View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by INRIA a CCSD electronic archive server

Unconstraint global polynomial optimization via Gradient Ideal

Marta Abril Bucero, Bernard Mourrain, Philippe Trébuchet

To cite this version:

Marta Abril Bucero, Bernard Mourrain, Philippe Trébuchet. Unconstraint global polynomial optimization via Gradient Ideal. 2013. hal-00779666v3

HAL Id: hal-00779666

https://hal.inria.fr/hal-00779666v3

Preprint submitted on 21 Mar 2013

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.

UNCONSTRAINT GLOBAL POLYNOMIAL OPTIMIZATION VIA GRADIENT IDEAL

MARTA ABRIL BUCERO, BERNARD MOURRAIN, PHILIPPE TREBUCHET

Abstract. In this paper, we describe a new method to compute the minimum of a real polynomial function and the ideal defining the points which minimize this polynomial function, assuming that the minimizer ideal is zero-dimensional. Our method is a generalization of Lasserre relaxation method and stops in a finite number of steps. The proposed algorithm combines Border Basis, Moment Matrices and Semidefinite Programming. In the case where the minimum is reached at a finite number of points, it provides a border basis of the minimizer ideal.

1. Introduction

Optimization appears in many areas of Scientific Computing, since the solution of a problem can often be described as the minimum of an optimization problem. Local methods such as gradient descent are often employed to handle global minimization problems. They 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.

In the case where the function f to minimize is a polynomial, it is possible to develop methods which ensure the computation of a global solution. Reformulating the problem as the computation of a (minimal) critical value of the polynomial f, different polynomial system solvers can be used to tackle it (see e.g. [27], [8]). 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 [29] 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 device 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 [14] (see also [26], [31]) 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^* [14]. 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, this approach suffers from two drawbacks:

(1) the hierarchy of optimization problems may not be *exact*, ie. it may not always convergence in a finite number of steps;
(2) the size of the SDP problems to be solved *grows exponentially* with t.

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 [14], [19]. 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 [17] shows that if the gradient ideal (generated by the first differentials of f) is zero-dimensional, then the hierarchy extended with constraints from the gradient ideal is exact. It is also proved that there is no duality gap if "good" generators of the gradient ideal are used. In [25], it is proved that the extended hierarchy is exact when the gradient ideal is radical and in [24] the extended hierarchy is proved to be exact for any polynomial f when the minimum is reached in $\mathbb{R}^n$ (see also [9]). In [30], the relaxation techniques are analyzed for functions for which the minimum is not reached and which satisfies some special properties "at infinity".

From an algorithmic point of view, this is not ending the investigations since a criteria is needed to determine at which step the minimum is reached. It is also important to known if all the minimizers can be recovered at that step.

Methods based on moment matrices have been proposed to compute generators of the ideal characterizing the real solutions of a polynomial system. In [16], the computation of generators of the (real) radical of an ideal is based on moment matrices which involve all monomials of a given degree and a stopping criterion related to Curto-Fialkow flat extension condition [5] is used. This method is improved in [13]. Combining the border basis algorithm of [21] and a weaker flat extension condition [18], a new algorithm which involves SDP problems of significantly smaller size is proposed to compute the (real) radical of an ideal, when this (real) radical ideal is zero-dimensional. In [25], an algorithm is also proposed to compute the global minimum of the polynomial function f based on techniques from [17]. It terminates when the gradient ideal is radical zero-dimensional.

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. [12] and reference therein). In [10, 11] 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 show that if the minimum f^* is reached in $\mathbb{R}^n$, a generalized hierarchy of relaxation problems which involve the gradient ideal $I_{grad}(f)$ is exact and yields the generators of the minimizer ideal $I_{min}(f)$ from a sufficiently high degree.

In the case where the minimizer ideal is zero dimensional, we give a criterion for deciding when the minimum is reached, based on the flat extension condition in [18].

This criterion is used in a new algorithm which computes a border basis of the minimizer ideal of a polynomial function, when this minimizer ideal is zero-dimensional.

The algorithm is an extension of the real radical algorithm described in [13]. The rows and columns of the matrices involved in the semi-definite programming problem are associated with the family of monomials candidates for being a basis of the quotient space, i.e, a subset of monomials of size much smaller than the number of monomials of the same degree. Thus, the size of the SDP problems involved in this computation is significantly smaller than the one in [14], [17] or [25].

We show that by solving this sequence of SDP problems, we obtain in a finite number of steps the minimum f^* of f, with no duality gap. When this minimum is reached, the kernel of the Hankel matrix associated to the solution of the SDP problem yields generators of the minimizer ideal which are not in the gradient ideal. Assuming that the minimizer ideal $I_{min}(f)$ is zero dimensional, computing the border basis of this kernel yields a representation

of the quotient algebra by $I_{min}(f)$ and thus a way to compute effectively the minimizer points, using eigenvector solvers.

An implementation of this method has been developed, which integrates a border basis implementation and a numerical SDP solver. It is demonstrated on typical examples.

Content. The paper is organized as follows. Section 2 recalls the concepts of algebraic tools as ideals, varieties, dual space, quotient algebra, the definitions and theorems about border basis, and the Hankel Operators involved in the computation of (real) radical ideals and in the computation of our minimizer ideal. In Section 3, we describe the main results on the exactness of the hierarchy of SDP problems on truncated Hankel operators and show that the minimizer ideal can be computed from the kernel of these Hankel. In Section 4, we analyze more precisely the case where the minimizer ideal is zero-dimensional. In section 5, we describe our algorithm and we prove its correctness. Finally in Section 6, we illustrate the algorithm on some classical examples.

2. Ideals, dual space, quotient algebra and border basis

In this section, we set our notation and recall the eigenvalue techniques for solving polynomial equations and the border basis method.

2.1. **Ideals and varieties.** Let $\mathbb{K}[\mathbf{x}]$ be the set of the polynomials in the variables $\mathbf{x} = (x_1, \ldots, x_n)$, with coefficients in the field $\mathbb{K}$. Hereafter, we will choose[1] $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let $\overline{\mathbb{K}}$ denotes the algebraic closure of $\mathbb{K}$. For $\alpha \in \mathbb{N}^n$, $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ is the monomial with exponent α and degree $|\alpha| = \sum_i \alpha_i$. The set of all monomials in $\mathbf{x}$ is denoted $\mathcal{M} = \mathcal{M}(\mathbf{x})$. We say that $\mathbf{x}^\alpha \leq \mathbf{x}^\beta$ if $\mathbf{x}^\alpha$ divides $\mathbf{x}^\beta$, i.e., if $\alpha \leq \beta$ coordinate-wise. For a polynomial $f = \sum_\alpha f_\alpha \mathbf{x}^\alpha$, its support is $supp(f) := \{\mathbf{x}^\alpha \mid f_\alpha \neq 0\}$, 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 $\leq t$,
- $\langle S \rangle = \left\{ \sum_{f \in S} \lambda_f f \mid f \in S, \lambda_f \in \mathbb{K} \right\}$ is the linear span of S,
- $(S) = \left\{ \sum_{f \in S} p_f f \mid p_f \in \mathbb{K}[\mathbf{x}], f \in S \right\}$ is the ideal in $\mathbb{K}[\mathbf{x}]$ generated by S,
- $\langle S \,|\, t \rangle = \left\{ \sum_{f \in S_t} p_f f \mid p_f \in \mathbb{K}[\mathbf{x}]_{t-\deg(f)} \right\}$ is the vector space spanned by $\{\mathbf{x}^\alpha f \mid f \in S_t, |\alpha| \leq t - \deg(f)\}$,
- $S^+ := S \cup x_1 S \cup \ldots \cup x_n S$ is the prolongation of S by one degree,
- $\partial S := S^+ \setminus S$ is the border of S,
- $S^{[t]} := S^{\overbrace{+\cdots+}^{t \text{ times}}}$ is the result of applying t times the prolongation operator '$+$' on S, with $S^{[1]} = S^+$ and, by convention, $S^{[0]} = S$.

Therefore, $S_t = S \cap \mathbb{K}[\mathbf{x}]_t$, $S^{[t]} = \{x^\alpha f \mid f \in S, |\alpha| \leq t\}$, $\langle S \,|\, t \rangle \subseteq (S) \cap \mathbb{K}[\mathbf{x}]_t = (S)_t$, but the inclusion may be strict.

If $\mathcal{B} \subseteq \mathcal{M}$ contains 1 then, for any monomial $m \in \mathcal{M}$, there exists an integer k for which $m \in \mathcal{B}^{[k]}$. The $\mathcal{B}$-*index* of m, denoted by $\delta_{\mathcal{B}}(m)$, is defined as the smallest integer k for which $m \in \mathcal{B}^{[k]}$.

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

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

The 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 $i \in [1, n]$.

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

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

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

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

its associated variety in $\mathbb{L}^n$. By convention $V(I) = V_{\overline{\mathbb{K}}}(I)$. For a set $V \subseteq \mathbb{K}^n$, we define its vanishing ideal

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

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

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

Obviously,

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

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

The following two famous theorems relate vanishing and radical ideals:

Theorem 2.1.

(i) **Hilbert's Nullstellensatz** *(see, e.g.,* [4, §4.1]*)* $\sqrt{I} = I(V_{\mathbb{C}}(I))$ *for an ideal* $I \subseteq \mathbb{C}[\mathbf{x}]$.

(ii) **Real Nullstellensatz** *(see, e.g.,* [1, §4.1]*)* $\sqrt[\mathbb{R}]{I} = I(V_{\mathbb{R}}(I))$ *for an ideal* $I \subseteq \mathbb{R}[\mathbf{x}]$.

