Exact relaxation for polynomial optimization on semi-algebraic sets

Marta Abril Bucero, Bernard Mourrain

EXACT RELAXATION FOR POLYNOMIAL OPTIMIZATION ON SEMI-ALGEBRAIC SETS

MARTA ABRIL BUCERO, BERNARD MOURRAIN

Abstract. In this paper, we study the problem of computing by relaxation hierarchies the infimum of a real polynomial function f on a closed basic semialgebraic set S and the points where this infimum is reached, if they exist. We show that when the infimum is reached, a relaxation hierarchy constructed from the Karush-Kuhn-Tucker ideal is always exact and that the vanishing ideal of the KKT minimizer points is generated by the kernel of the associated moment matrix in that degree, even if this ideal is not zero-dimensional. We also show that this relaxation allows to detect when there is no KKT minimizer. We prove that the exactness of the relaxation depends only on the real points which satisfy these constraints. This exploits representations of positive polynomials as elements of the preordering modulo the KKT ideal, which only involves polynomials in the initial set of variables. The approach provides a uniform treatment of different optimization problems considered previously. Applications to global optimization, optimization on semialgebraic sets defined by regular sets of constraints, optimization on finite semialgebraic sets, real radical computation are given.

1. Introduction

Optimization problems appear in many areas of Scientific Computing. Local methods such as gradient descent are often employed to handle them. These methods can be very efficient to compute a local minimum, but the output depends on the initial guess and they give no guarantee of a global solution.

The problem we consider in this paper, is how to compute and certify a global optimum of a polynomial function on semi-algebraic sets if it is reached. Reformulating the problem as the computation of a (minimal) critical value, polynomial system solvers can be used to tackle it (see e.g. [34], [6]). But in this case, the complex solutions of the underlying algebraic system come into play and additional computation efforts should be spent to remove these extraneous solutions. Semi-algebraic techniques such as Cylindrical Algebraic Decomposition or extensions [35] may also be considered here but suffer from similar issues. Though the global minimization problem is known to be NP-hard (see e.g. [28]), a practical challenge is to devise methods which take into account "only" the real solutions of the problem or which can approximate them efficiently.

Previous works. About a decade ago, a relaxation approach has been proposed in [16] (see also [33], [38]) to solve this difficult problem. Instead of searching points where the polynomial f reaches its minimum f^*, a probability measure which minimizes the function f is searched. This problem is relaxed into a hierarchy of finite dimensional convex minimization problems, that can be solved by Semi-Definite Programming (SDP) techniques, and which converges to the minimum f^* [16]. This hierarchy of SDP problems can be formulated in terms of linear matrix inequalities on moment matrices associated to the set of monomials of degree $\leq t \in \mathbb{N}$ for increasing values of t. The dual hierarchy can be described as a sequence of maximization problems over the cone of polynomials which are Sums of Squares (SoS). A feasibility condition is needed to prove that this dual hierarchy of maximization problems also converges to the minimum f^*, ie. that there is no duality gap.

From a computational point of view, the following issues need to be addressed:

(1) Is the hierarchy of optimization problems *exact*, ie. does it convergence in a finite number of steps?
(2) Is there a criterion to detect when the optimum is reached, that will eventually be satisfied?
(3) When the convergence criterion is satisfied, how can we recover all the points where the optimum is achieved?

To address the first issue, the following strategy has been considered: add polynomial inequalities or equalities satisfied by the points where the function f is minimum. Inequality constraints can for instance be added to restrict the optimization problem to a compact subset of $\mathbb{R}^n$ and to make the hierarchies exact [16], [25]. Natural constraints which do not require apriori bounds on the solutions are for instance the vanishing of the partial derivatives of f. A result in [20] shows that if the complex variety of the ideal generated by the equalities of the semialgebraic set is finite, then the hierarchy of relation problems introduced by Lasserre in [16] is exact. It is also proved that there is no duality gap if the generators of this ideal satisfy some conditions.

In [32], it is proved that a relaxation hierarchy using the gradient constraints is exact when the gradient ideal is radical. In [26], it is shown that this gradient hierarchy is exact, when the global minimizers satisfy the Boundary Hessian condition; In [5], it is proved that a relaxation hierarchy which involves the Karush-Kuhn-Tucker constraints is exact when the KKT ideal is radical. In [29], a relaxation hierarchy obtained by projection of the KKT constraints is proved to be exact under some regularity conditions. In [9], the exactness of a relaxation hierarchy related to the real KKT variety defined by a radical ideal is also analyzed under similar regularity conditions.

The case where the infimum value is not reached at points which satisfied the KKT constraints, has also been studied. In [36], relaxation techniques are studied for functions for which the minimum is not reached and which satisfies some special properties "at infinity". In [8], tangency constraints are used in a relaxation hierarchy which converges to the global minimum of a polynomial, when the polynomial is bounded by below over $\mathbb{R}^n$. In [7], generic changes of coordinates and a partial gradient ideal are used in a relaxation hierarchy which also converges to the global minimum of f on $\mathbb{R}^n$.

In the cases studied so far, the exactness of the relaxation is proved under a genericity condition or compactness property.

The second issue is also important from an algorithmic point of view, to certify when the minimum is reached. In [20], [19], a stopping criterion based on Curto-Fialkow flat extension condition [4] is used in a relaxation hierarchy for optimization on finite algebraic varieties and for real radical computation. This criterion is also used in [32], where an algorithm is described to compute the global minimum of the polynomial function f when this minimum is a reached and when the gradient ideal is zero-dimensional. An improvement of the flat extension criterion [23] is used in [15] to compute the real radical of an ideal when this real radical is zero-dimensional. The convergence criterion is combined with a normal form criterion [27] to detect when a border basis of the real radical is computed. In [30], it is proved that a Curto-Fialkow flat extension criterion is eventually satisfied on truncated moment matrices, in relaxation hierarchies [16], [32], [29] with a finite number of minimizers, under some regularity conditions or archimedean conditions.

The third issue is related to the problem of computing all the points which realize the minimal value, which is also important from a practical point of view. In [19], the kernel of moment matrices are used to compute generators of the real radical of an ideal. This method is improved in [15] to compute a border basis of the real radical, involving SDP problems of significantly smaller size, when the real radical ideal is zero-dimensional. The problem of computing the minimizer ideal has not been addressed so far in exact relaxation hierarchies, though it is mentioned in [30] for zero-dimensional minimizer ideals.

An interesting feature of these hierarchies of SDP problems is that, at any step they provide a lower bound of f^* and the SoS hierarchy gives certificates for these lower bounds (see e.g. [14]). In [10, 12] it is also shown how to obtain "good" upper bounds by perturbation techniques, which can directly be generalized to the approach we propose in this paper.

Contributions. The main contributions are the following:

- We prove that exact relaxation hierarchies depending on the variables $\mathbf{x}$ can be constructed for general polynomial optimization problems on basic closed semi-algebraic sets (Theorem 6.3).
- We also prove that the KKT minimizer ideal can be constructed from the Moment matrix of an optimal linear form, when the corresponding relaxation is exact, even if the ideal is not zero-dimensional (Theorem 5.10).
- We prove that the exactness of the relaxation depends only on the real points which satisfy these constraints (Theorem 5.10).
- We provide a general approach that allow us to treat in a uniform way and to extend results on the representation of polynomials which are positive (resp. non-negative) on the critical points in [5] (see Theorem 4.9) and on the exactness of relaxation hierarchies in [32], [8], [19], [29], [15], [31] (see Theorem 6.2, Theorem 6.4, Theorem 6.5, Theorem 6.6).

Content. The paper is organized as follows. In Section 2, we recall algebraic concepts, we define our minimization problem and describe the hierarchy of finite dimensional convex optimization problems that we consider. In Section 3 we analyze the varieties associated to the critical points of the minimization problem. Section 4 is devoted to the representation positive and non-negative polynomials on the critical points as sum of squares modulo the gradient ideal. In Section 5 we prove that when the order of relaxation is big enough, the sequence of finite dimensional convex optimization problems attains its limit and the minimizer ideal can be generated from the solution of our relaxation problem. In Section 6, we analyze some consequences of these results. Finally, Section 7 contains several examples which illustrate the approach.

2. Ideals, varieties, optimization and relaxation

In this section, we recall some algebraic concepts as ideals and varieties, we define our minimization problem and we set our notation.

2.1. **Ideals and varieties.** Let $\mathbb{K}[\mathbf{x}]$ be the set of the polynomials in the variables $\mathbf{x} = (x_1, \ldots, x_n)$, with coefficients in the field $\mathbb{K}$. Hereafter, we will choose[1] $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let $\overline{\mathbb{K}}$ denotes the algebraic closure of $\mathbb{K}$. For $\alpha \in \mathbb{N}^n$, $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ is the monomial with exponent α and degree $|\alpha| = \sum_i \alpha_i$. The set of all monomials in $\mathbf{x}$ is denoted $\mathcal{M} = \mathcal{M}(\mathbf{x})$.

For $t \in \mathbb{N} \cup \{\infty\}$ and $S \subseteq \mathbb{K}[\mathbf{x}]$, we introduce the following sets:

- S_t is the set of elements of S of degree $\leq t$,
- $\langle S \rangle = \left\{ \sum_{f \in S} \lambda_f \, f \mid f \in S, \lambda_f \in \mathbb{K} \right\}$ is the linear span of S,

[1]For notational simplicity, we will consider only these two fields in this paper, but $\mathbb{R}$ and $\mathbb{C}$ can be replaced respectively by any real closed field and any field containing its algebraic closure.

- $(S) = \left\{ \sum_{f\in S} p_f\, f \mid p_f \in \mathbb{K}[\mathbf{x}], f \in S \right\}$ is the ideal in $\mathbb{K}[\mathbf{x}]$ generated by S,
- $S_{\langle t\rangle} = \left\{ \sum_{f\in S_t} p_f\, f \mid p_f \in \mathbb{K}[\mathbf{x}]_{t-\deg(f)} \right\}$ is the vector space spanned by $\{\mathbf{x}^\alpha f \mid f \in S_t, |\alpha| \le t - \deg(f)\}$,
- $\mathcal{Q}_t^+ = \left\{ \sum_{i=1}^{l} p_i^2 \mid l \in \mathbb{N}, p_i \in \mathbb{R}[\mathbf{x}]_t \right\}$ is the set of finite sums of squares of polynomials of degree $\le t$; $\mathcal{Q}^+ = \mathcal{Q}_\infty^+$.

By definition $S_{\langle t\rangle} \subseteq (S) \cap \mathbb{K}[\mathbf{x}]_t = (S)_t$, but the inclusion may be strict.

By convention, a set of constrains $G = \{g_1^0, \dots, g_{n_1}^0;\ g_1^+, \dots, g_{n_2}^+\} \subset \mathbb{R}[\mathbf{x}]$ is a finite set of polynomials composed of a subset $G^0 = \{g_1^0, \dots, g_{n_1}^0\}$ corresponding to the equality constraints and a subset $G^+ = \{g_1^+, \dots, g_{n_2}^+\}$ corresponding to the non-negativity constraints. For two set of constraints $G, G' \subset \mathbb{R}[\mathbf{x}]$, we say that $G \subset G'$ if $G^0 \subset G'^0$ and $G^+ \subset G'^+$.

Definition 2.1. *For $t \in \mathbb{N} \cup \{\infty\}$ and a set of constraints $G = \{g_1^0, \dots, g_{n_1}^0;\ g_1^+, \dots, g_{n_2}^+\} \subset \mathbb{R}[\mathbf{x}]$, we define the (truncated) quadratic module of G by*

$$\mathcal{Q}_t(G) = \{\sum_{i=1}^{n_2} g_i^0\, h_i + s_0 + \sum_{j=1}^{n_2} g_j^+\, s_j \mid h_i \in \mathbb{R}[\mathbf{x}]_{2t-\deg(g_i^0)}, s_0 \in \mathcal{Q}_{2t}^+, s_i \in \mathcal{Q}^+_{t-\deg(g_i^+)/2}\}.$$

If $\tilde{G}$ is such that $\tilde{G}^0 = G^0$ and $\tilde{G}^+ = \{\prod (g_1^+)^{\epsilon_1} \cdots (g_{n_2}^+)^{\epsilon_{n_2}} \mid \epsilon_i \in \{0,1\}\}$, $\mathcal{Q}_t(\tilde{G})$ is also called the (truncated) preordering of G and denoted $\mathcal{P}_t(G)$. When $t = \infty$, $\mathcal{P}(G) := \mathcal{P}_\infty(G)$ is the preordering of G. The (truncated) preodering generated by the positive constraints is denoted $\mathcal{P}^+(G) = \mathcal{P}(G^+)$.

Definition 2.2. *Given $t \in \mathbb{N} \cup \{\infty\}$ and a set of constraints $G \subset \mathbb{R}[\mathbf{x}]$, we define*

$$\mathcal{N}_t(G) := \{\Lambda \in (\mathbb{R}[\mathbf{x}]_{2t})^* \mid \Lambda(p) \ge 0,\ \forall p \in \mathcal{Q}_t(G), \Lambda(1) = 1\}.$$

When we replace $\mathcal{Q}_t(G)$ by $\mathcal{P}_t(G)$ in this definition, we denote the corresponding set by $\mathcal{L}_t(G)$.

Given a set $I \subseteq \mathbb{K}[\mathbf{x}]$ and a field $\mathbb{L} \supseteq \mathbb{K}$, we denote by

$$\mathcal{V}^{\mathbb{L}}(I) := \{x \in \mathbb{L}^n \mid f(x) = 0\ \forall f \in I\}$$

its associated variety in $\mathbb{L}^n$. By convention $\mathcal{V}(I) = \mathcal{V}^{\overline{\mathbb{K}}}(I)$, where $\overline{\mathbb{K}}$ is the algebraic closure of $\mathbb{K}$. Hereafter, we will also consider sets of homogeneous equations I and the varieties $\mathbb{P}\mathcal{V}(I)$ (resp. $\mathbb{P}\mathcal{V}^{\mathbb{R}}(I)$) defined in the projective space $\mathbb{P}^n$ (resp. the real projective space $\mathbb{R}\mathbb{P}^n$).

