Math. Program., Ser. A (2014) 147:519–538
DOI 10.1007/s10107-013-0732-2

FULL LENGTH PAPER

A Frank–Wolfe type theorem for nondegenerate polynomial programs

Si Tiep Dinh · Huy Vui Ha · Tien Son Pham

Received: 24 January 2013 / Accepted: 20 November 2013 / Published online: 8 December 2013

Abstract In this paper, we study the existence of optimal solutions to a constrained polynomial optimization problem. More precisely, let f_0 and $f_1, \ldots, f_p \colon \mathbb{R}^n \to \mathbb{R}$ be convenient polynomial functions, and let $S := \{x \in \mathbb{R}^n : f_i(x) \leq 0, i = 1, \ldots, p\} \neq \emptyset$. Under the assumption that the map $(f_0, f_1, \ldots, f_p) \colon \mathbb{R}^n \to \mathbb{R}^{p+1}$ is non-degenerate at infinity, we show that if f_0 is bounded from below on S, then f_0 attains its infimum on S.

Keywords Existence of optimal solutions · Frank–Wolfe type theorem · Newton polyhedron · Nondegenerate polynomial programs

Mathematics Subject Classification (1991) 90C26 · 90C30

1 Introduction

Unboundedness and existence of solutions are important issues in optimization theory; the reader is invited to see [2] for an extensive survey. In this paper, we are interested in

Si Tiep Dinh and Huy Vui Ha: These authors were partially supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED) Grant Numbers 101.01-2011.44.
Tien Son Pham: This author was partially supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED) Grant Numbers 101.04-2013.07.

S. T. Dinh · H. V. Ha
Institute of Mathematics, 18, Hoang Quoc Viet Road, Cau Giay District, 10307 Hanoi, Vietnam
e-mail: dstiep@math.ac.vn

H. V. Ha
e-mail: hhvui@math.ac.vn

T. S. Pham (✉)
Department of Mathematics, University of Dalat, 1 Phu Dong Thien Vuong, Dalat, Vietnam
e-mail: sonpt@dlu.edu.vn

the question of existence of optimal solutions to a constrained polynomial optimization problem. More precisely, let f_0 and $f_1, \dots, f_p \colon \mathbb{R}^n \to \mathbb{R}$ be polynomial functions in the variable $x \in \mathbb{R}^n$. Denote by S the basic closed semi-algebraic set generated by $f_1, \dots, f_p$, i.e.,

$$S := \{x \in \mathbb{R}^n : \ f_1(x) \leq 0, \dots, f_p(x) \leq 0\},$$

and suppose throughout that S is nonempty. We consider the following constrained optimization problem

$$f^* := \inf f_0(x) \quad \text{such that} \quad x \in S \tag{1}$$

of minimizing the polynomial f_0 over S.

In the case when all $f_i, i = 0, \dots, p$, are *linear*, it is well known that the set of optimal solutions is nonempty, provided that the problem is bounded from below. In 1956, Frank and Wolfe [15] proved that if f_i remain *affine linear* functions for $i = 1, \dots, p$, and f_0 is an arbitrary *quadratic* polynomial, then the condition of f_0 being bounded from below on S implies that an optimal solution exists. If the statement holds for another class of polynomial functions $f_0, \dots, f_p$, we will speak of a *Frank–Wolfe type theorem.*

Many other authors generalized the Frank–Wolfe theorem to broader classes of functions. For example, Perold [40] generalized the Frank–Wolfe theorem for a class of non-quadratic objective functions and *linear* constraints. Andronov et al. [1] extended the Frank–Wolfe theorem to the case of a *cubic* polynomial objective function f_0 under linear constraints. Luo and Zhang [28] (see also [45]) also extended the Frank–Wolfe theorem to various classes of general convex or non-convex *quadratic* constraint systems. Belousov and Klatte [5] (see also [3,4,7,35]) showed that an optimal solution always exists if $f_0, f_1, \dots, f_p$ are *convex* polynomials of *arbitrary* degree. These results do not hold in general if we remove both assumptions on the convexity and on the degree (see, for example, [5,15,18,19,28,44]).

In this paper, we consider the class of polynomial maps which are *(Newton) non-degenerate at infinity.* This notion extends the definitions of non-degenerate for analytic functions, in the (local and at infinity) complex setting [23,24]. It is worth paying attention to the fact that non-degenerate at infinity polynomial maps have a number of remarkable properties which make them an attractive domain for various applications. Further, the class of polynomial maps (with fixed Newton polyhedra), which are non-degenerate at infinity, is generic in the sense that it is an open and dense semi-algebraic set.

The purpose of this paper is to establish a Frank–Wolfe type theorem for polynomial maps that are non-degenerate at infinity. Precisely, with the definitions in the next section, the main result of this paper is as follows.

Theorem 1.1 *Assume that the polynomial map $F := (f_0, f_1, \dots, f_p) \colon \mathbb{R}^n \to \mathbb{R}^{p+1}$ is convenient and non-degenerate at infinity. If f_0 is bounded from below on S, then f_0 attains its infimum on S.*

It should be emphasized that we do not require the polynomials f_i to be convex, and their degrees can be arbitrary. Although Theorem 1.1 does not imply the previous versions of the Frank–Wolfe theorem, however, by genericity of the condition of non-degeneracy at infinity, we conclude that Problem (1) has an optimal solution for almost all polynomial functions $f_0, f_1, \ldots, f_p$. In connection to this fact, we would like to mention that the second and third authors [20] have shown that almost every linear function, which is bounded from below on S, attains its infimum on S, and if, in addition, S is convex and compact, then there is a unique optimal solution, as shown recently by Bolte et al. [10].

As an application of Theorem 1.1, let us consider the problem of computing numerically the global infimum f^* of Problem (1). This is an NP-hard problem, even if f_0 is quadratic and the f_i are linear. For instance, the Maximum-Cut problem for graphs is of this form, and it is NP-hard [17]. Alternatively, Problem (1) contains *the partition problem* (see Example 4.1 below), which is known to be NP-complete [17]. A standard approach for solving Problem (1) is the hierarchy of semidefinite programming (SDP) relaxations proposed by Lasserre [25] (see also [37,38,43]). It is based on results about moment sequences and (the dual theory of) representations of nonnegative polynomials as sums of squares. For details about these methods and their applications, see [12,18–20,22,25–27,29,30,34,39,44].

In the papers [12,20,30,34], the authors have proposed semidefinite programming relaxations for finding the global infimum f^*, under the assumption that the objective function f_0 attains its infimum on the constrained set S. This assumption is non-trivial and the question of how to verify if a given polynomial on a given semi-algebraic set has this property is important and difficult (confer [34, Section 7]). Theorem 1.1 provides clearly a large class of polynomial optimization problems for which these methods can be applied.

The result presented in the paper, together with Hölder-type global error bound theorems in [21], suggests that the class of polynomial maps, which is non-degenerate at infinity, may offer an appropriate domain on which the machinery of polynomial optimization works with full efficiency.

The proof of the main theorem involves the theory of Newton polyhedra and semi-algebraic geometry, and relies heavily on Curve Selection Lemma at infinity (Lemma 3.3).

The paper is structured as follows. Section 2 introduces some basic notions and results of Newton polyhedra and non-degeneracy at infinity in the context of real polynomial maps. Section 3 contains some basic ingredients necessary for the proofs of the main theorem quoted above. Section 4 is devoted to prove Theorems 1.1. Finally, the proofs of the technical ingredients will be given in Sect. 5.

2 Newton polyhedra and non-degeneracy conditions

2.1 Newton polyhedra

Throughout this paper, $\mathbb{R}^n$ denotes the Euclidean space of dimension n, the canonical basis of $\mathbb{R}^n$ is denoted by $\{e_1, \ldots, e_n\}$. The corresponding inner product (resp., norm) in $\mathbb{R}^n$ is defined by $\langle x, y\rangle$ for any $x, y \in \mathbb{R}^n$ (resp., $\|x\| := \sqrt{\langle x, x\rangle}$ for any $x \in \mathbb{R}^n$).

Throughout the text, we consider a fixed coordinate system $x_1, \ldots, x_n \in \mathbb{R}^n$. Let $J \subset \{1, \ldots, n\}$, then we define

$$\mathbb{R}^J := \{x \in \mathbb{R}^n \ : \ x_j = 0, \text{for all } j \notin J\}.$$

We denote by $\mathbb{R}_+$ the set of non-negative real numbers. We also set $\mathbb{Z}_+ := \mathbb{R}_+ \cap \mathbb{Z}$. If $\kappa = (\kappa_1, \ldots, \kappa_n) \in \mathbb{Z}_+^n$, we denote by x^κ the monomial $x_1^{\kappa_1} \ldots x_n^{\kappa_n}$ and by $|\kappa|$ the sum $\kappa_1 + \cdots + \kappa_n$.