2.2. **The roots from the quotient algebra structure.** Given an ideal $I \subseteq \mathbb{K}[\mathbf{x}]$, the quotient set $\mathbb{K}[\mathbf{x}]/I$ consists of all cosets $[f] := f + I = \{f + q \mid q \in I\}$ for $f \in \mathbb{K}[\mathbf{x}]$, i.e., all equivalent classes of polynomials of $\mathbb{K}[\mathbf{x}]$ modulo the ideal I. The quotient set $\mathbb{K}[\mathbf{x}]/I$ is an algebra with addition $[f] + [g] := [f + g]$, scalar multiplication $\lambda[f] := [\lambda f]$ and with multiplication $[f][g] := [fg]$, for $\lambda \in \mathbb{R}$, $f, g \in \mathbb{K}[\mathbf{x}]$.

We will say that an ideal I is zero-dimensional if $0 < |V_{\overline{\mathbb{K}}}(I)| < \infty$. Then $\mathbb{K}[\mathbf{x}]/I$ is a finite-dimensional vector space and its dimension is the number of roots counted with multiciplicity (see e.g. [4], [6]). Thus $|V_{\overline{\mathbb{K}}}(I)| \leq \dim_{\mathbb{K}} \mathbb{K}[\mathbf{x}]/I$, with equality if and only if I is radical.

Assume that $0 < |V_{\overline{\mathbb{K}}}(I)| < \infty$ and set $N := \dim_{\mathbb{K}} \mathbb{K}[\mathbf{x}]/I$ so that $|V_{\overline{\mathbb{K}}}(I)| \leq N < \infty$. Consider a set $\mathcal{B} := \{b_1, \ldots, b_N\} \subseteq \mathbb{K}[\mathbf{x}]$ for which $\{[b_1], \ldots, [b_N]\}$ is a basis of $\mathbb{K}[\mathbf{x}]/I$; by abuse of language we also say that $\mathcal{B}$ itself is a basis of $\mathbb{K}[\mathbf{x}]/I$. Then every $f \in \mathbb{K}[\mathbf{x}]$ can be written in a unique way as $f = \sum_{i=1}^N c_i b_i + p$, where $c_i \in \mathbb{K}$, $p \in I$; the polynomial $\pi_{I,\mathcal{B}}(f) := \sum_{i=1}^N c_i b_i$ is called the remainder of f modulo I, or its *normal form*, with respect to the basis $\mathcal{B}$. In other words, $\langle \mathcal{B} \rangle$ and $\mathbb{K}[\mathbf{x}]/I$ are isomorphic vector spaces.

Given a polynomial $h \in \mathbb{K}[\mathbf{x}]$, we can define the *multiplication (by h) operator* as

$$\begin{array}{rccl} \mathcal{M}_h : & \mathbb{K}[\mathbf{x}]/I & \longrightarrow & \mathbb{K}[\mathbf{x}]/I \\ & [f] & \longmapsto & \mathcal{M}_h([f]) := [hf], \end{array} \tag{1}$$

Assume that $N := \dim_{\mathbb{K}} \mathbb{K}[\mathbf{x}]/I < \infty$. Then the multiplication operator $\mathcal{M}_h$ can be represented by its matrix, again denoted $\mathcal{M}_h$ for simplicity, with respect to a given basis $\mathcal{B} = \{b_1, \ldots, b_N\}$ of $\mathbb{K}[\mathbf{x}]/I$.

Namely, setting $\pi_{I,\mathcal{B}}(hb_j) := \sum_{i=1}^N a_{ij} b_i$ for some scalars $a_{ij} \in \mathbb{K}$, the jth column of $\mathcal{M}_h$ is the vector $(a_{ij})_{i=1}^N$.

Define the vector $\zeta_{\mathcal{B},v} := (b_j(v))_{j=1}^N \in \overline{\mathbb{K}}^N$, whose coordinates are the evaluations of the polynomials $b_j \in \mathcal{B}$ at the point $v \in \overline{\mathbb{K}}^n$. The following famous result (see e.g., [3, Chapter 2§4], [6]) relates the eigenvalues of the multiplication operators in $\mathbb{K}[\mathbf{x}]/I$ to the algebraic variety $V(I)$. This result underlies the so-called eigenvalue method for solving polynomial equations [7] and plays a central role in many algorithms, also in the present paper.

Theorem 2.2. *Let I be a zero-dimensional ideal in $\mathbb{K}[\mathbf{x}]$, $\mathcal{B}$ a basis of $\mathbb{K}[\mathbf{x}]/I$, and $h \in \mathbb{K}[\mathbf{x}]$. The eigenvalues of the multiplication operator $\mathcal{M}_h$ are the evaluations $h(v)$ of the polynomial h at the points $v \in V(I)$. Moreover, $(\mathcal{M}_h)^T \zeta_{\mathcal{B},v} = h(v)\zeta_{\mathcal{B},v}$ and the set of common eigenvectors of $(\mathcal{M}_h)_{h\in\mathbb{K}[\mathbf{x}]}$ are up to a non-zero scalar multiple the vectors $\zeta_{\mathcal{B},v}$ for $v \in V(I)$.*

Throughout the paper we also denote by $\mathcal{M}_i := \mathcal{M}_{x_i}$ the matrix of the multiplication operator by the variable x_i. By the above theorem, the eigenvalues of the matrices $\mathcal{M}_i$ are the ith coordinates of the points $v \in V(I)$. Thus the task of solving a system of polynomial equations is reduced to a task of numerical linear algebra once a basis of $\mathbb{K}[\mathbf{x}]/I$ and a normal form algorithm are available, permitting the construction of the multiplication matrices $\mathcal{M}_i$.

2.3. **Border bases.** The eigenvalue method for solving polynomial equations from the above section requires the knowledge of a basis of $\mathbb{K}[\mathbf{x}]/I$ and an algorithm to compute the normal form of a polynomial with respect to this basis. In this section we will recall a general method for computing such a basis and a method to reduce polynomials to their normal form.

Throughout $\mathcal{B} \subseteq \mathcal{M}$ is a finite set of monomials.

Definition 2.3. *A rewriting family F for a (monomial) set $\mathcal{B}$ is a set of polynomials $F = \{f_i\}_{i\in\mathcal{I}}$ such that*

- *$supp(f_i) \subseteq \mathcal{B}^+$,*
- *f_i has exactly* **one** *monomial in $\partial\mathcal{B}$, denoted as $\gamma(f_i)$ and called the leading monomial of f_i. (The polynomial f_i is normalized so that the coefficient of $\gamma(f_i)$ is 1.)*
- *if $\gamma(f_i) = \gamma(f_j)$ then $i = j$.*

Definition 2.4. *We say that the rewriting family F is* graded *if $\deg(\gamma(f)) = \deg(f)$ for all $f \in F$.*

Definition 2.5. *A rewriting family F for $\mathcal{B}$ is said to be* complete *in degree t if it is graded and satisfies $(\partial\mathcal{B})_t \subseteq \gamma(F)$; that is, each monomial $m \in \partial\mathcal{B}$ of degree at most t is the leading monomial of some (necessarily unique) $f \in F$.*

Definition 2.6. *Let F be a rewriting family for $\mathcal{B}$, complete in degree t. Let $\pi_{F,\mathcal{B}}$ be the projection on $\langle\mathcal{B}\rangle$ along F defined recursively on the monomials $m \in \mathcal{M}_t$ in the following way:*

- *if $m \in \mathcal{B}_t$, then $\pi_{F,\mathcal{B}}(m) = m$,*
- *if $m \in (\partial\mathcal{B})_t$ $(= (\mathcal{B}^{[1]} \setminus \mathcal{B}^{[0]})_t)$, then $\pi_{F,\mathcal{B}}(m) = m - f$, where f is the (unique) polynomial in F for which $\gamma(f) = m$,*
- *if $m \in (\mathcal{B}^{[k]} \setminus \mathcal{B}^{[k-1]})_t$ for some integer $k \geq 2$, write $m = x_{i_0} m'$, where $m' \in \mathcal{B}^{[k-1]}$ and $i_0 \in [1, n]$ is the smallest possible variable index for which such a decomposition exists, then $\pi_{F,\mathcal{B}}(m) = \pi_{F,\mathcal{B}}(x_{i_0}\, \pi_{F,\mathcal{B}}(m'))$.*

If F is a graded rewriting family, one can easily verify that $\deg(\pi_{F,\mathcal{B}}(m)) \leq \deg(m)$ for $m \in \mathcal{M}_t$. The map $\pi_{F,\mathcal{B}}$ extends by linearity to a linear map from $\mathbb{K}[\mathbf{x}]_t$ onto $\langle\mathcal{B}\rangle_t$. By construction, $f = \gamma(f) - \pi_{F,\mathcal{B}}(\gamma(f))$ and $\pi_{F,\mathcal{B}}(f) = 0$ for all $f \in F_t$. The next theorems show that, under some natural commutativity condition, the map $\pi_{F,\mathcal{B}}$ coincides with the linear projection from $\mathbb{K}[\mathbf{x}]_t$ onto $\langle\mathcal{B}\rangle_t$ along the vector space $\langle F \mid t\rangle$. It leads to the notion of border basis.