For a set $V \subseteq \mathbb{K}^n$, we define its vanishing ideal

$$\mathcal{I}(V) := \{p \in \mathbb{K}[\mathbf{x}] \mid p(v) = 0\ \forall v \in V\}.$$

For a set $V \subset \mathbb{L}^n$ with $\mathbb{L} \supseteq \mathbb{K}$, $V^{\mathbb{K}} = V \cap \mathbb{K}^n$. Hereafter, we will take $\mathbb{K} = \mathbb{R}$ and $\mathbb{L} = \mathbb{C}$, so that $\mathcal{V}(I) = \mathcal{V}^{\mathbb{C}}(I)$, $\mathcal{V}^{\mathbb{R}}(I) = \mathcal{V}(I)^{\mathbb{R}} = \mathcal{V}(I) \cap \mathbb{R}^n$.

Definition 2.3. *For a set of constrains $G = (G^0; G^+) \subset \mathbb{R}[\mathbf{x}]$,*

$$\begin{aligned} \mathcal{S}(G) &:= \{\mathbf{x} \in \mathbb{R}^n \mid g^0(\mathbf{x}) = 0\ \forall g^0 \in G^0,\ g^+(\mathbf{x}) \ge 0\ \forall g^+ \in G^+\}, \\ \mathcal{S}^+(G) &:= \{\mathbf{x} \in \mathbb{R}^n \mid g^+(\mathbf{x}) \ge 0\ \forall g^+ \in G\}. \end{aligned}$$

To describe the vanishing ideal of these sets, we introduce the following ideals:

Definition 2.4. *For a set of constraints $G = (G^0; G^+) \subset \mathbb{R}[\mathbf{x}]$,*

$$\begin{aligned} \sqrt{G^0} &= \{p \in \mathbb{R}[\mathbf{x}] \mid p^m \in (G^0) \ \text{ for some } m \in \mathbb{N} \setminus \{0\}\} \\ \sqrt[\mathbb{R}]{G^0} &= \{p \in \mathbb{R}[\mathbf{x}] \mid p^{2m} + q \in (G^0) \ \text{ for some } m \in \mathbb{N} \setminus \{0\}, q \in \mathcal{Q}^+\} \\ \sqrt[G^+]{G^0} &= \{p \in \mathbb{R}[\mathbf{x}] \mid p^{2m} + q \in (G^0) \ \text{ for some } m \in \mathbb{N} \setminus \{0\}, q \in \mathcal{P}^+(G)\} \end{aligned}$$

These ideals are called respectively the radical of G^0, the real radical of G^0, the G^+-radical of G^0.

Remark 2.5. *If $G^+ = \emptyset$, then $\sqrt[G^+]{G^0} = \sqrt[\mathbb{R}]{G^0}$.*

The following three famous theorems relate vanishing and radical ideals:

Theorem 2.6. *Let $G = (G^0; G^+)$ be a set of constraints of $\mathbb{R}[\mathbf{x}]$.*

(i) **Hilbert's Nullstellensatz** *(see, e.g.,* [3, §4.1]*)* $\sqrt{G^0} = \mathcal{I}(\mathcal{V}^{\mathbb{C}}(G^0))$.
(ii) **Real Nullstellensatz** *(see, e.g.,* [1, §4.1]*)* $\sqrt[\mathbb{R}]{G^0} = \mathcal{I}(\mathcal{V}^{\mathbb{R}}(G^0))$.
(iii) **Positivstllensatz** *(see, e.g.,* [1, §4.4]*)* $\sqrt[G^+]{G^0} = \mathcal{I}(\mathcal{S}(G)) = \mathcal{I}(\mathcal{V}^{\mathbb{R}}(G^0) \cap \mathcal{S}^+(G))$.

2.2. **Minimization problem.** Let $f, g_1^0, \ldots, g_{n_1}^0, g_1^+, \ldots, g_{n_2}^+ \in \mathbb{R}[\mathbf{x}]$ be polynomials functions. The minimization problem that we consider all along the paper is the following:

$$
\begin{aligned}
\inf_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\
s.t. \quad & g_1^0(\mathbf{x}) = \cdots = g_{n_1}^0(\mathbf{x}) = 0 \\
& g_1^+(\mathbf{x}) \geq 0, ..., g_{n_2}^+(\mathbf{x}) \geq 0
\end{aligned} \tag{1}
$$

More precisely, the objectives of the method we are going to describe are to compute the minimum value when f is bounded by below and the points where this minimum value is reached if they exists.

Hereafter, we fix the set of constraints $\mathbf{g} = \{g_1^0, \ldots, g_{n_1}^0; g_1^+, \ldots, g_{n_2}^+\}$ and denoted by $S := \mathcal{S}(\mathbf{g})$ the basic semi-algebraic set defined by the constraints of our minimization problem (1).

When $n_1 = n_2 = 0$, there is no constraint and $S = \mathbb{R}^n$. In this case, we are considering a global minimization problem.

The points $\mathbf{x}^* \in S$ which satisfy $f(\mathbf{x}^*) = \inf_{\mathbf{x}\in S} f(\mathbf{x})$ are called the *minimizers* of f on S. If the set of minimizers is not empty, we say that the *minimization problem is feasible.*

2.3. **Relaxation hierarchy.** The approach proposed by Lasserre in [16] to solve Problem (1) consists in approximating the optimization problem by a sequence of finite dimensional convex optimization problems which can be solved efficiently by Semi-Definite Programming tools. This sequence is called Lasserre hierarchy of relaxation problems. Hereafter, we are going the consider the relaxation hierarchy associated to preordering sequences:

$$\cdots \subset \mathcal{L}_{t+1}(G) \subset \mathcal{L}_t(G) \subset \cdots \text{ and } \cdots \subset \mathcal{P}_t(G) \subset \mathcal{P}_{t+1}(G) \subset \cdots$$

These convex sets are used to define extrema that approximate the solution of the minimization problem (1).

Definition 2.7. *Let $t \in \mathbb{N}$ and let G be set of constraints in $\mathbb{R}[\mathbf{x}]$. We define the following extrema:*

- $f_G^* = \inf_{\mathbf{x}\in\mathcal{S}(G)} f(\mathbf{x})$,
- $f_{t,G}^{\mu} = \inf \{\Lambda(f) \ s.t. \ \Lambda \in \mathcal{L}_t(G)\}$,
- $f_{t,G}^{sos} = \sup \{\gamma \in \mathbb{R} \ s.t. \ f - \gamma \in \mathcal{P}_t(G)\}$.

By convention if the corresponding sets are empty, $f_G^* = -\infty$, $f_{t,G}^{sos} = -\infty$ and $f_{t,G}^{\mu} = +\infty$.

Remark 2.8. *We have $f_{t,G}^{sos} \leq f_{t,G}^{\mu} \leq f_G^*$.*

Indeed, if there exists $\gamma \in \mathbb{R}$ such that $f - \gamma = q \in \mathcal{P}_t(G)$ then $\forall \Lambda \in \mathcal{L}_t(G)$, $\Lambda(f - \gamma) = \Lambda(f) - \gamma = \Lambda(q) \geq 0$, which proves the first inequality.

Since for any $\mathbf{s} \in S$, the evaluation $\mathbf{1_s} : p \in \mathbb{R}[\mathbf{x}] \mapsto p(\mathbf{s})$ is in $\mathcal{L}_t(G)$, we have $\mathbf{1_s}(f) = f(\mathbf{s}) \geq f_{t,G}^{\mu}$. This proves the second inequality.

As $\mathcal{L}_{t+1}(G) \subset \mathcal{L}_t(G)$ and $\mathcal{P}_t(G) \subset \mathcal{P}_{t+1}(G)$ we have the following increasing sequences for $t \in \mathbb{N}$:

$$\cdots f^{\mu}_{t,G} \leq f^{\mu}_{t+1,G} \leq \cdots \leq f^*_G \text{ and } \cdots f^{sos}_{t,G} \leq f^{sos}_{t+1,G} \leq \cdots \leq f^*_G.$$

The foundation of Lasserre relaxation method is to show that these sequences converge to f^*_G, see [16].

Hereafter, we are interested in constructing hierarchies for which, the minimum f^*_G is reached in a finite number of steps. Such hierarchies are called *exact*. We are also interested to compute the minimizers points. For that purpose, we introduce now the truncated Hankel operators, which will play a central role in the construction of the minimizer ideal of f on S.

Definition 2.9. *For a $t \in \mathbb{N}$ with $\mathbb{R}[\mathbf{x}]_t \subset \mathbb{R}[\mathbf{x}]$ and a linear form $\Lambda \in (\mathbb{R}[\mathbf{x}]_{2t})^*$, we define the map $M^t_\Lambda : \mathbb{R}[\mathbf{x}]_t \to (\mathbb{R}[\mathbf{x}]_t)^*$ by $M^t_\Lambda(p) = \Lambda(p\,q)$ for $p, q \in \mathbb{R}[\mathbf{x}]_t$. Thus M^t_Λ is called a truncated Hankel operator, defined on the subspace $\mathbb{R}[\mathbf{x}]_t$.*

Its matrix in monomial bases of $\mathbb{R}[\mathbf{x}]_t$ and $(\mathbb{R}[\mathbf{x}]_{2t})^*$ is also called the moment matrix of Λ.

Definition 2.10. *We define the kernel of the truncated Hankel operator:*

$$\ker M^t_\Lambda = \{p \in \mathbb{R}[\mathbf{x}]_t \mid \Lambda(p\,q) = 0 \ \forall q \in \mathbb{R}[\mathbf{x}]_t\}. \tag{2}$$

Given $t \in \mathbb{N}$ and $G = \{0\}$ and $\Lambda, \Lambda' \in \mathbb{R}[\mathbf{x}]^*_{2t}$, we easily check the following properties:

- $\forall p \in \mathbb{R}[\mathbf{x}]_t$, $\Lambda(p^2) = 0$ implies $p \in \ker M^t_\Lambda$.
- $\ker M^t_{\Lambda+\Lambda'} = \ker M^t_\Lambda \cap \ker M^t_{\Lambda'}$.

The kernel of these truncated Hankel operators will be used to compute generators of the minimizer ideal, as we will see.

3. Varieties of critical points

Before describing how to compute the minimizer points, we analyze the geometry of this minimization problem and the varieties associated to its critical points. In the following, we will denote by $\mathbf{y} = (\mathbf{x}, \mathbf{u}, \mathbf{v})$ and $\mathbf{z} = (\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s})$, the $n + n_1 + n_2$ and $n + n_1 + 2n_2$ variables of these problems. For any ideal $J \subset \mathbb{R}[\mathbf{z}]$, we will denote $J^{\mathbf{x}} = J \cap \mathbb{R}[\mathbf{x}]$. The projection of $\mathbb{C}^n \times \mathbb{C}^{n_1+2\,n_2}$ (resp. $\mathbb{C}^n \times \mathbb{C}^{n_1+n_2}$) on $\mathbb{C}^n$ is denoted $\pi^{\mathbf{x}}$.

3.1. The gradient variety.

A natural approach when dealing with constraints in optimization problems is to introduce Lagrangian multipliers. Replacing the inequalities $g^+_i \geq 0$ by the equalities $g^+_i - s^2_i = 0$ (adding new variables s_i) and introducing new parameters for all the equality constraints yields the following minimization problem:

$$\begin{array}{rl} \inf\limits_{(\mathbf{x},\mathbf{u},\mathbf{v},\mathbf{s}) \in \mathbb{R}^n \times \mathbb{R}^{n_1+2\,n_2}} & f(\mathbf{x}) \\ s.t. & \nabla F(\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s}) = 0 \end{array} \tag{3}$$

where $F(\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s}) = f(\mathbf{x}) - \sum_{i=1}^{n_1} u_i g_i(\mathbf{x}) - \sum_{j=1}^{n_2} v_j (g^+_j(\mathbf{x}) - s^2_j)$, $\mathbf{u} = (u_1, ..., u_{n_1})$, $\mathbf{v} = (v_1, ..., v_{n_2})$ and $\mathbf{s} = (s_1, ..., s_{n_2})$.

Definition 3.1. *We define the gradient ideal of $F(\mathbf{z})$:*

$$I_{grad} = (\nabla F(\mathbf{z})) = (F_1, ..., F_n, g^0_1, ..., g^0_{n_1}, g^+_1 - s^2_1, ..., g^+_{n_2} - s^2_{n_2}, v_1 s_1, ..., v_{n_2} s_{n_2}) \subset \mathbb{R}[\mathbf{z}]$$

where $F_i = \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_1} u_j \frac{\partial g^0_j}{\partial x_i} - \sum_{j=1}^{n_2} v_j \frac{\partial g^+_j}{\partial x_i}$. The gradient variety is $V_{grad} = \mathcal{V}(I_{grad})$ and we denote $V^{\mathbf{x}}_{grad} = \overline{\pi^{\mathbf{x}}(V_{grad})}$.

Definition 3.2. *For any $F \in \mathbb{R}[\mathbf{z}]$, the values of F at the (resp. real) points of $\mathcal{V}(\nabla F) = V_{grad}$ are called the (resp. real) critical values of F.*

We easily check the following property:

Lemma 3.3. $F\,|_{V_{grad}} = f\,|_{V_{grad}}$.

Thus minimizing F on V_{grad} is the same as minimizing f on V_{grad}, that is computing the minimal critical value of F.

3.2. **The Karush-Kuhn-Tucker variety.** In the case of a constrained problem, one usually introduce the Karush-Kuhn-Tucker (KKT) constraints:

Definition 3.4. *A point $\mathbf{x}^*$ is called a KKT point if there exists $u_1, \ldots, u_{n_1}, v_1, \ldots, v_{n_2} \in \mathbb{R}$ s.t.*

$$\nabla f(\mathbf{x}^*) - \sum_{i=1}^{n_1} u_i \nabla g_i^0(\mathbf{x}^*) - \sum_{j=0}^{n_2} v_j \nabla g_j^+(\mathbf{x}^*) = 0,\ g_i^0(\mathbf{x}^*) = 0,\ v_j g_j^+(\mathbf{x}^*) = 0.$$

The corresponding minimization problem is the following:

$$\begin{array}{lll} (4) \quad & \inf\limits_{(\mathbf{x},\mathbf{u},\mathbf{v}) \in \mathbb{R}^{n+n_1 n_2}} & f(\mathbf{x}) \\ & s.t. & F_1 = \cdots = F_n = 0 \\ & & g_1^0 = \cdots = g_{n_1}^0 = 0 \\ & & v_1\, g_1^+ = \cdots = v_{n_2}\, g_{n_2}^+ = 0 \\ & & g_1^+ \geq 0, \ldots, g_{n_2}^+ \geq 0 \end{array}$$

where $F_i = \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_1} u_j \frac{\partial g_j^0}{\partial x_i} - \sum_{j=1}^{n_2} v_j \frac{\partial g_j^+}{\partial x_i}$.