Definition 2.1 A subset $\Gamma \subset \mathbb{R}_+^n$ is said to be a *Newton polyhedron at infinity,* if there exists some finite subset $A \subset \mathbb{Z}_+^n$ such that Γ is equal to the convex hull in $\mathbb{R}^n$ of $A \cup \{0\}$. Hence we say that Γ is the Newton polyhedron at infinity determined by A and we write $\Gamma = \Gamma(A)$. We say that a Newton polyhedron at infinity $\Gamma \subset \mathbb{R}_+^n$ is *convenient* if it intersects each coordinate axis in a point different from the origin, that is, if for any $i \in \{1, \ldots, n\}$ there exists some integer $m_j > 0$ such that $m_j e_j \in \Gamma$.

Given a Newton polyhedron at infinity $\Gamma \subset \mathbb{R}_+^n$ and a vector $q \in \mathbb{R}^n$, we define

$$\begin{aligned} d(q, \Gamma) &:= \min\{\langle q, \kappa\rangle \ : \ \kappa \in \Gamma\}, \\ \Delta(q, \Gamma) &:= \{\kappa \in \Gamma \ : \ \langle q, \kappa\rangle = d(q, \Gamma)\}. \end{aligned}$$

We say that a subset Δ of Γ is a *face* of Γ if there exists a vector $q \in \mathbb{R}^n$ such that $\Delta = \Delta(q, \Gamma)$. The dimension of a face Δ is defined as the minimum of the dimensions of the affine subspaces containing Δ. The faces of Γ of dimension 0 are called the *vertices* of Γ. We denote by Γ_∞ the set of faces of Γ which do not contain the origin 0 in $\mathbb{R}^n$.

Let $\Gamma_1, \ldots, \Gamma_p$ be a collection of p Newton polyhedra at infinity in $\mathbb{R}_+^n$, for some $p \geq 1$. The *Minkowski sum* of $\Gamma_1, \ldots, \Gamma_p$ is defined as the set

$$\Gamma_1 + \cdots + \Gamma_p = \left\{\kappa^1 + \cdots + \kappa^p \ : \ \kappa^i \in \Gamma_i, \quad \text{for all i} = 1, \ldots, \text{p}\right\}.$$

By definition, $\Gamma_1 + \cdots + \Gamma_p$ is again a Newton polyhedron at infinity. Moreover, by applying the definitions given above, it is easy to check that

$$d(q, \Gamma_1 + \cdots + \Gamma_p) = d(q, \Gamma_1) + \cdots + d(q, \Gamma_p), \tag{2}$$

$$\Delta(q, \Gamma_1 + \cdots + \Gamma_p) = \Delta(q, \Gamma_1) + \cdots + \Delta(q, \Gamma_p), \tag{3}$$

for all $q \in \mathbb{R}^n$. The following result will be useful in the sequel.

Lemma 2.1 (i) *Assume that* Γ *is a convenient Newton polyhedron at infinity. Let* Δ *be a face of* Γ *and let* $q = (q_1, \ldots, q_n) \in \mathbb{R}^n$ *be such that* $\Delta = \Delta(q, \Gamma)$. *Then the following conditions are equivalent:*

(i1) $\Delta \in \Gamma_\infty$*;*
(i2) $d(q, \Gamma) < 0$*;*
(i3) $\min_{j=1,\ldots,n} q_j < 0$.

(ii) *Assume that* $\Gamma_1, \ldots, \Gamma_p$ *are some Newton polyhedra at infinity. Let* Δ *be a face of the Minkowski sum* $\Gamma := \Gamma_1 + \cdots + \Gamma_p$. *Then the following statements hold.*

(ii1) There exists a unique collection of faces $\Delta_1, \ldots, \Delta_p$ *of* $\Gamma_1, \ldots, \Gamma_p$, *respectively, such that*

$$\Delta = \Delta_1 + \cdots + \Delta_p.$$

(ii2) If $\Gamma_1, \ldots, \Gamma_p$ *are convenient, then* $\Gamma_\infty \subset \Gamma_{1,\infty} + \cdots + \Gamma_{p,\infty}$.

Proof (i) First of all, suppose that $\Delta \in \Gamma_\infty$. By definition, $\Delta(q, \Gamma) = \{\kappa \in \Gamma : \langle q, \kappa\rangle = d(q, \Gamma)\}$ and $\langle q, \kappa\rangle > d(q, \Gamma)$ for $\kappa \in \Gamma \backslash \Delta(q, \Gamma)$. Since $0 \notin \Delta = \Delta(q, \Gamma)$, it follows that $0 = \langle q, 0\rangle > d(q, \Gamma)$. Let $\kappa \in \Delta$, then $\langle q, \kappa\rangle = d(q, \Gamma) < 0$. Note that all the coordinates of κ are not negative, so at least one of the coordinates of q must be negative.

Now assume that $d(q, \Gamma) < 0$. By contradiction, suppose that $\Delta \notin \Gamma_\infty$. Hence $0 \in \Delta$, so by definition, $d(q, \Gamma) = 0$. This is a contradiction.

Finally, assume that $q_{j_*} := \min_{j=1,\ldots,n} q_j < 0$. Since Γ is convenient, there exists an integer $m_{j_*} > 0$ such that $m_{j_*} e_{j_*} \in \Gamma$. So $\langle q, e_{j_*}\rangle = q_{j_*} \cdot m_{j_*} < 0$, which implies that $d(q, \Gamma) < 0$.

(ii1) By definition and the relation (3), there exists a vector $q \in \mathbb{R}^n$ such that

$$\Delta = \Delta(q, \Gamma_1 + \cdots + \Gamma_p) = \Delta(q, \Gamma_1) + \cdots + \Delta(q, \Gamma_p).$$

It is clear that $\Delta_i := \Delta(q, \Gamma_i)$ is a face of Γ_i for $i = 1, \ldots, p$.

(ii2) Let $\Delta \in \Gamma_\infty$ and let $q = (q_1, \ldots, q_n) \in \mathbb{R}^n$ such that $\Delta = \Delta(q, \Gamma)$. Then by (i), $\min_{j=1,\ldots,n} q_j < 0$. By the relation (3), $\Delta = \Delta(q, \Gamma_1) + \cdots + \Delta(q, \Gamma_p)$. We need to prove that $\Delta(q, \Gamma_i) \in \Gamma_{i,\infty}$ for $i = 1, \ldots, p$. By contradiction, suppose that there exists an index i_0 such that $\Delta(q, \Gamma_{i_0}) \notin \Gamma_{i_0,\infty}$. Hence $0 \in \Delta(q, \Gamma_{i_0})$, so by definition, $d(q, \Gamma_{i_0}) = 0$. On the other side, by (i) and by the fact that $\min_{j=1,\ldots,n} q_j < 0$, it follows that $d(q, \Gamma_{i_0}) < 0$. This contradiction ends the proof of the lemma. □

2.2 Non-degeneracy at infinity

In [23] (see also [24]), Khovanskii introduced a condition of non-degeneracy for complex analytic maps $F : (\mathbb{C}^n, 0) \to (\mathbb{Z}^p, 0)$ in terms of Newton polyhedra of the component functions of F. This notion has been applied extensively to the study of several questions concerning isolated complete intersection singularities (see, for instance, [8,11,16,36]). We will apply this condition for real polynomial maps. First we need to introduce some notations and definitions.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a polynomial function. Suppose that f is written as $f = \sum_\kappa a_\kappa x^\kappa$. Then the support of f, denoted by $\mathrm{supp}(f)$, is defined as the set of those $\kappa \in \mathbb{Z}^n_+$ such that $a_\kappa \neq 0$. We denote the set $\Gamma(\mathrm{supp}(f))$ by $\Gamma(f)$. This set is called the *Newton polyhedron at infinity* of f. The polynomial f is said to be *convenient* if $\Gamma(f)$ is convenient. If $f \equiv 0$, then we set $\Gamma(f) = \emptyset$. Note that, if f is convenient, then for each nonempty subset J of $\{1, \ldots, n\}$, we have $\Gamma(f) \cap \mathbb{R}^J = \Gamma(f|_{\mathbb{R}^J})$. The

Newton boundary at infinity of f, denoted by $\Gamma_\infty(f)$, is defined as the union of the faces of $\Gamma(f)$ which do not contain the origin 0 in $\mathbb{R}^n$. For each face Δ of $\Gamma_\infty(f)$, we define the *principal part of f at infinity with respect to* Δ, denoted by f_Δ, as the sum of the terms $a_\kappa x^\kappa$ such that $\kappa \in \Delta$.