Definition 2.7. *Let $\mathcal{B} \subset \mathcal{M}$ be connected to 1. A family $F \subset \mathbb{K}[\mathbf{x}]$ is a border basis for $\mathcal{B}$ if it is a rewriting family for $\mathcal{B}$, complete in all degrees, and such that $\mathbb{K}[\mathbf{x}] = \langle \mathcal{B} \rangle \oplus (F)$.*

An algorithmic way to check that we have a border basis is based on the following result, that we recall from [21]:

Theorem 2.8. *Assume that $\mathcal{B}$ is connected to 1 and let F be a rewriting family for $\mathcal{B}$, complete in degree $t \in \mathbb{N}$. Suppose that, for all $m \in \mathcal{M}_{t-2}$,*

$$\pi_{F,\mathcal{B}}(x_i\, \pi_{F,\mathcal{B}}(x_j\, m)) = \pi_{F,\mathcal{B}}(x_j\, \pi_{F,\mathcal{B}}(x_i\, m)) \quad \text{for all } i,j \in [1,n]. \tag{2}$$

Then $\pi_{F,\mathcal{B}}$ coincides with the linear projection of $\mathbb{K}[\mathbf{x}]_t$ on $\langle \mathcal{B} \rangle_t$ along the vector space $\langle F \,|\, t \rangle$; that is, $\mathbb{K}[\mathbf{x}]_t = \langle \mathcal{B} \rangle_t \oplus \langle F \,|\, t \rangle$.

In order to have a simple test and effective way to test the commutation relations (2), we introduce now the commutation polynomials.

Definition 2.9. *Let F be a rewriting family and $f, f' \in F$. Let m, m' be the smallest degree monomials for which $m\,\gamma(f) = m'\,\gamma(f')$. Then the polynomial $C(f,f') := mf - m'f' = m'\pi_{F,\mathcal{B}}(f') - m\pi_{F,\mathcal{B}}(f)$ is called the* commutation polynomial *of f, f'.*

Definition 2.10. *For a rewriting family F with respecet to $\mathcal{B}$, we denote by $C^+(F)$ the set of polynomials of the form $m\,f - m'\,f'$, where $f, f' \in F$ and $m, m' \in \{0, 1, x_1, \ldots, x_n\}$ satisfy*

- *either $m\,\gamma(f) = m'\,\gamma(f')$,*
- *or $m\,\gamma(f) \in \mathcal{B}$ and $m' = 0$.*

Therefore, $C^+(F) \subset \langle \mathcal{B}^+ \rangle$ and $C^+(F)$ contains all commutation polynomials $C(f,f')$ for $f, f' \in F$ whose monomial multipliers m, m' are of degree ≤ 1. The next result can be deduced using Theorem 2.8.

Theorem 2.11. *Let $\mathcal{B} \subset \mathcal{M}$ be connected to 1 and let F be a rewriting family for $\mathcal{B}$, complete in degree t. If for all $c \in C^+(F)$ of degree $\leq t$, $\pi_{F,\mathcal{B}}(c) = 0$, then $\pi_{F,\mathcal{B}}$ is the projection of $\mathbb{K}_t$ on $\langle \mathcal{B} \rangle_t$ along $\langle F \,|\, t \rangle$, ie. $\mathbb{K}_t = \langle \mathcal{B} \rangle_t \oplus \langle F \,|\, t \rangle$.*

If such a property is satisfied we say that F is a border basis for B in degree $\leq t$.

Theorem 2.12 (border basis, [21])**.** *Let $\mathcal{B} \subset \mathcal{M}$ be connected to 1 and let F be a rewriting family for $\mathcal{B}$, complete in any degree. Assume that $\pi_{F,\mathcal{B}}(c) = 0$ for all $c \in C^+(F)$. Then $\mathcal{B}$ is a basis of $\mathbb{K}[\mathbf{x}]/(F)$, $\mathbb{K} = \langle \mathcal{B} \rangle \oplus (F)$, and $(F)_t = \langle F \,|\, t \rangle$ for all $t \in \mathbb{N}$; the set F is a* border basis *of the ideal $I = (F)$ with respect to $\mathcal{B}$.*

This implies the following characterization of border bases using the commutation property.

Corollary 2.13 (border basis, [20])**.** *Let $\mathcal{B} \subset \mathcal{M}$ be connected to 1 and let F be a rewriting family for $\mathcal{B}$, complete in any degree. If for all $m \in \mathcal{B}$ and all indices $i, j \in [1,n]$, we have:*

$$\pi_{F,\mathcal{B}}(x_i\, \pi_{F,\mathcal{B}}(x_j\, m)) = \pi_{F,\mathcal{B}}(x_j\, \pi_{F,\mathcal{B}}(x_i\, m)),$$

then $\mathcal{B}$ is a basis of $\mathbb{K}/(F)$, $\mathbb{K} = \langle \mathcal{B} \rangle \oplus (F)$, and $(F)_t = \langle F \,|\, t \rangle$ for all $t \in \mathbb{N}$.

2.4. **Hankel operators and positive linear forms.** This section is based on [13].

Definition 2.14. *For $\Lambda \in \mathbb{R}[\mathbf{x}]^*$, the Hankel operator H_Λ is the operator from $\mathbb{R}[\mathbf{x}]$ to $\mathbb{R}[\mathbf{x}]^*$ defined by*

$$H_\Lambda : p \in \mathbb{R}[\mathbf{x}] \mapsto p \cdot \Lambda \in \mathbb{R}[\mathbf{x}]^* \tag{3}$$

Definition 2.15. *We define the kernel of the Hankel operator:*

$$\ker H_\Lambda = \{p \in \mathbb{R}[\mathbf{x}] \mid p \cdot \Lambda = 0, i.e,\ \Lambda(pq) = 0\ \forall q \in \mathbb{R}[\mathbf{x}]\} \tag{4}$$

To analyse the properties of Λ, we study the quotient algebra $\mathbb{K}[\mathbf{x}]/\ker H_\Lambda = \mathcal{A}_\Lambda$. A first result is the following (see e.g. [13]):

Lemma 2.16. *The rank of the operator H_Λ is finite if and only if* $\ker H_\Lambda$ *is a zero-dimensional ideal, in which case* $\dim \mathbb{K}[\mathbf{x}]/\ker H_\Lambda = rank H_\Lambda$.

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

Proposition 2.17. *Let $\mathbb{K} = \mathbb{C}$ and assume that $rank H_\Lambda = r < \infty$. Then, the ideal* $\ker H_\Lambda$ *is radical if and only if*

$$\Lambda = \sum_{i=1}^{r} \lambda_i \mathbf{1}_{\zeta_i} \quad \text{with } \lambda_i \in \mathbb{K} - \{0\} \quad \text{and } \zeta_i \in \mathbb{K}^n \text{ pairwise distinct,} \tag{5}$$

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

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

Definition 2.18. *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 $\Lambda \succcurlyeq 0$ iff $H_\Lambda \succcurlyeq 0$. If moreover $\Lambda(1) = 1$, we say that Λ is a probability measure.*

This term is justified by a theorem of Riesz-Haviland [28], 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 μ 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 2.19. *Assume $rank H_\Lambda = r < \infty$. Then $\Lambda \succcurlyeq 0$ if and only if Λ has a decomposition (5) with $\lambda_i > 0$ and distinct $\zeta_i \in \mathbb{R}^n$, in which case $V(\ker H_\Lambda) = \{\zeta_1, \ldots, \zeta_r\} \subset \mathbb{R}^n$.*

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

3. Main Results

In this section, we give the main results which shows how the minimum of f and the ideal defining the points where this minimum is reached, can be computed.

Hereafter, we will assume that the minimum f^* of f is reached at a point $\mathbf{x}^* \in \mathbb{R}^n$.

Definition 3.1. *We define the gradient ideal of $f(x)$:*

$$I_{grad}(f) = (\nabla f(x)) = (\frac{\partial f}{\partial x_1}, .., \frac{\partial f}{\partial x_n}). \tag{6}$$

Definition 3.2. *We define the minimizer ideal of f(x):*

$$I_{min}(f) = I(\mathbf{x}^* \in \mathbb{R}^n \text{ s.t. } f(\mathbf{x}^*) \text{ is minimum}). \tag{7}$$

By construction, $I_{grad}(f) \subset I_{min}(f)$ and $I_{min}(f) \neq (1)$ if the minimum f^* is reached in $\mathbb{R}^n$. The objective of this section is to describe a method to compute generators of $I_{min}(f)$ from generators of $I_{grad}(f)$. For that purpose, first of all we need to restrict our analysis to matrices of finite size. For this reason, we consider here truncated Hankel operators, which will play a central role in the construction of the minimizer ideal of f.

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

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

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

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

Given a subspace $E_0 \subset E$, Λ induces a linear map on $\langle E_0 \cdot E_0 \rangle$ and we can consider the induced truncated Hankel operator $H_\Lambda^{E_0} : E_0 \to (E_0)^*$ as a restriction of H_Λ^E.

Below we will deal with the following sets and minimums, in order to define our primal-dual problems.

Definition 3.5. *Given a vector space $E \subset \mathbb{R}[\mathbf{x}]$ and $G \subset \langle E \cdot E \rangle$, we define*

$$\mathcal{L}_{G,E,\succcurlyeq} := \{\Lambda \in \langle E \cdot E \rangle^* \mid \Lambda \perp G,\ \Lambda(1) = 1,\ \Lambda(p^2) \geq 0\ \forall p \in E\}. \tag{9}$$

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

Notice that if $G \subset G'$ and $E \subset E'$ then $\mathcal{L}_{G',E',\succcurlyeq} \subset \mathcal{L}_{G,E,\succcurlyeq}$. When E and G are vector spaces of finite dimension, $\mathcal{L}_{G,E,\succcurlyeq}$ 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 4. We will need the following result:

Lemma 3.6. *For any vector space E and $G = \{0\}$ and $\Lambda, \Lambda' \in \mathcal{L}_{G,E,\succcurlyeq}$, we have:*

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

Proof. The first point is a consequence of the positivity of H_Λ^E.

