



**HAL**  
open science

# Exact relaxation for polynomial optimization on semi-algebraic sets

Marta Abril Bucero, Bernard Mourrain

► **To cite this version:**

Marta Abril Bucero, Bernard Mourrain. Exact relaxation for polynomial optimization on semi-algebraic sets. 2013. hal-00846977v1

**HAL Id: hal-00846977**

**<https://inria.hal.science/hal-00846977v1>**

Preprint submitted on 24 Jul 2013 (v1), last revised 1 Jul 2014 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# CERTIFIED RELAXATION FOR POLYNOMIAL OPTIMIZATION ON SEMI-ALGEBRAIC SETS

MARTA ABRIL BUCERO, BERNARD MOURRAIN

ABSTRACT. In this paper, we describe a relaxation method to compute the minimal critical value of a real polynomial function on a semialgebraic set  $S$  and the ideal defining the points at which the minimal critical value is reached. We show that any relaxation hierarchy which is the projection of the Karush-Kuhn-Tucker relaxation stops in a finite number of steps and the ideal defining the minimizers is generated by the kernel of the associated moment matrix in that degree. Assuming the minimizer ideal is zero-dimensional, we give a new criterion to detect when the minimum is reached and we prove that this criterion is satisfied for a sufficiently high degree. This exploits new representation of positive polynomials as elements of the preordering modulo the KKT ideal, which only involves polynomials in the initial set of variables.

## 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 reached. Reformulating the problem as the computation of a (minimal) critical values, polynomial system solvers can be used to tackle it (see e.g. [28], [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 [31] may also be considered here but suffer from similar issues. Though the global minimization problem is known to be NP-hard (see e.g. [23]), 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 [15] (see also [27], [35]) 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  $f^*$  [15]. 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 [15], [20]. 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 [18] 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 [15] is exact. It is also proved that there is no duality gap if the generators of this ideal satisfy some conditions.

In [26], it is proved that a relaxation hierarchy using the gradient constraints is exact when the gradient ideal is radical. In [21], 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-Khun-Tucker constraints is exact when the KKT ideal is radical. In [24], 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 minimum value is not reached at points which satisfied the KKT constraints, has also been studied. In [33], 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 [6], this approach is extended to polynomial minimization on smooth equidimensional algebraic varieties with constraints defining a radical ideal.

The second issue is also important from an algorithmic point of view, to certify when the minimum is reached. In [18], [17], a stopping criterion based on Curto-Fialkow flat extension condition [4] is used in a relaxation hierarchy for optimisation on finite algebraic varieties and for real radical computation. This criterion is also used in [26], where an algorithm is described to compute the global minimum of the polynomial function  $f$  when this minimum is reached and when the gradient ideal is zero-dimensional. An improvement of the flat extension criterion [19] is used in [14] 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 [22] to detect when a border basis of the real radical is computed. In [25], it is proved that a Curto-Fialkow flat extension criterion is eventually satisfied on truncated moment matrices, in relaxation hierarchies [15], [26], [24] 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 [17], the kernel of moment matrices are used to compute generators of the real radical of an ideal. This method is improved in [14] 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 [25] 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. [13]). 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.** We consider the problem of computing the minimal critical value of a polynomial  $f$  on a semi-algebraic set  $\mathcal{S} \subset \mathbb{R}^n$  defined by equations  $\mathbf{g} = 0$  and inequalities  $\mathbf{h} \geq 0$ .

The main results are the following (Theorem 5.15, Corollary 5.16):

- Any hierarchy of relaxation problems based on the KKT constraints or on their projection is exact.
- When the minimum is reached (for a sufficiently high degree), the kernel of the moment matrices associated to the solution of the corresponding SDP problem is a generating family of the minimizer ideal.

This applies to any polynomial constraints which real variety intersected with the set of real points where  $H \geq 0$  is between (the projection of) the real KKT variety and (the projection of) the minimizer variety. It generalizes results on the exactness of the hierarchy in [26], [5] or [24], [9] which all require regularity conditions. It also shows that the minimizer ideal can be constructed from the kernel of moment matrices associated to the corresponding semi-definite programming problem.

When  $f = 0$ , this yields a new method to compute  $\sqrt[H]{(G)}$  which extends the approach in [19], [17], [14].

For polynomials which are positive on  $\mathcal{S}$ , we give new decomposition of polynomials as elements in the preordering module of  $H$ , modulo the KKT constraints, which only involves polynomial in  $\mathbb{R}[\mathbf{x}]$ . This generalizes results in [5].

In the case where the minimizer ideal is zero dimensional, we give a new criterion to detect when the minimum is reached, based on the flat extension condition in [19]. This flat extension condition is weaker than the one described in [25], since it allows more freedom in the choice of the submatrices of the moment matrices. Finally, we prove that when the relaxation order is big enough, this criterion is satisfied. Hence, this approach yields a complete algorithm which increments the order of relaxation and check the convergence certificate until it is satisfied.

**Content.** The paper is organized as follows. Section 2 recalls the concepts of algebraic tools as ideals, varieties and we define our minimization problem and we set our notation. In section 3 we describe the sequence of finite dimensional optimization problems on which we apply SemiDefinite Programming methods. In section 4 we analyze the decomposition of polynomials 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 optimization problems attains its limit and the minimizer ideal can be generated from the solution of our relaxation problem. In section 6, we analyse some consequences of this result. Finally in the section 7, we describe a criterion to detect when we have found the minimum based on a flat extension property, when the minimizer ideal is zero-dimensional. We prove also that when the order of relaxation is big enough, this criterion is satisfied.

## 2. IDEALS, VARIETIES AND OPTIMIZATION

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  $\mathbf{x} = (x_1, \dots, x_n)$ , with coefficients in the field  $\mathbb{K}$ . Hereafter, we will choose<sup>1</sup>  $\mathbb{K} = \mathbb{R}$  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  $\alpha$  and degree  $|\alpha| = \sum_i \alpha_i$ . The set of all monomials in  $\mathbf{x}$  is denoted  $\mathcal{M} = \mathcal{M}(\mathbf{x})$ . For a polynomial  $f = \sum_\alpha f_\alpha \mathbf{x}^\alpha$ , its support is  $\text{supp}(f) := \{\mathbf{x}^\alpha \mid f_\alpha \neq 0\}$ , the set of monomials occurring with a nonzero coefficient in  $f$ .

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

- $S_t$  is the set of elements of  $S$  of degree  $\leq t$ ,
- $\langle S \rangle = \{ \sum_{f \in S} \lambda_f f \mid f \in S, \lambda_f \in \mathbb{K} \}$  is the linear span of  $S$ ,
- $(S) = \{ \sum_{f \in S} p_f f \mid p_f \in \mathbb{K}[\mathbf{x}], f \in S \}$  is the ideal in  $\mathbb{K}[\mathbf{x}]$  generated by  $S$ ,
- $S_{\langle t \rangle} = \{ \sum_{f \in S_t} p_f f \mid p_f \in \mathbb{K}[\mathbf{x}]_{t - \deg(f)} \}$  is the vector space spanned by  $\{\mathbf{x}^\alpha f \mid f \in S_t, |\alpha| \leq t - \deg(f)\}$ ,
- $S^+ := S \cup x_1 S \cup \dots \cup x_n S$  is the prolongation of  $S$  by one degree,
- $S \cdot S = \{ab \mid a, b \in S\}$ .

Therefore,  $S_t = S \cap \mathbb{K}[\mathbf{x}]_t$ ,

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

A set of monomials  $B$  is said to be *connected to 1* if  $1 \in B$  and for every monomial  $m \neq 1$  in  $B$ ,  $m = x_{i_0} m'$  for some  $i_0 \in [1, n]$  and  $m' \in B$ .

Given a vector space  $E \subseteq \mathbb{K}[\mathbf{x}]$ , its prolongation  $E^+ := E + x_1 E + \dots + x_n E$  is again a vector space.

A vector space  $E$  is said to be *connected to 1* if  $1 \in E$  and any non-constant polynomial  $p \in E$  can be written as  $p = p_0 + \sum_{i=1}^n x_i p_i$  for some polynomials  $p_0, p_i \in E$  with  $\deg(p_0) \leq \deg(p)$ ,  $\deg(p_i) \leq \deg(p) - 1$  for  $i \in [1, n]$ .

Obviously,  $E$  is connected to 1 when  $E = \langle B \rangle$  for some monomial set  $B \subseteq \mathcal{M}$  which is connected to 1. Moreover,  $E^+ = \langle B^+ \rangle$  if  $E = \langle B \rangle$ .

Given an ideal  $I \subseteq \mathbb{K}[\mathbf{x}]$  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  $\mathcal{V}(I) = \mathcal{V}^{\overline{\mathbb{K}}}(I)$ .

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

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

Furthermore, we denote by

$$\sqrt{I} := \{f \in \mathbb{K}[\mathbf{x}] \mid f^m \in I \text{ for some } m \in \mathbb{N} \setminus \{0\}\}$$

the radical of  $I$ .

For  $\mathbb{K} = \mathbb{R}$ , we have  $\mathcal{V}(I) = \mathcal{V}^{\mathbb{C}}(I)$ , but one may also be interested in the subset of real solutions, namely the real variety  $\mathcal{V}^{\mathbb{R}}(I) = \mathcal{V}(I) \cap \mathbb{R}^n$ . The corresponding vanishing ideal is  $\mathcal{I}(\mathcal{V}^{\mathbb{R}}(I))$  and the *real radical ideal* is

$$\sqrt[\mathbb{R}]{I} := \{p \in \mathbb{R}[\mathbf{x}] \mid p^{2m} + \sum_j q_j^2 \in I \text{ for some } q_j \in \mathbb{R}[\mathbf{x}], m \in \mathbb{N} \setminus \{0\}\}.$$

---

<sup>1</sup>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.

Obviously,

$$I \subseteq \sqrt{I} \subseteq \mathcal{I}(\mathcal{V}^{\mathbb{C}}(I)), \quad I \subseteq \sqrt[{\mathbb{R}}]{I} \subseteq \mathcal{I}(\mathcal{V}^{\mathbb{R}}(I)).$$

An ideal  $I$  is said to be *radical* (resp., *real radical*) if  $I = \sqrt{I}$  (resp.  $I = \sqrt[{\mathbb{R}}]{I}$ ). If  $I \subseteq \mathbb{R}$  is real radical, then  $I$  is radical and  $\mathcal{V}(I) = \mathcal{V}^{\mathbb{R}}(I) \subseteq \mathbb{R}^n$  if  $|\mathcal{V}^{\mathbb{R}}(I)| < \infty$ .

**Definition 2.1.** Given a vector space  $E \subset \mathbb{R}[\mathbf{x}]$  and  $H = \{h_1, \dots, h_{n_2}\} \subset \mathbb{R}[\mathbf{x}]$  we define the quadratic preordering of  $H$  on  $E$  by

$$(1) \quad \mathcal{P}_E(H) = \left\{ \sum_{\epsilon \in \{0,1\}^{n_2}} s_{\epsilon} h_1^{\epsilon_1} \cdots h_{n_2}^{\epsilon_{n_2}} \mid s_{\epsilon} \in \sum E^2, s_{\epsilon} h_1^{\epsilon_1} \cdots h_{n_2}^{\epsilon_{n_2}} \in \langle E \cdot E \rangle \right\}.$$

If  $E = \mathbb{R}[\mathbf{x}]_t$  (resp.  $E = \mathbb{R}[\mathbf{x}]$ ), we denote  $\mathcal{P}_E(H)$  by  $\mathcal{P}_t(H)$  (resp.  $\mathcal{P}(H)$ ). If  $\mathbf{x} = (\mathbf{x}', \mathbf{x}'')$  and  $E = \mathbb{R}[\mathbf{x}']_t$  (resp.  $E = \mathbb{R}[\mathbf{x}']$ ) and  $H \subset \mathbb{R}[\mathbf{x}']$ , we also denote  $\mathcal{P}_E(H)$  by  $\mathcal{P}_t(H)^{\mathbf{x}'}$  (resp.  $\mathcal{P}_H^{\mathbf{x}'}$ ). If  $H = \emptyset$ , we denote respectively  $\mathcal{P}_t, \mathcal{P}, \mathcal{P}_t^{\mathbf{x}'}$  or  $\mathcal{P}^{\mathbf{x}'}$ , where  $\mathcal{P}_t = \sum \mathbb{R}[\mathbf{x}]_t^2$ ,  $\mathcal{P} = \sum \mathbb{R}[\mathbf{x}]^2$ ,  $\mathcal{P}_t^{\mathbf{x}'} = \sum \mathbb{R}[\mathbf{x}']_t^2$  and  $\mathcal{P}^{\mathbf{x}'} = \sum \mathbb{R}[\mathbf{x}']^2$  and sets of sum of squares polynomials.