Let $F := (f_1, \ldots, f_p) : \mathbb{R}^n \to \mathbb{R}^p$, $1 \le p \le n$, be a polynomial map. We say that F is *convenient* if all its components f_i are convenient. Let $\Gamma(F)$ denote the Minkowski sum $\Gamma(f_1) + \cdots + \Gamma(f_p)$, and we denote by $\Gamma_\infty(F)$ the union of faces of $\Gamma(F)$ which do not contain the origin 0 in $\mathbb{R}^n$. Let Δ be a face of the $\Gamma(F)$. According to Lemma 2.1, we have the unique decomposition $\Delta = \Delta_1 + \cdots + \Delta_p$, where Δ_i is a face of $\Gamma(f_i)$, for all $i = 1, \ldots, p$. We denote by F_Δ the polynomial map $(f_{1,\Delta_1}, \ldots, f_{p,\Delta_p}) : \mathbb{R}^n \to \mathbb{R}^p$.

Definition 2.2 We say that F is *Khovanskii non-degenerate at infinity* if and only if for any face Δ of $\Gamma_\infty(F)$ and for all $x \in (\mathbb{R}\backslash\{0\})^n \cap F_\Delta^{-1}(0)$, we have

$$\operatorname{rank} \begin{pmatrix} x_1 \frac{\partial f_{1,\Delta_1}}{\partial x_1}(x) & \cdots & x_n \frac{\partial f_{1,\Delta_1}}{\partial x_n}(x) \\ \vdots & \cdots & \vdots \\ x_1 \frac{\partial f_{p,\Delta_p}}{\partial x_1}(x) & \cdots & x_n \frac{\partial f_{p,\Delta_p}}{\partial x_n}(x) \end{pmatrix} = p.$$

Definition 2.3 We say that F is *non-degenerate at infinity* if and only if for any face Δ of $\Gamma_\infty(F)$ and for all $x \in (\mathbb{R}\backslash\{0\})^n$, we have

$$\operatorname{rank} \begin{pmatrix} x_1 \frac{\partial f_{1,\Delta_1}}{\partial x_1}(x) & \cdots & x_n \frac{\partial f_{1,\Delta_1}}{\partial x_n}(x) & f_{1,\Delta_1}(x) & 0 & \cdots & 0 \\ x_1 \frac{\partial f_{2,\Delta_2}}{\partial x_1}(x) & \cdots & x_n \frac{\partial f_{2,\Delta_2}}{\partial x_n}(x) & 0 & f_{2,\Delta_2}(x) & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 \frac{\partial f_{p,\Delta_p}}{\partial x_1}(x) & \cdots & x_n \frac{\partial f_{p,\Delta_p}}{\partial x_n}(x) & 0 & 0 & \cdots & f_{p,\Delta_p}(x) \end{pmatrix} = p.$$

For each subset $I := \{i_1, \ldots, i_q\} \subset \{1, \ldots, p\}$, we define the polynomial map $F_I : \mathbb{R}^n \to \mathbb{R}^q$ by $F_I(x) = (f_{i_1}(x), \ldots, f_{i_q}(x))$.

The connection between non-degeneracy conditions is given by the following result.

Lemma 2.2 *Let $F := (f_1, \ldots, f_p) : \mathbb{R}^n \to \mathbb{R}^p$, $1 \le p \le n$, be a polynomial map. Then F is non-degenerate at infinity if and only if F_I is Khovanskii non-degenerate at infinity, for all subset $I \subset \{1, \ldots, p\}$.*

Proof The statement is straightforward from the definitions. □

Remark 2.1 The above lemma implies that if F is non-degenerate at infinity then F is Khovanskii non-degenerate at infinity. The converse does not hold. However, both conditions constitute generic conditions in the sense that the class of polynomial maps (with fixed Newton polyhedra), which is non-degenerate at infinity, is an open and dense semi-algebraic set. The proof of this genericity is skipped since it is quite long and is not the goal of the paper; the interested reader may find a proof in [13] (see also [24,36]).

3 Some technical ingredients

In this section, we present some important ingredients necessary to prove Theorem 1.1. The reader can find the proofs of these results in Sect. 5.

3.1 Semi-algebraic geometry

In this subsection, we recall some notions and results of semi-algebraic geometry, which can be found in [6,9,14].

Definition 3.1 (i) A subset of $\mathbb{R}^n$ is called *semi-algebraic* if it is a finite union of sets of the form

$$\left\{x \in \mathbb{R}^n \ : \ f_i(x) = 0, \quad i = 1, \ldots, k; \ f_i(x) > 0, \quad i = k+1, \ldots, p\right\},$$

where all f_i are polynomials.

(ii) Let $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^p$ be semi-algebraic sets. A map $F\colon A \to B$ is said to be *semi-algebraic* if its graph

$$\{(x, y) \in A \times B \ : \ y = F(x)\}$$

is a semi-algebraic subset in $\mathbb{R}^n \times \mathbb{R}^p$.

Semi-algebraic sets and functions have a number of remarkable properties as follows.

- The class of semi-algebraic sets is closed with respect to Boolean operators; a Cartesian product of semi-algebraic sets is a semi-algebraic set;
- The closure and the interior of a semi-algebraic set is a semi-algebraic set;
- A composition of semi-algebraic maps is a semi-algebraic map.

A major fact concerning the class of semi-algebraic sets is its stability under linear projections (see, for example, [6, Theorem 2.3.4], [9, Theorem 2.2.1 and Proposition 2.2.7]).

Theorem 3.1 (Tarski–Seidenberg Theorem) *Let $F\colon A \to B$ be a semi-algebraic map. Then the image $F(A) \subset B$ is a semi-algebraic subset.*

Remark 3.1 As an immediate consequence of Tarski–Seidenberg Theorem, we get the semi-algebraicity of any set of the form $\{x \in A \ : \ \exists y \in B, (x, y) \in S\}$, provided that A, B, and S are semi-algebraic sets in the corresponding spaces. It follows also that $\{x \in A \ : \ \forall y \in B, (x, y) \in S\}$ is a semi-algebraic set as its complement is the union of the complement of A and the set $\{x \in A \ : \ \exists y \in B, (x, y) \notin S\}$. Thus, if we have a finite collection of semi-algebraic sets, then any set obtained from them by a finite chain of quantifiers is also semi-algebraic.

In the sequel, we will need the following results (see, for example, [6,9,14,31]).

Lemma 3.1 (Monotonicity Lemma) *Let $a < b$ in $\mathbb{R}$. If $f\colon (a, b) \to \mathbb{R}$ is a semi-algebraic function, then there is a partition $t_0 := a < t_1 < \cdots < t_{l+1} := b$ of (a, b) such that $f|_{(t_i, t_{i+1})}$ is C^1, and either constant or strictly monotone, for $i \in \{0, \ldots, l\}$.*

Lemma 3.2 (Growth Dichotomy Lemma) *Let $f\colon (0,\epsilon) \to \mathbb{R}$ be a semi-algebraic function with $f(t) \neq 0$ for all $t \in (0,\epsilon)$. Then there exist some constants $c \neq 0$ and $q \in \mathbb{Q}$ such that $f(t) = ct^q + o(t^q)$ as $t \to 0^+$.*

Let us give a version of Curve Selection Lemma which will be used frequently in the paper. Milnor [32] has proved this lemma at points of the closure of a semi-algebraic set. Némethi and Zaharia [33] showed how to extend the result at infinity at some fiber of a polynomial map. We give here a more general statement, and for the sake of completeness we include a proof of this fact in Sect. 5.

Lemma 3.3 (Curve Selection Lemma at infinity) *Let $A \subset \mathbb{R}^n$ be a semi-algebraic set, and let $F := (f_1,\ldots,f_p)\colon \mathbb{R}^n \to \mathbb{R}^p$ be a semi-algebraic map. Assume that there exists a sequence $x^k \in A$ such that $\lim_{k\to\infty} \|x^k\| = \infty$ and $\lim_{k\to\infty} F(x^k) = y \in (\overline{\mathbb{R}})^p$, where $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$. Then there exists a smooth semi-algebraic curve $\varphi\colon (0,\epsilon) \to \mathbb{R}^n$ such that $\varphi(t) \in A$ for all $t \in (0,\epsilon)$, $\lim_{t\to 0} \|\varphi(t)\| = \infty$, and $\lim_{t\to 0} F(\varphi(t)) = y$.*

3.2 The set of asymptotic critical values

Let $F = (f_1,\ldots,f_p)\colon \mathbb{R}^n \to \mathbb{R}^p$ be a C^1-map, and define the Rabier function $\nu_F\colon \mathbb{R}^n \to \mathbb{R}$ by

$$\nu_F(x) := \min_{\sum_{i=1}^p |\lambda_i| = 1} \left\| \sum_{i=1}^p \lambda_i \nabla f_i(x) \right\|.$$

Remark 3.2 (i) By definition, $\nu_F(x) = 0$ if and only if the gradient vectors $\nabla f_1(x),\ldots,\nabla f_p(x)$ are linearly dependent.
(ii) If the map F is semi-algebraic then so is ν_F.

