9 have prescribed denominators, we are not required to make any assumption on the image of the set of critical points under $f$.

The following example illustrates the representations stated in Theorem 9:

**Example 19.** Consider Example 4 where $V(h_F) = \{(0, t, 0, \pm 1) : t \in \mathbb{R}\}$ which gives $\lambda_0(f - 0) = 0$ on $V(h_F)$ since $\lambda_0 = 0$ on $V(h_F)$. Similar consideration applies to Example 14.

The following lemma is utilized to prove Theorem 9:

**Lemma 27.** Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$ such that $f(C(g))$ with $g := (g_1, \ldots, g_m)$ is finite. Let $h_F$ be defined as in (14). Let $W$ be a semi-algebraically path connected component of $V(h_F) \setminus \{\lambda_0 = 0\}$. Then $f$ is constant on $W$.

**Proof.** Choose two arbitrary points $(x^{(0)}, \lambda^{(0)}), (x^{(1)}, \lambda^{(1)})$ in $W$. We claim that $f(x^{(0)}) = f(x^{(1)})$. By assumption, there exists a continuous piecewise-differentiable path $\phi(\tau) = (x(\tau), \lambda(\tau))$, for $\tau \in [0, 1]$, lying inside $W$ such that
\(\phi(0) = (x^{(0)}, \lambda^{(0)})\) and \(\phi(1) = (x^{(1)}, \lambda^{(1)})\) (see, e.g., \cite{21} Theorem 1.8.1). The Lagrangian function \(L(x, \lambda)\) defined in \cite{20} is equal to \(f(x)\) on \(V(h_{FJ})\backslash \{\lambda_0 = 0\}\), which contains \(\phi([0, 1])\). By Lemma 8, the function \(L \circ \phi\) is semi-algebraic. Moreover, the function \(L \circ \phi\) is continuous since \(L\) and \(\phi\) are continuous. It implies that \(L \circ \phi\) is a continuous piecewise-differentiable function thanks to Lemma 14. Note that the function \(L \circ \phi\) has zero subgradient on \([0, 1]\). From Lemma 15, it follows that \(f(x(0)) = (L \circ \phi)(0) = (L \circ \phi)(1) = f(x(1))\). We now obtain \(f(x(0)) = f(x(1))\) and hence that \(f\) is constant on \(W\).

**Proof of Theorem 9**

**Proof.** Using Lemma 10, we decompose \(V(h_{FJ})\backslash \{\lambda_0 = 0\}\) into semi-algebraically path connected components: \(Z_1, \ldots, Z_r\). From this, Lemma 10 yields that \(f\) is constant on \(Z_i\), which implies that \(f(V(h_{FJ})\backslash \{\lambda_0 = 0\})\). We write

\[f(V(h_{FJ})\backslash \{\lambda_0 = 0\}) = \{t_1, \ldots, t_r\} \subset \mathbb{R},\]  

where \(t_i \neq t_j\) if \(i \neq j\). For \(j = 1, \ldots, r\), set \(W_j := V_c(h_{FJ}, f - t_j)\). Then \(W_j\) is a complex variety defined by finitely many polynomials in \(\mathbb{R}[x]\). We claim that \(W_1, \ldots, W_r\) are pairwise disjoint. Otherwise, let \((x, \lambda) \in W_i \cap W_j\) with \(i \neq j\). It implies that \(t_i = f(x) = t_j\), which is impossible. Let \(U = W_1 \cup \cdots \cup W_r\). We now prove that

\[V(h_{FJ})\backslash \{\lambda_0 = 0\} = (U \backslash \{\lambda_0 = 0\}) \cap \mathbb{R}^{n+m+1}.\]  

Let \((x, \lambda) \in V(h_{FJ})\backslash \{\lambda_0 = 0\}\cap \mathbb{R}^{n}\). By \cite{21}, there exists \(j \in \{1, \ldots, r\}\) such that \(f(x) = t_j\). It implies that \((x, \lambda) \in W_j \subset U\) and so we get \((x, \lambda) \in U \cap \mathbb{R}^{n+m+1}\). Thus \(V(h_{FJ})\backslash \{\lambda_0 = 0\} \subset (U \backslash \{\lambda_0 = 0\}) \cap \mathbb{R}^{n+m+1}\) since \((x, \lambda)\) is arbitrary. Conversely, suppose that \((x, \lambda) \in (U \backslash \{\lambda_0 = 0\}) \cap \mathbb{R}^{n+m+1}\). Then there is \(j \in \{1, \ldots, r\}\) such that \(x \in W_j\). It implies that \((x, \lambda) \in V(h_{FJ})\) by the definition of \(W_j\). Then \((x, \lambda) \in V(h_{FJ})\backslash \{\lambda_0 = 0\}\). Thus \((U \backslash \{\lambda_0 = 0\}) \cap \mathbb{R}^{n+m+1} \subset V(h_{FJ})\backslash \{\lambda_0 = 0\}\) since \((x, \lambda)\) is arbitrary.