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

Conversely, let $p \in \ker H_{\Lambda+\Lambda'}$. 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 $p \in \ker H_\Lambda \cap \ker H_{\Lambda'}$. □

Definition 3.7. *Given a vector space $E \subset \mathbb{R}[\mathbf{x}]$ and $G \subset \langle E \cdot E \rangle$, we define*

$$\mathcal{S}_{G,E} := \{\ p \in \mathbb{R}[\mathbf{x}] \mid p = \sum_{i=1}^{s} h_i^2 + h,\ h_i \in E,\ h \in G\}. \tag{10}$$

If $E = \mathbb{R}[\mathbf{x}]_t$ and $G' = \langle G \,|\, 2t \rangle$, we also denote $\mathcal{S}_{G',E}$ by $\mathcal{S}_{G,t}$.

Notice that if $G \subset G'$ and $E \subset E'$ then $\mathcal{S}_{G,E} \subset \mathcal{S}_{G',E'}$. When E and G are vector spaces of finite dimension, $\mathcal{S}_{G,E}$ is the projection of the sum of a linear space and the convex cone of positive quadratic forms on $E^* \times E^*$.

Definition 3.8. *Let E be a subspace of $\mathbb{R}[\mathbf{x}]$ such that $1 \in E$ and $f \in \langle E \cdot E \rangle$ and let $G \subset \langle E \cdot E \rangle$. We assume that f attains its minimum at some points $\mathbf{x}^* \in \mathbb{R}^n$. We define the following extrema:*

- *$f^* = \min_{x \in \mathbb{R}^n} f(x)$,*
- *$f_{G,E}^\mu = \inf\ \{\Lambda(f)^2\ s.t.\ \Lambda \in \mathcal{L}_{G,E,\succcurlyeq}\}$,*
- *$f_{G,E}^{sos} = \sup\ \{\lambda \in \mathbb{R}\ s.t.\ f - \lambda \in \mathcal{S}_{G,E}\}$.*

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

Remark 3.9. *By convention if the sets are empty, $f_{G,E}^{sos} = -\infty$ and $f_{G,E}^\mu = +\infty$.*

[2]the notation $^\mu$ stands for measure

We now analyse the relations between these different extrema.

Remark 3.10. *Let $E \subset E'$ be two subspaces of $\mathbb{R}[\mathbf{x}]$ and $G \subset \langle E \cdot E \rangle$, $G' \subset \langle E' \cdot E' \rangle$ with $G \subset G'$ then we directly deduce that*

- $f^{\mu}_{G,E} \leq f^{\mu}_{G',E'}$,
- $f^{sos}_{G,E} \leq f^{sos}_{G',E'}$,

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

If we take $G \subset I_{min}(f)$, we have the following relations between the extrema.

Proposition 3.11. *Let $G \subset I_{min}(f)$. Then $f^{sos}_{G,E} \leq f^{\mu}_{G,E} \leq f^*$.*

Proof. We have $f^{sos}_{G,E} \leq f^{\mu}_{G,E}$ because if there exists $\lambda \in \mathbb{R}$ such that $f - \lambda \in \mathcal{S}_{G,E}$, i.e. $f - \lambda = \sum_i h_i^2 + g$ with $h_i \in E$ and $g \in G$ then $\forall \Lambda \in \mathcal{L}_{G,E,\succcurlyeq}$, $\Lambda(f - \lambda) = \Lambda(f) - \lambda = \sum_i \Lambda(h_i^2) \geq 0$. We deduce that $\Lambda(f) \geq \lambda$ and we conclude $f^{sos}_{G,E} \leq f^{\mu}_{G,E}$. For the second inequality, let $\mathbf{x}^*$ be a point of $\mathbb{R}^n$ such that $f(\mathbf{x}^*)$ is the mininum of f and let $\mathbf{1}_{x^*} \in \mathbb{R}[\mathbf{x}]^* : p \longmapsto p(\mathbf{x}^*)$ be the evaluation at $\mathbf{x}^*$. Then we have $H_{\mathbf{1}_{\mathbf{x}^*}} \succcurlyeq 0$, $\mathbf{1}_{\mathbf{x}^*}(1) = 1$ and $\mathbf{1}_{\mathbf{x}^*}(G) = 0$ since $G \subset I_{min}(f)$, so that $\mathbf{1}_{\mathbf{x}^*} \in \mathcal{L}_{G,E,\succcurlyeq}$. We deduce the inequality $f^{\mu}_{G,E} \leq f^*$. □

Following the relaxation approach proposed in [14], we are now going to consider a hierarchy of convex optimization problems and show that for such hierarchy, the minimum f^* is always reached in a finite number of steps. Let us consider the sequence of spaces

$$\cdots \subset \mathcal{L}_{G,t+1,\succcurlyeq} \subset \mathcal{L}_{G,t,\succcurlyeq} \subset \cdots \text{ and } \cdots \subset \mathcal{S}_{G,t} \subset \mathcal{S}_{G,t+1} \subset \cdots$$

for $t \in \mathbb{N}$, $G \subset I_{min}(f)$. Using Remark 3.10 and the fact that $(G) = \cup_{t \in \mathbb{N}} \langle G | t \rangle$, we check that

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

We are going to show that these limits are attained for some $t \in \mathbb{N}$.

The next result which is a slight variation of a result in [15] (and also used in [13]) shows that for a high enough degree, the kernel of some truncated Hankel operators allows us to compute generators of the real radical of an ideal.

Proposition 3.12. *For $G \subset \mathbb{R}[\mathbf{x}]$ with $I_{grad}(f) \subset (G)$, there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$, $\forall \Lambda \in \mathcal{L}_{G,t,\succcurlyeq}$, $\sqrt[\mathbb{R}]{I_{grad}(f)} \subset (\ker H^t_{\Lambda^*})$.*

Proof. Let $g_1, \ldots, g_n$ be generators of $I := I_{grad}(f)$, $d_s := deg(g_s)$, $d := \max_{s=1,\ldots,n} d_s$ and $h_1, \ldots, h_k$ be generators of the ideal $J := \sqrt[\mathbb{R}]{I}$. By the Real Nullstellensatz, for $l \in 1, \ldots, k$, there exist $m_l \in \mathbb{N}, m_l \geq 1$ and polynomials $u_j^{(l)}$ and sums of squares σ_l such that $h_l^{2m_l} + \sigma_l = \sum_{j=1}^n u_j^{(l)} g_j$. As $I \subset (G)$, there exist t'_0 such that $u_j^{(l)} g_j \in \langle G \,|\, t'_0 \rangle$. Set $t_0 := \max_{l \leq k, j \leq n}(t'_0, d, deg(h_l^{2m_l}), deg(\sigma_l))$ and let $t \geq t_0$. As $u_j^{(l)} g_j \in \langle G \,|\, t \rangle$, $u_j^{(l)} g_j \in \ker H^t_{\Lambda}$ for all $\Lambda \in \mathcal{L}_{G,t,\succcurlyeq}$. Hence $h_l^{2m_l} + \sigma_l \in \ker H^t_{\Lambda^*}$, which implies that $h_l \in \ker H^t_{\Lambda}$ since $H^t_{\Lambda^*} \succcurlyeq 0$. □

To compute generators of $I_{min}(f)$, we use the decomposition of $V_{\mathbb{R}}(I_{grad}(f))$ in components where f has a constant value.

Lemma 3.13. *Let $V_{\mathbb{R}}(I_{grad}(f)) = W_0 \cup W_1 \cup \ldots \cup W_s$ be the decomposition of the variety in disjoint real subvarieties, such that $f(W_j) = f_j \in \mathbb{R}$ with $f_i < f_j$ $\forall 0 \leq i < j \leq s$. Then there exist polynomials $p_0, \ldots, p_r \in \mathbb{R}[\mathbf{x}]$ such that $p_i(W_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta function.*

Proof. As in [25], we decompose $V_{\mathbb{C}}(I_{grad}(f))$ as an union of complex varieties $V_{\mathbb{C}}(I_{grad}(f)) = W_{\mathbb{C},0} \cup W_{\mathbb{C},1} \cup \ldots \cup W_{\mathbb{C},s} \cup W_{\mathbb{C},s+1}$ such that $W_{\mathbb{C},i}$, $i = 0, \ldots, s$ have real points and $f(W_{\mathbb{C},i}) = f_i \in \mathbb{R}$ is constant on $W_{\mathbb{C},i}$ and $W_{\mathbb{C},s+1}$ has no real point. We number

these varieties so that $f_0 < f_1 < \cdots < f_s$ and $f_0 = f^*$. By construction, the varieties $W_i := W_{\mathbb{C},i} \cap \mathbb{R}^n$ are disjoint, f is constant on W_i and $V_{\mathbb{R}}(I_{grad}(f)) = W_0 \cup W_1 \cup \ldots \cup W_s$. Let us take $p_i = L_i(f(\mathbf{x}))$ where $L_0, \ldots, L_s$ are the Lagrange interpolation polynomials at the value $f_0, \ldots, f_n \in \mathbb{R}$. They satisfy $p_i(W_j) = \delta_{ij}$. □

Remark 3.14. *By definition the polynomials p_i have the following properties:*

- $p_0 + \ldots + p_s \equiv 1$ *modulo* $\sqrt[\mathbb{R}]{I_{grad}(f)}$.
- $(p_i)^2 \equiv p_i$ *modulo* $\sqrt[\mathbb{R}]{I_{grad}(f)}$ $\forall i = 0, \ldots, s$.

The next results shows that in the sequence of optimization problems that we consider, the minimum of f is reached from a certain degree.