**Definition 2.2.** For sets of polynomials  $G, H \subset \mathbb{R}[\mathbf{x}]$ , we denote

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

**Definition 2.3.**  $\sqrt[{\mathbb{H}}]{I} = \{p \in \mathbb{R}[\mathbf{x}] \mid \exists m \in \mathbb{N} \text{ s.t. } p^{2m} + \mathcal{P}(H) \in I\}$

**Remark 2.4.** If  $H = \emptyset$ , then  $\sqrt[{\mathbb{H}}]{I} = \sqrt[{\mathbb{R}}]{I}$ .

The following three famous theorems relate vanishing and radical ideals:

**Theorem 2.5.**

- (i) **Hilbert's Nullstellensatz** (see, e.g., [3, §4.1])  $\sqrt{I} = \mathcal{I}(\mathcal{V}^{\mathbb{C}}(I))$  for an ideal  $I \subseteq \mathbb{C}[\mathbf{x}]$ .
- (ii) **Real Nullstellensatz** (see, e.g., [1, §4.1])  $\sqrt[{\mathbb{R}}]{I} = \mathcal{I}(\mathcal{V}^{\mathbb{R}}(I))$  for an ideal  $I \subseteq \mathbb{R}[\mathbf{x}]$ .
- (iii) **Positivstellensatz** (see, e.g., [1, §4.4])  $\sqrt[{\mathbb{H}}]{I} = \mathcal{I}(\mathcal{W}(I, H))$  for an ideal  $I \subseteq \mathbb{R}[\mathbf{x}]$ .

**2.2. Minimization problem.** Let  $f, g_1, \dots, g_{n_1}, h_1, \dots, h_{n_2} \in \mathbb{R}[\mathbf{x}]$  be polynomial functions. The minimization problem that we consider all along the paper is the following: compute

$$(2) \quad \begin{aligned} &\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \\ &\text{s.t. } g_1(\mathbf{x}) = \cdots = g_{n_1}(\mathbf{x}) = 0 \\ &h_1(\mathbf{x}) \geq 0, \dots, h_{n_2}(\mathbf{x}) \geq 0 \end{aligned}$$

We denote  $\mathcal{S}$  the basic semi-algebraic set  $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) = \cdots = g_{n_1}(\mathbf{x}) = 0, h_1(\mathbf{x}) \geq 0, \dots, h_{n_2}(\mathbf{x}) \geq 0\}$  and by  $H = \{h_1, \dots, h_{n_2}\}$ .

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

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

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

**Definition 2.6.** A point  $\mathbf{x}^*$  is called a KKT point if there exists  $u_1, \dots, u_{n_1}, v_1, \dots, v_{n_2} \in \mathbb{R}$  s.t.

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

The corresponding minimization problem is the following:

$$(3) \quad \begin{aligned} & \min_{(\mathbf{x}, \mathbf{u}, \mathbf{v}) \in \mathbb{R}^{n+n_1+n_2}} f(x) \\ & \text{s.t. } F_1 = \dots = F_n = 0 \\ & \quad g_1 = \dots = g_{n_1} = 0 \\ & \quad v_1 h_1 = \dots = v_{n_2} h_{n_2} = 0 \\ & \quad h_1 \geq 0, \dots, h_{n_2} \geq 0 \end{aligned}$$

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

This leads to the definitions of the following ideal and varieties:

**Definition 2.7.** The Karush-Kuhn-Tucker (KKT) ideal associated to problem (2) is

$$(4) \quad I_{KKT} = (F_1, \dots, F_n, g_1, \dots, g_{n_1}, v_1 h_1, \dots, v_{n_2} h_{n_2}) \subset \mathbb{R}[\mathbf{y}].$$

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

The set of KKT points of  $\mathcal{S}$  is denoted  $\mathcal{S}_{KKT}$  and a KKT-minimizer of  $f$  on  $\mathcal{S}$  is a point  $\mathbf{x}^* \in \mathcal{S}_{KKT}$  such that  $f(\mathbf{x}^*) = \min_{\mathbf{x} \in \mathcal{S}_{KKT}} f(\mathbf{x})$ . A minimizer of  $f$  on  $\mathcal{S}$  is not necessarily a KKT point, but often it is possible to reformulate the minimization problem so that the minimizers become KKT points.

*The gradient variety.* Replacing the inequalities  $h_i \geq 0$  by the equalities  $h_i - s_i^2 = 0$  (adding new variables  $s_i$ ) and introducing Lagrange multipliers yields the following minimization problem:

$$(5) \quad \begin{aligned} & \min_{(\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s}) \in \mathbb{R}^{n+n_1+2n_2}} f(\mathbf{x}) \\ & \text{s.t. } \nabla F(\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s}) = 0 \end{aligned}$$

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

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  $J^{\mathbf{x}} = J \cap \mathbb{R}[\mathbf{x}]$  and  $J^{\mathbf{y}} = J \cap \mathbb{R}[\mathbf{y}]$ . The projection of  $\mathbb{C}^{n+n_1+2n_2}$  on  $\mathbb{C}^n$  (resp.  $\mathbb{C}^{n+n_1+n_2}$ ) is denoted  $\pi^{\mathbf{x}}$  (resp.  $\pi^{\mathbf{y}}$ ).

**Definition 2.8.** We define the gradient ideal of  $F(\mathbf{z})$ :

$$I_{grad} = (\nabla F(\mathbf{z})) = (F_1, \dots, F_n, g_1, \dots, g_{n_1}, h_1 - s_1^2, \dots, h_{n_2} - s_{n_2}^2, v_1 s_1, \dots, v_{n_2} s_{n_2}) \subset \mathbb{R}[\mathbf{z}]$$

where  $F_i = \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_1} u_j \frac{\partial g_j}{\partial x_i} - \sum_{j=1}^{n_2} v_j \frac{\partial h_j}{\partial x_i}$ . The gradient variety is  $V_{grad} = \mathcal{V}(I_{grad})$  and we denote  $V_{grad}^{\mathbf{x}} = \pi^{\mathbf{x}}(V_{grad})$ ,  $V_{grad}^{\mathbf{y}} = \pi^{\mathbf{y}}(V_{grad})$ .

**Definition 2.9.** For any  $F \in \mathbb{R}[\mathbf{z}]$ , the values of  $F$  at the (real) points of  $\mathcal{V}(\nabla F)$  are called the (real) critical values of  $F$ .

If  $F = f - \sum_{i=1}^{n_1} u_i g_i - \sum_{j=1}^{n_2} v_j (h_j - s_j^2)$ , they will also be called the (real) KKT critical values of  $f$ .

**Remark 2.10.** 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,  $h_i(\mathbf{x}) \geq 0$  for  $i = 1, \dots, n_2$ , so that  $V_{KKT}^{\mathbb{R}} \cap \mathcal{W}_+(H) = \pi^{\mathbf{y}}(V_{grad}^{\mathbb{R}})$ .

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

**Proposition 2.11.**  $I_{KKT} = I_{grad} \cap \mathbb{R}[\mathbf{y}]$ .

*Proof.* As  $s_i(s_i v_i) + v_i(h_i - s_i^2) = v_i h_i \forall i = 1, \dots, n_2$ , 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  $s_i(s_i v_i) + v_i(h_i - s_i^2) = v_i h_i$  ( $i = 1, \dots, n_2$ ) are the only S-polynomials involving the variables  $s_1, \dots, s_{n_2}$  which may have a non-trivial reduction. Thus  $K \cap \mathbb{R}[\mathbf{y}]$  is also the Groebner basis of  $F_1, \dots, F_n, g_1, \dots, g_{n_1}, v_1 h_1, \dots, v_{n_2} h_{n_2}$  and we have  $(K) \cap \mathbb{R}[\mathbf{y}] = I_{grad} \cap \mathbb{R}[\mathbf{y}] = I_{KKT}$ .  $\square$

*The minimizer variety.* We easily check the following property:

**Lemma 2.12.**  $F|_{V_{grad}} = f|_{V_{grad}}$ .

Thus minimizing  $F$  on  $V_{grad}$  is the same as minimizing  $f$  on  $V_{grad}$ .

**Definition 2.13.** We define the minimizer variety and ideal of  $f$  on  $\mathcal{S}$  as:

$$\begin{aligned} V_{min} &= \{\mathbf{z}^* = (\mathbf{x}^*, \mathbf{u}^*, \mathbf{v}^*, \mathbf{s}^*) \in \mathbb{R}^{n+n_1+2n_2} \text{ s.t. } \nabla F(\mathbf{x}^*, \mathbf{u}^*, \mathbf{v}^*, \mathbf{s}^*) = 0 \text{ and } f(\mathbf{x}^*) \text{ is minimum}\} \\ I_{min} &= \mathcal{I}(V_{min}). \end{aligned}$$

We will also denote  $V_{min}^{\mathbf{y}} = \pi^{\mathbf{y}}(V_{min})$  and  $V_{min}^{\mathbf{x}} = \pi^{\mathbf{x}}(V_{min})$ .

Notice that  $I_{min}$  is a real radical ideal.

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

By definition, a point  $\mathbf{z}^* = (\mathbf{x}^*, \mathbf{u}^*, \mathbf{v}^*, \mathbf{s}^*)$  in  $V_{min}$  is such that  $\mathbf{x}^* = \pi^{\mathbf{x}}(\mathbf{z}^*)$  is a KKT-minimizer.

By construction,  $I_{KKT} \subset I_{min}$  and  $I_{min} \neq (1)$  if the minimum  $f^*$  is reached in  $V_{KKT}^{\mathbb{R}} \cap \mathcal{W}_+(H)$ .

**Remark 2.14.** If we take  $f = 0$  in the minimization problem (2), then all the points of  $\mathcal{S}$  are KKT-minimizers and  $I_{min} = \mathcal{I}(\mathcal{S}) = \sqrt[\mathbb{H}]{(G)}$ . Moreover,  $I_{KKT} \cap \mathbb{R}[\mathbf{x}] = (g_1, \dots, g_{n_1})$  since  $F_1, \dots, F_n, v_1 h_1, \dots, v_{n_2} h_{n_2}$  are homogeneous of degree 1 in the variables  $\mathbf{u}, \mathbf{v}$ .

### 3. HIERARCHY OF RELAXATION PROBLEMS

In this section, we describe the sequence of finite dimensional optimization problems that we consider. It is a generalization of Lasserre hierarchy of relaxation problems, in which we control the vector space of polynomials on which we apply Semi-Definite Programming methods.

**Definition 3.1.** Given a vector space  $E \subset \mathbb{R}[\mathbf{y}]$  and finite subsets  $G, H \subset \mathbb{R}[\mathbf{y}]$ , we define

$$(6) \quad \mathcal{P}_{E,G,H} := \{g + h \mid g \in \langle G \rangle \cap \langle E \cdot E \rangle, h \in \mathcal{P}_E(H)\}.$$

If  $E = \mathbb{R}[\mathbf{y}]_t$  and  $G' = G_{(2t)}$ , we denote  $\mathcal{P}_{E,G',H}$  by  $\mathcal{P}_{t,G,H}$  and  $\mathcal{P}_{E,G,H}^{\mathbf{x}} = \mathcal{P}_{E \cap \mathbb{R}[\mathbf{x}],G,H}$ .

By definition,  $\mathcal{P}_{E,G,H}$  is a cone of polynomials included in the cone of polynomials of  $\langle E \cdot E \rangle$  which are positive on the semi-algebraic set  $\mathcal{S}$ . When  $E$  is a vector space of finite dimension,  $\mathcal{P}_{E,G,H}$  is the projection of the sum of a linear space and the convex cone of positive quadratic forms on  $E^* \times E^*$ .

Notice that if  $E \subset E', G \subset G'$  and  $H \subset H'$  then  $\mathcal{P}_{E,G,H} \subset \mathcal{P}_{E',G',H'}$ .

We consider now its dual cone.



**Definition 3.2.** Given a vector space  $E \subset \mathbb{R}[\mathbf{y}]$  and finite subsets  $G, H \subset \mathbb{R}[\mathbf{y}]$ , we define

$$(7) \quad \mathcal{L}_{E,G,H} := \{\Lambda \in \langle E \cdot E \rangle^* \mid \Lambda(p) \geq 0, \forall p \in \mathcal{P}_{E,G,H}\}.$$

If  $E = \mathbb{R}[\mathbf{y}]_t$  and  $G' = G_{\langle 2t \rangle}$ , we also denote  $\mathcal{L}_{E,G',H}$  by  $\mathcal{L}_{t,G,H}$  and  $\mathcal{L}_{E,G,H}^{\mathbf{x}} = \mathcal{L}_{E \cap \mathbb{R}[\mathbf{x}],G,H}$ .

By this definition, for any element  $\Lambda \in \mathcal{L}_{E,G,H}$  and any  $g \in G$ , we have  $\Lambda(g) = 0$ .

When  $E$  is a vector space of finite dimension,  $\mathcal{L}_{E,G,H}$  is the intersection of the closed convex cone of semi-definite positive quadratic forms on  $E^* \times E^*$  with a linear space, thus it is a convex closed semi-algebraic set. More details on its description will be given in Section 7.

Notice that if  $E \subset E'$ ,  $G \subset G'$  and  $H \subset H'$  then  $\mathcal{L}_{E',G',H} \subset \mathcal{L}_{E,G,H}$ .

We consider now truncated Hankel operators, which will play a central role in the construction of the minimizer ideal of  $f$  on  $\mathcal{S}$ .

**Definition 3.3.** For a vector space  $E \subset \mathbb{R}[\mathbf{y}]$  and a linear form  $\Lambda \in \langle E \cdot E \rangle^*$ , we define the map  $M_{\Lambda}^E : E \rightarrow E^*$  by  $M_{\Lambda}^E(p) = p \cdot \Lambda$  for  $p \in E$ . Thus  $M_{\Lambda}^E$  is called a truncated Hankel operator, defined on the subspace  $E$ .