This leads to the following definitions:

Definition 3.5. *The Karush-Kuhn-Tucker (KKT) ideal associated to Problem* (1) *is*

$$I_{KKT} = (F_1, ..., F_n, g_1^0, ..., g_{n_1}^0, v_1 g_1^+, ..., v_{n_2} g_{n_2}^+) \subset \mathbb{R}[\mathbf{y}]. \tag{5}$$

The KKT variety is $V_{KKT} = \mathcal{V}(I_{KKT}) \subset \mathbb{C}^n \times \mathbb{C}^{n_1+n_2}$ and the real KKT variety is $V_{KKT}^{\mathbb{R}} = V_{KKT} \cap \mathbb{R}^n \times \mathbb{R}^{n_1+n_2}$. Its projection on $\mathbf{x}$ is $V_{KKT}^{\mathbf{x}} = \overline{\pi^{\mathbf{x}}(V_{KKT})}$, where $\pi^{\mathbf{x}}$ is the projection of $\mathbb{C}^n \times \mathbb{C}^{n_1+n_2}$ onto $\mathbb{C}^n$.

The set of KKT points of S is denoted S_{KKT} and a KKT-minimizer of f on S is a point $\mathbf{x}^* \in S_{KKT}$ such that $f(\mathbf{x}^*) = \min_{\mathbf{x} \in S_{KKT}} f(\mathbf{x})$.

Notice that $V_{KKT}^{\mathbf{x},\mathbb{R}} = \overline{\pi^{\mathbf{x}}(V_{KKT})}^{\mathbb{R}} = \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$, since any linear dependency relation between real vectors can be realized with real coefficients.

The KKT ideal is related to the gradient ideal as follows:

Proposition 3.6. $I_{KKT} = I_{grad} \cap \mathbb{R}[\mathbf{y}]$.

Proof. As $s_i(s_i v_i) + v_i(g_i^+ - s_i^2) = v_i g_i^+\ \forall i = 1, ..., n_2$, we have $I_{KKT} \subset I_{grad} \cap \mathbb{R}[\mathbf{y}]$.

In order to prove the equality, we use the property that if K is a Groebner basis of I_{grad} for an elimination ordering such that $\mathbf{s} \gg \mathbf{x}, \mathbf{u}, \mathbf{v}$ then $K \cap \mathbb{R}[\mathbf{y}]$ is the Groebner basis of $I_{grad} \cap \mathbb{R}[\mathbf{y}]$ (see [3]). Notice that $s_i(s_i v_i) + v_i(g_i^+ - s_i^2) = v_i g_i^+$ $(i = 1, ..., n_2)$ are the only S-polynomials involving the variables $s_1, \ldots, s_{n_2}$ which may have a non-trivial reduction. Thus $K \cap \mathbb{R}[\mathbf{y}]$ is also the Gröbner basis of $F_1, ..., F_n, g_1^0, ..., g_{n_1}^0, v_1 g_1^+, ..., v_{n_2} g_{n_2}^+$ and we have $(K) \cap \mathbb{R}[\mathbf{y}] = I_{grad} \cap \mathbb{R}[\mathbf{y}] = I_{KKT}$. □

The KKT points on S are related to the real points of the gradient variety as follows:

Lemma 3.7. $S_{KKT} = \pi^{\mathbf{x}}(V_{grad}^{\mathbb{R}}) = V_{grad}^{\mathbf{x},\mathbb{R}} \cap \mathcal{S}^{+}(\mathbf{g})$.

Proof. A real point $\mathbf{y} = (\mathbf{x}, \mathbf{u}, \mathbf{v})$ of $V_{KKT}^{\mathbb{R}}$ lifts to a point $\mathbf{z} = (\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s})$ in $V_{grad}^{\mathbb{R}}$, if and only if, $g_i^+(\mathbf{x}) \geq 0$ for $i = 1, \ldots, n_2$. This implies that $V_{KKT}^{\mathbb{R}} = \pi^{\mathbf{y}}(V_{grad}^{\mathbb{R}}) \cap \mathcal{S}^{+}(\mathbf{g})$, which gives by projection the equalities $S_{KKT} = \pi^{\mathbf{x}}(V_{grad}^{\mathbb{R}}) \cap \mathcal{S}^{+}(\mathbf{g}) = \pi^{\mathbf{x}}(V_{grad}^{\mathbb{R}})$ since a point $\mathbf{x}$ of $V_{grad}^{\mathbb{R}}$ satisfies $\mathbf{g}_j^{+}(\mathbf{x}) \geq 0$ for $j \in [1, n_2]$. □

This shows that if a minimizer point of f on S is a KKT point, then it is the projection of a real critical point of F.

3.3. **The Fritz John variety.** A minimizer of f on S is not necessarily a KKT point. More general conditions that are satisfied by minimizers were given by F. John for polynomial non-negativity constraints and further refined for general polynomial constraints [13, 24]. To describe these conditions, we introduce a new variable u_0 and denote by $\mathbf{y}'$ the set of variables $\mathbf{y}' = (\mathbf{x}, u_0, \mathbf{u}, \mathbf{v})$. Let $F_i^{u_0} = u_0 \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_1} u_j \frac{\partial g_j^0}{\partial x_i} - \sum_{j=1}^{n_2} v_j \frac{\partial g_j^+}{\partial x_i}$.

Definition 3.8. *For any $\gamma \subset [1, n_1]$, let*

$$I_{FJ}^{\gamma} = (F_1^{u_0}, ..., F_n^{u_0}, g_1^0, ..., g_{n_1}^0, v_1 g_1^+, ..., v_{n_2} g_{n_2}^+, u_i, i \notin \gamma) \subset \mathbb{R}[\mathbf{y}']. \quad (6)$$

For $m \in \mathbb{N}$, the m^{th} Fritz-John (FJ) ideal associated to Problem (1) *is*

$$I_{FJ}^{m} = \cap_{|\gamma|=m} I_{FJ}^{\gamma}. \quad (7)$$

Let $V_{FJ}^{\gamma} = \mathcal{V}(I_{FJ}^{\gamma}) \subset \mathbb{C}^n \times \mathbb{P}^{n_1+n_2}$. The m^{th} FJ variety is $V_{FJ}^{m} = \mathcal{V}(I_{FJ}^{m}) = \cup_{|\gamma|=m} V_{FJ}^{\gamma}$, and the real FJ variety is $V_{FJ}^{m,\mathbb{R}} = V_{FJ}^{m} \cap \mathbb{R}^n \times \mathbb{RP}^{n_1+n_2}$. Its projection on $\mathbf{x}$ is $V_{FJ}^{m,\mathbf{x}} = \pi^{\mathbf{x}}(V_{FJ}^{m}) = \overline{\pi^{\mathbf{x}}(V_{FJ}^{m})}$. When $m = max_{\mathbf{x} \in S}$ $\text{rank}([\nabla g_1^0(\mathbf{x}), \ldots, \nabla g_{n_1}^0(\mathbf{x})])$, the m^{th} FJ variety will be denoted V_{FJ}.

Notice that this definition slightly differs from the classical one [13, 17, 24], which does not provide any information when the gradient vectors $\nabla g_i^0(\mathbf{x}), i = 1 \ldots n_1$ are linearly dependent on S.

Proposition 3.9. *Any minimizer $\mathbf{x}^*$ of f on S is the projection of a real point of $V_{FJ}^{\mathbb{R}}$*

Proof. The proof is similar to Theorem 4.3.2 of [24]. At a minimizer point $\mathbf{x}^*$ (if it exists) We consider a maximal set of linearly independent gradients $\nabla g_j^0(\mathbf{x}^*)$ for $j \in \gamma$ (with $|\gamma| \leq m$) and apply the same proof as [24][Theorem 4.3.2]. This shows that $\mathbf{x}^* \in V_{FJ}^{\gamma,\mathbb{R}} \subset V_{FJ}^{\mathbb{R}}$. □

Definition 3.10. *We denote by $V_{sing} = V_{FJ} \cap \mathcal{V}(u_0)$ the intersection of V_{FJ} with the hyperplan $u_0 = 0$.*

We easily check that the "affine part" of V_{FJ} corresponding to $u_0 \neq 0$ is the variety V_{KKT}. Thus, we have the decomposition

$$V_{FJ} = V_{sing} \cup V_{KKT},$$

Its projection on $\mathbb{C}^n$ decomposes as

$$V_{FJ}^{\mathbf{x}} = V_{sing}^{\mathbf{x}} \cup V_{KKT}^{\mathbf{x}}. \quad (8)$$

Let us describe more precisely the projection $V_{FJ}^{\mathbf{x}}$ onto $\mathbb{C}^n$. For $\nu = \{j_1, \dots, j_k\} \subset [1, n_2]$, we define

$$\begin{aligned} A_\nu &= [\nabla f, \nabla g_{t_1}^0, \dots, \nabla g_{t_s}^0, \nabla g_{j_1}^+, \dots, \nabla g_{j_k}^+] \\ V_\nu &= \{\mathbf{x} \in \mathbb{C}^n \mid g_1^0(\mathbf{x}) = 0, i = 1 \dots n_1,\ g_j^+(\mathbf{x}) = 0, j \in \nu, \operatorname{rank}(A_\nu) \le m + |\nu|\}. \end{aligned}$$

Let $\Delta_1^\nu, \dots, \Delta_{l_\nu}^\nu$ be polynomials defining the variety $\{\mathbf{x} \in \mathbb{C}^n \mid \operatorname{rank}(A_\nu) \le m + |\nu|\}$. If $n > m+|\nu|$, these polynomials can be chosen as linear combinations of $(m+|\nu|+1)$-minors of the matrix A_ν, as described in [2, 29]. If $n \le m + |\nu|$, we take $l_\nu = 0$, $\Delta_i^\nu = 0$. Let Γ_{FJ} be the union of $\mathbf{g}^0$ and the set of polynomials

$$g_{\nu,i} := \Delta_i^\nu \prod_{j \notin \nu} g_j^+, \tag{9}$$

for $i = 1, \dots, l_\nu, \nu \subset [0, n_2]$.

Lemma 3.11. $V_{FJ}^{\mathbf{x}} = \cup_{\nu \subset [0,n_2]} V_\nu = \mathcal{V}(\Gamma_{FJ})$.

Proof. For any $\mathbf{x} \in \mathbb{C}^n$, let $\nu(\mathbf{x}) = \{j \in [1, n_2] \mid g_j^+(\mathbf{x}) = 0\}$.

Let $\mathbf{y}'$ be a point of V_{FJ}, $\mathbf{x}$ its projection on $\mathbb{C}^n$ and $\nu(\mathbf{x}) = \nu = \{j_1, \dots, j_k\}$. We have $g_j^+(\mathbf{x}) \ne 0$, $v_j = 0$ for $j \notin \nu$ and $\Delta_i^\nu = 0$ for $i = 1, \dots, l_\nu$. This implies that $\operatorname{rank}(A_\nu(\mathbf{x})) \le m + |\nu|$ and there exists $(u_0, u_1, \dots, u_{n_1}, v_1, \dots, v_{n_2}) \ne 0$ and $\gamma \subset [1, n_1]$ of size $|\gamma| \le m$ such that

$$u_0 \nabla f + u_1 \nabla g_1^0 + \dots + u_{n_1} \nabla g_{n_1}^0 + v_1 \nabla g_{j_1}^+ + \dots + v_{n_2} \nabla g_{j_k}^+ = 0,$$

with $u_i = 0$, $i \notin \gamma \subset [1, n_1]$. Therefore $\mathbf{x} \in \pi^{\mathbf{x}}(V_{FJ})$, which proves that $\mathcal{V}(\mathbf{g}^0, g_{\nu,i}, \nu \subset [0, n_2], i = 1 \dots l_\nu) \subset \pi^{\mathbf{x}}(V_{FJ})$.

Conversely, if $\mathbf{x} \in \pi^{\mathbf{x}}(V_{FJ})$ then $\mathbf{x} \in V_{\nu(\mathbf{x})} \subset \cup_\nu V_\nu$ which is defined by the polynomials $g_1^0, \dots, g_{n_1}^0$ and $g_{\nu,i} := \Delta_i^\nu \prod_{j \notin \nu} g_j^+$, for $i = 1, \dots, l_\nu, \nu \subset [0, n_2]$. □

Remark 3.12. *The real variety* $\pi^{\mathbf{x}}(V_{FJ}^{\mathbb{R}}) = V_{FJ}^{\mathbf{x}} \cap \mathbb{R}^n$ *can also be defined by the union of* $\mathbf{g}^0$ *and the set* Φ_{FJ} *of polynomials*

$$g_\nu := \Delta^\nu \prod_{j \notin \nu} g_j^+ \text{ where } \Delta^\nu = \det(A_\nu A_\nu^T), \tag{10}$$

for $\nu \subset [0, n_2]$ *and* $n > m + |\nu|$*, as described in* [9].

Similarly the projection $V_{sing}^{\mathbf{x}}$ onto $\mathbb{C}^n$ can be described as follows. For $\nu = \{j_1, \dots, j_k\} \subset [1, n_2]$,

$$\begin{aligned} B_\nu &= [\nabla g_{t_1}^0, \dots, \nabla g_{t_s}^0, \nabla g_{j_1}^+, \dots, \nabla g_{j_k}^+] \\ W_\nu &= \{\mathbf{x} \in \mathbb{C}^n \mid g_1^0(\mathbf{x}) = 0, i = 1 \dots n_1,\ g_j^+(\mathbf{x}) = 0, j \in \nu, \operatorname{rank}(B_\nu) \le m + |\nu| - 1\}. \end{aligned}$$

Let $\Theta_1^\nu, \dots, \Theta_{l_\nu}^\nu$ be polynomials defining the variety $\{\mathbf{x} \in \mathbb{C}^n \mid \operatorname{rank}(B_\nu) \le m + |\nu| - 1\}$ and let Γ_{sing} be the union of $\mathbf{g}^0$ and the set of polynomials

$$\sigma_{\nu,i} := \Theta_i^\nu \prod_{j \notin \nu} g_j^+, \tag{11}$$

for $\nu \subset [0, n_2], i = 1 \dots l_\nu$.