Theorem 3.15. *For $G \subset \mathbb{R}[\mathbf{x}]$ with $I_{grad}(f) \subset (G) \subset I_{min}(f)$, there exists $t_1 \geq 0$ such that $\forall t \geq t_1$, $f^{sos}_{G,t} = f^{\mu}_{G,t} = f^*$ and $\forall \Lambda^* \in \mathcal{L}_{G,t,\succcurlyeq}$ with $\Lambda^*(f) = f^{\mu}_{G,t}$, we have $p_i \in \ker H^t_{\Lambda^*}$ $\forall i = 1, \ldots, s$.*

Proof. By [24][Theorem 2.3], there exists $t'_1 \in \mathbb{N}$ such that $\forall t \geq t'_1$, $f^{sos}_{D,t} = f^{\mu}_{D,t} = f^*$ where $D = \{\frac{\partial f}{\partial x_1}, .., \frac{\partial f}{\partial x_n}\}$. As $(D) = I_{grad}(f) \subset (G) = (g_1, \ldots, g_s)$, there exists $q_{i,j} \in \mathbb{R}[\mathbf{x}]$ such that $\frac{\partial f}{\partial x_i} = \sum_{j=1}^s q_{i,j} g_j$. Let $d = \max\{\deg(q_{i,j}), i = 1, \ldots, n, j = 1, \ldots, s\}$. Then $\langle D \,|\, t\rangle \subset \langle G \,|\, t+d\rangle$. By Remark 3.10 and Proposition 3.11, for $t \geq t'_1$, $f^* = f^{sos}_{D,t} \leq f^{sos}_{G,t+d} \leq f^{\mu}_{G,t+d} \leq f^*$. Thus for $t \geq t'_1 + d$, we have

$$f^{sos}_{G,t} = f^{\mu}_{G,t} = f^*. \tag{11}$$

Let $J = \sqrt[\mathbb{R}]{I_{grad}(f)}$. By Lemma 3.13 and Remark 3.14, we can write

$$f \equiv \sum_{i=0}^{s} f_i\, p_i^2 \ \textit{modulo}\ J, \tag{12}$$

where $f_i = f(W_i) \in \mathbb{R}$ and $f_0 = f(W_0) = f^*$. Then

$$f \equiv f^* p_0^2 + \sum_{i=1}^{s} f_i\, p_i^2 \equiv f^*(1 - \sum_{r=1}^{s} p_r^2) + \sum_{i=1}^{s} f_i\, p_i^2 \equiv f^* + \sum_{i=1}^{s} (f_i - f^*)\, p_i^2 \ \textit{modulo}\ J. \tag{13}$$

Hence

$$f - f^* \equiv \sum_{i=1}^{s} (f_i - f^*)\, p_i^2 + h. \tag{14}$$

with $h \in J$. By Theorem 3.12, there exists $t''_1 \geq t_0$ such that $\Lambda(h) = 0$ for all $\Lambda \in \ker \mathcal{L}_{G,t''_1,\succcurlyeq}$.

Let us take $t_1 = max\{t'_1 + d, t''_1, \deg(p_1), \ldots, \deg(p_s)\}$, $t \geq t_1$ and $\Lambda^* \in \mathcal{L}_{G,t,\succcurlyeq}$ such that $\Lambda^*(f) = f^{\mu}_{G,t}$. From Equation (11), we deduce that

$$\Lambda^*(f - f^*) = 0 = \sum_{i=1}^{s} (f_i - f^*) \Lambda^*(p_i^2) \tag{15}$$

This implies that $\Lambda^*(p_i^2) = 0$ and $p_i \in \ker H^t_{\Lambda^*}$ for $i = 1, \ldots, s$, since $f_i - f^* > 0$ and $\Lambda^* \succcurlyeq 0$ on $\mathbb{R}[\mathbf{x}]_t$. □

The example 3.4 in [25] shows that it may not always be possible to write $f - f^*$ as a sum of squares modulo $I_{grad}(f)$ but, from the previous result we see that $f - f^*$ is the limit of sums of squares modulo $I_{grad}(f) \cap \mathbb{R}[\mathbf{x}]_t$ for a fixed $t \geq t_1$. Under the assumption that there exists $\mathbf{x}^* \in \mathbb{R}^n$ such that $f(\mathbf{x}^*) = f^*$, we can construct $\Lambda^* = \mathbf{1}_{\mathbf{x}^*} \in \mathcal{L}_{G,t,\succcurlyeq}$ such that $\Lambda^*(f) = f^*$. This means that $f^{\mu}_{G,t}$ is reached for $t \geq t_1$.

A direct corollary of the previous theorem is that for any $G \subset \mathbb{R}[\mathbf{x}]$ with $I_{grad}(f) \subset (G) \subset I_{min}(f)$, we have

$$f^{sos}_{I_{grad}(f)} = f^{sos}_{(G)} = f^{\mu}_{I_{grad}(f)} = f^{\mu}_{(G)} = f^*.$$

As for the construction of generators of the real radical $\sqrt[\mathbb{R}]{I_{grad}(f)}$ (Proposition 3.12), we can construct generators of $I_{min}(f)$ from the kernel of a truncated Hankel operator associated to any linear form which minimizes f:

Theorem 3.16. *For $G \subset \mathbb{R}[\mathbf{x}]$ with $I_{grad}(f) \subset (G) \subset I_{min}(f)$, there exists $t_2 \in \mathbb{N}$ such that $\forall t \geq t_2$, for $\Lambda^* \in \mathcal{L}_{G,t,\succcurlyeq}$ with $\Lambda^*(f) = f^{\mu}_{G,t}$, we have $I_{min}(f) \subset (\ker H^t_{\Lambda^*})$.*

Proof. Let $I = I_{grad}(f)$ and $J = \sqrt[\mathbb{R}]{I}$. We consider again the above decomposition $V_{\mathbb{R}}(I) = W_0 \cup W_1 \cup \ldots \cup W_s$ with $f_0 = f(W_0) < f_1 = f(W_1) < \cdots < f_s = f(W_s)$. We denote by $h_1, \ldots, h_k$ a family of generators of the ideal $I_{min}(f) = I(W_0)$ and $d = \max\{\deg(h_1), \ldots, \deg(h_k)\}$.

Let us fix $0 \leq j \leq k$ and show that $h_j \in \ker H^t_{\Lambda^*}$ for t sufficiently large. We know that $p_0 h_j^2 \mid_{W_0} = 0$ and that for $i = 1, \ldots, s$, $p_0 h_j^2 \mid_{W_i} = 0$ which implies that $p_0 h_j^2 \in J := \sqrt[\mathbb{R}]{I_{grad}(f)}$. By Theorem 3.12, there exists $t'_{2,j}$ such that $\Lambda(p_0 h_j^2) = 0$ for all $\Lambda \in \mathcal{L}_{G,t,\succcurlyeq}$ with $t \geq t'_{2,j}$.

By Theorem 3.15, $p_i \in \ker H^t_{\Lambda^*}$ for $t \geq t_1$. By Remark 3.14, $p_0 + \ldots + p_s \equiv 1$ modulo J. By Theorem 3.12, there exists t''_2 such that $\Lambda(p_0 + \ldots + p_s - 1) = 0$ for all $\Lambda \in \mathcal{L}_{G,t,\succcurlyeq}$ with $t \geq t''_2$.

Let us take $t_2 := \max\{t_1 + 2\,d, t''_2 + 2\,d, t'_{2,1}, \ldots, t'_{2,k}\}$ and $t \geq t_2$. Then for $\Lambda^* \in \mathcal{L}_{G,t,\succcurlyeq}$ with $\Lambda^*(f) = f^{\mu}_{G,t}$, we have

$$\Lambda^*(h_j^2) = \Lambda^*(h_j^2 p_0) + \Lambda^*(h_j^2 p_1) + \ldots + \Lambda^*(h_j^2 p_s) = 0.$$

Hence, $h_j \in \ker H^t_{\Lambda^*}$ for $j = 1, \ldots, k$ and $I_{min}(f) \subset (\ker H^t_{\Lambda^*})$. □

We introduce now the notion of *generic linear form for $f \in \mathbb{R}[\mathbf{x}]$*. Such a linear form will allow us to compute $I_{min}(f)$ as we will see.

Proposition 3.17. *For $\Lambda^* \in \mathcal{L}_{G,E,\succcurlyeq}$ and $p \in \mathbb{R}[\mathbf{x}]$, the following assertions are equivalent:*

(i) $rank H^E_{\Lambda^*} = max_{\Lambda \in \mathcal{L}_{G,E,\succcurlyeq}, \Lambda(p) = p^{\mu}_{G,E}} rank H^E_{\Lambda}$.

(ii) $\forall \Lambda \in \mathcal{L}_{G,E,\succcurlyeq}$ *such that* $\Lambda(p) = p^{\mu}_{G,E}$, $\ker H^E_{\Lambda^*} \subset \ker H^E_{\Lambda}$.

(iii) $rank H^{E_0}_{\Lambda^*} = max_{\Lambda \in \mathcal{L}_{G,E,\succcurlyeq}, \Lambda(p) = p^{\mu}_{G,E}} rank H^{E_0}_{\Lambda}$ *for any subspace* $E_0 \subset E$.

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

Proof. (i) $\Longrightarrow$ (ii): Note that $\frac{1}{2}(\Lambda + \Lambda^*) \in \mathcal{L}_{G,E,\succcurlyeq}$ and $\frac{1}{2}(\Lambda + \Lambda^*)(p) = p^{\mu}_{G,E}$ and $\ker H^E_{\frac{1}{2}(\Lambda+\Lambda^*)} = \ker H^E_{\Lambda} \cap \ker H^E_{\Lambda^*}$ (using Lemma 3.6). Hence, $rank H^E_{\frac{1}{2}(\Lambda+\Lambda^*)} \geq rank H^E_{\Lambda^*}$ and thus equality holds. This implies that $\ker H^E_{\frac{1}{2}(\Lambda+\Lambda^*)} = \ker H^E_{\Lambda^*}$ is thus contained in $\ker H^E_{\Lambda}$.