Let \(b = 1 - \lambda_0^2 - \cdots - \lambda_r^2\). By the definition of \(U\), Lemma 12 shows that there exists \(p \in P(g, b)[x, \lambda]\) such that \(f - p\) vanishes on \(U \cap \mathbb{R}^{n+m+1}\). We write \(p\) as in \cite{20} for some \(\sigma, \psi \in S^2[x, \lambda]\). Let \(q = \sum_{\alpha \in \{0, 1\}^m} \sigma_\alpha g^\alpha \in P(g)[x, \lambda]\). Since \(b = 0\) on \(V(h_{FJ})\), \(f = p = q\) on \(V(h_{FJ})\backslash \{\lambda_0 = 0\}\) thanks to \cite{21}. Thus \(\lambda_0(f - q)\) vanishes on \(V(h_{FJ})\).

Assume that \(S(g)\) satisfies the Archimedean condition. Then there exists \(R > 0\) such that \(g_{m+1} = R - x_1^2 - \cdots - x_n^2 \in Q(g)[x]\). It implies that \(S(g, b)\) with \(b = 1 - \lambda_0^2 - \cdots - \lambda_r^2\) satisfies the Archimedean condition due to \cite{20}. By the definition of \(U\) Lemma 12 yields that there exists \(p \in Q(g, b)[x, \lambda]\) such that \(f - p\) vanishes on \(U \cap \mathbb{R}^{n+m+1}\). We write \(p\) as in \cite{20} for some \(\sigma_\alpha \in S^2[x, \lambda]\). Let \(q = \sigma_0 + \sum_{j=1}^m \sigma_j g_j \in Q(g)[x, \lambda]\). Since \(b = 0\) on \(V(h_{FJ})\), \(f = p = q\) on \(V(h_{FJ})\backslash \{\lambda_0 = 0\}\) thanks to \cite{21}. Thus \(\lambda_0(f - q)\) vanishes on \(V(h_{FJ})\).

**Acknowledgements.** The author was supported by the MESRI funding from EDMITT.
References

[1] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013.

[2] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[3] M. A. Bucero and B. Mourrain. Exact relaxation for polynomial optimization on semi-algebraic sets. *arXiv preprint arXiv:1307.6426*, 2013.

[4] J. Demmel, J. Nie, and V. Powers. Representations of positive polynomials on noncompact semialgebraic sets via KKT ideals. *Journal of pure and applied algebra*, 209(1):189–200, 2007.

[5] R. M. Freund. Optimality conditions for constrained optimization problems. *Massachusetts Institute of Technology (available at this link)*, 2004.

[6] D. Hilbert. Über die Darstellung Definiter Formen als Summe von Formen-quadrate. *Mathematische Annalen*, 32(3):342–350, 1888.

[7] T. Jacobi and A. Prestel. Distinguished representations of strictly positive polynomials. *Journal für die Reine und Angewandte Mathematik*, 2001.

[8] F. Jhon. Extremum problems with inequalities as side conditions, studies and essays, courant anniversary volume, 1948.

[9] W. Karush. Minima of functions of several variables with inequalities as side constraints. *M. Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago*, 1939.

[10] J.-L. Krivine. Anneaux préordonnés. *Journal d’analyse mathématique*, 12(1):307–326, 1964.

[11] H. Kuhn. A. W. TUCKER, Nonlinear Programming. In *2nd Berkeley Symposium*, pages 481–492, 1951.

[12] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on optimization*, 11(3):796–817, 2001.

[13] N. H. A. Mai, J.-B. Lasserre, V. Magron, and J. Wang. Exploiting constant trace property in large-scale polynomial optimization. *arXiv preprint arXiv:2012.08873*, 2020.

[14] M. Marshall. Representations of non-negative polynomials having finitely many zeros. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 15, pages 599–609, 2006.

[15] M. Marshall. Representations of non-negative polynomials, degree bounds and applications to optimization. *Canadian Journal of Mathematics*, 61(1):205–221, 2009.

[16] A. Mosek. The mosek optimization software. *Online at http://www. mosek. com*, 54(2-1):5, 2010.
[17] J. Nie. Discriminants and nonnegative polynomials. *Journal of Symbolic Computation*, 47(2):167–191, 2012.

[18] J. Nie. An exact Jacobian SDP relaxation for polynomial optimization. *Mathematical Programming*, 137(1):225–255, 2013.

[19] J. Nie. Optimality conditions and finite convergence of Lasserre’s hierarchy. *Mathematical programming*, 146(1-2):97–121, 2014.

[20] J. Nie, J. Demmel, and B. Sturmfels. Minimizing polynomials via sum of squares over the gradient ideal. *Mathematical programming*, 106(3):587–606, 2006.

[21] T. S. Pham and H. H. Vui. *Genericity in polynomial optimization*, volume 3. World Scientific, 2016.

[22] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.

[23] C. Scheiderer. Sums of squares of regular functions on real algebraic varieties. *Transactions of the American Mathematical Society*, 352(3):1039–1069, 2000.

[24] C. Scheiderer. Sums of squares on real algebraic curves. *Mathematische Zeitschrift*, 245(4):725–760, 2003.

[25] C. Scheiderer. Distinguished representations of non-negative polynomials. *Journal of Algebra*, 289(2):558–573, 2005.

[26] C. Scheiderer. Sums of squares on real algebraic surfaces. *Manuscripta mathematica*, 119(4):395–410, 2006.

[27] K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289(1):203–206, 1991.

[28] J. Wang, V. Magron, and J.-B. Lasserre. TSSOS: A Moment-SOS hierarchy that exploits term sparsity. *SIAM Journal on Optimization*, 31(1):30–58, 2021.