We similar arguments, we prove the following

Lemma 3.13. $V_{sing}^{\mathbf{x}} = \cup_{\nu \subset [0,n_2]} W_\nu = \mathcal{V}(\Gamma_{sing})$.

3.4. **The minimizer variety.** By the decomposition (8) and Proposition 3.9, we known that the minimizer points of f on S are in

$$S_{FJ} = S_{KKT} \cup S_{sing} \tag{12}$$

where $S_{FJ} = \pi^{\mathbf{x}}(V_{FJ}^{\mathbb{R}}) \cap S = \pi^{\mathbf{x}}(V_{FJ}^{\mathbb{R}}) \cap \mathcal{S}^+(\mathbf{g})$, $S_{KKT} = \pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}}) \cap S = \pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}}) \cap \mathcal{S}^+(\mathbf{g})$, $S_{sing} = \pi^{\mathbf{x}}(V_{sing}^{\mathbb{R}}) \cap S = \pi^{\mathbf{x}}(V_{sing}^{\mathbb{R}}) \cap \mathcal{S}^+(\mathbf{g})$. Therefore, we can decompose the initial optimization problem (1) into two subproblems:

(1) find the infimum of f on S_{KKT};
(2) find the infimum of f on S_{sing};

and take the least of these two infima. Since the second problem is of the same type as (1) but with the additional constraints $\sigma_{i,\nu} = 0$ described in (11), we will analyze only the first subproblem. The approach developed for this first sub-problem will be applied recursively to the second subproblem, in order to obtain the solution of Problem (1).

Definition 3.14. *We define the KKT-minimizer set and ideal of f on S as:*

$$\begin{aligned} S_{min} &= \{\mathbf{x}^* \in S_{KKT} \text{ s.t. } \forall \mathbf{x} \in S_{KKT}, f(\mathbf{x}^*) \le f(\mathbf{x})\} \\ I_{min} &= \mathcal{I}(S_{min}) \subset \mathbb{R}[\mathbf{x}]. \end{aligned}$$

A point $\mathbf{x}^*$ in S_{min} is called a KKT-minimizer. Notice that $I_{KKT} \subset I_{min}$ and that I_{min} is a real radical ideal.

We have $I_{min} \neq (1)$, if and only if, the KKT-minimum f^* is reached in S_{KKT}.

If $n_1 = n_2 = 0$, I_{min} is the vanishing ideal of the *critical points* $\mathbf{x}^*$ of f (satisfying $\nabla f(\mathbf{x}^*) = 0$) where $f(\mathbf{x}^*)$ reaches its minimal critical value.

Remark 3.15. *If we take $f = 0$ in the minimization problem (1), then all the points of S are KKT-minimizers and $I_{min} = \mathcal{I}(S) = \sqrt[\mathbf{g}^+]{\mathbf{g}^0}$. Moreover, $I_{KKT} \cap \mathbb{R}[\mathbf{x}] = (g_1^0, \ldots, g_{n_1}^0) = (\mathbf{g}^0)$ since $F_1, \ldots, F_n, v_1 g_1^+, \ldots, v_{n_2} g_{n_2}^+$ are homogeneous of degree 1 in the variables $\mathbf{u}, \mathbf{v}$.*

4. Representation of positive polynomials

In this section, we are going to analyze the decomposition of polynomials as sum of squares modulo the gradient ideal. Hereafter, J_{grad} is an ideal of $\mathbb{R}[\mathbf{z}]$ such that $\mathcal{V}(J_{grad}) = V_{grad}$ and G is a set of constraints in $\mathbb{R}[\mathbf{x}]$ such that $\mathcal{S}^+(G) = \mathcal{S}^+(\mathbf{g})$.

The first steps consists in decomposing V_{grad} in components on which f has a constant value. We recall here a result, which also appears (with slightly different hypotheses) in [32, Lemma 3.3][2].

Lemma 4.1. *Let $f \in \mathbb{R}[\mathbf{x}]$ and let V be an irreducible subvariety contained in $\mathcal{V}^{\mathbb{C}}(\nabla f)$. Then $f(x)$ is constant on V.*

Proof. If V is irreducible in the Zariski topology induced from $\mathbb{C}[\mathbf{x}]$, then it is connected in the strong topology on $\mathbb{C}^n$ and even piecewise smoothly path-connected [37]. Let x, y be two arbitrary points of V. There exists a piecewise smooth path $\varphi(t)$ $(0 \le t \le 1)$ lying inside V such that $x = \varphi(0)$ and $y = \varphi(1)$. Without loss of generality, we can assume that φ is smooth between x and y in order to prove that $f(x) = f(y)$. By the Mean Value Theorem, it holds that for some $t_1 \in (0, 1)$

$$\mathrm{Re}(f(y) - f(x)) = \mathrm{Re}(f(\varphi(t)))'(t_1) = \mathrm{Re}((\nabla f(\varphi(t_1)) * \varphi'(t_1))) = 0$$

[2] In its proof, the Mean Value Theorem is applied for a complex valued function, which is not valid. We correct the problem in the proof of Lemma 4.1.

since ∇f vanishes on V. Then $\mathrm{Re}(f(y)) = \mathrm{Re}(f(x))$. We have the same result for the imaginary part: for some $t_2 \in (0,1)$

$$\mathrm{Im}(f(y)) - \mathrm{Im}(f(x)) = \mathrm{Im}(f(\varphi(t)))'(t_2) = \mathrm{Im}((\nabla f(\varphi(t_2)) * \varphi'(t_2))) = 0$$

since ∇f vanishes on V. Then $\mathrm{Im}(f(y)) = \mathrm{Im}(f(x))$. We conclude that $f(y) = f(x)$ and hence f is constant on V. □

Lemma 4.2. *The ideal J_{grad} can be decomposed as $J_{grad} = J_0 \cap J_1 \cap \cdots \cap J_s$ with $V_i = \mathcal{V}(J_i)$ and $W_i = \overline{\pi^{\mathbf{x}}(V_i)}$ where $\pi^{\mathbf{x}}(V_i)$ is the projection of V_i on $\mathbb{C}^n$ such that*

- $f(V_j) = f_j \in \mathbb{C}$, $f_i \neq f_j$ *if* $i \neq j$,
- $W_i^{\mathbb{R}} \cap \mathcal{S}^+(G) \neq \emptyset$ *for* $i = 0, \ldots, r$,
- $W_i^{\mathbb{R}} \cap \mathcal{S}^+(G) = \emptyset$ *for* $i = r+1, \ldots, s$,
- $f_0 < \cdots < f_r$.

Proof. Consider a minimal primary decomposition of J_{grad}:

$$J_{grad} = Q_0 \cap \cdots \cap Q_{s'},$$

where Q_i is a primary component, and $\mathcal{V}(Q_i)$ is an irreducible variety in $\mathbb{C}^{n+n_1+2n_2}$ included in V_{grad}. By Lemma 4.1, F is constant on $\mathcal{V}(Q_i)$. By Lemma 3.3, it coincides with f on each variety $\mathcal{V}(Q_i)$. We group the primary components Q_i according to the values $f_0, \ldots, f_s$ of f on these components, into $J_0, \ldots, J_s$ so that $f(\mathcal{V}(J_j)) = f_i$ with $f_i \neq f_j$ if $i \neq j$.

We can number them so that $\overline{\pi^{\mathbf{x}}(V_i)}^{\mathbb{R}} \cap \mathcal{S}^+(G)$ is empty for $i = r+1, \ldots, s$ and contains a real point $\mathbf{x}_i$ for $i = 0, \ldots, r$. Notice that such a point $\mathbf{x}_i$ is in $\mathcal{S}$, since it satisfies $g^0(\mathbf{x}_i) = 0\ \forall g^0 \in G^0$ and $g^+(\mathbf{x}_i) \geq 0\ \forall g^+ \in G^+$. As it is the limit of the projection of points in $\mathcal{V}(J_i)$ on which f is constant, we have $f_i = f(\mathbf{x}_i) \in \mathbb{R}$ for $i = 0, \ldots, r$. We can then order $J_0, \ldots, J_r$ so that $f_0 < \cdots < f_r$. □

Remark 4.3. *If the minimum of f on S is reached at a KKT-point, then we have $f_0 = \min_{\mathbf{x} \in S} f(\mathbf{x})$.*

Remark 4.4. *If $V_{grad}^{\mathbb{R}} = \emptyset$, then for all $i = 0, \ldots, s$, $W_i^{\mathbb{R}} \cap \mathcal{S}^+(G) = \emptyset$ and by convention, we take $r = -1$.*

Lemma 4.5. *There exist $p_0, \ldots, p_s \in \mathbb{C}[\mathbf{x}]$ such that*

- $\sum_{i=0}^{s} p_i = 1 \mod J_{grad}$,
- $p_i \in \bigcap_{j \neq i} J_j$,
- $p_i \in \mathbb{R}[\mathbf{x}]$ *for* $i = 0, \ldots, r$.

Proof. Let $(L_i)_{i=0,\ldots,s}$ be the univariate Lagrange interpolation polynomials at the values $f_0, \ldots, f_s \in \mathbb{C}$ and let $q_i(\mathbf{x}) = L_i(f(\mathbf{x}))$.

The polynomials q_i are constructed so that

- $q_i(V_j) = 0$ if $j \neq i$,
- $q_i(V_i) = 1$,

where $V_i = \mathcal{V}(J_i)$. As the set $\{f_{r+1}, \ldots, f_s\}$ is stable by conjugation and $f_0, \ldots, f_r \in \mathbb{R}$, by construction of the Lagrange interpolation polynomials we deduce that $q_0, \ldots, q_r \in \mathbb{R}[\mathbf{x}]$.

By Hilbert's Nullstellensatz, there exists $N \in \mathbb{N}$ such that $q_i^N \in \bigcap_{j \neq i} J_j$. As $\sum_{j=0}^{s} q_j^N = 1$ on V_{grad} and $q_i^N q_j^N = 0 \mod \bigcap_i J_i = J_{grad}$ for $i \neq j$, we deduce that there exists $N' \in \mathbb{N}$

such that

$$\begin{aligned} 0 &= (1-\sum_{j=0}^{s} q_j^N)^{N'} \mod J_{grad} \\ &= 1-\sum_{j=0}^{s}(1-(1-q_j^N)^{N'}) \mod J_{grad}. \end{aligned}$$

As the polynomial $p_i = 1-(1-q_j^N)^{N'} \in \mathbb{C}[\mathbf{x}]$ is divisible by q_j^N, it belongs to $\bigcap_{j\neq i} J_j$. Since $q_j \in \mathbb{R}[\mathbf{x}]$ for $j=0,\ldots,r$, we have $p_j \in \mathbb{R}[\mathbf{x}]$ for $j=0,\ldots,r$, which ends the proof of this lemma. □

Lemma 4.6. $-1 \in \mathcal{P}^+(G) + (\bigcap_{i>r} J_i^{\mathbf{x}})$.

Proof. As $\bigcup_{i>r} \overline{\pi^{\mathbf{x}}(V_i)}^{\mathbb{R}} \cap \mathcal{S}^+(G) = \mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i \cap \mathbb{R}[\mathbf{x}]) \cap \mathcal{S}^+(G) = \mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i^{\mathbf{x}}) \cap \mathcal{S}^+(G) = \emptyset$, we have $\mathcal{I}(\mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i^{\mathbf{x}}) \cap \mathcal{S}^+(G)) = \mathbb{R}[\mathbf{x}] \ni 1$ and by the Positivstellensatz (Theorem 2.6 (iii)),

$$-1 \in \mathcal{P}^+(G) + (\bigcap_{i>r} J_i^{\mathbf{x}}).$$

□

Corollary 4.7. *If $S_{min} = \emptyset$, then $-1 \in \mathcal{P}^+(G) + J_{grad}^{\mathbf{x}}$.*

Proof. If $S_{min} = \emptyset$, then f has no real KKT critical value on $S(G)$ and $r=-1$. Lemma 4.6 implies that $-1 \in \mathcal{P}^+(G) + (\bigcap_{i=0}^{s} J_i^{\mathbf{x}}) = \mathcal{P}^+(G) + J_{grad}^{\mathbf{x}}$. □

In this case, $\forall p \in \mathbb{R}[\mathbf{x}]$, $p = \frac{1}{4}((p+1)^2 - (p-1)^2) \in \mathcal{P}^+(G) + J_{grad}^{\mathbf{x}}$. If G^0 is chosen such that $V(G^0) \subset V_{grad}^{\mathbf{x}}$ then $S_{min} = \emptyset$ if and only if $-1 \in \mathcal{P}(G)$.

We recall another useful result on the representation of positive polynomials (see for instance [5]):

Lemma 4.8. *Let $J \subset \mathbb{R}[\mathbf{z}]$ and $V = \mathcal{V}(J)$ such that $f(V) = f^*$ with $f^* \in \mathbb{R}^+$. There exists $t \in \mathbb{N}$, s.t. $\forall \epsilon > 0$, $\exists\, q \in \mathbb{R}[\mathbf{x}]$ with $\deg(q) \le t$ and $f + \epsilon = q^2 \mod J$.*

Proof. We know that $\frac{f+\epsilon}{f^*+\epsilon} - 1$ vanishes on V. By Hilbert's Nullstellensatz $(\frac{f+\epsilon}{f^*+\epsilon} - 1)^l \in J$ for some $l \in \mathbb{N}$. From the binomial theorem, it follows that

$$(1+(\frac{f+\epsilon}{f^*+\epsilon}-1))^{1/2} \equiv \sum_{i}^{l-1} \binom{1/2}{k} (\frac{f+\epsilon}{f^*+\epsilon}-1)^k \stackrel{def}{=} \frac{q}{\sqrt{f^*+\epsilon}} \; mod\ J$$

Then $f+\epsilon = q^2\ mod\ J$. □

In particular, if $f^* > 0$ this lemma implies that $f = (f - \frac{1}{2}f^*) + \frac{1}{2}f^* = q^2 \mod J$ for some $q \in \mathbb{R}[\mathbf{x}]$.