(ii) $\Longrightarrow$ (iii): Given $E_0 \subset E$, we show that $\ker H^{E_0}_{\Lambda^*} \subset \ker H^{E_0}_{\Lambda}$. By Lemma 3.6, we have $\ker H^{E_0}_{\Lambda^*} \subset \ker H^E_{\Lambda^*}$ and, by the above, we have $\ker H^E_{\Lambda^*} \subset \ker H^E_{\Lambda}$.

(iii) $\Longrightarrow$ (i) This implication is obvious. □

The next result, which refines Theorem 3.16, shows only elements in $I_{min}(f)$ are involved in the kernel of a truncated Hankel operator associated to a generic linear form for f.

Theorem 3.18. *Let E, G be as in Definition 3.8 with $G \subset I_{min}(f)$. If $\Lambda^* \in \mathcal{L}_{G,E,\succcurlyeq}$ is generic for f and such that $\Lambda^*(f) = f^*$, then $\ker H^E_{\Lambda^*} \subset I_{min}(f)$.*

Proof. Let $\mathbf{x}^* \in \mathbb{R}^n$ such that $f(\mathbf{x}^*) = f^*$ is minimum. Let $\underline{\mathbf{1}}_{\mathbf{x}^*}$ denotes the evaluation at $\mathbf{x}^*$ restricted to $\langle E \cdot E \rangle$ and $h \in \ker H^E_{\Lambda^*}$. Our objective is to show that $h(\mathbf{x}^*) = 0$. Suppose for contradiction that $h(\mathbf{x}^*) \neq 0$. We know that $\underline{\mathbf{1}}_{\mathbf{x}^*} \in \mathcal{L}_{G,E,\succcurlyeq}$ since $G \subset I_{min}(f)$

and $\underline{\mathbf{1}}_{\mathbf{x}^*}(f) = f(\mathbf{x}^*) = f^*$. We define $\tilde{\Lambda} = \frac{1}{2}(\Lambda^* + \underline{\mathbf{1}}_{\mathbf{x}^*})$. By definition $\tilde{\Lambda} \in \mathcal{L}_{G,E,\succcurlyeq}$ and $\tilde{\Lambda}(f) = \frac{1}{2}(\Lambda^*(f) + \underline{\mathbf{1}}_{\mathbf{x}^*}(f)) = \frac{1}{2}(\Lambda^*(f) + f(\mathbf{x}^*)) = f^*$. As $h \in \ker H^E_{\Lambda^*}$,

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

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

$$0 = H^E_{\tilde{\Lambda}}(\tilde{h})(h) = \tilde{\Lambda}(h\tilde{h}) = \frac{1}{2}(\Lambda^*(h\tilde{h}) + h(\mathbf{x}^*)\tilde{h}(\mathbf{x}^*)) = \frac{1}{2}(H^E_{\Lambda^*}(h)(\tilde{h}) + h(\mathbf{x}^*)\tilde{h}(\mathbf{x}^*)).$$

As $h \in H^E_{\Lambda^*}$, then

$$0 = H^E_{\tilde{\Lambda}}(\tilde{h})(h) = \frac{1}{2}(h(\mathbf{x}^*)\tilde{h}(\mathbf{x}^*)).$$

As $\tilde{h}(\mathbf{x}^*) \neq 0$, since we have supposed that $h(\mathbf{x}^*) \neq 0$ it yields a contradiction. □

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

Theorem 3.19. *For $G \subset \mathbb{R}[\mathbf{x}]$ with $I_{grad}(f) \subset (G) \subset I_{min}(f)$, there exists $t_2 \in \mathbb{N}$ (defined in Theorem 3.16) such that such that $\forall t \geq t_2$, for $\Lambda^* \in \mathcal{L}_{G,t,\succcurlyeq}$ generic for f, we have $\Lambda^*(f) = f^*$ and $(\ker H^t_{\Lambda^*}) = I_{min}(f)$.*

Proof. We obtain the result as consequence of Theorem 3.15, Theorem 3.16 and Theorem 3.18. □

As a consequence, in a finite number of steps, the sequence of optimization problems that we consider gives the minimum of f and the generators of $I_{min}(f)$.

4. Zero-dimensional Case

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 [18] and applies to global polynomial optimization problems where the minimizer ideal $I_{min}(f)$ is zero-dimensional.

Definition 4.1. *Given vector subspaces $E_0 \subset E \subset \mathbb{K}[\mathbf{x}]$ and $\Lambda \in \langle E \cdot E \rangle^*$, H^E_Λ is said to be a flat extension of its restriction $H^{E_0}_\Lambda$ if $rank H^E_\Lambda = rank H^{E_0}_\Lambda$.*

We recall here a result from [18], which gives a rank condition for the existence of a flat extension of a truncated Hankel operator.

Theorem 4.2. *Consider a monomial set $B \subset \mathcal{M}$ connected to 1 and a linear function Λ defined on $\langle B^+ \cdot B^+ \rangle$. Let $E = \langle B \rangle$ and $E^+ = \langle B^+ \rangle$. Assume that $\operatorname{rank} H^{E^+}_\Lambda = \operatorname{rank} H^E_\Lambda = |B|$. Then there exists a (unique) linear form $\tilde{\Lambda} \in \mathbb{K}[\mathbf{x}]^*$ which extends Λ, i.e., $\tilde{\Lambda}(p) = \Lambda(p)$ for all $p \in \langle B^+ \cdot B^+ \rangle$, satisfying $\operatorname{rank} H_{\tilde{\Lambda}} = \operatorname{rank} H^{E^+}_\Lambda$. Moreover, we have $\ker H_{\tilde{\Lambda}} = (\ker H^{E^+}_\Lambda)$.*

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

Proposition 4.3. *Let E, G be as in Definition 3.8 with $G \subset I_{min}(f)$. If $\Lambda^* \in \mathcal{L}_{G,E,\succcurlyeq}$ coincides with a probability measure μ on $\langle E \cdot E \rangle$ and satisfies $\Lambda^*(f) = f^\mu_{G,E}$, then $\Lambda^*(f) = f^\mu_{G,E} = f^*$.*

Proof. By Proposition 3.11, $f^\mu_{G,E} \leq f^*$. Conversely as $f(x) \geq f^*$ for all $x \in \mathbb{R}^n$, we have $\Lambda^*(f) = \int f d\mu \geq \int f^* d\mu = f^*$. □

Proposition 4.4. *If there exists* $\Lambda \in \mathcal{L}_{G,E,\succcurlyeq}$ *with* $\ker H^E_\Lambda = \{0\}$*, then* $f^{sos}_{G,E} = f^{\mu}_{G,E}$.

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

Theorem 4.5. *Let* B *is a monomial set connected to* 1*,* $E = \langle B^+ \rangle$ *and* $G \subset \langle B^+ \cdot B^+ \rangle \cap I_{min}(f)$*. Let* $\Lambda^* \in \mathcal{L}_{G,E,\succcurlyeq}$ *such that* Λ^* *is generic for* f *and satisfies the flat extension property:* $\operatorname{rank} H^{B^+}_{\Lambda^*} = \operatorname{rank} H^B_{\Lambda^*} = |B|$*. Then there is no duality gap,* $f^* = f^{\mu}_{G,E} = f^{sos}_{G,E}$ *and* $(\ker H^{B^+}_{\Lambda^*}) = I_{min}(f)$.

Proof. As $\operatorname{rank} H^{B^+}_{\Lambda^*} = \operatorname{rank} H^B_{\Lambda^*} = |B|$, Theorem 4.2 implies that there exists a (unique) linear function $\tilde{\Lambda}^* \in \mathbb{K}[\mathbf{x}]^*$ which extends Λ^*. As $\operatorname{rank} H_{\tilde{\Lambda}^*} = \operatorname{rank} H^B_{\Lambda^*} = |B|$ and $\ker H_{\tilde{\Lambda}^*} = (\ker H^{B^+}_{\Lambda})$, any polynomial $p \in \mathbb{R}[\mathbf{x}]$ can be reduced modulo $\ker H_{\tilde{\Lambda}^*}$ to a polynomial $b \in \langle B \rangle$ so that $p - b \in \ker H_{\tilde{\Lambda}^*}$. Then $\tilde{\Lambda}^*(p^2) = \tilde{\Lambda}^*(b^2) = \Lambda^*(b^2) \geq 0$ since $\Lambda^* \in \mathcal{L}_{G,E,\succcurlyeq}$. This implies that $\tilde{\Lambda}^* \succcurlyeq 0$. By Proposition 2.19, $\tilde{\Lambda}^*$ has a decomposition $\tilde{\Lambda}^* = \sum_{i=1}^r \lambda_i \mathbf{1}_{\zeta_i}$ with $\lambda_i > 0$ and $\zeta_i \in \mathbb{R}^n$.

As $\tilde{\Lambda}^*(1) = \Lambda^*(1) = 1$, $\tilde{\Lambda}^*$ is a probability measure. By Proposition 4.3, $\tilde{\Lambda}^*(f) = \Lambda^*(f) = f^{\mu}_{G,E} = f^*$.

As $\tilde{\Lambda}^* = \sum_{i=1}^r \lambda_i \mathbf{1}_{\zeta_i}$ with $\lambda_i > 0$ and $\zeta_i \in \mathbb{R}^n$ and as $\tilde{\Lambda}^*(1) = \sum_{i=1}^r \lambda_i = 1$ and $\tilde{\Lambda}^*(f) = \sum_{i=1}^r \lambda_i f(\zeta_i) = f^*$, we deduce that $f(\zeta_i) = f^*$ for $i = 1, \ldots, r$ so that $\{\zeta_1, \ldots, \zeta_r\} \subset V(I_{min}(f))$.