**Definition 3.4.** We define the kernel of the truncated Hankel operator:

$$(8) \quad \ker M_{\Lambda}^E = \{p \in E \mid p \cdot \Lambda = 0, \text{ i.e., } \Lambda(pq) = 0 \forall q \in E\}.$$

When  $E = \mathbb{R}[\mathbf{y}]_t$  for  $t \in \mathbb{N}$ ,  $M_{\Lambda}^E$  is also denoted  $M_{\Lambda}^t$ .

Given a subspace  $E_0 \subset E$ ,  $\Lambda$  induces a linear map on  $\langle E_0 \cdot E_0 \rangle$  and we can consider the induced truncated Hankel operator  $M_{\Lambda}^{E_0} : E_0 \rightarrow (E_0)^*$  as the ‘‘restriction’’ of  $M_{\Lambda}^E$ . Below we will deal with the following sets and minimums, in order to define our primal-dual problems.

We will need the following result:

**Lemma 3.5.** For any vector space  $E$  and  $G = \{0\}$  and  $\Lambda, \Lambda' \in \langle E \cdot E \rangle^*$ , we have:

- $\forall p \in E, \Lambda(p^2) = 0$  implies  $p \in \ker M_{\Lambda}^E$ .
- $\ker M_{\Lambda+\Lambda'}^E = \ker M_{\Lambda}^E \cap \ker M_{\Lambda'}^E$ .

*Proof.* The first point is a consequence of the positivity of  $M_{\Lambda}^E$ .

For the second point,  $\forall p, q \in E, \forall t \in \mathbb{R}, \Lambda((p+ tq)^2) = t^2 \Lambda(q^2) + 2t \Lambda(pq) \geq 0$ . Dividing by  $t$  and letting  $t$  go to zero yields  $\Lambda(pq) = 0$ , thus showing  $p \in \ker M_{\Lambda}^E$ . The inclusion  $\ker M_{\Lambda}^E \cap \ker M_{\Lambda'}^E \subset \ker M_{\Lambda+\Lambda'}^E$  is obvious.

Conversely, let  $p \in \ker M_{\Lambda+\Lambda'}^E$ . In particular,  $(\Lambda + \Lambda')(p^2) = 0$ , which implies  $\Lambda(p^2) = \Lambda'(p^2) = 0$  (since  $\Lambda(p^2), \Lambda'(p^2) \geq 0$ ) and thus  $p \in \ker M_{\Lambda}^E \cap \ker M_{\Lambda'}^E$ .  $\square$

**Definition 3.6.** Let  $E$  be a subspace of  $\mathbb{R}[\mathbf{y}]$  such that  $1 \in E$  and  $f \in \langle E \cdot E \rangle$  and finite subsets  $G, H \subset \mathbb{R}[\mathbf{y}]$ . We define the following extrema:

- $f^* = \inf_{\mathbf{x} \in \mathcal{S}_{KKT}} f(\mathbf{x})$ ,
- $f_{E,G,H}^{\mu} = \inf \{\Lambda(f) \text{ s.t. } \Lambda \in \mathcal{L}_{E,G,H}, \Lambda(1) = 1\}$ ,
- $f_{E,G,H}^{sos} = \sup \{\gamma \in \mathbb{R} \text{ s.t. } f - \gamma \in \mathcal{P}_{E,G,H}\}$ .

If  $E = \mathbb{R}[\mathbf{y}]_t$  (resp.  $E = \mathbb{R}[\mathbf{y}]$ ) and  $G' = G_{\langle 2t \rangle}$  we also denote  $f_{E,G',H}^{\mu}$  by  $f_{t,G,H}^{\mu}$  (resp.  $f_{(G),H}^{\mu}$ ) and  $f_{E,G',H}^{sos}$  by  $f_{t,G,H}^{sos}$  (resp.  $f_{(G),H}^{sos}$ ).

**Remark 3.7.** By convention if the sets are empty,  $f_{E,G,H}^{sos} = -\infty$  and  $f_{E,G,H}^{\mu} = +\infty$ .

We now analyse the relations between these different extrema.

**Remark 3.8.** Let  $E \subset E'$  be two subspaces of  $\mathbb{R}[\mathbf{y}]$  and  $G \subset \langle E \cdot E \rangle, G' \subset \langle E' \cdot E' \rangle$  with  $G \subset G'$  then we directly deduce that

- $f_{E,G,H}^\mu \leq f_{E',G',H}^\mu$ ,
- $f_{E,G,H}^{sos} \leq f_{E',G',H}^{sos}$ ,

from the fact that  $\mathcal{L}_{E',G',H} \subset \mathcal{L}_{E,G,H}$  and  $\mathcal{P}_{E,G,H} \subset \mathcal{P}_{E',G',H}$ .

If we suppose  $V_{min} \neq \emptyset$  and we take  $V_{min}^{\mathbf{y}} \subset \mathcal{W}^{\mathbb{R}}(G, H)$ , we have the following relations between the extrema.

**Proposition 3.9.** *Let  $G \subset \mathbb{R}[\mathbf{y}]$  such that  $V_{min}^{\mathbf{y}} \subset \mathcal{W}(G, H)$ . Then  $f_{E,G,H}^{sos} \leq f_{E,G,H}^\mu \leq f^*$ .*

*Proof.* We have  $f_{E,G,H}^{sos} \leq f_{E,G,H}^\mu$  because if there exists  $\gamma \in \mathbb{R}$  such that  $f - \gamma \in \mathcal{P}_{E,G,H}$ , i.e.  $f - \gamma = s + g$  with  $s \in \mathcal{P}_{E,H}$  and  $g \in \langle G \rangle \cap \langle E \cdot E \rangle$ , then  $\forall \Lambda \in \mathcal{L}_{E,G,H}$ ,  $\Lambda(f - \gamma) = \Lambda(f) - \gamma \geq 0$ . We deduce that  $\Lambda(f) \geq \gamma$  and we conclude  $f_{E,G,H}^{sos} \leq f_{E,G,H}^\mu$ . For the second inequality, let  $\mathbf{y}^*$  be a point of  $V_{min}^{\mathbf{y}}$  such that  $f(\mathbf{x}^*)$  is the minimum and let  $\mathbf{1}_{\mathbf{y}^*} \in \mathbb{R}[\mathbf{y}]^* : p \mapsto p(\mathbf{y}^*)$  be the evaluation at  $\mathbf{y}^*$ . Then we have  $\mathbf{1}_{\mathbf{y}^*}(1) = 1$  and  $\mathbf{1}_{\mathbf{y}^*}(p) \geq 0 \forall p \in \mathcal{P}_{E,G,H}$  such that  $p = g + q$  with  $g \in \langle G \rangle \cap \langle E \cdot E \rangle$  and  $h \in \mathcal{P}_{E,H}$  since  $\mathbf{1}_{\mathbf{y}^*}(p) = p(\mathbf{y}^*) = g(\mathbf{y}^*) + h(\mathbf{y}^*)$ ,  $g(\mathbf{y}^*) = 0$  and  $h(\mathbf{y}^*) \geq 0$  due to  $V_{min}^{\mathbf{y}} \subset \mathcal{W}(G, H)$ , so that  $\mathbf{1}_{\mathbf{y}^*} \in \mathcal{L}_{E,G,H}$ . We deduce the inequality  $f_{E,G,H}^\mu \leq f^*$ .  $\square$

Following the relaxation approach proposed in [15], we are now going to consider a hierarchy of convex optimization problems and show that for such a hierarchy, the minimum  $f^*$  is always reached in a finite number of steps. Let us consider the sequence of spaces

$$\cdots \subset \mathcal{L}_{t+1,G,H} \subset \mathcal{L}_{t,G,H} \subset \cdots \quad \text{and} \quad \cdots \subset \mathcal{P}_{t,G,H} \subset \mathcal{P}_{t+1,G,H} \subset \cdots$$

for  $t \in \mathbb{N}$ . Using Remark 3.8 and  $V_{min}^{\mathbf{y}} \subset \mathcal{W}(G, H)$  the fact that  $(G) = \cup_{t \in \mathbb{N}} G_{(t)}$ , we check that

- the increasing sequence  $\cdots f_{t,G,H}^\mu \leq f_{t+1,G,H}^\mu \leq \cdots \leq f^*$ , for  $t \in \mathbb{N}$  converges to  $f_{(G),H}^\mu \leq f^*$ ,
- the increasing sequence  $\cdots f_{t,G,H}^{sos} \leq f_{t+1,G,H}^{sos} \leq \cdots \leq f^*$ , for  $t \in \mathbb{N}$  converges to  $f_{(G),H}^{sos} \leq f^*$ .

#### 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  $\mathcal{V}(J_{grad}) = V_{grad}$ .

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 [26, Lemma 3.3]<sup>2</sup>.

**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 [34]. Let  $x, y$  be two arbitrary points of  $V$ . There exists a piecewise smooth path  $\varphi(t)$  ( $0 \leq t \leq 1$ ) lying inside  $V$  such that  $x = \varphi(0)$  and  $y = \varphi(1)$ . Without loss of generality, we can assume that  $\varphi$  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  $t_1 \in (0, 1)$

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

<sup>2</sup>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  $\nabla f$  vanishes on  $V$ . Then  $\operatorname{Re}(f(y)) = \operatorname{Re}(f(x))$ . We have the same result for the imaginary part: for some  $t_2 \in (0, 1)$

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

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

**Lemma 4.2.** *The ideal  $J_{grad}$  can be decomposed as  $J_{grad} = J_0 \cap J_1 \cap \dots \cap J_s$  with  $V_i = \mathcal{V}(J_i)$  and  $W_i = \overline{\pi^{\mathbf{x}}(V_i)}^{\mathbb{R}}$  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  $i \neq j$ ,
- $W_i^{\mathbb{R}} \cap \mathcal{W}_+(H) \neq \emptyset$  for  $i = 0, \dots, r$ ,
- $W_i^{\mathbb{R}} \cap \mathcal{W}_+(H) = \emptyset$  for  $i = r+1, \dots, s$ ,
- $f_0 < \dots < f_r$ .

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

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

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 2.12, it coincides with  $f$  on each  $\mathcal{V}(Q_i)$ . We group the primary components  $Q_i$  according to the values  $f_0, \dots, f_s$  of  $f$  on these components, into  $J_0, \dots, J_s$  so that  $f(\mathcal{V}(J_j)) = f_j$  with  $f_i \neq f_j$  if  $i \neq j$ .

We can number them so that  $\overline{\pi^{\mathbf{x}}(V_i)}^{\mathbb{R}} \cap \mathcal{W}_+(H)$  is empty for  $i = r+1, \dots, s$  and contains a real point  $\mathbf{x}_i$  for  $i = 0, \dots, r$ . Notice that such a point  $\mathbf{x}_i$  is in  $\mathcal{S}$ , since it satisfies  $g_j(\mathbf{x}_i) = 0$  for  $j = 1, \dots, n_1$ ,  $h_k(\mathbf{x}_i) \geq 0$  for  $k = 1, \dots, n_2$ . As it is the limit of projections of points in  $\mathcal{V}(J_i)$  on which  $f$  is constant, we have  $f_i = f(\mathbf{x}_i) \in \mathbb{R}$  for  $i = 0, \dots, r$ . We can then order  $J_0, \dots, J_r$  so that  $f_0 < \dots < f_r$ .  $\square$

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

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

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

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

*Proof.* Let  $(L_i)_{i=0, \dots, s}$  be the univariate Lagrange interpolation polynomials at the values  $f_0, \dots, f_s \in \mathbb{C}$  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  $\{f_{r+1}, \dots, f_s\}$  is stable by conjugation and  $f_0, \dots, f_r \in \mathbb{R}$ , by construction of the Lagrange interpolation polynomials we deduce that  $q_0, \dots, q_r \in \mathbb{R}[\mathbf{x}]$ .

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  $q_i^N q_j^N = 0 \pmod{\bigcap_i J_i = J_{grad}}$  for  $i \neq j$ , we deduce that there exists  $N' \in \mathbb{N}$

such that

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

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  $j = 0, \dots, r$ , we have  $p_j \in \mathbb{R}[\mathbf{x}]$  for  $j = 0, \dots, r$ , which ends the proof of this lemma.  $\square$

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

*Proof.* As  $\bigcup_{i>r} \overline{\pi^{\mathbf{x}}(V_i)}^{\mathbb{R}} \cap \mathcal{W}_+(H) = \mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i \cap \mathbb{R}[\mathbf{x}]) \cap \mathcal{W}_+(H) = \mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i^{\mathbf{x}}) \cap \mathcal{W}_+(H) = \emptyset$ , we have  $\mathcal{I}(\mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i^{\mathbf{x}}) \cap \mathcal{W}_+(H)) = \mathbb{R}[\mathbf{x}] \ni 1$  and by the Positivstellensatz (Theorem 2.5 (iii)),

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

$\square$

**Remark 4.7.** If  $V_{grad}^{\mathbb{R}} = \emptyset$ , then  $-1 \in \mathcal{P}^{\mathbf{x}}(H) + (\bigcap_{i=0}^s J_i^{\mathbf{x}}) = \mathcal{P}^{\mathbf{x}}(H) + J_{grad}^{\mathbf{x}}$  and  $\forall p \in \mathbb{R}[\mathbf{x}]$ ,  $p = \frac{1}{4}((p+1)^2 - (p-1)^2) \in \mathcal{P}^{\mathbf{x}}(H) + J_{grad}^{\mathbf{x}}$ .