Theorem 4.9. *Let $G \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that $\mathcal{S}^+(G) = \mathcal{S}^+(\mathbf{g})$, let $f \in \mathbb{R}[\mathbf{x}]$, let $f_0 < \cdots < f_r$ be the real KKT critical values of f on S and let $p_0,\ldots,p_r$ be the associated polynomials defined in Lemma 4.5.*

(1) $f - \sum_{i=0}^{r} f_i\, p_i^2 \in \mathcal{P}^+(G) + \sqrt{J_{grad}^{\mathbf{x}}}$.

(2) *If $f \ge 0$ on S_{KKT}, then $f \in \mathcal{P}^+(G) + \sqrt{J_{grad}^{\mathbf{x}}}$.*

(3) *If $f > 0$ on S_{KKT}, then $f \in \mathcal{P}^+(G) + J_{grad}^{\mathbf{x}}$.*

Proof. By Lemma 4.5, we have

$$1 = (\sum_{i=0}^{s} p_i)^2 = \sum_{i=0}^{s} p_i^2 \mod J_{grad}.$$

Thus $f = \sum_{i=0}^{s} f\, p_i^2 \mod J_{grad}$.

By Lemma 4.6, $-1 \in \mathcal{P}^+(G) + (\bigcap_{j>r} J_j^{\mathbf{x}})$ so that $f = \frac{1}{4}((f+1)^2 - (f-1)^2) \in \mathcal{P}^+(G) + \cap_{j>r} J_j^{\mathbf{x}}$ and

$$\sum_{i>r} f\, p_i^2 \in \mathcal{P}^+(G) + \bigcap_{j=0}^{s} J_j^{\mathbf{x}} = \mathcal{P}^+(G) + J_{grad}^{\mathbf{x}}. \tag{13}$$

As the polynomial $(f - f_i)\, p_i^2$ vanishes on V_{grad}, we deduce that

$$f = \sum_{i=0}^{r} f_i\, p_i^2 + \sum_{i=r+1}^{s} f\, p_i^2 + \sqrt{J_{grad}^{\mathbf{x}}} = \sum_{i=0}^{r} f_i\, p_i^2 + \mathcal{P}^+(G) + \sqrt{J_{grad}^{\mathbf{x}}},$$

which proves the first point.

If $f \geq 0$ on S_{KKT}, then $f_i \geq 0$ for $i = 0, \ldots, r$ and $\sum_{i=0}^{r} f_i\, p_i^2 \in \mathcal{P}^+(G)$ so that

$$f \in \mathcal{P}^+(G) + \sqrt{J_{grad}^{\mathbf{x}}},$$

which proves the second point.

If $f > 0$ on S_{KKT} by Lemma 4.8, we have $f p_i^2 = q_i^2 \mod J_{grad}^{\mathbf{x}}$ with $q_i \in \mathbb{R}[\mathbf{x}]$, which shows that

$$\sum_{i=0}^{r} f_i\, p_i^2 = \sum_{i=0}^{r} q_i^2 \mod J_{grad}^{\mathbf{x}}$$

Therefore, $\sum_{i=0}^{r} f_i\, p_i^2 \in \mathcal{P}^+(G) + J_{grad}^{\mathbf{x}}$ and $f \in \mathcal{P}^+(G) + J_{grad}^{\mathbf{x}}$ by (13), which proves the third point. □

This theorem involves only polynomials in $\mathbb{R}[\mathbf{x}]$ and the points (2) and (3) generalize results of [5] on the representation of positive polynomials.

Let us give now a refinement of Theorem 4.9 with a control of the degrees of the polynomials involved in the representation of f as an element of $\mathcal{P}^+(G) + J_{grad}^{\mathbf{x}}$.

Theorem 4.10. *Let $G \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that $\mathcal{V}(G^0) \subset V_{grad}^{\mathbf{x}}$ and $\mathcal{S}^+(G) = \mathcal{S}^+(\mathbf{g})$. If $f \geq 0$ on S_{KKT}, then there exists t_0 such that $\forall \epsilon > 0$,*

$$f + \epsilon \in \mathcal{P}_{t_0}(G).$$

Proof. Let $J_{grad} = (G^0) \cap I_{grad} \subset \mathbb{R}[\mathbf{z}]$, so that $\mathcal{V}(J_{grad}) = V_{grad}$ since $\mathcal{V}(G^0) \subset V_{grad}^{\mathbf{x}}$. Using the decomposition (13) obtained in the proof of Theorem 4.9, we can choose $t_0' \in \mathbb{N}$ and $t_0 \geq t_0' \in \mathbb{N}$ big enough such that $\deg(p_i) \leq t_0/2$ and

$$\sum_{i>r} f\, p_i^2 \in \mathcal{P}_{t_0'}^+(G) + \ J_{grad} \cap \mathbb{R}[\mathbf{x}]_{t_0'} \subset \mathcal{P}_{t_0}(G),$$

since $J_{grad}^{\mathbf{x}} = (G^0) \cap I_{grad}^{\mathbf{x}} \subset (G^0)$. Then $\forall \epsilon > 0$,

$$\sum_{i>r} (f + \epsilon)\, p_i^2 = \sum_{i>r} f\, p_i^2 + \sum_{i>r} \epsilon\, p_i^2 \in \mathcal{P}_{t_0}(G). \tag{14}$$

As $\forall\ \epsilon > 0,\ f + \epsilon > 0$ on S_{KKT}, i.e, $f_i + \epsilon > 0$ for $i = 0, \ldots, r$, we deduce from Lemma 4.8 that if t_0 is big enough, we have

$$(f+\epsilon)\, p_i^2 = q_i^2 \mod G^0_{\langle t_0\rangle} \cap \mathbb{R}[\mathbf{x}] \tag{15}$$

with $deg(q_i) \le t_0/2$ for $i = 0, \ldots, r$.

Since $1 - \sum_{i=0}^{s} p_i^2 = 0 \mod (G^0)$, we can choose t_0 big enough so that

$$(f+\epsilon) - \sum_{i=0}^{s} (f+\epsilon)\, p_i^2 \in G^0_{\langle t_0\rangle} \cap \mathbb{R}[\mathbf{x}]. \tag{16}$$

From Equations (14), (15), (16), we deduce that if $t_0 \in \mathbb{N}$ is big enough, $\forall \epsilon > 0$

$$f + \epsilon \in \mathcal{P}_{t_0}(G),$$

which concludes the proof of the theorem. □

5. Finite convergence

In this section, we show that the sequence of relaxation problems attains its limit in a finite number of steps and that the minimizer ideal can be recovered from an optimal solution of the corresponding relaxation problem. We will use hereafter the following notation:

- $f^* = \inf_{\mathbf{x} \in S_{KKT}} f(\mathbf{x})$
- $S_{min} = \{\mathbf{x}^* \in S_{KKT} \mid f(\mathbf{x}^*) = f^*\}$

We first show that $S_{min} = \emptyset$ can be detected from a adapted relaxation sequence:

Proposition 5.1. *Let $G = (G^0; G^+)$ be a set of constraints of $\mathbb{R}[\mathbf{x}]$, such that $S_{min} \subset \mathcal{S}(G)$ and $\mathcal{V}(G^0) \subset V^{\mathbf{x}}_{KKT}$ and $G^+ = \mathbf{g}^+$. Then $S_{min} = \emptyset$, if and only if, there $t_0 \in \mathbb{N}$ such that $\forall t \ge t_0$, $\mathcal{L}_t(G) = \emptyset$.*

Proof. Let $J_{grad} = (G^0) \cap I_{grad}$ and let G' be a set of constraints such that $(G'^0) = J_{grad} \cap \mathbb{R}[\mathbf{x}] = J^{\mathbf{x}}_{grad}$ and $G'^+ = \mathbf{g}^+$ be a finite set. By hypothesis, $\mathcal{V}(J_{grad}) = V_{grad}$. We deduce from Corollary 4.7 that if $S_{min} = \emptyset$, then

$$-1 \in \mathcal{P}^+(G') + (G'^0) \subset \mathcal{P}(G) = \cup_{t \in \mathbb{N}} \mathcal{P}_t(G).$$

Thus there exists t_0 such that $-1 \in \mathcal{P}_t(G)$ for $t \ge t_0$, which implies that $\mathcal{L}_t(G) = \emptyset$, since if there exists $\Lambda \in \mathcal{L}_t(G)$, then $\Lambda(1) = 1$ and $\Lambda(-1) \ge 0$.

Conversely, suppose that $S_{min} \neq \emptyset$ contains a point $\mathbf{x}^*$. As $S_{min} \subset \mathcal{S}(G)$, for all $t \in \mathbb{N}$ the evaluation $\underline{\mathbf{1}}_{\mathbf{x}^*}$ at $\mathbf{x}^*$ restricted to $\mathbb{R}[\mathbf{x}]_{2t}$ is an element of $\mathcal{L}_t(G) \neq \emptyset$. □

This proposition gives a way to check whether $S_{min} = \emptyset$, using the relaxation sequence $\mathcal{L}_t(G)$. We are now going to analyse the case where f has KKT minimizers on S.

From now on, we assume that $S_{min} \neq \emptyset$.

First, we recall a property similar to [18][Claim 4.7]:

Proposition 5.2. *Let $G = (G^0; G^+)$ be a set of constraints of $\mathbb{R}[\mathbf{x}]$. There exists $t_0 \in \mathbb{N}$ such that $\forall t \ge t_0$, $\forall \Lambda \in \mathcal{L}_t(G)$, $\sqrt[G^+]{G^0} \subset (\ker M^t_\Lambda)$.*

Proof. Let $G^0 = \{g_1, \ldots, g_l\}$ and let $q_1, \ldots, q_k$ be generators of $J := \sqrt[G^+]{G^0}$. By the Positivestllensatz, for $j \in 1, \ldots, k$, there exist $m_j \in \mathbb{N}^*$ and polynomials $u_r^{(j)} \in \mathbb{R}[\mathbf{x}]$ and $\sigma_j \in \mathcal{P}^+(G)$ such that

$$q_j^{2m_j} + \sigma_j = \sum_{r=1}^{l} u_r^{(j)} g_r.$$

Let us take $t_0 \in \mathbb{N}$ big enough such that $u_r^{(j)} g_r \in G_{\langle t_0 \rangle}$ and $\sigma_j \in \mathcal{P}_{t_0}^+(G)$. Then for all $t \geq t_0$ and all $\Lambda \in \mathcal{L}_t(G)$, we have $\Lambda(u_r^{(j)} g_r) = 0$, $\Lambda(q_j^{2m_j}) \geq 0$, $\Lambda(\sigma_j) \geq 0$ and $\Lambda(q_j^{2m_j}) + \Lambda(\sigma_j) = 0$, which implies that $\Lambda(q_j^{2m_j}) = 0$ and $q_j \in \ker M_\Lambda^t$. This proves that $(q_1, \ldots, q_l) = J \subset (\ker M_\Lambda^t)$. □

Remark 5.3. *With the same arguments, we can show that for any $t' \in \mathbb{N}$, there exists $t'_0 \geq t'$ such that $\forall t \geq t'_0$, $\forall \Lambda \in \mathcal{L}_t(G)$,*

$$Q_{\langle t' \rangle} \subset \ker M_\Lambda^t,$$

where $Q = \{q_1, \ldots, q_k\}$ generates $J = \sqrt[G^+]{G^0}$.

The next result shows that in the sequence of optimization problems that we consider, the minimum of f on S_{KKT} is reached from some degree.

Theorem 5.4. *Let G be a set of constraints of $\mathbb{R}[\mathbf{x}]$ such that $S_{min} \subset \mathcal{S}(G) \subset V_{KKT}^{\mathbf{x},\mathbb{R}}$. There exists $t_1 \geq 0$ such that $\forall t \geq t_1$,*

1. *$f_{t,G}^\mu = f^*$ is reached for some $\Lambda^* \in \mathcal{L}_t(G)$,*
2. *$\forall \Lambda^* \in \mathcal{L}_t(G)$ with $\Lambda^*(f) = f_{t,G}^\mu = f^*$, we have $p_i \in \ker M_{\Lambda^*}^t$, $\forall i = 1, \ldots, r$,*
3. *if $\mathcal{V}(G^0) \subset V_{KKT}^{\mathbf{x}}$ then $f_{t,G}^{sos} = f_{t,G}^\mu = f^*$.*

Proof. By Theorem 4.9(1) applied to $f - f^*$, we can write

$$f - f^* \equiv \sum_{i=1}^r (f_i - f^*)\, p_i^2 + h + g.$$

with $h \in \mathcal{P}^+(G)$ and $g \in \sqrt{I_{grad}} \cap \mathbb{R}[\mathbf{x}] = \sqrt{I_{KKT}} \cap \mathbb{R}[\mathbf{x}] \subset \sqrt[\mathbb{R}]{I_{KKT}} \cap \mathbb{R}[\mathbf{x}]$ (by Proposition 3.6). Since $\mathcal{S}(G) \subset V_{KKT}^{\mathbf{x},\mathbb{R}} = \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$, we have $\sqrt[\mathbb{R}]{I_{KKT}} \cap \mathbb{R}[\mathbf{x}] \subset \mathcal{I}(\mathcal{S}(G)) = \sqrt[G^+]{(G^0)}$ by the Positivstellensatz. We deduce that $g \in \sqrt[G^+]{(G^0)}$. By proposition 5.2, there exists $t_1 \geq t_0$ such that for all $t \geq t_1$, for all $\Lambda \in \mathcal{L}_t(G)$, $\Lambda(g) = 0$, $\Lambda(h) \geq 0$.