By Proposition 2.17, Theorem 4.2 and Theorem 3.18, we have

$$\ker H_{\tilde{\Lambda}^*} = I(\zeta_1, \ldots, \zeta_r) = (\ker H^{B^+}_{\Lambda^*}) \subset I_{min}(f).$$

We deduce that $\{\zeta_1, \ldots, \zeta_r\} = V(I_{min}(f))$ so that $(\ker H^{B^+}_{\Lambda^*}) = I_{min}(f)$.

As we have the flat extension condition, $ker H^B_{\Lambda^*} = \{0\}$ and by Proposition 4.4 there is not duality gap: $f^* = f^{\mu}_{G,E} = f^{sos}_{G,E}$. □

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

5. Minimizer border basis algorithm

In this section we describe the algorithm to compute the global minimum of a polynomial, assuming f^* is reached in $\mathbb{R}^n$ and $I_{min}(f)$ is zero dimensional. It can be seen as a type of border basis algorithm, for we insert an additional step in the main loop. It is closely connected to the real radical border basis algorithm presented in [13] but instead of "minimizing zero" to generate new elements in the real radical, we minimize f to compute generators of the minimizer ideal $I_{min}(f)$.

5.1. **Description.** The convex optimization problems that we consider are the following:

Algorithm 5.1: Optimal Linear Form

Input: $f \in \mathbb{R}[\mathbf{x}]$, $M = (\mathbf{x}^\alpha)_{\alpha \in A}$ a monomial set containing 1 with $f = \sum_{\alpha \in A+A} f_\alpha \mathbf{x}^\alpha \in \langle M \cdot M \rangle$, $G \subset \mathbb{R}[\mathbf{x}]$.

Output: the minimum $f^*_{G,M}$ of $\sum_{\alpha \in A+A} \lambda_\alpha f_\alpha$ subject to:

- $H^M_{\Lambda^*} = (h_{\alpha,\beta})_{\alpha,\beta \in A} \succcurlyeq 0$,
- $H^M_{\Lambda^*}$ satisfies the Hankel constraints $h_{0,0} = 1$, and $h_{\alpha,\beta} = h_{\alpha',\beta'} = \lambda_{\alpha+\beta}$ if $\alpha + \beta = \alpha' + \beta'$,
- $\Lambda^*(g) = \sum_{\alpha \in A+A} g_\alpha \lambda_\alpha = 0$ for all $g = \sum_{\alpha \in A+A} g_\alpha \mathbf{x}^\alpha \in G \cap \langle M \cdot M \rangle$.

and $\Lambda^* \in \langle M \cdot M \rangle^*$ represented by the vector $[\lambda_\alpha]_{\alpha \in A+A}$.

This optimization problem is a Semi-Definite Programming problem, corresponding to the optimization of a linear functional on a linear subspace of the cone of Positive Semi-Definite matrices. It is a convex optimization problem, which can be solved efficiently by SDP solvers. If an Interior Point Methods is used, the solution Λ^* is in the interior of a face on which the minimum $\Lambda^*(f)$ is reached so that Λ^* is generic for f. This is the case for tools such as `csdp` or `sdpa`, that we will use in the experimentations.

In the case, where $M = B$ is a monomial set connected to 1, F is a complete rewriting family for B in degree $\leq 2t$, and $G = \{b - \pi_{F,B}(b); b \in B \cdot B\}$, we will also denote Optimal Linear Form $(f, B_t, \pi_{F,B})$:=Optimal Linear Form (f, B_t, G).

Algorithm 5.2: Minimizer Ideal of f

Input: A real polynomial function f with $I_{min}(f) \neq (1)$ and zero-dimensional.
$F := \{\frac{\partial f}{\partial x_1}, .., \frac{\partial f}{\partial x_n}\}$; $t := \lceil deg(f)/2 \rceil$; $B :=$ set of monomials of degree $\leq t$;
$\tilde{f} := -\infty$; stop:=false;
While not stop

(1) Compute the commuting relations for F with respect to B in degree $\leq 2\,t$.
(2) Reduce them by the existing relations F.
(3) Add the non-zero reduced relations to F and update B.
(4) Let $[f^*_{F,B_t}, \Lambda^*] :=$ Optimal Linear Form(f, B_t, π_{F,B_t})
  (a) If there is a duality gap then go to 1 with $t := t + 1$.
  (b) If $\tilde{f} \neq f^*_{F,B_t}$ then $\tilde{f} := f^*_{F,B_t}$; go to 1 with $t := t + 1$.
  (c) Compute the border basis F' of $F + \ker H_{\Lambda^*}^{B_{t-1}}$ and the basis B' in degree $\leq 2\,(t-1)$.
  (d) If there exists $b' \in B'$ with $\deg(b') \geq t - 1$ then go to (1) with $t := t + 1$.
  (e) Let $[f^*_{F',B'}, \Lambda'] :=$ Optimal Linear Form$(f, B', \pi_{F',B'})$.
  (f) If $f^*_{F',B'} = \tilde{f}$ and $\ker H_{\Lambda'}^{B'_{t-1}} = \{0\} \Rightarrow$ stop=true, $F = F'$, $B = B'$, $\tilde{f} = f^*_{F,B_{t-1}}$. else go to 1 with $t := t + 1$.

Output: A basis B of $\mathcal{A} = \mathbb{R}[\mathbf{x}]/I_{min}(f)$, a border basis F of $I_{min}(f)$ for B and the minimum $\tilde{f}$.

5.2. **Correctness of the algorithm.** In this subsection, we analyse the correctness of the algorithm.

Lemma 5.1. *Let $t \in \mathbb{N}$, $B \subset \mathbb{R}[\mathbf{x}]_{2t}$ be a monomial set connected to 1, $F \subset \mathbb{R}[\mathbf{x}]$ be a border basis for B in degree $\leq 2t$ and $G = \langle B_t \cdot B_t \rangle \cap \langle F \,|\, 2t \rangle$. Let $E \subset \mathbb{R}[\mathbf{x}]_t$ be a vector space containing $B^+ \cap \mathbb{R}[\mathbf{x}]_t$ and $G' = \langle E \cdot E \rangle \cap \langle F \,|\, 2t \rangle$. For all $\Lambda \in \mathcal{L}_{G,B_t,\succcurlyeq}$, there exists a unique $\tilde{\Lambda} \in \mathcal{L}_{G',E,\succcurlyeq}$ which extends Λ. Moreover, $\tilde{\Lambda}$ satisfies* $\operatorname{rank} H_{\tilde{\Lambda}}^E = \operatorname{rank} H_{\Lambda}^{B_t}$ *and* $\ker H_{\tilde{\Lambda}}^E = \ker H_{\Lambda}^{B_t} + \langle F \,|\, t \rangle \cap E$.

Proof. Suppose that $F \subset \mathbb{R}[\mathbf{x}]$ be a border basis for B in degree $\leq 2t$, that is, all boundary polynomials of $C^+(F_{2\,t})$ reduces to 0 by $F_{2\,t}$. Then by Theorem 2.11, we have $\mathbb{R}[\mathbf{x}]_{2\,t} = \langle B \rangle_{2\,t} \oplus \langle F \,|\, 2\,t \rangle$ and $\langle E \cdot E \rangle = \langle B \rangle_{2\,t} \oplus G'$, $\langle B_t \cdot B_t \rangle = \langle B_{2\,t} \cap B_t \cdot B_t \rangle \oplus G$. Thus for all $\Lambda \in \mathcal{L}_{G,B_t,\succcurlyeq}$, there exists a unique $\tilde{\Lambda} \in \langle E \cdot E \rangle^*$ such that $\tilde{\Lambda}(G') = 0$ and $\tilde{\Lambda}(b) = \Lambda(b)$ for all $b \in B_t \cdot B_t$.

As $E \cdot (E \cap \langle F \,|\, t \rangle) \subset \langle E \cdot E \rangle \cap \langle F \,|\, 2t \rangle = G'$, we have $\tilde{\Lambda}(E \cdot (\langle F \,|\, t \rangle \cap E)) = 0$ so that

$$\langle F \,|\, t \rangle \cap E \subset \ker H_{\tilde{\Lambda}}^E. \tag{16}$$

For any element $b \in \ker H_{\Lambda}^{B_t}$ we have $\forall b' \in B_t, \Lambda(b\,b') = \tilde{\Lambda}(b\,b') = 0$. As $\tilde{\Lambda}(B_t \cdot (\langle F \,|\, t \rangle \cap E)) = 0$ and $E = \langle B_t \rangle \oplus (\langle F \,|\, t \rangle \cap E)$, for any element $e \in E$, $\tilde{\Lambda}(b\,e) = 0$. This proves that

$$\ker H_{\Lambda}^{B_t} \subset \ker H_{\tilde{\Lambda}}^{E}. \tag{17}$$

Conversely as $E = \langle B_t \rangle \oplus (\langle F \,|\, t \rangle \cap E)$, any element of E can be reduced modulo $\langle F \,|\, t \rangle \cap E$ to an element of $\langle B_t \rangle$, which shows that

$$\ker H_{\tilde{\Lambda}}^{E} \subset \ker H_{\Lambda}^{B_t} + \langle F \,|\, t \rangle \cap E. \tag{18}$$

From the inclusions (16), (17) and (18), we deduce that $\ker H_{\tilde{\Lambda}}^{E} = \ker H_{\Lambda}^{B_t} + (\langle F \,|\, t \rangle \cap E)$ and that $\operatorname{rank} H_{\tilde{\Lambda}}^{E} = \operatorname{rank} H_{\Lambda}^{B_t}$.