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  $\deg(q) \leq t$  and  $f + \epsilon = q^2 \pmod{J}$ .

*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

$$\left(1 + \left(\frac{f+\epsilon}{f^*+\epsilon} - 1\right)\right)^{1/2} \equiv \sum_i^{l-1} \binom{1/2}{i} \left(\frac{f+\epsilon}{f^*+\epsilon} - 1\right)^i \stackrel{def}{=} \frac{q}{\sqrt{f^*+\epsilon}} \pmod{J}$$

Then  $f + \epsilon = q^2 \pmod{J}$ .  $\square$

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

**Theorem 4.9.** Let  $f \in \mathbb{R}[\mathbf{x}]$ , let  $f_0 < \dots < f_r$  be its real KKT critical values of  $f$  on  $S$  and let  $p_0, \dots, p_r$  be the associated polynomials defined in Lemma 4.5.

- (1)  $f - \sum_{i=0}^r f_i p_i^2 \in \mathcal{P}^{\mathbf{x}}(H) + \sqrt{J_{grad}^{\mathbf{x}}}$ .
- (2) If  $f \geq 0$  on  $S_{KKT}$ , then  $f \in \mathcal{P}^{\mathbf{x}}(H) + \sqrt{J_{grad}^{\mathbf{x}}}$ .
- (3) If  $f > 0$  on  $S_{KKT}$ , then  $f \in \mathcal{P}^{\mathbf{x}}(H) + J_{grad}^{\mathbf{x}}$ .

*Proof.* By Lemma 4.5, we have

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

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

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

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

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

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

which proves the first point.

If  $f \geq 0$  on  $\mathcal{S}_{KKT}$ , then  $f_i \geq 0$  for  $i = 0, \dots, r$  and  $\sum_{i=0}^r f_i p_i^2 \in \mathcal{P}^{\mathbf{x}}(H)$  so that

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

which proves the second point.

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

$$\sum_{i=0}^r f_i p_i^2 = \sum_{i=0}^r q_i^2 \pmod{J_{grad}^{\mathbf{x}}} \in \mathcal{P}^{\mathbf{x}}(H) + J_{grad}^{\mathbf{x}}.$$

This proves the third point.  $\square$

Notice that the points (2) and (3) of this theorem are generalization of results [5] on the representation of positive polynomials, which only involves polynomials in  $\mathbb{R}[\mathbf{x}]$ .

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  $\mathcal{P}^{\mathbf{x}}(H) + J_{grad}^{\mathbf{x}}$ .

**Theorem 4.10.** *let  $H = \{h_1, \dots, h_{n_2}\}$  and  $G \subset \mathbb{R}[\mathbf{z}]$  such that  $J_{grad} = (G)$ . If  $f \geq 0$  on  $\mathcal{S}_{KKT}$ , then there exists  $t_0$  such that  $\forall \epsilon > 0$ ,*

$$f + \epsilon \in \mathcal{P}_{t_0, G, H}^{\mathbf{x}}.$$

*Proof.* Let us look more carefully at the decompositions obtained in the proof of Theorem 4.9. As  $J_{grad} = (G)$ , we can choose  $t_0 \in \mathbb{N}$  big enough such that

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

and  $\deg(p_i) < t_0/2$ . Then  $\forall \epsilon > 0$ ,

$$(9) \quad \sum_{i>r} (f + \epsilon) p_i^2 = \sum_{i>r} f p_i^2 + \sum_{i>r} \epsilon p_i^2 \in \mathcal{P}_{t_0, G, H}^{\mathbf{x}}.$$

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

$$(10) \quad (f + \epsilon) p_i^2 = q_i^2 \pmod{G_{(t_0)} \cap \mathbb{R}[\mathbf{x}]}$$

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

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

$$(11) \quad (f + \epsilon) - \sum_{i=0}^s (f + \epsilon) p_i^2 \in G_{(t_0)} \cap \mathbb{R}[\mathbf{x}].$$

From Equations (9), (10), (11), we deduce that if  $t_0 \in \mathbb{N}$  is big enough,  $\forall \epsilon > 0$

$$f + \epsilon \in \mathcal{P}_{t_0, G, H}^{\mathbf{x}},$$

which concludes the proof of the theorem.  $\square$

## 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 a solution of the relaxation problem. Two types of minimization problems are considered. One involves the KKT-ideal and polynomials in the variables  $\mathbf{y}$  and the other one is obtained by intersection with the ring of polynomials in  $\mathbf{x}$ . In this section, we assume that  $V_{min} \neq \emptyset$ , that is  $f^*$  is reached for some point  $\mathbf{x}^* \in \mathcal{S}_{KKT}$ .

**Proposition 5.1.** *Let  $G, H$  be finite subsets of  $\mathbb{R}[\mathbf{y}]$ . There exists  $t_0 \in \mathbb{N}$  such that  $\forall t \geq t_0$ ,  $\forall \Lambda \in \mathcal{L}_{t,G,H}$ ,  $\sqrt[t]{G} \subset (\ker M_\Lambda^t)$ .*

*Proof.* Let  $G = \{g_1, \dots, g_l\}$  and let  $q_1, \dots, q_k$  be generators of  $J := \sqrt[t]{G}$ . By the Positivstellensatz, for  $j \in 1, \dots, k$ , there exist  $m_j \in \mathbb{N}^*$  and polynomials  $u_r^{(j)} \in \mathbb{R}[\mathbf{y}]$  and  $\sigma_j \in \mathcal{P}(H)$  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  $\sigma_j \in \mathcal{P}_{t_0, H}$ . Then for all  $t \geq t_0$  and all  $\Lambda \in \mathcal{L}_{t,G,H}$ , 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  $(q_1, \dots, q_l) = J \subset (\ker M_\Lambda^t)$ .  $\square$

**Remark 5.2.** *With the same arguments, we can show that for any  $t' \in \mathbb{N}$ , there exists  $t'_0 \geq t'$  such that  $\forall t \geq t'_0$ ,  $\forall \Lambda \in \mathcal{L}_{t,G,H}$ ,*

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

where  $Q = \{q_1, \dots, q_k\}$  is a generating family of  $J = \sqrt[t]{G}$ .

The next result shows that in the sequence of optimization problems that we consider, the minimum of  $f$  on  $\mathcal{S}_{KKT}$  is reached from a certain degree.

**Theorem 5.3.** *For  $G \subset \mathbb{R}[\mathbf{y}]$  with  $V_{min}^{\mathbf{y}} \subset \mathcal{W}(G, H) \subset V_{KKT}^{\mathbb{R}}$ , there exists  $t_1 \geq 0$  such that  $\forall t \geq t_1$ ,*

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

*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  $h \in \mathcal{P}(H)^{\mathbf{x}}$  and  $g \in \sqrt{I_{grad}^{\mathbf{x}}} \subset \sqrt[t]{I_{grad}^{\mathbf{y}}} = \sqrt[t]{I_{KKT}^{\mathbb{R}}}$  (by Proposition 2.11). Since  $\mathcal{W}(G, H) \subset V_{KKT}^{\mathbb{R}}$ , by the Positivstellensatz, we have  $g \in \sqrt[t]{I_{KKT}^{\mathbb{R}}} \subset \sqrt[t]{(G)}$ . By proposition 5.1, there exists  $t_1 \geq t_0$  such that for all  $t \geq t_1$ , for all  $\Lambda \in \mathcal{L}_{t,G,H}$ ,  $\Lambda(g) = 0$ ,  $\Lambda(h) \geq 0$ .

Let us fix  $t \geq t_1$  and  $\Lambda^* \in \mathcal{L}_{t,G,H}$  such that  $\Lambda^*(f) = f_{t,G,H}^\mu$ . 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  $\Lambda^*(h) \geq 0$  ( $h \in \mathcal{P}_{t,H}$ ), we deduce that  $\Lambda^*(f - f^*) = \Lambda^*(f) - f^* \geq 0$ .

As  $V_{min} \neq \emptyset$ ,  $V_{min}^y \subset \mathcal{V}^{\mathbb{R}}(G)$ , by Proposition 3.9, we have  $\Lambda^*(f) \leq f^*$ , so that  $\Lambda^*(f) = f_{t,G,H}^\mu = f^*$ , which proves the first point. Hence for  $i = 1, \dots, r$ ,  $\Lambda^*(p_i^2) = 0$  and  $p_i \in \ker M_{\Lambda^*}^t$ , which proves the second point.

To prove that  $f_{t,G,H}^{sos} = f^*$  when  $\mathcal{V}(G) \subset V_{KKT}$ , we apply Theorem 4.10 to  $f - f^*$  which is positive on  $\mathcal{S}_{KKT}$ . Let us take  $J_{grad} = (G) \cap I_{grad} \subset \mathbb{R}[\mathbf{z}]$  so that  $\mathcal{V}(J_{grad}) = \mathcal{V}(G) \cup V_{grad} = V_{grad}$  since  $\mathcal{V}(G) \subset V_{KKT}$ . We denote by  $G'$  a finite family of generators of  $J_{grad}$ .

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

$$f - f^* + \epsilon \in \mathcal{P}_{t_0,H}^{\mathbf{x}} + G'_{\langle t_0 \rangle} \cap \mathbb{R}[\mathbf{y}].$$

As  $G' \subset (G)$ , we can choose  $t_1 \geq t'_0$  such that  $G'_{\langle t_0 \rangle} \subset G_{\langle t_1 \rangle}$ .

Then  $\forall t \geq t_1$ ,  $f - f^* + \epsilon \in \mathcal{P}_{t,H}^{\mathbf{x}} + G_{\langle t \rangle} = \mathcal{P}_{t,G,H}$ . Hence by maximality,  $\forall \epsilon > 0$ ,  $f^* - \epsilon \leq f_{t,G,H}^{sos}$ . We deduce that  $f^* \leq f_{t,G,H}^{sos}$ , and by Proposition 3.9, we have  $f_{t,G,H}^{sos} = f_{t,G,H}^\mu = f^*$ .  $\square$

A similar result holds for polynomials in the variables  $\mathbf{x}$ :

**Corollary 5.4.** *For  $G \subset \mathbb{R}[\mathbf{x}]$  with  $V_{min}^{\mathbf{x}} \subset \mathcal{W}(G, H) \subset \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$ , there exists  $t_1 \geq 0$  such that  $\forall t \geq t_1$ ,*

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

*Proof.* Let us take  $G' \subset \mathbb{R}[\mathbf{y}]$  such that  $(G') = (G) + I_{KKT} \in \mathbb{R}[\mathbf{y}]$ . Then  $V_{min}^y \subset \mathcal{V}^{\mathbb{R}}(G') \cap \mathcal{W}_+(H) \subset V_{KKT}^{\mathbb{R}}$  and by Theorem 5.3,  $f_{t,G',H}^\mu = f^*$  is reached for some  $\Lambda^* \in \mathcal{L}_{t,G',H}$  and  $p_i \in \ker M_{\Lambda^*}^t$ .

Let  $\Lambda'$  be the restriction of  $\Lambda^* \in \mathcal{L}_{t,G',H}$  to  $\mathbb{R}[\mathbf{x}]_{2t}$ . By construction,  $\Lambda^*(f) = \Lambda'(f) = f^*$ ,  $\Lambda' \in \mathcal{L}_{t,G,H}$  and  $p_i \in \ker M_{\Lambda'}^t$ .

If moreover  $\mathcal{V}(G) \subset \overline{\pi^{\mathbf{x}}(V_{grad})}$ , then  $\mathcal{V}(G') \subset \overline{\pi^y(V_{grad})} = V_{KKT}$ . By application of Theorem 4.10,  $\forall \epsilon > 0$ ,  $f - f^* + \epsilon \in \mathcal{P}_{t,G',H}^{\mathbf{x}} \subset \mathcal{P}_{t,G,H}$  and we deduce that  $f_{t,G,H}^{sos} = f^*$ .  $\square$

**Remark 5.5.** *Under some conditions such as  $\mathcal{P}_{G,H}$  is archimedean or Boundary Hessian conditions are satisfied, it is possible to prove  $f - f^* \in \mathcal{P}_{G,H}$  [29, 32, 21, 11] and there is no duality gap:  $f_{t,G,H}^{sos} = f_{t,G,H}^\mu = f^*$  if  $t$  is high enough degree. The exactness of the hierarchy and generators of the minimizer ideal can then be deduced in a similar way.*

As for the construction of generators of  $\sqrt[H]{I_{KKT}}$  (Proposition 5.1), we can construct generators of  $I_{min}^{\mathbf{x}}$  from the kernel of a truncated Hankel operator associated to any linear form which minimizes  $f$ , using the following propositions:

**Proposition 5.6.**  $I_{min} = (p_1, \dots, p_r) + \sqrt[H]{I_{grad}} = (p_1, \dots, p_r) + \sqrt[\mathbb{R}]{I_{grad}}$ .