Let us fix $t \geq t_1$ and $\Lambda^* \in \mathcal{L}_t(G)$ such that $\Lambda^*(f) = f_{t,G}^\mu$. Then

$$\Lambda^*(f - f^*) = \sum_{i=1}^r (f_i - f^*) \Lambda^*(p_i^2) + \Lambda^*(h).$$

As $f_i - f^* = f_i - f_0 > 0$, $\Lambda^*(p_i^2) \geq 0$ and $\Lambda^*(h) \geq 0$ ($h \in \mathcal{P}_t^+(G)$), we deduce that $\Lambda^*(f - f^*) = \Lambda^*(f) - f^* \geq 0$.

As $\emptyset \neq S_{min} \subset \mathcal{S}(G)$, we have $\Lambda^*(f) \leq f^*$ (by Remark 2.8), so that $\Lambda^*(f) = f_{t,G}^\mu = f^*$, which proves the first point. Hence for $i = 1, \ldots, r$, $\Lambda^*(p_i^2) = 0$ and $p_i \in \ker M_{\Lambda^*}^t$, which proves the second point.

To prove that $f_{t,G}^{sos} = f^*$ when $\mathcal{V}(G^0) \subset V_{KKT}^{\mathbf{x}}$, we apply Theorem 4.10 to $f - f^*$ which is positive on S_{KKT}. Let us take $J_{grad} = (G^0) \cap I_{grad} \subset \mathbb{R}[\mathbf{z}]$. We denote by $\tilde{G}$ the set of constraints such that $\tilde{G}^0$ is a finite family of generators of $J_{grad} \cap \mathbb{R}[\mathbf{x}]$ and $\tilde{G}^+ = G^+$.

By Theorem 4.10, there exists t_0 such that $\forall \epsilon > 0$,

$$f - f^* + \epsilon \in \mathcal{P}_{t_0}(\tilde{G}).$$

As $(\tilde{G}^0) = (G^0) \cap I_{grad} \subset (G^0)$, we can choose $t_1 \geq t_0$ such that $\tilde{G}_{\langle t_0 \rangle} \subset G_{\langle t_1 \rangle}$ and $\mathcal{P}_{t_0}(\tilde{G}) \subset \mathcal{P}_{t_1}(G)$.

Then $\forall t \geq t_1$, $f - f^* + \epsilon \in \mathcal{P}_t(G)$. Hence by maximality, $\forall \epsilon > 0, f^* - \epsilon \leq f_{t,G}^{sos}$. We deduce that $f^* \leq f_{t,G}^{sos}$, which implies that $f_{t,G}^{sos} = f_{t,G}^\mu = f^*$ and proves the third point. □

As for the construction of generators of $\sqrt[G+]{I_{KKT}}$ (Proposition 5.2), we can construct generators of I_{min} from the kernel of a truncated Hankel operator associated to any linear form which minimizes f, using the following propositions:

Proposition 5.5. $I_{min} = (p_1, ..., p_r) + \sqrt[G+]{I^{\mathbf{x}}_{KKT}}$.

Proof. First of all, we proof that $I^{\mathbf{z}}_{min} = (p_1, ..., p_r) + \sqrt[G+]{I_{grad}} = (p_1, ..., p_r) + \sqrt[\mathbb{R}]{I_{grad}}$. Using the decomposition of Lemma 4.2 and the polynomials p_i of Lemma 4.5, we have

$$V^{\mathbb{R}}_{grad} = (V_0 \cup V_1 \cup \cdots \cup V_s) \cap \mathbb{R}^{n+n_1+2\,n_2} = V^{\mathbb{R}}_0 \cup \cdots \cup V^{\mathbb{R}}_r,$$

By construction, $\mathcal{I}(V^{\mathbb{R}}_0) = I_{min}$, $p_i(V^{\mathbb{R}}_0) = 0$ for $i = 1, \ldots, s$ and $p_i \in \mathbb{R}[\mathbf{x}]$ for $i = 0, \ldots, r$. This implies that $p_i \in I_{min}$ for $i = 1, \ldots, r$.

As $V^{\mathbb{R}}_0 \subset V^{\mathbb{R}}_{grad}$, we also have $\sqrt[G+]{I_{grad}} \subset I_{min}$.

We have proved so far that $(p_1, \ldots, p_r) + \sqrt[G+]{I_{grad}} \subset I_{min}$. In order to prove the reverse inclusion, we denote by $q_1, \ldots, q_m$ a family of generators of the ideal I_{min}. Take one of these generators q_j $(1 \le j \le m)$. By construction, $q_j\, p_0(V^{\mathbb{R}}_0) = 0$ and $q_j p_0(V^{\mathbb{R}}_i) = 0$ for $i = 1, \ldots, r$, which implies that $q_j p_0 \in \sqrt[G+]{I_{grad}}$.

By Lemma 4.5, we have the decomposition

$$q_j \equiv q_j(p_0 + p_1 + \cdots + p_s) \mod I_{grad} \subset \sqrt[G+]{I_{grad}}.$$

Moreover $(p_{r+1} + \cdots + p_s) \in \mathbb{R}[\mathbf{z}]$ and vanishes on $V^{\mathbb{R}}_k$ for $k = 0, \ldots, r$. Thus $(p_{r+1} + \cdots + p_s) \in \sqrt[G+]{I_{grad}}$ and we deduce that $q_j \in (p_1, \ldots, p_r) + \sqrt[G+]{I_{grad}}$. This proves the other inclusion and the first equality.

As $V^{\mathbb{R}}_{grad} = V^{\mathbb{R}}_{grad} \cap \mathcal{S}^+(G)$ (Remark 3.7), by the Positivstellensatz, $\sqrt[G+]{I_{grad}} = \sqrt[\mathbb{R}]{I_{grad}}$, which proves the second equality.

By the Positivstellensatz and Remark 3.7, we have

$$\sqrt[G+]{I_{grad}} \cap \mathbb{R}[\mathbf{x}] = \sqrt[\mathbb{R}]{I_{grad}} \cap \mathbb{R}[\mathbf{x}] = \mathcal{I}(\pi^{\mathbf{x}}(V^{\mathbb{R}}_{grad})) = \mathcal{I}(\pi^{\mathbf{x}}(V_{KKT})^{\mathbb{R}} \cap \mathcal{S}^+(G)) = \sqrt[G+]{I^{\mathbf{x}}_{KKT}}.$$

and

$$I_{min} = I^{\mathbf{z}}_{min} \cap \mathbb{R}[\mathbf{x}] = (p_1, ..., p_r) \cap \mathbb{R}[\mathbf{x}] + \sqrt[G+]{I_{grad}} \cap \mathbb{R}[\mathbf{x}] = (p_1, ..., p_r) + \sqrt[G+]{I^{\mathbf{x}}_{KKT}}.$$

which proves the equality. □

Theorem 5.6. *For $G \subset \mathbb{R}[\mathbf{x}]$ with $S_{min} \subset \mathcal{S}(G) \subset V^{\mathbf{x},\mathbb{R}}_{KKT}$, there exists $t_2 \in \mathbb{N}$ such that $\forall t \ge t_2$, for $\Lambda^* \in \mathcal{L}_t(G)$ with $\Lambda^* = f^{\mu}_{t,G}$, we have $I_{min} \subset (\ker M^t_{\Lambda^*})$.*

Proof. To prove the inclusion we take $t_2 = max\{t_0, t_1\}$ and we combine Corollary 5.5 with Proposition 5.2 for $G \subset \mathbb{R}[\mathbf{x}]$ and Theorem 5.4. □

We introduce now the notion of *optimal linear form for f*. Such a linear form will allow us to compute I_{min} as we will see.

Proposition 5.7. *For $\Lambda^* \in \mathcal{L}_t(G)$ and $p \in \mathbb{R}[\mathbf{x}]$, the following assertions are equivalent:*

(i) $\mathrm{rank} M^t_{\Lambda^*} = max_{\Lambda \in \mathcal{L}_t(G), \Lambda(p) = p^{\mu}_{t,G}} \mathrm{rank} M^t_{\Lambda}$.

(ii) $\forall \Lambda \in \mathcal{L}_t(G)$ *such that* $\Lambda(p) = p^{\mu}_{t,G}$, $\ker M^t_{\Lambda^*} \subset \ker M^t_{\Lambda}$.

We say that $\Lambda^* \in \mathcal{L}_t(G)$ is optimal for p if it satisfies one of the equivalent conditions (i)-(ii).

A proof of this proposition can be found in [15](Proposition 4.7).

Remark 5.8. *A linear form $\Lambda^* \in \mathcal{L}_t(G)$ optimal for p can be computed by solving a Semi-Definite Programming problem by an interior point method* [17]. *In this case, the solution Λ^* obtained by convex optimization will be in the interior of the face of linear forms which minimize f.*

The next result, which refines Theorem 5.6, shows that only elements in I_{min} are involved in the kernel of a truncated Hankel operator associated to an optimal linear form for f.

Theorem 5.9. *Let $t \in \mathbb{N}$ such that $f \in \mathbb{R}[\mathbf{x}]_{2t}$ and let $G \subset \mathbb{R}[\mathbf{x}]_{2t}$ with $S_{min} \subset \mathcal{S}(G)$. If $\Lambda^* \in \mathcal{L}_t(G)$ is optimal for f and such that $\Lambda^*(f) = f^*$, then $\ker M^t_{\Lambda^*} \subset I_{min}$.*

Proof. It is similar to proof of Theorem 4.9 in [15]. □

The last result of this section shows that an optimal linear form for f yields the generators of the minimizer ideal I_{min} in high enough degree.

Theorem 5.10. *Let $\mathbf{g} \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints with $S_{min} \neq \emptyset$. For a set of constraints $G \subset \mathbb{R}[\mathbf{x}]$ with $S_{min} \subset \mathcal{S}(G) \subset V^{\mathbf{x},\mathbb{R}}_{KKT}$, there exists $t_2 \in \mathbb{N}$ (defined in Theorem 5.6) such that $\forall t \geq t_2$,*

- *$f^{\mu}_{t,G} = \min_{\mathbf{x} \in S_{KKT}} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(G)$,*
- *$\forall \Lambda^* \in \mathcal{L}_t(G)$ optimal for f, we have $\Lambda^*(f) = f^*$ and $(\ker M^t_{\Lambda^*}) = I_{min}$,*
- *if $\mathcal{V}(G^0) \subset V^{\mathbf{x}}_{KKT}$ then $f^{sos}_{t,G} = f^{\mu}_{t,G} = f^*$.*

Proof. We obtain the result as a consequence of Theorem 5.4, Theorem 5.6 and Theorem 5.9. □

The same results hold if we replace G by any other finite set defining a *real* variety such that $S_{min} \subset \mathcal{S}(G) \subset V^{\mathbf{x},\mathbb{R}}_{KKT}$.

Remark 5.11. *We can also replace the initial set of constraints $\mathbf{g}$ by any other set $\tilde{\mathbf{g}}$ defining the same semi-algebraic set $S = \mathcal{S}(\mathbf{g}) = \mathcal{S}(\tilde{\mathbf{g}})$ and consider the KKT variety associated to $\tilde{\mathbf{g}}$.*

6. Consequences

Let us describe now some consequences of these results in specific cases, which have been previously studied.

6.1. **Global optimization.** We consider here the case $n_1 = n_2 = 0$. Theorem 4.9 implies the following result (compare with [32]):

Theorem 6.1. *Let $f \in \mathbb{R}[\mathbf{x}]$.*

(1) *If the real critical values of f are positive, then $f \in \mathcal{Q}^+ + \sqrt{(\frac{\partial f}{\partial_{x_1}}, \dots, \frac{\partial f}{\partial_{x_n}})}$.*
(2) *If the real critical values of f are strictly positive, then $f \in \mathcal{Q}^+ + (\frac{\partial f}{\partial_{x_1}}, \dots, \frac{\partial f}{\partial_{x_n}})$.*

In particular, if there is no real critical value, then $f \in \mathcal{Q}^+ + (\frac{\partial f}{\partial_{x_1}}, \dots, \frac{\partial f}{\partial_{x_n}})$.

A consequence of Proposition 5.1 and Theorem 5.10 is the following:

Theorem 6.2. *Let $f \in \mathbb{R}[\mathbf{x}]$ and $G = \{\frac{\partial f}{\partial_{x_1}}, \dots, \frac{\partial f}{\partial_{x_n}}\}$. Then, there exists $t_0 \in \mathbb{N}$, such that $\forall t \geq t_0$ either $\mathcal{L}_t(G) = \emptyset$ and $S_{min} = \emptyset$ or*

(1) *$f^{sos}_{t,G} = f^{\mu}_{t,G} = f^* = \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(G)$,*
(2) *$\forall \Lambda^* \in \mathcal{L}_t(G)$ optimal for f, $\ker M^t_{\Lambda^*}$ generates I_{min}.*

The first point of this theorem can also be found in [32].

6.2. **General case.** A direct consequence of Proposition 5.1 and Theorem 5.10 is the following:

Theorem 6.3. *Let $G \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that*
- $(G^0) = I_{KKT} \cap \mathbb{R}[\mathbf{x}]$,
- $G^+ = \mathbf{g}^+$.

Then there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$, either $\mathcal{L}_t(G) = \emptyset$ and $S_{min} = \emptyset$ or
- $f^{sos}_{t,G} = f^{\mu}_{t,G} = min_{\mathbf{x} \in S_{KKT}} f(\mathbf{x})$ *is reached for some* $\Lambda^* \in \mathcal{L}_t(G)$,
- $\forall \Lambda^* \in \mathcal{L}_t(G)$ *optimal for* f, *we have* $\Lambda^*(f) = f^*$ *and* $(\ker M^t_{\Lambda^*}) = I_{min}$.

The set G^0 is constructed so that $\mathcal{V}(G^0) = V^{\mathbf{x}}_{KKT}$. As we have seen, the weaker condition $S_{min} \subset \mathcal{S}(G) \subset V^{\mathbf{x}}_{KKT}$ is sufficient to have an exact relaxation sequence.

The generators G^0 of $I_{KKT} \cap \mathbb{R}[\mathbf{x}]$ can be computed by elimination techniques (for instance by Gröbner basis computation with a product order on monomials [3]).

6.3. **Regular case.** We consider here a semi-algebraic set S such that its defining constraints intersect properly. For any $\mathbf{x} \in \mathbb{C}^n$, let $\nu(\mathbf{x}) = \{j \in [1, n_2] \mid g^+_j(\mathbf{x}) = 0\}$.