Definition 3.2 [41] We define the set of *asymptotic critical values* of F as

$$\widetilde{K}_\infty(F) := \Big\{ y \in \mathbb{R}^p : \exists \{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n \text{such that} \\ \lim_{k\to\infty} \|x^k\| = \infty, \quad \lim_{k\to\infty} F\left(x^k\right) = y, \quad \text{and} \quad \lim_{k\to\infty} \nu_F\left(x^k\right) = 0 \Big\}.$$

Clearly the set $\widetilde{K}_\infty(F)$ is closed, and $\widetilde{K}_\infty(F) = \emptyset$ if F is a proper map in the sense that

$$\lim_{\|x\|\to\infty} \|F(x)\| = \infty.$$

The following result will be useful for our later analysis. We include its proof in Sect. 5.

Theorem 3.2 *Let $F = (f_1,\ldots,f_p)\colon \mathbb{R}^n \to \mathbb{R}^p$, $1 \le p \le n$, be a convenient polynomial map. Suppose that F is Khovanskii non-degenerate at infinity. Then $\widetilde{K}_\infty(F) = \emptyset$.*

Corollary 3.1 *Let* $F := (f_1, \ldots, f_p)\colon \mathbb{R}^n \to \mathbb{R}^p, 1 \leq p \leq n$, *be a convenient polynomial map. If* F *is non-degenerate at infinity then* $\widetilde{\widetilde{K}}(F_I) = \emptyset$ *for all nonempty subset* $I \subset \{1, \ldots, p\}$.

Proof The statement follows immediately from Lemma 2.2 and Theorem 3.2. □

3.3 Regularity at infinity

Let us consider the optimization problem (1) from the introduction:

$$f^* := \inf f_0(x) \quad \text{such that} \quad x \in S := \{x \in \mathbb{R}^n \ : \ f_1(x) \leq 0, \ldots, f_p(x) \leq 0\}.$$

As is known that most numerical optimization methods targeting local (including global) minimizers are often based on one optimality condition which is the Karush–Kuhn–Tucker (KKT) system. Sometimes the KKT system fails to hold at some minimizers. Hence, we usually make an assumption called a *constraint qualification* to ensure that the KKT system holds. Since the restriction $f_0|_S$ may attain its infimum "at infinity", a constraint qualification "at infinity" is needed and defined as follows.

Definition 3.3 For each $x \in S := \{x \in \mathbb{R}^n \ : \ f_1(x) \leq 0, \ldots, f_p(x) \leq 0\}$, let $I(x)$ be the set of indices i for which f_i vanishes at x. The set S is called *regular at infinity* if there exists a real number $R_0 > 0$ such that for each $x \in S$, $\|x\| \geq R_0$, the gradient vectors $\nabla f_i(x)$, $i \in I(x)$, are linearly independent.

The following lemmas follow from Curve Selection Lemma at infinity. The reader may find the proofs in Sect. 5.

Lemma 3.4 *Suppose that the closed semi-algebraic set*

$$S := \{x \in \mathbb{R}^n \ : \ f_1(x) \leq 0, \ldots, f_p(x) \leq 0\}$$

is unbounded and regular at infinity. Then there exists a real number $R_0 > 0$ *such that for all* $R \geq R_0$, *the set*

$$S_R := \left\{x \in S \ : \ \|x\|^2 = R^2\right\}$$

is a nonempty compact set. Moreover S_R *is regular, i.e., for each* $x \in S_R$, *the vectors* x *and* $\nabla f_i(x)$, $i \in I(x)$, *are linearly independent.*

Lemma 3.5 *Assume that the polynomial map* $F := (f_1, \ldots, f_p)\colon \mathbb{R}^n \to \mathbb{R}^p$ *is convenient and non-degenerate at infinity. Then the set* $S := \{x \in \mathbb{R}^n \ : \ f_1(x) \leq 0, \ldots, f_p(x) \leq 0\}$ *is regular at infinity.*

4 Proof of the main result

Now we are ready to prove the Frank–Wolfe type Theorem 1.1. The intuition behind the proof is as follows. By assumption, the restriction $f_0|_S$ is bounded from below, so it attains its infimum at some points in S or "at infinity" (see, for example, [18,19,44]). If $f_0|_S$ attains its infimum at infinity, it means that $f_0|_S$ has "singularities at infinity", so this possibility is eliminated by the condition of non-degeneracy at infinity (see Theorem 3.2).

Proof of Theorem 1.1 First of all, we prove that f_0 is coercive on S in the sense that

$$\lim_{r\to\infty} \min_{x\in S, \|x\|^2=r^2} f_0(x) = +\infty.$$

Suppose that it is not so; i.e., there exists a sequence $\{x^k\}_{k\in\mathbb{N}} \subset S$ such that

(a1) $\lim_{k\to\infty} \|x^k\| = \infty$, $\lim_{k\to\infty} f_0(x^k) = y \in \mathbb{R}$; and
(a2) x^k is a solution of the following constrained polynomial optimization problem:

$$\min_{x\in S,\ \|x\|^2=r_k^2} f_0(x),$$

where $\{r_k\}_{k\in\mathbb{N}}$ is strictly increasing with $r_k \geq k$.

By Lemma 3.5, the set S is regular at infinity, so the set $S_k := S \cap \{\|x\|^2 = k^2\}$, $k \gg 1$, is regular, in view of Lemma 3.4. It follows from Lagrange's multipliers theorem that there exist some real numbers λ_i^k, $i = 1, \ldots, p$, and μ_k such that

(a3) $\nabla f_0(x^k) + \sum_{i=1}^p \lambda_i^k \nabla f_i(x^k) + \mu_k x^k = 0$; and
(a4) $\lambda_i^k f_i(x^k) = 0$ for $i = 1, \ldots, p$.

Let

$$A := \Big\{(x, \lambda_0, \ldots, \lambda_p, \mu) \in \mathbb{R}^n \times \mathbb{R}^{p+2} \ | \ x \in S,\ \lambda_0 > 0,\ \|(\lambda_0, \ldots, \lambda_p, \mu)\| = 1,$$
$$\lambda_0 \nabla f_0(x) + \sum_{i=1}^p \lambda_i \nabla f_i(x) + \mu x = 0,$$
$$\lambda_i f_i(x) = 0, \text{ for } i = 1, \ldots, p\Big\}.$$

Then A is a semi-algebraic set and the sequence $(x^k, 1, \lambda_1^k, \ldots, \lambda_p^k, \mu_k) \in A$ tends to infinity as $k \to \infty$. By applying Curve Selection Lemma at infinity (Lemma 3.3) for the semi-algebraic function $A \to \mathbb{R}$, $(x, \lambda_0, \ldots, \lambda_p, \mu) \mapsto f_0(x)$, we get a smooth semi-algebraic curve

$$(\varphi, \lambda_0, \ldots, \lambda_p, \mu)\colon (0, \epsilon) \to \mathbb{R}^n \times \mathbb{R}^{p+2}, \quad t \mapsto (\varphi(t), \lambda_0(t), \ldots, \lambda_p(t), \mu(t)),$$

satisfying the following conditions

(a5) $\varphi(t) \in S$, $\lambda_0(t) > 0$ for $t \in (0, \epsilon)$, and $\|(\lambda_0(t), \ldots, \lambda_p(t), \mu(t))\| \equiv 1$;
(a6) $\lim_{t\to 0} \|\varphi(t)\| = \infty$ and $\lim_{t\to 0} f_0(\varphi(t)) = y$;
(a7) $\lambda_0(t)\nabla f_0(\varphi(t)) + \sum_{i=1}^p \lambda_i(t)\nabla f_i(\varphi(t)) + \mu(t)\varphi(t) \equiv 0$; and

(a8) $\lambda_i(t) f_i(\varphi(t)) \equiv 0$, for $i = 1, \ldots, p$.