*Proof.* Using the decomposition of Lemma 4.2 and the polynomials  $p_i$  of Lemma 4.5, we have

$$V_{grad}^{\mathbb{R}} = (V_0 \cup V_1 \cup \dots \cup V_s) \cap \mathbb{R}^{n+n_1+2n_2} = V_0^{\mathbb{R}} \cup \dots \cup V_r^{\mathbb{R}},$$

By construction,  $\mathcal{I}(V_0^{\mathbb{R}}) = I_{min}$ ,  $p_i(V_0^{\mathbb{R}}) = 0$  for  $i = 1, \dots, s$  and  $p_i \in \mathbb{R}[\mathbf{x}]$  for  $i = 0, \dots, r$ . This implies that  $p_i \in I_{min}$  for  $i = 1, \dots, r$ .

As  $V_0^{\mathbb{R}} \subset V_{grad}^{\mathbb{R}}$ , we also have  $\sqrt[H]{I_{grad}} \subset I_{min}$ .

We have proved so far that  $(p_1, \dots, p_r) + \sqrt[H]{I_{grad}} \subset I_{min}$ . In order to prove the reverse inclusion, we denote by  $q_1, \dots, q_m$  a family of generators of the ideal  $I_{min}$ . Take one of these generators  $q_j$  ( $1 \leq j \leq m$ ). By construction,  $q_j p_0(V_0^{\mathbb{R}}) = 0$  and  $q_j p_0(V_i^{\mathbb{R}}) = 0$  for  $i = 1, \dots, r$ , which implies that  $q_j p_0 \in \sqrt[H]{I_{grad}}$ .

By Lemma 4.5, we have the decomposition

$$q_j \equiv q_j(p_0 + p_1 + \dots + p_s) \pmod{I_{grad}} \subset \sqrt[H]{I_{grad}},$$

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

As  $V_{grad}^{\mathbb{R}} = V_{grad}^{\mathbb{R}} \cap \mathcal{W}_+(H)$  (Remark 2.10), by the Positivstellensatz,  $\sqrt[H]{I_{grad}} = \sqrt[\mathbb{R}]{I_{grad}}$ , which proves the second equality.  $\square$

**Corollary 5.7.**  $I_{min}^{\mathbf{y}} = (p_1, \dots, p_r) + \sqrt[\mathbb{Y}]{I_{KKT}}$ .

*Proof.* We obtain the result as a consequence of that  $I_{min}^{\mathbf{y}} = I_{min} \cap \mathbb{R}[\mathbf{y}]$  and  $\sqrt[\mathbf{y}]{I_{grad}} = \sqrt[H]{I_{grad}} \cap \mathbb{R}[\mathbf{y}] = \mathcal{I}(\pi^{\mathbf{y}}(V_{grad}^{\mathbb{R}})) = \mathcal{I}(\pi^{\mathbf{y}}(V_{KKT}^{\mathbb{R}}) \cap \mathcal{W}_+(H)) = \sqrt[\mathbf{y}]{I_{grad}^{\mathbf{y}}}$ . By remark 2.10,  $\pi^{\mathbf{y}}(V_{grad}^{\mathbb{R}}) = V_{KKT}^{\mathbb{R}} \cap \mathcal{W}_+(H)$  then  $\mathcal{I}(\pi^{\mathbf{y}}(V_{grad}^{\mathbb{R}})) = I(V_{KKT}^{\mathbb{R}} \cap \mathcal{W}_+(H)) = \sqrt[\mathbf{y}]{I_{KKT}}$ .  $\square$

**Corollary 5.8.**  $I_{min}^{\mathbf{x}} = (p_1, \dots, p_r) + \sqrt[\mathbf{x}]{I_{KKT}^{\mathbf{x}}}$ .

*Proof.* The proof is similar to 5.7.  $\square$

**Theorem 5.9.** For  $G \subset \mathbb{R}[\mathbf{y}]$  with  $V_{min}^{\mathbf{y}} \subset \mathcal{V}^{\mathbb{R}}(G) \cap \mathcal{W}_+(H) \subset V_{KKT}^{\mathbb{R}}$ , there exists  $t_2 \in \mathbb{N}$  such that  $\forall t \geq t_2$ , for  $\Lambda^* \in \mathcal{L}_{t,G,H}$  with  $\Lambda^* = f_{t,G,H}^{\mu}$ , we have  $I_{min}^{\mathbf{y}} \subset (\ker M_{\Lambda^*}^t)$ .

*Proof.* To prove the inclusion we take  $t_2 = \max\{t_0, t_1\}$  and we combine Corollary 5.7 with Proposition 5.1 and Theorem 5.3.  $\square$

**Corollary 5.10.** For  $G \subset \mathbb{R}[\mathbf{x}]$  with  $V_{min}^{\mathbf{x}} \subset \mathcal{V}^{\mathbb{R}}(G) \cap \mathcal{W}_+(H) \subset \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$ , there exists  $t_2 \in \mathbb{N}$  such that  $\forall t \geq t_2$ , for  $\Lambda^* \in \mathcal{L}_{t,G,H}$  with  $\Lambda^* = f_{t,G,H}^{\mu}$ , we have  $I_{min}^{\mathbf{x}} \subset (\ker M_{\Lambda^*}^t)$ .

*Proof.* To prove the inclusion we take  $t_2 = \max\{t_0, t_1\}$  and we combine Corollary 5.8 with Proposition 5.1 for  $G \subset \mathbb{R}[\mathbf{x}]$  and Corollary 5.4.  $\square$

We introduce now the notion of *generic linear form for  $f$* . Such a linear form will allow us to compute  $I_{min}^{\mathbf{y}}$  as we will see.

**Proposition 5.11.** For  $\Lambda^* \in \mathcal{L}_{E,G,H}$  and  $p \in \mathbb{R}[\mathbf{y}]$ , the following assertions are equivalent:

- (i)  $\text{rank} M_{\Lambda^*}^E = \max_{\Lambda \in \mathcal{L}_{E,G,H}, \Lambda(p) = p_{E,G,H}^{\mu}} \text{rank} M_{\Lambda}^E$ .
- (ii)  $\forall \Lambda \in \mathcal{L}_{E,G,H}$  such that  $\Lambda(p) = p_{E,G,H}^{\mu}$ ,  $\ker M_{\Lambda^*}^E \subset \ker M_{\Lambda}^E$ .
- (iii)  $\text{rank} M_{\Lambda^*}^{E_0} = \max_{\Lambda \in \mathcal{L}_{E,G,H}, \Lambda(p) = p_{E,G,H}^{\mu}} \text{rank} M_{\Lambda}^{E_0}$  for any subspace  $E_0 \subset E$ .

We say that  $\Lambda^* \in \mathcal{L}_{E,G,H}$  is *generic for  $p$*  if it satisfies one of the equivalent conditions (i)-(iii).

*Proof.* (i)  $\implies$  (ii): Note that  $\frac{1}{2}(\Lambda + \Lambda^*) \in \mathcal{L}_{E,G,H}$  and  $\frac{1}{2}(\Lambda + \Lambda^*)(p) = p_{E,G,H}^{\mu}$  and  $\ker M_{\frac{1}{2}(\Lambda + \Lambda^*)}^E = \ker M_{\Lambda}^E \cap \ker M_{\Lambda^*}^E$  (using Lemma 3.5). Hence,  $\text{rank} M_{\frac{1}{2}(\Lambda + \Lambda^*)}^E \geq \text{rank} M_{\Lambda^*}^E$  and thus equality holds. This implies that  $\ker M_{\frac{1}{2}(\Lambda + \Lambda^*)}^E = \ker M_{\Lambda^*}^E$  is thus contained in



$\ker M_{\Lambda}^E$ .

(ii)  $\implies$  (iii): Given  $E_0 \subset E$ , we show that  $\ker M_{\Lambda^*}^{E_0} \subset \ker M_{\Lambda}^{E_0}$ . By Lemma 3.5, we have  $\ker M_{\Lambda^*}^{E_0} \subset \ker M_{\Lambda^*}^E$  and, by the above, we have  $\ker M_{\Lambda^*}^E \subset \ker M_{\Lambda}^E$ .

(iii)  $\implies$  (i) This implication is obvious.  $\square$

**Remark 5.12.** *When  $E$  is a finite dimensional vector space, a linear form  $\Lambda^* \in \mathcal{L}_{E,G,H}$  generic for  $p$  can be computed by solving a Semi-Definite Programming problem by an interior point method [16]. In this case, the solution  $\Lambda^*$  obtained by convex optimization will be in the interior of the face of linear forms which minimize  $p$ .*

The next result, which refines Theorem 5.9, shows that only elements in  $I_{min}^{\mathbf{y}}$  are involved in the kernel of a truncated Hankel operator associated to a generic linear form for  $f$ .

**Theorem 5.13.** *Let  $E \subset \mathbb{R}[\mathbf{y}]$  such that  $1 \in E$  and  $f \in \langle E \cdot E \rangle$  and let  $G, H \subset \langle E \cdot E \rangle$  with  $\emptyset \neq V_{min}^{\mathbf{y}} \subset \mathcal{W}(G, H)$ . If  $\Lambda^* \in \mathcal{L}_{E,G,H}$  is generic for  $f$  and such that  $\Lambda^*(f) = f^*$ , then  $\ker M_{\Lambda^*}^E \subset I_{min}^{\mathbf{y}}$ .*

*Proof.* Let  $p \in \ker M_{\Lambda^*}^E$  and  $\mathbf{y}^* \in V_{min}^{\mathbf{y}}$  with  $\mathbf{x}^* = \pi^{\mathbf{x}}(\mathbf{y}^*)$ . Then  $\mathbf{y}^* \in V_{KKT}^{\mathbb{R}} \cap \mathcal{W}_+(H)$  and  $f(\mathbf{x}^*) = f^*$  is minimum on  $\mathcal{S}_{KKT}$ . Let  $\underline{\mathbf{1}}_{\mathbf{y}^*}$  denotes the evaluation at  $\mathbf{y}^*$  restricted to  $\langle E \cdot E \rangle$ .

Our objective is to show that  $p(\mathbf{y}^*) = 0$ . Suppose for contradiction that  $p(\mathbf{y}^*) \neq 0$ . We know that  $\underline{\mathbf{1}}_{\mathbf{y}^*} \in \mathcal{L}_{E,G,H}$  since  $V_{min}^{\mathbf{y}} \subset \mathcal{V}^{\mathbb{R}}(G) \cap \mathcal{W}_+(H)$  and  $\underline{\mathbf{1}}_{\mathbf{y}^*}(f) = f(\mathbf{x}^*) = f^*$ . We define  $\tilde{\Lambda} = \frac{1}{2}(\Lambda^* + \underline{\mathbf{1}}_{\mathbf{y}^*})$ . By construction,  $\tilde{\Lambda} \in \mathcal{L}_{E,G,H}$  and  $\tilde{\Lambda}(f) = \frac{1}{2}(\Lambda^*(f) + \underline{\mathbf{1}}_{\mathbf{y}^*}(f)) = \frac{1}{2}(\Lambda^*(f) + f(\mathbf{x}^*)) = f^*$ . As  $p \in \ker M_{\Lambda^*}^E$ ,

$$\tilde{\Lambda}(p^2) = \frac{1}{2}(\Lambda^*(p^2) + \underline{\mathbf{1}}_{\mathbf{y}^*}(p^2)) = \frac{1}{2}p^2(\mathbf{y}^*) \neq 0$$

thus  $p \in \ker M_{\Lambda^*}^E \setminus \ker M_{\tilde{\Lambda}}^E$  and by the maximality of the rank of  $M_{\Lambda^*}^E$ ,  $\ker M_{\tilde{\Lambda}}^E \not\subset \ker M_{\Lambda^*}^E$ . Hence there exists  $\tilde{p} \in \ker M_{\tilde{\Lambda}}^E \setminus \ker M_{\Lambda^*}^E$ . Then  $0 = M_{\tilde{\Lambda}}^E(\tilde{p}) = \frac{1}{2}(M_{\Lambda^*}^E(\tilde{p}) + M_{\underline{\mathbf{1}}_{\mathbf{y}^*}}^E(\tilde{p})) = \frac{1}{2}(M_{\Lambda^*}^E(\tilde{p}) + \tilde{p}(\mathbf{y}^*) \cdot \underline{\mathbf{1}}_{\mathbf{y}^*})$ . As  $M_{\Lambda^*}^E(\tilde{p}) \neq 0$  implies  $\tilde{p}(\mathbf{y}^*) \neq 0$ . On the other hand,

$$\begin{aligned} 0 &= M_{\tilde{\Lambda}}^E(\tilde{h})(p) = \tilde{\Lambda}(p\tilde{p}) = \frac{1}{2}(\Lambda^*(p\tilde{p}) + p(\mathbf{y}^*)\tilde{p}(\mathbf{y}^*)) = \\ &= \frac{1}{2}(M_{\Lambda^*}^E(p)(\tilde{p}) + p(\mathbf{y}^*)\tilde{p}(\mathbf{y}^*)). \end{aligned}$$

As  $p \in \ker M_{\Lambda^*}^E$ , we have

$$0 = M_{\tilde{\Lambda}}^E(\tilde{p})(p) = \frac{1}{2}(p(\mathbf{y}^*)\tilde{p}(\mathbf{y}^*)).$$

As  $\tilde{p}(\mathbf{y}^*) \neq 0$  and we have supposed that  $p(\mathbf{y}^*) \neq 0$ , it yields a contradiction.  $\square$

**Corollary 5.14.** *Let  $E \subset \mathbb{R}[\mathbf{x}]$  such that  $1 \in E$  and  $f \in \langle E \cdot E \rangle$  and let  $G, H \subset \langle E \cdot E \rangle$  with  $V_{min}^{\mathbf{x}} \subset \mathcal{V}^{\mathbb{R}}(G) \cap \mathcal{W}_+(H)$ . If  $\Lambda^* \in \mathcal{L}_{E,G,H}$  is generic for  $f$  and such that  $\Lambda^*(f) = f^*$ , then  $\ker M_{\Lambda^*}^E \subset I_{min}^{\mathbf{x}}$ .*

*Proof.* It is a consequence of Theorem 5.13, since  $V_{min}^{\mathbf{x}} = \pi^{\mathbf{x}}(V_{min}^{\mathbf{y}})$ .  $\square$

The last result of this section shows that a generic linear form for  $f$  yields the generators of the minimizer ideal  $I_{min}^{\mathbf{y}}$  in high enough degree.