Definition 6.4. *We say that a set of constraints $\mathbf{g} = (g^0_1, \ldots g^0_{n_1};\ g^+_1, \ldots, g^+_{n_2})$ is regular if for all points $\mathbf{x} \in \mathcal{S}(\mathbf{g})$ with $\nu(\mathbf{x}) = \{j_1, \ldots, j_k\}$, the vectors $\nabla g^0_1(\mathbf{x}), \ldots, \nabla g^0_{n_1}(\mathbf{x})$, $\nabla g^+_{j_1}(\mathbf{x}), \ldots, \nabla g^+_{j_k}(\mathbf{x})$ are linearly independent.*

This condition is used for instance in [9]. It implies that $\forall \mathbf{x} \in S$, $|\nu(\mathbf{x})| \leq n - n_1$ and that $B_{\nu(\mathbf{x})}(\mathbf{x})$ is of rank $n_1 + |\nu(\mathbf{x})|$. A stronger condition, called the $\mathbb{C}$-regularity, corresponds to sets of constraints such that $\forall \mathbf{x} \in \mathbb{C}^n$, $B_{\nu(\mathbf{x})}(\mathbf{x})$ is of rank $n_1 + |\nu(\mathbf{x})|$. This condition is used for instance in [29]. It is satisfied for semi-algebraic sets defined by "generic" constraints when $n_1 \leq n$ as shown in [29].

If $\mathbf{g}$ is regular, then for all points $\mathbf{x}$ in S the rank of $B_{\nu(\mathbf{x})}(\mathbf{x})$ is $n_1 + |\nu(\mathbf{x})|$ and $S_{sing} = \emptyset$. The decomposition (12) implies that $S_{FJ} = S_{KKT}$ and that all minimizer points of f on S are KKT points. If moreover $\mathbf{g}$ is $\mathbb{C}$-regular, then $V^{\mathbf{x}}_{FJ} = \mathcal{V}(\mathbf{g}^0 \cup \Gamma_{FJ}) = V^{\mathbf{x}}_{KKT}$.

We deduce from Theorem 5.10 the following result:

Theorem 6.5. *Let $\mathbf{g} \subset \mathbb{R}[\mathbf{x}]$ be a regular set of constraints and let $G \subset \mathbb{R}[\mathbf{x}]$ be the set of constraints such that*
- $G^0 = \Gamma_{FJ}$ *defined in* (9) *(resp.* $G^0 = \Phi_{FJ}$ *defined in* (10)*)*,
- $G^+ = \mathbf{g}^+$.

Suppose that $\min_{\mathbf{x} \in \mathcal{S}(\mathbf{g})} f(\mathbf{x})$ is reached at some point of $\mathcal{S}(\mathbf{g})$. Then, there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$,
1. $f^{\mu}_{t,G} = f^* = \min_{\mathbf{x} \in \mathcal{S}(\mathbf{g})} f(\mathbf{x})$ *is reached for some* $\Lambda^* \in \mathcal{L}_t(G)$,
2. $\forall \Lambda^* \in \mathcal{L}_t(G)$ *optimal for* f, $\ker M^t_{\Lambda^*}$ *generates* $I^{\mathbf{x}}_{min}$,
3. *If* $\mathbf{g}$ *is* $\mathbb{C}$*-regular and* $G^0 = \mathbf{g}^0 \cup \Gamma_{FJ}$, *then* $f^{sos}_{t,G} = f^{\mu}_{t,G} = f^*$.

By Lemma 3.11 and Remark 3.12, G is constructed so that $S_{min} \subset \mathcal{S}(G) = S_{KKT} \subset V^{\mathbf{x},\mathbb{R}}_{KKT}$.

Points (1) and (3) are proved for $G^0 = \mathbf{g}^0 \cup \Gamma_{FJ}$ in [29] under the condition that $\mathbf{g}$ is $\mathbb{C}$-regular. These points can also be found in [9] for $G^0 = \mathbf{g}^0 \cup \Psi_{FJ}$ under the condition that $\mathbf{g}$ regular (but a problem appears in the proof: the vanishing of the polynomials Ψ_{FJ} at a point $\mathbf{x} \in \mathbb{C}^n$ does not imply that rank $A_{\nu(\mathbf{x})}(\mathbf{x}) < n_1 + |\nu(\mathbf{x})|$).

In this case, the relaxation constructed with Γ_{FJ} (or Φ_{FJ}) is exact and can be used to compute the minimizer ideals of f on the semi-algebraic set S.

6.4. **Zero dimensional real variety.** Let $\mathbf{g} \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that $\mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is finite and let $S := \mathcal{S}(\mathbf{g})$. By remark 5.11, we can assume that S is defined by a set of constraints $\tilde{\mathbf{g}}$ such that $(\tilde{\mathbf{g}}^0)$ is radical. Then $\forall \mathbf{x} \in \mathcal{V}(\mathbf{g}^0) = \mathcal{V}(\tilde{\mathbf{g}}^0)$, the Jacobian matrix $\tilde{B}_{\nu(\mathbf{x})}(\mathbf{x})$ associated to $\tilde{\mathbf{g}}^0$ is of rank n. Therefore we have $\mathcal{V}(\mathbf{g}^0) = \mathcal{V}(\tilde{\mathbf{g}}^0) = V^{\mathbf{x}}_{KKT}$ and any point of S is a KKT-point: $S = S_{FJ} = S_{KKT}$. Consequently, we deduce from Theorem 5.10 the following result:

Theorem 6.6. *Let $\mathbf{g} = (\mathbf{g}^0, \mathbf{g}^+) \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that $\mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is finite. Then there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$,*

1. *$f^{sos}_{t,\mathbf{g}} = f^{\mu}_{t,\mathbf{g}} = f^* = \min_{\mathbf{x}\in\mathcal{S}(\mathbf{g})} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(\mathbf{g})$,*
2. *$\forall \Lambda^* \in \mathcal{L}_t(\mathbf{g})$ optimal for f, $\ker M^t_{\Lambda^*}$ generates I_{min}.*

This answers an open question in [22]. The first point was also solved in [31] using dedicated techniques.

6.5. **Smooth real variety.** We consider a set of constraints $\mathbf{g} = \{g^0_1, \ldots, g^0_{n_1}\} \subset \mathbb{R}[\mathbf{x}]$ such that $\mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is equidimensional smooth and $\mathbf{g}^+ = \emptyset$. This means that $S = \mathcal{S}(\mathbf{g}) = \mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is the union of irreducible components of the same dimension d and that for any point $\mathbf{x} \in S$, $B_\emptyset(\mathbf{x}) = [\nabla g^0_1(\mathbf{x}), \ldots, \nabla g^0_{n_1}(\mathbf{x})]$ is of rank $m = \dim S = n - d$. Therefor, $S_{sing} = \emptyset$. In this case, $\nabla f(\mathbf{x})$ is a linear combination of $\nabla g^0_1(\mathbf{x}), \ldots, \nabla g^0_{n_1}(\mathbf{x})$, if and only if, $\mathrm{rank} A_\emptyset(\mathbf{x}) \leq r$.

The set Γ_{FJ} defined in (9) (or $G^0 = \mathbf{g}^0 \cup \Phi_{FJ}$ defined in (10)), or the union Δ^{n-d} of gb^0 and the set of $(n-d+1) \times (n-d+1)$ minors of the Jacobian matrix of $\{f, g^0_1, \ldots, g^0_{n_1}\}$, which contain the first column ∇f define the variety S_{KKT}.

We deduce from Theorem 5.10, the following result:

Theorem 6.7. *Let $\mathbf{g} = \{g^0_1, \ldots, g^0_{n_1}\} \subset \mathbb{R}[\mathbf{x}]$ such that $S = \mathcal{V}^{\mathbb{R}}(\mathbf{g})$ is an equidimensional and smooth variety of dimension d.*

Let $G \subset \mathbb{R}[\mathbf{x}]$ be the set of constraints such that $G^0 = \Gamma_{FJ}$ defined in (9) (or $G^0 = \Phi_{FJ}$ defined in (10), $G^0 = \Delta^{n-d}$) Then there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$, either $\mathcal{L}_t(G) = \emptyset$ and $S_{min} = \emptyset$ or

1. *$f^{\mu}_{t,G} = f^* = \min_{\mathbf{x}\in S} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(G)$,*
2. *$\forall \Lambda^* \in \mathcal{L}_t(G)$ optimal for f, $\ker M^t_{\Lambda^*}$ generates I_{min}.*

6.6. **Known minimum.** In the case where we know the minimimun f^* of f on the basic closed semi-algebraic set S, we take $\mathbf{g}'$ with $\mathbf{g}'^0 = \{\mathbf{g}^0, f - f^*\}$ and $\mathbf{g}'^+ = \mathbf{g}^+$. Let $S = \mathcal{S}(\mathbf{g})$, $S' = \mathcal{S}(\mathbf{g}')$. By construction $S_{min} \subset S'$ and $S' = S'_{KKT}$ and $\mathcal{V}(\mathbf{g}^0) \subset V^{\mathbf{x}}_{KKT}(\mathbf{g}'^0)$. Theorem 5.10 applied to $\mathbf{g}'$ implies the following result:

Theorem 6.8. *Let $\mathbf{g} = \{g^0_1, \ldots, g^0_{n_1}; g^+_1, \ldots, g^+_{n_2}\} \subset \mathbb{R}[\mathbf{x}]$. Let f^* be the minimum of f and $G \subset \mathbb{R}[\mathbf{x}]$ the set of constraints such that $G^0 = \{\mathbf{g}^0, f - f^*\}$ and $G^+ = \mathbf{g}^+$. Then there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$,*

1. *$f^{sos}_{t,G} = f^{\mu}_{t,G} = f^* = \min_{\mathbf{x}\in\mathcal{S}(G)} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_{t,G}$,*
2. *$\forall \Lambda^* \in \mathcal{L}_t(G)$ optimal for f, $\ker M^t_{\Lambda^*}$ generates I_{min}.*

6.7. **G^+-radical computation.** In the case where $f = 0$, by Remark 3.15 all the points of S are KKT points and minimizers of f so that $S_{min} = S = S_{KKT}$. Moreover, $I^{\mathbf{x}}_{KKT} = (g^0_1, \ldots, g^0_{n_1})$ since $F_1, \ldots, F_n$, $v_1 g^+_1, \ldots, v_{n_2} g^+_{n_2}$ are homogeneous of degree 1 in the variables $u_1, \ldots, u_{n_1}, v_1, \ldots, v_{n_2}$. We deduce the following result:

Theorem 6.9. *Let $G = \{g^0_1, \ldots, g^0_{n_1}; g^+_1, \ldots, g^+_{n_2}\} \subset \mathbb{R}[\mathbf{x}]$. There exists $t_2 \in \mathbb{N}$ such that $\forall t \geq t_2$, $\forall \Lambda^* \in \mathcal{L}_t(G)$ optimal for 0, we have $(\ker M^t_{\Lambda^*}) = \mathcal{I}(\mathcal{S}) = \sqrt[G^+]{(G^0)}$.*

This gives a way to compute $\sqrt[G^+]{(G^0)}$, which generalizes the approach of [19] or [15] to compute the real radical of an ideal.

7. Examples

This section contains examples which illustrate different aspects of our method. In the case of a finite number of minimizers for a function f on the semi-algebraic set S defined by the set of constraints $\mathbf{g}$, the approach we describe leads to the following algorithm:

(1) Compute $G \subset \mathbb{R}[\mathbf{x}]$ such that G^0 generates $I_{KKT} \cap \mathbb{R}[\mathbf{x}]$ and $G^+ = \mathbf{g}^+$;
(2) $t := \lceil \frac{1}{2} \max\{\deg(f), \deg(g_i^0), \deg(g_j^+)\} \rceil$;
(3) Compute $\Lambda^* \in \mathcal{L}_t(G)$ optimal for f (solving a finite dimensional SDP problem by an interior point method);
(4) Check the convergence certificate for $M_{\Lambda^*}^t$ (by flat extension [11, 23]);
(5) If it is not satisfied, then $t := t + 1$ and repeat from step (2);
Otherwise compute $K := \ker M_{\Lambda^*}^t$.

Output $f^* = \Lambda^*(f)$ and the generators K of I_{min}.

Example 7.1. *We consider the "ill-posed" problem*

$$\min\ x\ s.t\ x^3 \geq 0.$$

The ideal I_{KKT} is $I_{KKT} = (1 - 3v_1x^2, v_1x^3) = (1)$. Thus $V_{KKT} = \emptyset$. According to the decomposition (12), $S_{FJ} = S_{sing}$ and we compute the minimum of x on S_{sing}, which is defined by $x^2 = 0$:

$$\min\ x\ s.t\ x^2 = 0.$$

Now according to section 6.4, the relaxation associated to this problem is exact and yields the solution $x = 0$.

Example 7.2. *We consider the following problem*

$$\begin{array}{rl} \min & f(x,y,z) = x^2 + y^2 + z^2; \\ s.t & \operatorname{rank} \begin{pmatrix} x+z+1 & x+y & y+z \\ x+y & y+z & x+z+1 \end{pmatrix} \leq 1 \end{array}$$

or equivalently

$$\begin{array}{rl} \min & f(x,y,z) = x^2 + y^2 + z^2; \\ s.t & (x+z+1)(y+z) - (x+y)^2 = 0; \\ & (x+z+1)^2 - (y+z)(x+y) = 0; \\ & (x+z+1)(x+y) - (y+z)^2 = 0; \end{array}$$

This corresponds to computing the closest on a twisted cubic defined by 2×2 minors. The set of constraints $\mathbf{g}$ is not regular but $\mathcal{S}(\mathbf{g}) = \mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is a smooth real variety.

In the first iteration of the algorithm, the order is 1, the size of the Hankel matrix M_Λ^1 is 3, $\min \Lambda(f) = 1$, there is no duality gap. The flat extension condition is satisfied for M_Λ^1 and we have found the minimum. The algorithm stops and we obtain $I_{min} = (x, y-1, z)$ and the points which minimize f, $\{(x = 0, y = 1, z = 0)\}$.