Since the (smooth) functions λ_i and $f_i \circ \varphi$ are semi-algebraic, for $\epsilon > 0$ small enough, these functions are either constant or strictly monotone (see Monotonicity Lemma 3.1). Then, by Condition (a8), we can see that either $\lambda_i(t) \equiv 0$ or $f_i \circ \varphi(t) \equiv 0$; in particular,

$$\lambda_i(t) \frac{d}{dt}(f_i \circ \varphi)(t) \equiv 0, \quad i = 1, \ldots, p.$$

Replacing $\lambda_1(t), \ldots, \lambda_p(t)$, and $\mu(t)$ by $\lambda_1(t)/\lambda_0(t), \ldots, \lambda_p(t)/\lambda_0(t)$, and $\mu(t)/\lambda_0(t)$, respectively, we may assume that $\lambda_0(t) \equiv 1$. Let $I := \{i \in \{1, \ldots, p\} : f_i \circ \varphi(t) \equiv 0\}$. Then $\lambda_i(t) \equiv 0$ for $i \notin I$. It follows from Condition (a7) that

$$\begin{aligned}
\frac{\mu(t)}{2} \frac{d\|\varphi(t)\|^2}{dt} &= \mu(t) \left\langle \varphi(t), \frac{d\varphi}{dt} \right\rangle \\
&= -\left\langle \nabla f_0(\varphi(t)), \frac{d\varphi}{dt} \right\rangle - \sum_{i \in I} \lambda_i(t) \left\langle \nabla f_i(\varphi(t)), \frac{d\varphi}{dt} \right\rangle \\
&= -\frac{d}{dt}(f_0 \circ \varphi)(t) - \sum_{i \in I} \lambda_i(t) \frac{d}{dt}(f_i \circ \varphi)(t) = -\frac{d}{dt}(f_0 \circ \varphi)(t).
\end{aligned}$$

Therefore, by Condition (a7) again,

$$\left| \frac{d}{dt}(f_0 \circ \varphi)(t) \right| = \left| \frac{\mu(t)}{2} \frac{d\|\varphi(t)\|^2}{dt} \right| \tag{4}$$

$$= \frac{\|\nabla f_0(\varphi(t)) + \sum_{i \in I} \lambda_i(t) \nabla f_i(\varphi(t))\|}{2\|\varphi(t)\|} \left| \frac{d\|\varphi(t)\|^2}{dt} \right|. \tag{5}$$

Note that $f_0 \circ \varphi(t) \not\equiv 0$ since otherwise we have $\mu(t) \equiv 0$, and by Condition (a7), the constrained set S is not regular at infinity, which is a contradiction. Now, by Growth Dichotomy Lemma 3.2, we may write

$$\begin{aligned}
\|\varphi(t)\| &= c_1 t^{\alpha} + \text{higher order terms in } t, \\
f_0(\varphi(t)) &= c_2 t^{\beta} + \text{higher order terms in } t,
\end{aligned}$$

where $c_1 \neq 0$, $c_2 \neq 0$. By Condition (a6), $\alpha < 0$, $\beta \geq 0$. Then it follows from the relations (4) and (5) that

$$\left\| \nabla f_0(\varphi(t)) + \sum_{i \in I} \lambda_i(t) \nabla f_i(\varphi(t)) \right\| = c t^{\beta - \alpha} + \text{higher order terms in } t,$$

for some constant $c \neq 0$. Consequently, we get

$$\lim_{t\to 0}\left\|\nabla f_0(\varphi(t)) + \sum_{i\in I}\lambda_i(t)\nabla f_i(\varphi(t))\right\| = 0.$$

Hence, $\lim_{t\to 0}\nu_{F_I}(\varphi(t)) = 0$. Therefore, if we write $I = \{i_1, \ldots, i_q\} \subset \{1, \ldots, p\}$, then $(y, 0, \ldots, 0) \in \widetilde{K}(f_0, f_{i_1}, \ldots, f_{i_q})$, which contradicts Corollary 3.1. So we have proved that

$$\lim_{r\to\infty}\ \min_{x\in S,\|x\|^2=r^2} f_0(x) = +\infty.$$

Set

$$l_r := \min_{x\in S,\|x\|^2=r^2} f_0(x).$$

Let $r_0 \geq 0$ be such that $\{x \in S : \|x\|^2 = r_0^2\} \neq \emptyset$. Since f_0 is coercive on S, there exists $r_1 > r_0$ such that $l_r > l_{r_0} \geq \inf_{x\in S} f_0(x)$ for $r > r_1$. Thus

$$\inf_{x\in S} f_0(x) = \inf_{r\geq 0}\min_{x\in S,\|x\|^2=r^2} f_0(x) = \inf_{r\geq 0} l_r = \inf_{r_1\geq r\geq 0} l_r = \inf_{x\in S, r_1^2\geq\|x\|^2} f_0(x).$$

It is clear that the set $\{x \in S : r_1^2 \geq \|x\|^2\}$ is compact, so f_0 attains its infimum on S. The Theorem 1.1 is proved. □

Remark 4.1 The assumption in the Frank–Wolfe type Theorem that the polynomial map $(f_0, f_1, \ldots, f_p)$ are convenient cannot be removed. A counterexample is $f_0(x_1, x_2) := x_1^2 + (x_1x_2 - 1)^2$ and $S := \mathbb{R}^2$. It is easy to check that f_0 is non-degenerate at infinity. However, f_0 has 0 as unattainable infimum.

Remark 4.2 Let $F := (f_0, \ldots, f_p)\colon \mathbb{R}^n \to \mathbb{R}^{p+1}$ satisfy the assumptions of Theorem 1.1 and let M be a real number such that $M \geq f_0(x^0)$ for some $x^0 \in S$. Then the restriction $f_0|_S$ is coercive, and so the semi-algebraic set

$$S_M := \left\{x \in \mathbb{R}^n\ :\ M - f_0(x) \geq 0, f_1(x) \leq 0, \ldots, f_p(x) \leq 0\right\}$$

is compact. Moreover, it is clear that

$$\inf_{x\in S} f_0(x) = \inf_{x\in S_M} f_0(x).$$

Therefore we may construct a sequence of semidefinite programmings whose optimal values tend to the minimum of f_0 on S. Indeed it suffices to replace S by S_M and apply the associated standard hierarchy of semidefinite relaxations defined for the compact case [25].

In the rest of this section we give some illustrative examples of Theorem 1.1.

Example 4.1 Let us consider the problem of deciding whether an integer sequence $a_1, \ldots, a_n$ can be partitioned, which is known to be NP-complete [17], with $a_1, \ldots, a_n$ being partitionable if there exists $x \in \{\pm 1\}^n$ such that $\sum_{j=1}^n a_j x_j = 0$. Equivalently, the sequence can be partitioned if the infimum f^* of the polynomial $f_0 := (\sum_{j=1}^n a_j x_j)^2 + \sum_{j=1}^n (x_j^2 - 1)^2$ on $\mathbb{R}^n$ is equal to 0, in this case, a global minimizer is ± 1-valued and thus provides a partition of the sequence.

By the definition, the Newton polyhedron at infinity $\Gamma(f_0)$ of f_0 is the convex hull of the origin and the vertices $(4, 0, \ldots, 0), \ldots, (0, \ldots, 4)$, and hence is convenient. For any face Δ of $\Gamma_\infty(f_0)$, the principal part of f_0 at infinity with respect to Δ is of the form

$$f_{0,\Delta} = \sum_{j \in J} x_j^4,$$

for some nonempty subset $J \subset \{1, \ldots, n\}$. It is clear that

$$\operatorname{rank}\left(x_1 \frac{\partial f_{0,\Delta}}{\partial x_1}(x), \ldots, x_n \frac{\partial f_{0,\Delta}}{\partial x_n}(x) \right) = 1$$

for all $x \in (\mathbb{R} \backslash \{0\})^n$. So the condition of non-degeneracy at infinity is satisfied. In view of Theorem 1.1, f_0 attains its infimum on $\mathbb{R}^n$. Further, thanks to Remark 4.2, the global infimum f^* can be approximated as close as desired by solving a hierarchy of semidefinite programmings defined for the compact case [25]. Confer [22, Example 10].

Example 4.2 Let $n = 3$ and consider the polynomial

$$f_0(x, y, z) := x^8 + y^8 + z^8 + M(x, y, z),$$

where $M(x, y, z) := x^4 y^2 + x^2 y^4 + z^6 - 3x^2 y^2 z^2$ is the Motzkin polynomial, which is nonnegative but not sum of squares [29,42]. Set $f_1(x, y, z) := 1 - x^2 - 2y^2 - 3z^2$.