**Theorem 5.15.** *For  $G \subset \mathbb{R}[\mathbf{y}]$  with  $V_{min}^{\mathbf{y}} \subset \mathcal{W}(G, H) \subset V_{KKT}^{\mathbb{R}}$ , there exists  $t_2 \in \mathbb{N}$  (defined in Theorem 5.9) such that  $\forall t \geq t_2$ ,*

- $f_{t,G,H}^{\mu} = f^*$  is reached for some  $\Lambda^* \in \mathcal{L}_{t,G,H}$ ,

- for  $\Lambda^* \in \mathcal{L}_{t,G,H}$  generic for  $f$ , we have  $\Lambda^*(f) = f^*$  and  $(\ker M_{\Lambda^*}^t) = I_{min}^y$ .

*Proof.* We obtain the result as a consequence of Theorem 5.3, Theorem 5.9 and Theorem 5.13.  $\square$

Similarly, the combination of Corollary 5.4, Corollary 5.10 and Corollary 5.14 yields the following result:

**Corollary 5.16.** *For  $G \subset \mathbb{R}[\mathbf{x}]$  with  $V_{min}^{\mathbf{x}} \subset \mathcal{V}^{\mathbb{R}}(G) \cap \mathcal{W}_+(H) \subset \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$ , there exists  $t_2 \in \mathbb{N}$  (defined in Corollary 5.10) such that  $\forall t \geq t_2$ ,*

- $f_{t,G,H}^{\mu} = f^*$  is reached for some  $\Lambda^* \in \mathcal{L}_{t,G,H}$ ,
- $\forall \Lambda^* \in \mathcal{L}_{t,G,H}$  generic for  $f$ , we have  $\Lambda^*(f) = f^*$  and  $(\ker M_{\Lambda^*}^t) = I_{min}^{\mathbf{x}}$ .

Direct consequences of these results are the following:

- If  $G = \{F_1, \dots, F_n, g_1, \dots, g_{n_1}, v_1 h_1, \dots, v_{n_2} h_2\}$ , the hierarchies of relaxation problems associated to  $\mathcal{P}_{t,G,H}$  and  $\mathcal{L}_{t,G,H}$  are exact.
- For  $t \in \mathbb{N}$  high enough and for  $\Lambda^* \in \mathcal{L}_{t,G,H}$  generic for  $f$ , we have  $\Lambda^*(f) = f^*$  and  $\ker M_{\Lambda^*}^t$  generates  $I_{min}^y$ .

The same results hold if we replace  $G$  by any other finite set defining a *real* variety  $\mathcal{V}^{\mathbb{R}}(G)$  such that  $\overline{V_{min}^y} \subset \mathcal{V}^{\mathbb{R}}(G) \cap \mathcal{W}_+(H) \subset \overline{V_{KKT}^{\mathbb{R}}}$  or for  $G \subset \mathbb{R}[\mathbf{x}]$  such that  $V_{min}^{\mathbf{x}} \subset \mathcal{V}^{\mathbb{R}}(G) \cap \mathcal{W}_+(H) \subset \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$ .

## 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 [26]):

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

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

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

A consequence of Corollary 5.16 is the following:

**Theorem 6.2.** *Let  $f \in \mathbb{R}[\mathbf{x}]$  and  $G = \{\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\}$ . Suppose that  $\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$  is reached at some point of  $\mathbb{R}^n$ . Then, there exists  $t_0 \in \mathbb{N}$ , such that  $\forall t \geq t_0$*

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

The first point of this theorem can also be found in [26].

**6.2. Generic case.** For any  $\mathbf{x} \in \mathbb{C}^n$ , we denote  $J(\mathbf{x}) = \{j \in [1, n_2] \mid h_j(\mathbf{x}) = 0\}$ .

Let  $(\mathbf{x}, \mathbf{u}, \mathbf{v})$  be a point of  $V_{KKT}$  and  $J(\mathbf{x}) = J = \{j_1, \dots, j_k\}$ . Then  $v_j = 0$  if  $j \notin J$  and the kernel of the matrix

$$A_J := [\nabla f, \nabla g_1, \dots, \nabla g_{n_1}, \nabla h_{j_1}, \dots, \nabla h_{j_k}].$$

contains the vector  $(1, -u_1, \dots, -u_{n_1}, -v_{j_1}, \dots, -v_{j_k})$ . This implies that the matrix  $A_J$  is of rank  $\leq n_1 + |J|$ . Let

$$\begin{aligned} W_J &= \{\mathbf{x} \in \mathbb{C}^n \mid g_1(\mathbf{x}) = 0, \dots, g_{n_1}(\mathbf{x}) = 0, h_{j_1}(\mathbf{x}) = 0, \dots, h_{j_k}(\mathbf{x}) = 0\} \\ V_J &= W_J \cap \{\mathbf{x} \in \mathbb{C}^n \mid \text{rank}(A_J) \leq n_1 + |J|\}. \end{aligned}$$

We have  $\pi^{\mathbf{x}}(V_{KKT}) \subset \cup_J V_J$ .

Suppose moreover that  $W_J$  is smooth, that is,

$$[\nabla g_1, \dots, \nabla g_{n_1}, \nabla h_{j_1}, \dots, \nabla h_{j_k}]$$

is of maximal rank  $\max(n, n_1 + |J|)$ . We use the convention that if  $W_J$  is empty, then it is smooth. Let  $\mathbf{x} \in W_J$ . The condition  $\text{rank}(A_J) \leq n_1 + |J|$  implies that there exists  $u_1, \dots, u_{n_1}, v_{j_1}, \dots, v_{j_k} \in \mathbb{C}$  such that

$$\nabla f(\mathbf{x}) - u_1 \nabla g_1(\mathbf{x}) - \dots - u_{n_1} \nabla g_{n_1}(\mathbf{x}) - v_{j_1} \nabla h_{j_1}(\mathbf{x}) - \dots - v_{j_k} \nabla h_{j_k}(\mathbf{x}) = 0.$$

and  $\mathbf{x} \in \pi^{\mathbf{x}}(V_{KKT})$ . Thus, if  $\forall J \subset [1, n_2]$ ,  $W_J$  is smooth, then  $\pi^{\mathbf{x}}(V_{KKT}) = \cup_J V_J$ .

**Definition 6.3.** *We say the  $(g_1, \dots, g_{n_1}; h_1, \dots, h_{n_2})$  is  $\mathbb{C}$ -regular if  $\forall J \subset [1, n_2]$   $W_J$  is smooth, or equivalently for all points  $\mathbf{x} \in \mathbb{C}^n$  with  $J(\mathbf{x}) = J = \{j_1, \dots, j_k\}$ , the vectors  $\nabla g_1(\mathbf{x}), \dots, \nabla g_{n_1}(\mathbf{x}), \nabla h_{j_1}(\mathbf{x}), \dots, \nabla h_{j_k}(\mathbf{x})$  are linearly independent.*

Let  $\Delta_1^J, \dots, \Delta_{l_J}^J$  be polynomials defining the variety  $\{\mathbf{x} \in \mathbb{C}^n \mid \text{rank}(A_J) \leq n_1 + |J|\}$ . Notice that if  $n_1 + |J| \geq n$ , we can take  $\Delta^J = 0$ .

We easily check that  $\cup_J V_J$  is defined by the polynomials  $g_1, \dots, g_{n_1}$  and

$$g_{J,i} := \Delta_i^J \prod_{j \notin J} h_j,$$

for  $i = 1, \dots, l_J, J \subset [0, n_2]$ .

Thus when  $(g_1, \dots, g_{n_1}; h_1, \dots, h_{n_2})$  is  $\mathbb{C}$ -regular (ie. all  $W_J$  are smooth),  $\pi^{\mathbf{x}}(V_{KKT}) = \mathcal{V}(G)$ . If moreover  $f$  reaches its minimum at  $\mathbf{x}^* \in \mathcal{S}$ , then  $\text{rank}(A_{J(\mathbf{x}^*)}) \leq n_1 + |J(\mathbf{x}^*)|$  and  $\mathbf{x}^*$  is a KKT-point. We deduce from Corollary 5.16 the following result:

**Theorem 6.4.** *Suppose that  $(g_1, \dots, g_{n_1}; h_1, \dots, h_{n_2})$  is  $\mathbb{C}$ -regular and that  $f \in \mathbb{R}[\mathbf{x}]$  reaches its minimum on  $\mathcal{S}$ . Let  $G$  be the set  $g_i, g_{J,j}$  for  $i = 1, \dots, n_1, j = 1, \dots, l_J, J \subset [0, n_2]$ . Then there exists  $t_0 \in \mathbb{N}$  such that  $\forall t \geq t_0$ ,*

- (1)  $f_{t,G,H}^{\text{sos}} = f_{t,G,H}^{\mu} = f^* = \min_{\mathbf{x} \in \mathcal{S}} f(\mathbf{x})$  is reached for some  $\Lambda^* \in \mathcal{L}_{t,G}$ ,
- (2)  $\forall \Lambda^* \in \mathcal{L}_{t,G}$  with  $\Lambda^*(f) = f_{t,G,H}^{\mu}$  generic for  $f$ ,  $\ker M_{\Lambda^*}^t$  generates  $I_{\min}$ .

The first point can also be found in [24, Theorem 2.3] for special generators  $\Delta_i^J$  under a similar genericity condition ([24, Assumption 2.2]). This condition is generically satisfied for semi-algebraic sets when  $n_1 \leq n$ .

**6.3. Regular semi-algebraic sets.** Another type of regularity condition over  $\mathcal{S}$  is frequently used:

**Definition 6.5.** *The semi algebraic set  $\mathcal{S} = \mathcal{W}(G, H)$  is regular if for all  $\mathbf{x} \in \mathcal{S}$ , the vectors  $\nabla g_1(\mathbf{x}), \dots, \nabla g_{n_1}(\mathbf{x}), \nabla h_{j_1}(\mathbf{x}), \dots, \nabla h_{j_k}(\mathbf{x})$  such that  $h_{j_i}(\mathbf{x}) = 0$  are linearly independent.*

Over the real, the condition that  $\text{rank}(A_J) \leq n_1 + |J|$  can be rewritten as

$$\Delta_J = \det(A_J A_J^T) = 0.$$

Let  $G$  be the set of polynomials  $g_1, \dots, g_{n_1}$  and  $g_J = \Delta_J \prod_{j \notin J} h_j$  for all non-empty subset  $J$  of  $\{1, \dots, n_2\}$ .

As in the previous section, we easily check that if  $\mathcal{S}$  is regular, then  $\mathcal{S} = \mathcal{W}(G, H) \subset \pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})$ . Similarly, if  $f$  reaches its minimum at  $\mathbf{x}^* \in \mathcal{S}$ , then  $\text{rank}(A_{J(\mathbf{x}^*)}) \leq n_1 + |J(\mathbf{x}^*)|$  and  $\mathbf{x}^*$  is a KKT-point. Corollary 5.16 implies the following result:

**Theorem 6.6.** *Suppose that  $\mathcal{S}$  is regular and  $f \in \mathbb{R}[\mathbf{x}]$  reaches its minimum on  $\mathcal{S}$ . Let  $G$  be the set of  $g_i$ ,  $i = 1, \dots, n_1$ ,  $g_J$ ,  $J \subset \{1, \dots, n_2\}$ . Then there exists  $t_0 \in \mathbb{N}$  such that  $\forall t \geq t_0$ ,*

- (1)  $f_{t,G,H}^{\mu} = f^* = \min_{\mathbf{x} \in \mathcal{S}} f(\mathbf{x})$  is reached for some  $\Lambda^* \in \mathcal{L}_{t,G,H}$ ,
- (2)  $\forall \Lambda^* \in \mathcal{L}_{t,G,H}$  with  $\Lambda^*(f) = f_{t,G,H}^{\mu}$  generic for  $f$ ,  $\ker M_{\Lambda^*}^t$  generates  $I_{min}$ ,
- (3) If moreover  $(g_1, \dots, g_{n_1}; h_1, \dots, h_{n_2})$  is  $\mathbb{C}$ -regular, then  $f_{t,G,H}^{sos} = f^*$ .

This theorem provides a generalization of a result in [9, Theorem 3.4], where the condition that  $(G)$  is radical is removed and the  $\mathbb{C}$ -regularity is replaced by the regularity on  $\mathcal{S}$ .