Example 7.3. *We consider the Motzkin polynomial,*

$$\min\ f(x,y) = 1 + x^4y^2 + x^2y^4 - 3x^2y^2$$

which is non negative on $\mathbb{R}^2$ but not a sum of squares in $\mathbb{R}[x,y]$. We compute its gradient ideal, $I_{grad}(f) = (-6xy^2 + 2xy^4 + 4x^3y^2, -6yx^2 + 2yx^4 + 4y^3x^2)$, which is not zero-dimensional.

In the first iteration of the algorithm, the order is 3, the size of the Hankel matrix M_Λ^3 *is 10,* $\min \Lambda(f) = -216$. *The flat extension condition is not satisfied hence we try with degree 4.*

In the second iteration the order is 4, the size of the Hankel matrix M_Λ^4 *is 15,* $\min \Lambda(f) = 0$, *there is no duality gap. The flat extension condition is satisfied for* M_Λ^4 *and we have found the minimum. The algorithm stops and we obtain* $I_{min} = (x^2-1, y^2-1)$ *and the points which minimize f,* $\{(x=1,y=1),(x=1,y=-1),(x=-1,y=1),(x=-1,y=-1)\}$.

For this example Gloptipoly must go until order 9 in order to satisfy the flat extension condition.

Example 7.4. *We consider the Robinson polynomial*

$$\min \ f(x,y) = 1 + x^6 - x^4 - x^2 + y^6 - y^4 - y^2 - x^4y^2 - x^2y^4 + 3x^2y^2$$

which is non negative on $\mathbb{R}^2$ *but not a sum of squares in* $\mathbb{R}[x,y]$. *We compute its gradient ideal,*
$I_{grad}(f) = (6x^5 - 4x^3 - 2x - 4x^3y^2 - 2xy^4 + 6xy^2, 6y^5 - 4y^3 - 2y - 4y^3x^2 - 2yx^4 + 6yx^2)$
which is not zero-dimensional.

In the first iteration, the order is 3, the size of the Hankel matrix M_Λ^3 *is 10,* $\min \Lambda(f) = -0.93$. *The flat extension condition is not satisfied hence we try with degree 4.*

In the second iteration the degree is 4, the size of the Hankel matrix M_Λ^4 *is 15,* $\min \Lambda(f) = 0$. *There is no duality gap. The flat extension condition is satisfied for* M_Λ^3 *and we have found the minimum.*

The algorithm stops and we obtain $I_{min} = (x^3 - x, y^3 - y, x^2y^2 - x^2 - y^2 + 1)$ *and the points which minimize f,* $\{(x=1,y=1),(x=1,y=-1),(x=-1,y=1),(x=-1,y=-1),(x=1,y=0),(x=-1,y=0),(x=0,y=1),(x=0,y=-1)\}$.

For this example, Gloptipoly must go until order 7 in order to satisfy the flat extension condition.

Example 7.5. *We consider the homogeneous Motzkin polynomial with a perturbation* $\epsilon = 0.005$,

$$\begin{array}{rl} \min & f(x,y,z) = x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6 + \epsilon(x^2+y^2+z^2); \\ s.t & h(x,y,z) = 1 - x^2 - y^2 - z^2 \geq 0 \end{array}$$

This example coming from [21]*[Example 6.25] is a case where the constraints* $\mathbf{g}$ *define a compact semi-algebraic set, but the direct relaxation using the associated quadratic module or preordering is not exact.*

We add the projection of the KKT ideal and we have the similiar problem

$$\begin{array}{rl} \min & x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6; \\ s.t & -4zx^4y - 20zx^2y^3 + 12x^2yz^3 - 0.06zy^5 + 12.06yz^5 = 0; \\ & -20zx^3y^2 - 4zxy^4 + 12xy^2z^3 - 0.06zx^5 + 12.06xz^5 = 0; \\ & (4x^3y^2 + 2xy^4 - 6xy^2z^2 + 0.03x^5)(-x^2 - y^2 - z^2 + 1) = 0; \\ & (2x^4y + 4x^2y^3 - 6x^2yz^2 + 0.03y^5)(-x^2 - y^2 - z^2 + 1) = 0; \\ & (-6x^2y^2z + 6.03z^5)(-x^2 - y^2 - z^2 + 1) = 0; \end{array}$$

where the first three equations are the 2×2 *minors of the Jacobian matrix of* f *and* h *and the last three equations are the gradient ideal of* f *multiplied by* h.

In the first iteration the order is 5, the size of the Hankel matrix M_Λ^5 *is 167,* $\min \Lambda(f) = 0$, *there is no duality gap. The flat extension condition is satisfied for* M_Λ^5 *and we have found the minimum. The algorithm stops and we obtain* $I_{min} = (x,y,z)$ *and the point which minimize* f *is* $(0,0,0)$.

For this example, the flat extension condition does not hold with Gloptipoly if $\epsilon \leq 0.01$.

Finally with these two last examples we show that even the minimizer ideal I_{min} is not zero-dimensional we can recover it from a solution of the relaxation problem.

Example 7.6. *We consider Motzkin polynomial over the unit ball:*

$$\begin{array}{rl} \min & f(x,y,z) = x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6; \\ s.t & h(x,y,z) = 1 - x^2 - y^2 - z^2 \geq 0 \end{array}$$

The polynomial f is homogeneous and non negative on $\mathbb{R}^3$ but not a sum of squares in $\mathbb{R}[x,y,z]$.

We add the projections of KKT ideal and we have the similiar problem

$$\begin{array}{rl} \min & x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6; \\ s.t & -4xy^5 + 12xy^3z^2 + 4yx^5 - 12x^3yz^2 = 0; \\ & -4zx^4y - 20zx^2y^3 + 12x^2yz^3 + 12yz^5 = 0; \\ & -20zx^3y^2 - 4zxy^4 + 12xy^2z^3 + 12xz^5 = 0; \\ & (4x^3y^2 + 2xy^4 - 6xy^2z^2)(-x^2 - y^2 - z^2 + 1) = 0; \\ & (2x^4y + 4x^2y^3 - 6x^2yz^2)(-x^2 - y^2 - z^2 + 1) = 0; \\ & (-6x^2y^2z + 6z^5)(-x^2 - y^2 - z^2 + 1) = 0; \end{array}$$

where the first three equations are the 2×2 minors of the Jacobian matrix of f and h and the last three equations are the gradient ideal of f multiplied by h.

In the first iteration the order is 5, the size of the Hankel matrix M_Λ^5 is 156, $\min \Lambda(f) = 0$*, there is no duality gap. We compute the kernel of this matrix:* $\ker M_\Lambda^5 = \langle z(y^2 - z^2), x(y^2 - z^2), z(x^2 - z^2), y(x^2 - z^2)\rangle$*. It generates the minimizer ideal $I_{min} = (z(y^2 - z^2), x(y^2 - z^2), z(x^2 - z^2), y(x^2 - z^2))$ defining 6 lines: $(x \pm y, x \pm z), (x,z), (y,z)$. Here $\mathcal{V}(I_{min})$ is not included in S.*

Example 7.7. *We consider minimization of a linear function on a torus:*

$$\begin{array}{rl} \min & f(x,y,z) = z \\ s.t & 9 - 10x^2 - 10y^2 + 6z^2 + x^4 + 2x^2y^2 + 2x^2z^2 + 2y^2z^2 + y^4 + z^4 = 0 \end{array}$$

In the first iteration, the order is 2, the size of the Hankel matrix M_Λ^2 is 10*, min $\Lambda(f) = -1$, there is no duality gap. We compute the kernel of this matrix:* $\ker M_\Lambda^2 = \langle x^2 + y^2 - 4, x(z+1), y(z+1), z(z+1), (z+1)\rangle$ *which generates the minimizer ideal $I_{min} = (x^2+y^2-4, z+1)$, defining a circle which is the intersection of the torus with a tangent plane. Notice that the multiplicity of this intersection has been removed in I_{min}.*

References

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

[2] W. Bruns and U. Vetter. *Determinantal rings*, volume 1327 of *Lecture Notes in Math.* Springer, Berlin, 1988.

[3] D.A. Cox, J.B. Little, and D.B. O'Shea. *Ideals, Varieties, and Algorithms : An Introduction to Computational Algebraic Geometry and Commutative Algebra (Undergraduate Texts in Mathematics).* Springer, 2005.

[4] R.E. Curto and L. Fialkow. Solution of the truncated complex moment problem for flat data. *Memoirs of the American Mathematical Society*, 119(568):1–62, 1996.

[5] James Demmel, Jiawang Nie, and Victoria Powers. Representations of positive polynomials on non-compact semialgebraic sets via kkt ideals. *Journal of Pure and Applied Algebra*, 209(1):189 – 200, 2007.

[6] A. Greuet and M. Safey El Din. Deciding reachability of the infimum of a multivariate polynomial. In *Proceedings of the 36th international symposium on Symbolic and algebraic computation*, ISSAC '11, pages 131–138, New York, NY, USA, 2011. ACM.

[7] F. Guo, M. Safey El Din, and L. Zhi. Global optimization of polynomials using generalized critical values and sums of squares. In *Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation*, ISSAC '10, pages 107–114, New York, NY, USA, 2010. ACM.

[8] H. V. Ha and T.S. Pham. Global optimization of polynomials using the truncated tangency variety. *SIAM Journal on Optimization*, 19(2):941–951, 2008.

[9] H. V. Ha and T.S. Pham. Representation of positive polynomials and optimization on noncompact semialgebraic sets. *SIAM Journal on Optimization*, 20(6):3082–3103, 2010.

[10] B. Hanzon and D. Jibetean. Global minimization of a multivariate polynomial using matrix methods. *J. Global Optim*, 27:1–23, 2001.

[11] D. Henrion and J.B. Lasserre. *Positive Polynomials in Control*, chapter Detecting Global Optimality and Extracting Solutions in GloptiPoly., pages 293–310. Lectures Notes in Control and Information Sciences. Springer, 2005.

[12] D. Jibetean and M. Laurent. Semidefinite approximations for global unconstrained polynomial optimization. *SIAM J. on Optimization*, 16(2):490–514, June 2005.

[13] F. John. Extremum problems with inequalities as side conditions. In *Studies and Essays, Courant Anniversary Volume*, pages 187–204. Wiley (Interscience), New York, 1948.

[14] E. L. Kaltofen, B. Li, Z. Yang, and L. Zhi. Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients. *J. Symb. Comput.*, 47(1):1–15, January 2012.

[15] J.-B. Lasserre, M. Laurent, B. Mourrain, P. Rostalski, and P. Trébuchet. Moment matrices, border bases and real radical computation. *Journal of Symbolic Computation*, 2012.

[16] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11:796–817, 2001.

[17] J.B Lasserre. *Moments, positive polynomials and their applications.* Imperial College Press, 210.

[18] J.B. Lasserre, M. Laurent, and P. Rostalski. Semidefinite characterization and computation of real radical ideals. *Foundations of Computational Mathematics*, 8(5):607–647, 2008.

[19] J.B. Lasserre, M. Laurent, and P. Rostalski. A unified approach for real and complex zeros of zero-dimensional ideals. In M. Putinar and S. Sullivant, editors, *Emerging Applications of Algebraic Geometry.*, volume 149, pages 125–156. Springer, 2009.

[20] M. Laurent. Semidefinite representations for finite varieties. *Math. Progr*, 109:1–26, 2007.

[21] M. Laurent. *Emerging Applications of Algebraic Geometry*, chapter Sums of squares, moment matrices and optimization over polynomials. Springer Verlag, 2008.

[22] M. Laurent. *Sums of squares, moment matrices and optimization over polynomials*, volume 149 of *IMA Volumes in Mathematics and its Applications*, pages 157–270. Springer, 2009.

[23] M. Laurent and B. Mourrain. A generalized flat extension theorem for moment matrices. *Arch. Math. (Basel)*, 93(1):87–98, July 2009.

[24] O.L. Mangasarian and S. Fromovitz. The fritz john necessary optimality conditions in the presence of equality and inequality constraints. *Journal of Mathematical Analysis and Applications*, 17:37–47, 1967.

[25] M. Marshall. Optimization of polynomial functions. *Canad. Math. Bull.*, 46:575–587, 2003.

[26] M. Marshall. Representations of non-negative polynomials, degree bounds and applications to optimization. *Can. J. Math.*, 61(1):205–221, 2009.

[27] B. Mourrain. A new criterion for normal form algorithms. In M. Fossorier, H. Imai, Shu Lin, and A. Poli, editors, *Proc. AAECC*, volume 1719 of *LNCS*, pages 430–443. Springer, Berlin, 1999.

[28] Y. Nesterov. Squared functional systems and optimization problems. In H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, *High performance optimization*, chapter 17, pages 405–440. Kluwer academic publishers, Dordrecht, The Netherlands, 2000.

[29] J. Nie. An exact jacobian sdp relaxation for polynomial optimization. *Mathematical Programming*, pages 1–31, 2011.

[30] J. Nie. Certifying convergence of Lasserre's hierarchy via flat truncation. *Mathematical Programming*, pages 1–26, 2012.

[31] J. Nie. Polynomials optimization with real variety. *SIAM Journal On Optimization*, 23(3):1634–1646, 2013.

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

[33] P.A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming Ser. B*, 96(2):293–320, 2003.

[34] P.A. Parrilo and B. Sturmfels. Minimizing polynomial functions. In *Proceedings of the DIMACS Workshop on Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science*, pages 83–100. American Mathematical Society, 2003.
[35] M. Safey El Din. Computing the global optimum of a multivariate polynomial over the reals. In *Proceedings of the twenty-first international symposium on Symbolic and algebraic computation*, ISSAC '08, pages 71–78, New York, NY, USA, 2008. ACM.
[36] M. Schweighofer. Global optimization of polynomials using gradient tentacles and sums of squares. *SIAM Journal on Optimization*, 17(3):920–942, 2006.
[37] Shafarevich. *Basic algebraic geometry.* Springer-Verlag, 1974.
[38] N.Z. Shor. Class of global minimum bounds of polynomial functions. *Cybernetics*, 23:731–734, 1987.

Marta Abril Bucero, Bernard Mourrain: Galaad, Inria Méditerranée, 06902 Sophia Antipolis