The Newton polyhedra at infinity of f_0 and f_1 are, respectively, the tetrahedra

$$\Gamma(f_0) = \Gamma\{(8, 0, 0), (0, 8, 0), (0, 0, 8)\},$$
$$\Gamma(f_1) = \Gamma\{(2, 0, 0), (0, 2, 0), (0, 0, 2)\}.$$

Hence, f_0 and f_1 are convenient. Let us check that the polynomial map $F := (f_0, f_1)$ is non-degenerate at infinity. The Minkowski sum

$$\Gamma(F) = \Gamma(f_0) + \Gamma(f_1) = \Gamma\{(10, 0, 0), (0, 10, 0), (0, 0, 10)\}$$

is again a tetrahedron. For simplicity, denote the convex hull of a set of points $a_1, \ldots, a_m \in \mathbb{R}^n$ by $[a_1, \ldots, a_m]$. Then $\Gamma_\infty(F)$ has seven faces which are

$$\begin{aligned}\Delta^1 &:= [(10, 0, 0), (0, 10, 0), (0, 0, 10)] = [(8, 0, 0), (0, 8, 0), (0, 0, 8)] \\ &\quad + [(2, 0, 0), (0, 2, 0), (0, 0, 2)],\end{aligned}$$

$$\begin{aligned}
\Delta^2 &:= [(10,0,0),(0,10,0)] = [(8,0,0),(0,8,0)] + [(2,0,0),(0,2,0)],\\
\Delta^3 &:= [(10,0,0),(0,0,10)] = [(8,0,0),(0,0,8)] + [(2,0,0),(0,0,2)],\\
\Delta^4 &:= [(0,10,0),(0,0,10)] = [(0,8,0),(0,0,8)] + [(0,2,0),(0,0,2)],\\
\Delta^5 &:= [(10,0,0)] = [(8,0,0)] + [(2,0,0)],\\
\Delta^6 &:= [(0,10,0)] = [(0,8,0)] + [(0,2,0)],\\
\Delta^7 &:= [(0,0,10)] = [(0,0,8)] + [(0,0,2)].
\end{aligned}$$

It is clear that the following corresponding matrices have rank 2 on $(\mathbb{R}\backslash\{0\})^3$:

$$\begin{aligned}
A_{\Delta^1} &:= \begin{pmatrix} 8x^8 & 8y^8 & 8z^8 & x^8+y^8+z^8 & 0 \\ -2x^2 & -4y^2 & -6z^2 & 0 & -x^2-2y^2-3z^2 \end{pmatrix},\\
A_{\Delta^2} &:= \begin{pmatrix} 8x^8 & 8y^8 & 0 & x^8+y^8 & 0 \\ -2x^2 & -4y^2 & 0 & 0 & -x^2-2y^2 \end{pmatrix},\\
A_{\Delta^3} &:= \begin{pmatrix} 8x^8 & 0 & 8z^8 & x^8+z^8 & 0 \\ -2x^2 & 0 & -6z^2 & 0 & -x^2-3z^2 \end{pmatrix},\\
A_{\Delta^4} &:= \begin{pmatrix} 0 & 8y^8 & 8z^8 & y^8+z^8 & 0 \\ 0 & -4y^2 & -6z^2 & 0 & -2y^2-3z^2 \end{pmatrix},\\
A_{\Delta^5} &:= \begin{pmatrix} 8x^8 & 0 & 0 & x^8 & 0 \\ -2x^2 & 0 & 0 & 0 & -x^2 \end{pmatrix},\\
A_{\Delta^6} &:= \begin{pmatrix} 0 & 8y^8 & 0 & y^8 & 0 \\ 0 & -4y^2 & 0 & 0 & -2y^2 \end{pmatrix},\\
A_{\Delta^7} &:= \begin{pmatrix} 0 & 0 & 8z^8 & z^8 & 0 \\ 0 & 0 & -6z^2 & 0 & -3z^2 \end{pmatrix}.
\end{aligned}$$

Hence the map F is non-degenerate at infinity. By Theorem 1.1, f_0 attains its infimum on the semi-algebraic set $\{f_1(x, y, z) \leq 0\} \subset \mathbb{R}^3$. Further, thanks to Remark 4.2, the global infimum f^* can be approximated as close as desired by solving a hierarchy of semidefinite programmings defined for the compact case [25].

5 Proof of the technical ingredients

In the rest of the paper, we give the proofs of the results stated in Sect. 3. Let us begin with the following.

Proof of Lemma 3.3 By replacing if necessary f_i by $\frac{\pm 1}{1+(f_i(x))^2}$, there is no loss of generality to assume that $y \in \mathbb{R}^p$.

We consider the semi-algebraic map $\Phi\colon \mathbb{R}^n \to \mathbb{R}^{n+1} \times \mathbb{R}^p$ given by

$$\Phi(x) := \left(\frac{x_1}{\sqrt{1+\|x\|^2}}, \ldots, \frac{x_n}{\sqrt{1+\|x\|^2}}, \frac{1}{\sqrt{1+\|x\|^2}}, F(x) \right).$$

Without loss of generality, we can suppose that the sequence $\Phi(x^k)$ is convergent to some point $(u, y) \in \mathbb{S}^n \times \mathbb{R}^p$. By Tarski–Seidenberg Theorem 3.1, $B := \Phi(A)$ is a semi-algebraic set. Thus we can apply Curve Selection Lemma from [6, Proposition 2.6.19] for the point $(u, y) \in \overline{B}$. There exists a continuous semi-algebraic curve

$$\psi(t)\colon [0, \epsilon) \to \mathbb{R}^{n+1} \times \mathbb{R}^p, \quad t \mapsto (\psi_1(t), \ldots, \psi_n(t), \psi_{n+1}(t), \ldots, \psi_{n+1+p}(t)),$$

such that

(b1) $\psi(0) = (u, y)$;
(b2) $\psi|_{(0,\epsilon)}$ is analytic; and
(b3) $\psi(t) \in B$ for all $t \in (0, \epsilon)$.

Note that $\psi_{n+1}(t) > 0$ for all $t \in (0, \epsilon)$. Define the curve $\varphi\colon (0, \epsilon) \to \mathbb{R}^n$, $t \mapsto \varphi(t)$, by

$$\varphi(t) := \left(\frac{\psi_1(t)}{\psi_{n+1}(t)}, \ldots, \frac{\psi_n(t)}{\psi_{n+1}(t)}\right).$$

Then it is clear that φ has the required properties. □

Proof of Theorem 3.2 By contradiction, suppose that $\widetilde{K}_\infty(F) \neq \emptyset$; i.e., there exist a point $y \in \mathbb{R}^p$ and a sequence $\{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ such that

$$\lim_{k\to\infty} \|x^k\| = \infty, \quad \lim_{k\to\infty} F(x^k) = y, \quad \text{and} \quad \lim_{k\to\infty} \nu_F(x^k) = 0.$$

By definition, there exists a sequence $\lambda^k := (\lambda_1^k, \ldots, \lambda_p^k) \in \mathbb{R}^p$, with $\sum_{i=1}^p |\lambda_i^k| = 1$, such that we have for all $k \geq 1$,

$$\nu_F\left(x^k\right) = \left\|\sum_{i=1}^p \lambda_i^k \nabla f_i\left(x^k\right)\right\|.$$

By applying Curve Selection Lemma at infinity (Lemma 3.3) with the following setup: the set $A := \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^p : \sum_{i=1}^p |\lambda_i| = 1, \nu_F(x) = \|\sum_{i=1}^p \lambda_i \nabla f_i(x)\|\}$ which is clearly semi-algebraic, the sequence $(x^k, \lambda^k) \in A$ which tends to infinity as $k \to \infty$, and the semi-algebraic map $G\colon A \to \mathbb{R}^{p+1}$, $(x, \lambda) \mapsto (F(x), \nu_F(x))$, it follows that there exist a positive constant ϵ and some smooth semi-algebraic curves $\varphi(t) := (\varphi_1(t), \ldots, \varphi_n(t))$ and $\lambda(t) := (\lambda_1(t), \ldots, \lambda_p(t))$, $0 < t < \epsilon$, such that

(c1) $\lim_{t\to 0} \|\varphi(t)\| = \infty$;
(c2) $\lim_{t\to 0} F(\varphi(t)) = y \in \mathbb{R}^p$;
(c3) $\sum_{i=1}^p |\lambda_i(t)| = 1$; and
(c4) $\lim_{t\to 0} \|\sum_{i=1}^p \lambda_i(t) \nabla f_i(\varphi(t))\| = 0$.

Let $J := \{j : \varphi_j \not\equiv 0\}$. By Condition (c1), $J \neq \emptyset$. By Growth Dichotomy Lemma (Lemma 3.2), for each $j \in J$, we can expand the coordinate φ_j as follows

$$\varphi_j(t) = x_j^0 t^{q_j} + \text{higher order terms in } t,$$

where $x_j^0 \neq 0$ and $q_j \in \mathbb{Q}$. From Condition (c1), we get $\min_{j \in J} q_j < 0$.