**6.4.  $H$ -radical computation.** In the case where  $f = 0$ , by Remark 2.14 all the points of  $\mathcal{S}$  are KKT points and minimizers of  $f$  so that  $V_{min}^{\mathbf{x}} = V_{KKT}^{\mathbb{R}} = \mathcal{V}(\sqrt[H]{(g_1, \dots, g_{n_1})})$ . Moreover,  $I_{KKT}^{\mathbf{x}} = (g_1, \dots, g_{n_1})$  since  $F_1, \dots, F_n, v_1 h_1, \dots, v_{n_2} h_2$  are homogeneous of degree 1 in the variables  $u_1, \dots, u_{n_1}, v_1, \dots, v_{n_2}$ . We deduce the following result:

**Theorem 6.7.** *Let  $G = \{g_1, \dots, g_{n_2}\}$ ,  $H = \{h_1, \dots, h_{n_2}\} \subset \mathbb{R}[\mathbf{x}]$ . There exists  $t_2 \in \mathbb{N}$  such that  $\forall t \leq t_2$ ,  $\forall \Lambda^* \in \mathcal{L}_{t,G,H}$  generic for 0, we have  $(\ker M_{\Lambda^*}^t) = \sqrt[H]{(G)}$ .*

This gives a way to compute  $\sqrt[H]{(G)}$ , which generalizes the approach of [17] or [14] to compute the real radical of an ideal. In the case where  $\sqrt[H]{(G)}$  is zero-dimensional, a convergence certificate will be described in the next section.

## 7. CONVERGENCE CERTIFICATION

In this section, we describe a criterion to detect when the kernel of a truncated Hankel operator associated to a generic linear form for  $f$  yields the generators of the minimizer ideal. It is based on a flat extension property [19] and applies to global polynomial optimization problems where the minimizer ideal  $I_{min}$  is zero-dimensional. In this section, we assume that  $V_{min} \neq \emptyset$ , that is  $f^*$  is reached for some point  $\mathbf{x}^* \in \mathcal{S}_{KKT}$ .

**7.1. Hankel operators and positive linear forms.** This section is based on [14].

**Definition 7.1.** *For  $\Lambda \in \mathbb{R}[\mathbf{x}]^*$ , the Hankel operator  $M_{\Lambda}$  is the operator from  $\mathbb{R}[\mathbf{x}]$  to  $\mathbb{R}[\mathbf{x}]^*$  defined by*

$$(12) \quad M_{\Lambda} : p \in \mathbb{R}[\mathbf{x}] \mapsto p \cdot \Lambda \in \mathbb{R}[\mathbf{x}]^*$$

To analyse the properties of  $\Lambda$ , we study the quotient algebra  $\mathbb{K}[\mathbf{x}]/\ker M_{\Lambda} = \mathcal{A}_{\Lambda}$ . A first result is the following (see e.g. [14]):

**Lemma 7.2.** *The rank of the operator  $M_{\Lambda}$  is finite if and only if  $\ker M_{\Lambda}$  is a zero-dimensional ideal, in which case  $\dim \mathbb{K}[\mathbf{x}]/\ker M_{\Lambda} = \text{rank } M_{\Lambda}$ . Moreover, if  $B$  is a finite subset of  $\mathbb{K}[\mathbf{x}]$  such that  $|B| = \text{rank } M_{\Lambda}^B = \dim \mathbb{K}[\mathbf{x}]/\ker M_{\Lambda}$ , then  $B$  is a basis of  $\mathbb{K}[\mathbf{x}]/\ker M_{\Lambda}$ .*

For a zero-dimensional ideal  $I \subset \mathbb{K}[\mathbf{x}]$  with simple zeros  $V(I) = \{\zeta_1, \dots, \zeta_r\} \subset \mathbb{K}^n$ , we have  $I^{\perp} = \langle \mathbf{1}_{\zeta_1}, \dots, \mathbf{1}_{\zeta_r} \rangle$  and the ideal  $I$  is radical as a consequence of Hilbert's Nullstellensatz. When  $I = \ker M_{\Lambda}$ , this yields the following property.

**Proposition 7.3.** *Let  $\mathbb{K} = \mathbb{C}$  and assume that  $\text{rank} M_\Lambda = r < \infty$ . Then, the ideal  $\ker M_\Lambda$  is radical if and only if*

$$(13) \quad \Lambda = \sum_{i=1}^r \lambda_i \mathbf{1}_{\zeta_i} \quad \text{with } \lambda_i \in \mathbb{K} - \{0\} \quad \text{and } \zeta_i \in \mathbb{K}^n \text{ pairwise distinct,}$$

in which case  $\ker M_\Lambda = I(\zeta_1, \dots, \zeta_r)$  is the vanishing ideal of the points  $\zeta_1, \dots, \zeta_r$ .

As a corollary, we deduce that when  $\mathbb{K} = \mathbb{R}$  and  $\text{rank} M_\Lambda = r < \infty$ , the ideal  $\ker M_\Lambda$  is real radical if and only if the points  $\zeta_i$  are in  $\mathbb{R}^n$ .

**Definition 7.4.** *We say that  $\Lambda \in \mathbb{R}[\mathbf{x}]^*$  is positive, which we denote  $\Lambda \succcurlyeq 0$ , if  $\Lambda(p^2) \geq 0$  for all  $p \in \mathbb{R}[\mathbf{x}]$ . We have that  $\Lambda \succcurlyeq 0$  if and only if  $M_\Lambda \succcurlyeq 0$ . If moreover  $\Lambda(1) = 1$ , we say that  $\Lambda$  is a probability measure.*

This term is justified by a theorem of Riesz-Haviland [30], which states that a Lebesgue measure on  $\mathbb{R}^n$  is uniquely determined by its value on the polynomials  $\in \mathbb{R}[\mathbf{x}]$ . In particular, if  $\Lambda \succcurlyeq 0$  and  $\Lambda(1) = 1$ , there exists a unique probability measure  $\mu$  such that  $\forall p \in \mathbb{R}[\mathbf{x}], \Lambda(p) = \int p d\mu$ .

An important property of positive forms is the following:

**Proposition 7.5.** *Assume  $\text{rank} M_\Lambda = r < \infty$ . Then  $\Lambda \succcurlyeq 0$  if and only if  $\Lambda$  has a decomposition (13) with  $\lambda_i > 0$  and distinct  $\zeta_i \in \mathbb{R}^n$ , in which case  $V(\ker M_\Lambda) = \{\zeta_1, \dots, \zeta_r\} \subset \mathbb{R}^n$ .*

In particular, it shows that if  $\Lambda \succcurlyeq 0$ , then  $\ker M_\Lambda$  is a real radical ideal.

## 7.2. Flat extensions.

**Definition 7.6.** *Given vector subspaces  $E_0 \subset E \subset \mathbb{K}[\mathbf{y}]$  and  $\Lambda \in \langle E \cdot E \rangle^*$ ,  $M_\Lambda^E$  is said to be a flat extension of its restriction  $M_\Lambda^{E_0}$  if  $\text{rank} M_\Lambda^E = \text{rank} M_\Lambda^{E_0}$ .*

We recall here a result from [19], which gives a rank condition for the existence of a flat extension of a truncated Hankel operator.

**Theorem 7.7.** *Consider a monomial set  $B \subset \mathcal{M}$  connected to 1, a linear space  $E$  with  $B^+ \subset E$  and a linear function  $\Lambda$  defined on  $\langle E \cdot E \rangle$ . Assume that  $\text{rank} M_\Lambda^E = \text{rank} M_\Lambda^B = |B|$ . Then there exists a (unique) linear form  $\tilde{\Lambda} \in \mathbb{K}[\mathbf{y}]^*$  which extends  $\Lambda$ , i.e.,  $\tilde{\Lambda}(p) = \Lambda(p)$  for all  $p \in \langle E \cdot E \rangle$ , satisfying  $\text{rank} M_{\tilde{\Lambda}}^E = \text{rank} M_{\tilde{\Lambda}}^B$ . Moreover, we have  $\ker M_{\tilde{\Lambda}}^E = (\ker M_\Lambda^E)$ .*

In other words, the condition  $\text{rank} M_\Lambda^E = \text{rank} M_\Lambda^B = |B|$  implies that the truncated Hankel operator  $M_\Lambda^E$  has a (unique) flat extension to a (full) Hankel operator  $M_{\tilde{\Lambda}}$ .

**Proposition 7.8.** *If there exists  $\Lambda \in \mathcal{L}_{E,G,H}$  with  $\ker M_\Lambda^E = \{0\}$ , then  $f_{E,G,H}^{\text{sos}} = f_{E,G,H}^\mu$ .*

*Proof.* If  $\ker M_\Lambda^E = \{0\}$  implies that  $\Lambda \succ 0$ , i.e.,  $M_\Lambda^E > 0$ . Hence by Slater's Theorem of [2] we have the strong duality then  $f_{E,G,H}^{\text{sos}} = f_{E,G,H}^\mu$ .  $\square$

**Theorem 7.9.** *Suppose that  $V_{\min} \neq \emptyset$ . Let  $B$  be a monomial set connected to 1,  $E \subset \mathbb{R}[\mathbf{y}]$  with  $B^+ \subset E$ ,  $G \cdot B \subset \langle E \cdot E \rangle$ ,  $H \cdot B \cdot B \subset \langle E \cdot E \rangle$  and  $V_{\min}^{\mathbf{y}} \subset \mathcal{V}^{\mathbb{R}}(G) \cap \mathcal{W}_+(H) \subset V_{KKT}^{\mathbb{R}}$ .*

*Let  $\Lambda^* \in \mathcal{L}_{E,G,H}$  such that  $\Lambda^*$  is generic for  $f$  and satisfies the flat extension property:  $\text{rank} M_{\Lambda^*}^E = \text{rank} M_{\Lambda^*}^B = |B|$ . Then  $f^* = f_{E,G,H}^\mu = f_{E,G,H}^{\text{sos}}$  and  $(\ker M_{\Lambda^*}^E) = I_{\min}^{\mathbf{y}}$ .*

*Proof.* As  $\text{rank} M_{\Lambda^*}^E = \text{rank} M_{\Lambda^*}^B = |B|$ , Theorem 7.7 implies that there exists a (unique) linear function  $\tilde{\Lambda}^* \in \mathbb{K}[\mathbf{y}]^*$  which extends  $\Lambda^*$ . As  $\text{rank} M_{\tilde{\Lambda}^*}^E = \text{rank} M_{\tilde{\Lambda}^*}^B = |B|$  and  $\ker M_{\tilde{\Lambda}^*}^E = (\ker M_{\Lambda^*}^E)$ , any polynomial  $p \in \mathbb{R}[\mathbf{y}]$  can be reduced modulo  $\ker M_{\tilde{\Lambda}^*}^E$  to a

polynomial  $b \in \langle B \rangle$  so that  $p - b \in \ker M_{\tilde{\Lambda}^*}$ . Then  $\tilde{\Lambda}^*(p^2) = \tilde{\Lambda}^*(b^2) = \Lambda^*(b^2) \geq 0$  since  $\Lambda^* \in \mathcal{L}_{E,G,H}$ . This implies that  $\tilde{\Lambda}^* \succneq 0$ . By Proposition 7.5,  $\tilde{\Lambda}^*$  has a decomposition  $\tilde{\Lambda}^* = \sum_{i=1}^r \lambda_i \mathbf{1}_{\zeta_i}$  with  $\lambda_i > 0$  and  $\zeta_i \in \mathbb{R}^{n+n_1+n_2}$ .

By Lemma 7.2,  $B$  is a basis of  $\mathbb{R}[\mathbf{y}]/\mathcal{I}(\zeta_1, \dots, \zeta_r)$  and there exist (interpolation) polynomials  $b_1, \dots, b_r \in \langle B \rangle$  satisfying  $b_i(\zeta_j) = 1$  if  $i = j$  and  $b_i(\zeta_j) = 0$  otherwise. We deduce that for  $i = 1, \dots, r$  and for all elements  $g \in G$ ,

$$\Lambda^*(b_i g) = 0 = \tilde{\Lambda}^*(b_i g) = \lambda_i g(\zeta_i).$$

As  $\lambda_i > 0$  then  $g(\zeta_i) = 0$ . Similarly, for all  $h \in H$ ,

$$\Lambda^*(b_i^2 h) = \tilde{\Lambda}^*(b_i^2 h) = \lambda_i h(\zeta_i) \geq 0$$

and  $h(\zeta_i) \geq 0$ , hence  $\zeta_i \in \mathcal{W}(G, H) \subset V_{KKT}^{\mathbb{R}} \cap \mathcal{W}_+(H) = \mathcal{S}_{KKT}$ . We deduce that  $f(\zeta_i) \geq f^* = \min_{\mathbf{x} \in \mathcal{S}_{KKT}} f(\mathbf{x})$ .