Recall that $\mathbb{R}^J := \{\kappa := (\kappa_1, \kappa_2, \ldots, \kappa_n) \in \mathbb{R}^n \ : \ \kappa_j = 0 \text{ for } j \notin J\}$. Since the map $F = (f_1, \ldots, f_p)$ is convenient, $\Gamma(f_i) \cap \mathbb{R}^J \neq \emptyset$ for all $i = 1, \ldots, p$. Let d_i be the minimal value of the linear function $\sum_{j \in J} q_j \kappa_j$ on $\Gamma(f_i) \cap \mathbb{R}^J$, and let Δ_i be the (unique) maximal face of $\Gamma(f_i) \cap \mathbb{R}^J$ where the linear function takes this value. Since f_i is convenient and $\min_{j \in J} q_j < 0$, by Lemma 2.1(i), we have $d_i < 0$ and Δ_i is a face of $\Gamma_\infty(f_i)$. Note that f_{i,Δ_i} does not depend on x_j for all $j \notin J$. Now suppose that f_i is written as $f_i = \sum_\kappa a_{i,\kappa} x^\kappa$. Then

$$
\begin{aligned}
f_i(\varphi(t)) &= \sum_{\kappa \in \Gamma(f_i) \cap \mathbb{R}^J} a_{i,\kappa} (\varphi(t))^\kappa \\
&= \sum_{\kappa \in \Gamma(f_i) \cap \mathbb{R}^J} a_{i,\kappa} (\varphi_1(t))^{\kappa_1} \ldots (\varphi_n(t))^{\kappa_n} \\
&= \sum_{\kappa \in \Gamma(f_i) \cap \mathbb{R}^J} \left(a_{i,\kappa} (x_1^0 t^{q_1})^{\kappa_1} \ldots \left(x_n^0 t^{q_n} \right)^{\kappa_n} + \text{higher order terms in } t \right) \\
&= \sum_{\kappa \in \Gamma(f_i) \cap \mathbb{R}^J} \left(a_{i,\kappa} (x^0)^\kappa t^{\sum_{j \in J} q_j \kappa_j} + \text{higher order terms in } t \right) \\
&= \sum_{\kappa \in \Delta_i} a_{i,\kappa} (x^0)^\kappa t^{d_i} + \text{higher order terms in } t \\
&= f_{i,\Delta_i}(x^0) t^{d_i} + \text{higher order terms in } t,
\end{aligned}
$$

where $x^0 := (x_1^0, \ldots, x_n^0)$ with $x_j^0 := 1$ for $j \notin J$. By Condition (c2) and $d_i < 0$, we have

$$
f_{i,\Delta_i}\left(x^0\right) = 0, \quad \text{for all} \quad i = 1, \ldots, p. \tag{6}
$$

Let $I := \{i \ : \ \lambda_i \not\equiv 0\}$. It follows from Condition (c3) that $I \neq \emptyset$. For $i \in I$, expand the coordinate λ_i in terms of the parameter (cf. Lemma 3.2) as follows

$$
\lambda_i(t) = \lambda_i^0 t^{\theta_i} + \text{higher order terms in } t,
$$

where $\lambda_i^0 \neq 0$ and $\theta_i \in \mathbb{Q}$.

For $i \in I$ and $j \in J$, by some similar calculations as with $f_i(\varphi(t))$, we have

$$
\frac{\partial f_i}{\partial x_j}(\varphi(t)) = \frac{\partial f_{i,\Delta_i}}{\partial x_j}(x^0) t^{d_i - q_j} + \text{higher order terms in } t.
$$

It implies that

$$\sum_{i\in I}\lambda_i(t)\frac{\partial f_i}{\partial x_j}(\varphi(t)) = \sum_{i\in I}\left(\lambda_i^0\frac{\partial f_{i,\Delta_i}}{\partial x_j}(x^0)t^{d_i+\theta_i-q_j} + \text{higher order terms in } t\right)$$
$$= \left(\sum_{i\in I'}\lambda_i^0\frac{\partial f_{i,\Delta_i}}{\partial x_j}(x^0)\right)t^{\ell-q_j} + \text{higher order terms in } t,$$

where $\ell := \min_{i\in I}(d_i+\theta_i)$ and $I' := \{i \in I : d_i+\theta_i = \ell\} \neq \emptyset$. We note, by Condition (c4), that for all $j \in J$,

$$\left\|\sum_{i\in I}\lambda_i(t)\frac{\partial f_i}{\partial x_j}(\varphi(t))\right\| = \left\|\sum_{i=1}^{p}\lambda_i(t)\frac{\partial f_i}{\partial x_j}(\varphi(t))\right\| \to 0, \quad \text{as} \quad t \to 0.$$

There are two cases to be considered.

Case 1 $\ell \le q_{j_*} := \min_{j\in J} q_j$ We have for all $j \in J$,

$$\sum_{i\in I'}\lambda_i^0\frac{\partial f_{i,\Delta_i}}{\partial x_j}\left(x^0\right) = 0.$$

On the other hand, for each $j \notin J$, the polynomial f_{i,Δ_i} does not depend on the variable x_j, so $\frac{\partial f_{i,\Delta_i}}{\partial x_j} \equiv 0$. Therefore

$$\sum_{i\in I'}\lambda_i^0\frac{\partial f_{i,\Delta_i}}{\partial x_j}\left(x^0\right) = 0, \quad \text{for all } j = 1,\ldots,n.$$

Consequently, we obtain

$$\text{rank}\begin{pmatrix} x_1^0\frac{\partial f_{1,\Delta_1}}{\partial x_1}(x^0) & \cdots & x_n^0\frac{\partial f_{1,\Delta_1}}{\partial x_n}(x^0) \\ \vdots & \cdots & \vdots \\ x_1^0\frac{\partial f_{p,\Delta_p}}{\partial x_1}(x^0) & \cdots & x_n^0\frac{\partial f_{p,\Delta_p}}{\partial x_n}(x^0) \end{pmatrix} < p$$

since the matrix has $\#I'$ linearly dependent rows. This, together with (6), contradicts the assumption that the polynomial map $F = (f_1,\ldots,f_p)$ is Khovanskii non-degenerate at infinity.

Case 2 $\ell > q_{j_*} := \min_{j\in J} q_j$. It follows from Condition (c3) that $\theta_i \ge 0$ for all $i \in I$ and $\theta_i = 0$ for some $i \in I$. Without lost of generality, we may assume that $1 \in I$ and $\theta_1 = 0$. Since f_1 is convenient, for any $j = 1,\ldots,n$, there exists a natural number $m_j \ge 1$ such that $m_je_j \in \Gamma_\infty(f_1)$. Then, by definition, it is clear that

$$q_jm_j \ge d_1, \quad \text{for all } j \in J.$$

On the other hand, we have

$$d_1 = d_1 + \theta_1 \geq \min_{i \in I}(d_i + \theta_i) = \ell.$$

Therefore

$$q_{j_*} m_{j_*} \geq d_1 \geq \ell > q_{j_*}.$$

Since $q_{j_*} = \min_{j \in J} q_j < 0$, it implies that $m_{j_*} < 1$, which is a contradiction. □

Proof of Lemma 3.4 It suffices to show that the semi-algebraic set S_R is regular for R large enough. Indeed, by contradiction and by Curve Selection Lemma at infinity (Lemma 3.3), there exist a smooth semi-algebraic curve $\varphi(t)$ and some smooth semi-algebraic functions $\lambda_i(t), \mu(t), t \in (0, \epsilon)$, such that

(d1) $\varphi(t) \in S$ for $t \in (0, \epsilon)$;
(d2) $\lim_{t \to 0} \|\varphi(t)\| = \infty$;
(d3) $\sum_{i=1}^{p} \lambda_i(t) \nabla f_i(\varphi(t)) + \mu(t)\varphi(t) = 0$; and
(d4) $\lambda_i(t) f_i(\varphi(t)) \equiv 0$, for $i = 1, \ldots, p$.

Since the functions λ_i and $f_i \circ \varphi$ are semi-algebraic, for $\epsilon > 0$ small enough, these functions are either constant or strictly monotone (see Monotonicity Lemma 3.1). Then, by Condition (d4), we can see that either $\lambda_i(t) \equiv 0$ or $f_i \circ \varphi(t) \equiv 0$; in particular,

$$\lambda_i(t) \frac{d}{dt}(f_i \circ \varphi)(t) \equiv 0, \quad i = 1, \ldots, p.$$

Hence, it follows from Condition (d3) that

$$\begin{aligned} 0 &= \sum_{i=1}^{p} \lambda_i(t) \left\langle \nabla f_i(\varphi(t)), \frac{d\varphi(t)}{dt} \right\rangle + \mu(t) \left\langle \varphi(t), \frac{d\varphi(t)}{dt} \right\rangle \\ &= \sum_{i=1}^{p} \lambda_i(t) \frac{d}{dt}(f_i \circ \varphi)(t) + \frac{\mu(t)}{2} \frac{d\|\varphi(t)\|^2}{dt} \\ &= \frac{\mu(t)}{2} \frac{d\|\varphi(t)\|^2}{dt}. \end{aligned}$$

So $\mu(t) \equiv 0$, which contradicts the regularity of the set S. □

Proof of Lemma 3.5 Suppose that the lemma does not hold. Then the set

$$S' := \{x \in S : I(x) \neq \emptyset, \text{ the gradient vectors } \nabla f_i(x), i \in I(x), \text{ are linearly dependent}\}$$

is not bounded. Since the number of subsets of $\{1, \ldots, p\}$ is finite, there exists a non empty subset $I := \{i_1, \ldots, i_q\} \subset \{1, \ldots, p\}$ such that

$$S'' := \{x \in S : I(x) = I, \quad \text{the gradient vectors } \nabla f_i(x), i \in I, \text{ are linearly dependent}\}$$

is unbounded. It is clear that the set S'' is semi-algebraic. By Curve Selection Lemma at infinity (Lemma 3.3), there exist a positive number ϵ and a smooth semi-algebraic curve $\varphi(t) := (\varphi_1(t), \ldots, \varphi_n(t)) \in S''$, $0 < t < \epsilon$, such that

(e1) $\lim_{t\to 0} \|\varphi(t)\| = \infty$;
(e2) $f_i(\varphi(t)) \equiv 0$ for $i \in I$ and $f_i(\varphi(t)) < 0$ for $i \notin I$;
(e3) The gradient vectors $\nabla f_i(\varphi(t))$, $i \in I$, are linearly dependent.

Then, by definition, $\nu_{F_I}(\varphi(t)) \equiv 0$ for $0 < t < \epsilon$, where F_I stands for the polynomial map $x \mapsto (f_{i_1}(x), \ldots, f_{i_q}(x))$. Consequently, we have $0 \in \widetilde{K}(F_I)$, which contradicts Corollary 3.1. □

Acknowledgments The authors would like to thank the referees, whose helpful comments and suggestions much improved the original manuscript. This research was performed while the authors had been visiting the Vietnam Institute for Advanced Study in Mathematics (VIASM). The authors would like to thank the Institute for hospitality and support.

References

1. Andronov, V.G., Belousov, E.G., Shironin, V.M.: On solvability of the problem of polynomial programming, Izvestija Akadem. Nauk SSSR, Tekhnicheskaja Kibernetika, no. 4, 194–197 (1982) (in Russian). Translation Appeared in News of the Academy of Science of USSR, Department of Technical Sciences, Technical. Cybernetics no. 4, 194–197 (1982)
2. Auslender, A.: How to deal with the unbounded in optimization: theory and algorithms. Math. Program. Ser. B **79**, 3–8 (1997)
3. Bank, B., Mandel, R.: Parametric Integer Optimization, Mathematical Research, vol. 39, edn. Academie-Verlag, Berlin (1988)
4. Belousov, E.G.: Introduction to Convex Analysis and Integer Programming. Moscow University Publ, Moscow (1977). (in Russian)
5. Belousov, E.G., Klatte, D.: A Frank–Wolfe type theorem for convex polynomial programs. Comput. Optim. Appl. **22**(1), 37–48 (2002)
6. Benedetti, R., Risler, J.: Real Algebraic and Semi-algebraic Sets. Hermann, Paris (1991)
7. Bertsekas, D.P., Tseng, P.: Set intersection theorems and existence of optimal solutions. Math. Program. Ser. A **110**(2), 287–314 (2007)
8. Bivia-Ausina, C.: Mixed Newton numbers and isolated complete intersection singularities. Proc. Lond. Math. Soc. **94**(3), 749–771 (2007)
9. Bochnak, J., Coste, M., Roy, M.-F.: Real Algebraic Geometry, vol. 36. Springer, Berlin (1998)
10. Bolte, J., Daniilidis, A., Lewis, A.S.: Generic optimality conditions for semialgebraic convex programs. Math. Oper. Res. **36**(1), 55–70 (2011)
11. Damon, J.: Topological invariants of μ-constant deformations of complete intersection singularities. Q. J. Math. Oxf. Ser. **40**(2), 139–159 (1989)
12. Demmel, J., Nie, J.W., Powers, V.: Representations of positive polynomials on noncompact semi-algebraic sets via KKT ideals. J. Pure Appl. Algebra **209**(1), 189–200 (2007)
13. Dinh, S.T., Hà, H.V., Phạm, T.S.: A Frank-Wolfe type theorem and Holder-type global error bound for generic polynomial systems, preprint 2012, VIASM. Available online from http://viasm.edu.vn/wp-content/uploads/2012/11/Preprint_1227.pdf
14. van den Dries, L., Miller, C.: Geometric categories and o-minimal structures. Duke Math. J. **84**, 497–540 (1996)

15. Frank, M., Wolfe, P.: An algorithm for quadratic programming. Naval Res. Logist. Q. **3**, 95–110 (1956)
16. Gaffney, T.: Integral closure of modules and Whitney equisingularity. Invent. Math. **107**, 301–322 (1992)
17. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. Publishers W.H. Freeman, San Francisco (1979)
18. Hà, H.V., Phạm, T.S.: Global optimization of polynomials using the truncated tangency variety and sums of squares. SIAM J. Optim. **19**, 941–951 (2008)
19. Hà, H.V., Phạm, T.S.: Solving polynomial optimization problems via the truncated tangency variety and sums of squares. J. Pure Appl. Algebra **213**, 2167–2176 (2009)
20. Hà, H.V., Phạm, T.S.: Representations of positive polynomials and optimization on noncompact semi-algebraic sets. SIAM J. Optim. **20**, 3082–3103 (2010)
21. Hà, H.V.: Global Hölderian error bound for non-degenerate polynomials. SIAM J. Optim. **23**(2), 917–933 (2013)
22. Jibetean, D., Laurent, M.: Semidefinite approximations for global unconstrained polynomial optimization. SIAM J. Optim. **16**, 490–514 (2005)
23. Khovanskii, A.G.: Newton polyhedra and toroidal varieties. Funct. Anal. Appl. **11**, 289–296 (1978)
24. Kouchnirenko, A.G.: Polyhedres de Newton et nombre de Milnor. Invent. Math. **32**, 1–31 (1976)
25. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**(3), 796–817 (2001)
26. Lasserre, J.B.: Moments, Positive Polynomials and Their Applications. Imperial College Press, London (2009)
27. Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. In: Putinar, M., Sullivant, S. (eds.) Emerging Applications of Algebraic Geometry, IMA Volumes in Mathematics and its Applications, vol. 149, pp. 157–270. Springer, Berlin (2009)
28. Luo, Z.-Q., Zhang, S.: On extensions of the Frank-Wolfe theorems. Comput. Optim. Appl. **13**(1–3), 87–110 (1999)
29. Marshall, M.: Positive Polynomials and Sums of Squares, Mathematical Surveys and Monographs, vol. 146. American Mathematical Society, Providence, RI (2008)
30. Marshall, M.: Representations of non-negative polynomials, degree bounds and applications to optimization. Can. J. Math. **61**(1), 205–221 (2009)
31. Miller, C.: Exponentiation is hard to avoid. Proc. Am. Math. Soc. **122**, 257–259 (1994)
32. Milnor, J.: Singular Points of Complex Hypersurfaces, Annals of Mathematics Studies, vol. 61. Princeton University Press, Princeton (1968)
33. Némethi, A., Zaharia, A.: Milnor fibration at infinity. Indag. Math. **3**, 323–335 (1992)
34. Nie, J., Demmel, J., Sturmfels, B.: Minimizing polynomials via sum of squares over the gradient ideal. Math. Program. Ser. A **106**(3), 587–606 (2006)
35. Obuchowska, W.T.: On generalizations of the Frank-Wolfe theorem to convex and quasi-convex programmes. Comput. Optim. Appl. **33**(2–3), 349–364 (2006)
36. Oka, M.: Non-degenerate Complete Intersection Singularity. Actualités Mathématiques, Hermann, Paris (1997)
37. Parrilo, P.A.: Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization, Ph.D. thesis, California Institute of Technology, May 2000
38. Parrilo, P.A., Sturmfels, B.: Minimizing polynomial functions. In: Algorithmic and Quantitative Real Algebraic Geometry (Piscataway, NJ, 2001). DIMACS: Series in Discrete Mathematics and Theoretical Computer Science, vol. 60, pp. 83–99. American Mathematical Society, Providence, RI (2003)
39. Parrilo, P.A.: Semidefinite Programming relaxations for semialgebraic problems. Math. Program. Ser. B **96**(2), 293–320 (2003)
40. Perold, A.F.: Generalization of the Frank–Wolfe Theorem. Math. Program. **18**, 215–227 (1980)
41. Rabier, P.J.: Ehresmann fibrations and Palais-Smale conditions for morphisms of Finsler manifolds. Ann. Math. **146**, 647–691 (1997)
42. Reznick, B.: Some concrete aspects of Hilbert's 17th problem. Contemp. Math. **253**, 251–272 (2000)
43. Shor, N.Z.: An approach to obtaining global extremums in polynomial mathematical programming problems. Kibernetika **5**, 102–106 (1987)
44. Schweighofer, M.: Global optimization of polynomials using gradient tentacles and sums of squares. SIAM J. Optim. **17**(3), 920–942 (2006)
45. Terlaky, T.: On l_p programming. Eur. J. Oper. Res. **22**, 70–100 (1985)