As  $\tilde{\Lambda}^* = \sum_{i=1}^r \lambda_i \mathbf{1}_{\zeta_i}$  with  $\lambda_i > 0$  and  $\zeta_i \in \mathbb{R}^{n+n_1+n_2}$ , and as  $\tilde{\Lambda}^*(1) = \Lambda^*(1) = \sum_{i=1}^r \lambda_i = 1$ , we have  $\tilde{\Lambda}^*(f) = \Lambda^*(f) = f_{E,G,H}^{\mu} = \sum_{i=1}^r \lambda_i f(\zeta_i) \geq f^*$ . By Proposition 3.9, we deduce that  $\Lambda^*(f) = f^*$  and  $f(\zeta_i) = f^*$  for  $i = 1, \dots, r$  so that  $\{\zeta_1, \dots, \zeta_r\} \subset V_{min}^{\mathbf{y}}$ .

By Proposition 7.5, Theorem 7.7 and Theorem 5.13, we have

$$\ker M_{\tilde{\Lambda}^*} = \mathcal{I}(\zeta_1, \dots, \zeta_r) = (\ker M_{\tilde{\Lambda}^*}^E) \subset I_{min}^{\mathbf{y}}.$$

We deduce that  $\{\zeta_1, \dots, \zeta_r\} = \mathcal{V}(I_{min}^{\mathbf{y}})$  and that  $(\ker M_{\tilde{\Lambda}^*}^E) = I_{min}^{\mathbf{y}}$ .

As we have the flat extension condition,  $\ker M_{\tilde{\Lambda}^*}^B = \{0\}$  and by Proposition 7.8, there is no duality gap:  $f_{B,G,H}^{\mu} = f_{B,G,H}^{sos}$ . As  $\Lambda^* \in \mathcal{L}_{E,G,H} \subset \mathcal{L}_{B,G,H}$  then  $\Lambda^*(f) = f_{E,G,H}^{\mu} = f_{B,G,H}^{\mu} = f^*$ . We know that  $f_{E,G,H}^{sos} \leq f_{E,G,H}^{\mu}$  and also  $f_{B,G,H}^{sos} \leq f_{E,G,H}^{sos}$ . Hence  $f_{E,G,H}^{\mu} = f_{E,G,H}^{sos} = f^*$ .  $\square$

**Corollary 7.10.** *Suppose that  $V_{min} \neq \emptyset$ . Let  $B$  be a monomial set connected to 1,  $E \subset \mathbb{R}[\mathbf{x}]$  with  $B^+ \subset E$ ,  $G \cdot B \subset \langle E \cdot E \rangle$ ,  $H \cdot B \cdot B \subset \langle E \cdot E \rangle$  and  $V_{min}^{\mathbf{x}} \subset \mathcal{W}(G, H) \subset \pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})$  and  $B \cdot G \subset \langle E \cdot E \rangle$ . Let  $\Lambda^* \in \mathcal{L}_{E,G,H}$  such that  $\Lambda^*$  is generic for  $f$  and satisfies the flat extension property:  $\text{rank } M_{\tilde{\Lambda}^*}^E = \text{rank } M_{\tilde{\Lambda}^*}^B = |B|$ . Then there is no duality gap,  $f^* = f_{E,G,H}^{\mu} = f_{E,G,H}^{sos}$  and  $(\ker M_{\tilde{\Lambda}^*}^E) = I_{min}^{\mathbf{x}}$ .*

Notice that if the hypotheses of this theorem are satisfied, then necessarily  $I_{min}$  is zero-dimensional.

This theorem gives a convergence certificate to check when the minimum  $f^*$  is reached and when a generating family of the minimizer ideal is obtained. It generalizes the flat truncation certificate given in [25].

**7.3. Certificate verification.** Finally it remains to prove that when the order of the relaxation is big enough, the convergence certificate is satisfied. Therefore, this approach yields a complete algorithm that we describe hereafter.

**Theorem 7.11.** *Let  $G \subset \mathbb{R}[\mathbf{y}]$  with  $V_{min}^{\mathbf{y}} \subset \mathcal{W}(G, H) \subset V_{KKT}^{\mathbb{R}}$ . Suppose that  $V_{min}^{\mathbf{y}}$  is finite and not empty. Then, there exists  $t_3 \in \mathbb{N}$  s.t.  $\forall t \geq t_3$ , and  $\forall \Lambda^* \in \mathcal{L}_{t,G,H}$  generic for  $f$ , there exists a vector space  $E \subset \mathbb{R}[\mathbf{y}]_t$  and  $B$  connected to 1 such that*

- $f \in \langle E \cdot E \rangle$ ,
- $B^+ \subset E$ ,  $G \cdot B \subset \langle E \cdot E \rangle$ ,  $H \cdot B \cdot B \subset \langle E \cdot E \rangle$ ,
- $\text{rank } M_{\tilde{\Lambda}^*}^E = \text{rank } M_{\tilde{\Lambda}^*}^B = |B|$ .

*Proof.* Let  $G' = \{g_1, \dots, g_s\} \in \mathbb{R}[\mathbf{y}]$  be a generating family of  $I_{min}^{\mathbf{y}}$ . Since  $(G') = I_{min}^{\mathbf{y}}$  is zero dimensional, there exists a finite set  $B$  of monomials connected to 1, which is a basis of

$\dim R/I_{min}^{\mathbf{y}}$ . Let  $t' \in \mathbb{N}$  be chosen so that  $B^+ \subset \mathbb{R}[\mathbf{y}]_{t'}$ ,  $f \in \mathbb{R}[\mathbf{y}]_{2t'}$ ,  $G \cdot B \subset \mathbb{R}[\mathbf{y}]_{t'}$  and  $H \cdot B \cdot B \subset \mathbb{R}[\mathbf{y}]_{t'}$ . We have  $\mathbb{R}[\mathbf{y}]_{t'} = \langle B \rangle \oplus (I_{min}^{\mathbf{y}})_{t'}$  and  $B \cdot G \subset \mathbb{R}[\mathbf{y}]_{2t'}$ .

Let  $t'' \in \mathbb{N}$  be chosen so that  $(I_{min}^{\mathbf{y}})_{t''} = G'_{\langle t'' \rangle} \cap \mathbb{R}[\mathbf{y}]_{t''}$  and we have

$$\mathbb{R}[\mathbf{y}]_{t''} = \langle B \rangle \oplus (G'_{\langle t'' \rangle} \cap \mathbb{R}[\mathbf{y}]_{t''}).$$

By Theorem 5.15 and Remark 5.2, there exists  $t_3 \in \mathbb{N}$  such that  $\forall t \geq t_3$ ,  $\forall \Lambda^* \in \mathcal{L}_{t,G,H}$  generic for  $f$ , we have  $\Lambda^*(f) = f^*$  and  $G'_{\langle t \rangle} \subset \ker M_{\Lambda^*}^t$ .

Let us take  $E = \mathbb{R}[\mathbf{y}]_{t'}$ . By construction,  $f \in \langle E \cdot E \rangle = \mathbb{R}[\mathbf{y}]_{2t'}$ ,  $B^+ \subset E$ ,  $G \cdot B \subset \langle E \cdot E \rangle$  and  $H \cdot B \cdot B \subset \langle E \cdot E \rangle$ . As

$$(G'_{\langle t'' \rangle} \cap \mathbb{R}[\mathbf{y}]_{t''}) \subset (\ker M_{\Lambda^*}^t \cap \mathbb{R}[\mathbf{y}]_{t''}) = \ker M_{\Lambda^*}^E,$$

we have  $\text{rank } M_{\Lambda^*}^E \leq \text{rank } M_{\Lambda^*}^B \leq |B|$ .

As  $(G') = (\ker M_{\Lambda^*}^E) = I_{min}^{\mathbf{y}}$ , we have  $\text{rank } M_{\Lambda^*}^E \geq \mathbb{R}[\mathbf{y}]_{t'}/(I_{min}^{\mathbf{y}})_{t'} = |B|$ . Hence we conclude that  $\text{rank } M_{\Lambda^*}^E = \text{rank } M_{\Lambda^*}^B = |B|$  and that the convergence certificate is satisfied for all  $t \geq t_3$  for some vector space  $E = \mathbb{R}[\mathbf{y}]_{t'}$ .  $\square$

**Corollary 7.12.** *Let  $G \subset \mathbb{R}[\mathbf{x}]$  with  $V_{min}^{\mathbf{x}} \subset \mathcal{W}(G, H) \subset \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$ . Suppose that  $V_{min}^{\mathbf{x}}$  is finite. Then, there exists  $t_3 \in \mathbb{N}$  s.t.  $\forall t \geq t_3$ , and  $\forall \Lambda^* \in \mathcal{L}_{t,G,H}$  generic for  $f$ , there exists a vector space  $E \subset \mathbb{R}[\mathbf{x}]_t$  and  $B$  connected to 1 such that*

- $f \in \langle E \cdot E \rangle$ ,
- $B^+ \subset E$ ,  $B \cdot G \subset \langle E \cdot E \rangle$ ,  $H \cdot B \cdot B \subset \langle E \cdot E \rangle$
- $\text{rank } M_{\Lambda^*}^E = \text{rank } M_{\Lambda^*}^B = |B|$ .

*Proof.* Direct consequence of the fact that  $I_{min}^{\mathbf{x}} = I_{min}^{\mathbf{y}} \cap \mathbb{R}[\mathbf{x}]$ .  $\square$

This leads to the following algorithm to compute  $f^*$  and the minimizer ideal when minimizer variety is finite, which increments the order of relaxation and check the convergence certificate until it is satisfied:

**Algorithm.**

**Input:**  $f \in \mathbb{R}[\mathbf{x}]$ ,  $G \subset \mathbb{R}[\mathbf{x}]$  such that  $V_{min}^{\mathbf{y}} \subset \mathcal{W}(G, H) \subset V_{KKT}^{\mathbb{R}}$ ,  $H = \{h_1, \dots, h_{n_2}\}$ .

- (1)  $t := \lceil \frac{1}{2} \max\{\deg(f), \deg(g_i), \deg(h_j)\} \rceil$ ;
- (2) Compute  $\Lambda^* \in \mathcal{L}_{t,G,H}$  generic for  $f$  (solving a finite dimensional SDP problem by an interior point method);
- (3) Check the convergence certificate for  $M_{\Lambda^*}^t$ ;
- (4) If it is not satisfied,  $t := t + 1$  and repeat from step (2);  
Otherwise compute  $K := \ker M_{\Lambda^*}^t$ .

**Output:**  $f^* = \Lambda^*(f)$ ,  $K$  generators of  $I_{min}^{\mathbf{x}}$ .

REFERENCES

- [1] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry*. Springer, 1998.
- [2] S. Boyd and L. Vandenberg. *Convex Optimization*. Cambridge University Press., New York, 2004.
- [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] J. Demmel, J. Nie, and V. Powers. Representations of positive polynomials on noncompact semialgebraic sets via {KKT} ideals. *Journal of Pure and Applied Algebra*, 209(1):189 – 200, 2007.

- [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) (2008), 941-951.
- [9] H. V. Ha and T. S. Pham Representation of positive polynomials and optimization on noncompact semialgebraic sets. *SIAM Journal on Optimization*, 20 (6) (2010), 3082-3103.
- [10] B. Hanzon and D. Jibeteau. Global minimization of a multivariate polynomial using matrix methods. *J. Global Optim.*, 27:1–23, 2001.
- [11] D. T. Hiep Representations of non-negatives polynomials via KKT ideals. *Ann. Polon. Math* 102 (2011), 101-109.
- [12] D. Jibeteau and M. Laurent. Semidefinite approximations for global unconstrained polynomial optimization. *SIAM J. on Optimization*, 16(2):490–514, June 2005.
- [13] 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.
- [14] 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.
- [15] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11:796–817, 2001.
- [16] J.B. Lasserre Moments, positive polynomials and their applications. *Imperial College Press* 2010.
- [17] 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.
- [18] M. Laurent. Semidefinite representations for finite varieties. *Math. Progr*, 109:1–26, 2007.
- [19] M. Laurent and B. Mourrain. A generalized flat extension theorem for moment matrices. *Arch. Math. (Basel)*, 93(1):87–98, July 2009.
- [20] M. Marshall. Optimization of polynomial functions. *Canad. Math. Bull.*, 46:575–587, 2003.
- [21] M. Marshall Representations of non-negatives polynomials via critical, degree bounds and applications to optimization. *Can. J. Math.* 61 (1)(2009), 205-221.
- [22] 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.
- [23] 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.
- [24] J. Nie. An exact jacobian sdp relaxation for polynomial optimization. *Mathematical Programming*, pages 1–31, 2011.
- [25] J. Nie. Certifying convergence of Lasserre’s hierarchy via flat truncation. *Math. Progr* 10.1007/s10107-012-0589-9, 2012
- [26] J. Nie, J. Demmel, and B. Sturmfels. Minimizing polynomials via sum of squares over gradient ideal. *Math. Program.*, 106(3):587–606, 2006.
- [27] P.A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming Ser. B*, 96(2):293–320, 2003.
- [28] 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.
- [29] M. Putinar Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math J.* 42 (1993), 969-984.
- [30] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.
- [31] 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.
- [32] K. Schmüdgen The K-moment problem for compact semi-algebraic sets. *Math Ann.* 289 (1991), 529-543.



- [33] M. Schweighofer. Global optimization of polynomials using gradient tentacles and sums of squares. *SIAM Journal on Optimization*, 17(3):920–942, 2006.
- [34] Shafarevich. Basic algebraic geometry. *Die Grundlehren der mathematischen Wissenschaften*. Band 213. Springer-Verlag, 1974.
- [35] 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