As $\Lambda \succcurlyeq 0$, by projection along $(\langle F \,|\, t \rangle \cap E) \subset \ker H_{\tilde{\Lambda}}^{E}$, $\forall p \in E$ there exists $b \in \langle B_t \rangle$ such that $\tilde{\Lambda}(p^2) = \Lambda(b^2) \geq 0$. Thus $\tilde{\Lambda} \succcurlyeq 0$ which ends the proof of this lemma. □

Lemma 5.2. *If the algorithm terminates, then F' is a border basis of $I_{min}(f)$ for B'.*

Proof. If the algorithme stops, this can happen only in step (4) in some degree $t+1$ such that

- $[f^*_{F,B_{t+1}}, \Lambda^*] :=$ Optimal Linear Form(f, B_{t+1}, F),
- F' is a border basis of $F + \ker H_{\Lambda^*}^{B_t}$ for B' in degree $\leq 2\,t$,
- the monomials in B' are of degree $< t$ so that $B'_t = B'$,
- $[f^*_{F',B'}, \Lambda'] :=$ Optimal Linear Form(f, B', F') with $\tilde{f} = \Lambda'(f) = f^*_{F',B'} = \Lambda^*(f) = f^*_{F,B_t}$ and $\ker H_{\Lambda'}^{B'} = \{0\}$.

Then $\langle B'^{+} \rangle = \langle B' \rangle \oplus \langle F' \rangle$. By Lemma 5.1 with $E = \langle B'^{+} \rangle$, the linear form $\Lambda' \in \mathcal{L}_{F',B',\succcurlyeq}$ can be extended to a linear form $\Lambda'' \in \mathcal{L}_{F'',B'^{+},\succcurlyeq}$ with $F'' = \langle F' \,|\, 2t \rangle \cap \langle B'^{+} \cdot B'^{+} \rangle$ such that $\ker H_{\Lambda''}^{B'^{+}} = \langle F' \,|\, t \rangle \cap \langle B^{+} \rangle = \langle F' \rangle$. As $\ker H_{\Lambda'}^{B'} = \{0\}$ and $\langle B'^{+} \rangle = \langle B' \rangle \oplus \langle F' \rangle$, we have $\operatorname{rank} H_{\Lambda''}^{B'^{+}} = \operatorname{rank} H_{\Lambda'}^{B'} = |B'|$ and the flat extension theorem (Theorem 4.2) applies: Λ'' is the restriction to $\langle B' \cdot B' \rangle$ of a positive linear form

$$\tilde{\Lambda}' = \sum_{i=1}^{r} \lambda_i \, \mathbf{1}_{\zeta_i}$$

with $r = |B'|$, $\zeta_i \in \mathbb{R}^n$ distinct, $\lambda_i > 0$ and $\sum_{i=1}^{r} \lambda_i = 1$. Then $\Lambda'(f) = \sum_{i=1}^{r} \lambda_i f(\zeta_i) \geq f^*$. By hypothesis, $\Lambda'(f) = f^*_{F_{2\,t},B_t} \leq f^*$ since $F \subset I_{grad}(f) \subset I_{min}(f)$. We deduce that $\Lambda'(f) = \Lambda^*(f) = f^*$, that $(F_{2t} + \ker H_{\Lambda^*}^{B_t}) = (F') \subset I_{min}(f)$ by Theorem 3.18 and that $(F') = I_{min}(f)$ by Theorem 4.5. Therefore F' is a border basis of $I_{min}(f)$ for B'. □

Proposition 5.3. *Assume that $I_{min}(f)$ is zero-dimensional. Then the algorithm terminates. It outputs a border basis F for B connected to 1, such that $\mathbb{R}[\mathbf{x}] = \langle B \rangle \oplus (F)$ and $(F) = I_{min}(f)$.*

Proof. First, we are going to prove by contradiction that when $I_{min}(f)$ is zero-dimensional, the algorithm terminates. Suppose that the loop goes for ever. Notice that at each step either F is extended by adding new linearly independent polynomials or it moves to $t+1$. Since the number of linearly independent polynomials added to F in degree $\leq 2\,t$ is finite, there is a step in the loop from which F is not modified any more in degree ≤ 2 t. In this case, all boundary C-polynomials of elements of F of degree $\leq 2\,t$ are reduced to 0 by $F_{2\,t}$ and F_{2t} is a border basis for B_{2t} in degree $\leq 2t$. By Lemma 5.1 with $E = R_t$, Λ^* extends to a linear form $\tilde{\Lambda}^* \in \mathbb{R}[\mathbf{x}]^*_{2\,t}$ such that

$$K^t := \ker H^t_{\tilde{\Lambda}^*} = \ker H_{\Lambda^*}^{B_t} + \langle F \,|\, t \rangle.$$

By Theorem 3.19, for $t \geq t_2$ we have $(K^t) = I_{min}(f)$ and $\Lambda^*(f) = f^*$. As $I_{min}(f)$ is zero-dimensional, for t high enough, a border basis F' of K^t in degree $\leq 2\,t$ is a border basis of $(K^t) = I_{min}(f)$. Let B' be the corresponding monomial basis. Then $r := |B'|$ is the number

of minimizers (ie. the points in $V(I_{min}(f))$). By Lemma 5.1 with $E = \langle B'^{+} \rangle$ and Theorem 4.5, any linear form $\Lambda' \in \mathcal{L}_{F',B,\succcurlyeq}$ generic for f is the restriction of a positive linear form

$$\tilde{\Lambda}' = \sum_{i=1}^{r} \lambda_i \, \mathbf{1}_{\zeta_i}$$

with $\{\zeta_1, \ldots, \zeta_r\} = V(I_{min}(f))$ and $\lambda_i > 0$ and $\sum_{i=1}^{r} \lambda_i = 1$. In this case, $\Lambda'(f) = \sum_{i=1}^{r} \lambda_i \, f(\zeta_i) = f^*$ and $\ker H_{\Lambda'}^{B'} = \{0\}$. We arrive at a contradiction, which shows that the algorithm should stops in step (4), for some degree t.

By Lemma 5.2, F' is a border basis of $I_{min}(f)$ with basis B' and $\Lambda'(f) = f^*$. □

6. Examples

This section contains examples which illustrate the behavior of the algorithm. In the first example of Motzkin polynomial, the gradient ideal is not zero-dimensional whereas that in the second example of Robinson polynomial, the gradient ideal is zero-dimensional. In all the examples the minimizer ideal is zero-dimensional hence when we apply our algorithm we obtain the good result in a finite number of steps.

The implementation of the previous algorithm has been performed using the BORDERBASIX[3] package of the MATHEMAGIX[4] software. BORDERBASIX is a `C++` implementation of the border basis algorithm of [32].

Semidefinite positive Hankel operators are computed using the semi definite programming routine of SDPA[5] software. For the link with SDPA we use a file interface since SDPA is not distributed as a library.

For the computation of border basis, we use as a choice function that is tolerant to numerical instability i.e. a choice function that chooses as leading monomial a monomial whose coefficient is maximal among the choosable monomials. This, according to [22], makes the border basis computation stable with respect to numerical perturbations. This property is fundamental as we use results from SDP solvers to get new equations. And as these equations are computed numerically, the computation is subject to numerical errors.

Once the border basis of the minimizer is computed, the roots are obtained using the numerical routines described in [7].

Experiments are made on an Intel Corei7 2.30GHz with 8Gb of RAM.

In the following examples, we use the natation $\overline{H}_{\Lambda}^{t} = H_{\Lambda}^{B_t}$

Example 6.1. *We consider the Motzkin polynomial*

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

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

- *In the first iteration the degree is 3, the size of the Hankel matrix $\overline{H}_{\Lambda}^{3}$ is 10,* $\min \Lambda(f) = -216$*, there is a duality gap hence we try with degree 4.*
- *In the second iteration the degree is 4, the size of the Hankel matrix $\overline{H}_{\Lambda}^{4}$ is 15, min $\Lambda(f) = 0$, there is no duality gap. The minimum differs from the previous minimum, so a new iteration is needed.*
- *In the third iteration the degree is 5, the size of the Hankel matrix $\overline{H}_{\Lambda}^{5}$ is 19 and* $\min \Lambda(f) = 0$*. The minimums are equal hence we compute the kernel of $\overline{H}_{\Lambda}^{4}$, which*

[3] www-sop.inria.fr/galaad/mmx/borderbasix

[4] www.mathemagix.org

[5] sdpa.sourceforge.net

is generated by 5 polynomials. We compute the border basis and obtain the basis $B = \{1, x, y, xy\}$. *All the elements of* B *have degree* < 4.

- *We compute a generic form for* f *with this border basis, the size of the Hankel matrix* $\overline{H}^4_\Lambda$ *is 4,* $\min \Lambda(f)$ *=0 and* $\ker \overline{H}^4_\Lambda = \{0\}$.

After the fourth iteration the algorithm stops and we obtain

(1) $I_{min} = (x^2 - 1, y^2 - 1)$.
(2) *The basis* $B = \{1, x, y, xy\}$.
(3) *The points which minimize f,* $\{(x = 1, y = 1), (x = 1, y = -1), (x = -1, y = 1), (x = -1, y = -1)\}$.

The complete process of resolution took $0.286s$ *on which* $0.154s$ *is spent computing SDP.*

Example 6.2. *We consider the Robinson polynomial,*

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

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

- *In the first iteration the degree is 3, the size of the Hankel matrix* $\overline{H}^3_\Lambda$ *is 10, min* $\Lambda(f) = -0.93$, *there is no duality gap hence we compare the minimum with degree 4.*
- *In the second iteration the degree is 4, the size of the Hankel matrix* $\overline{H}^4_\Lambda$ *is 15,* $\min \Lambda(f) = 0$. *As the minimums are different, another iteration is needed.*
- *In the third iteration the degree is 5, the size of the Hankel matrix* $\overline{H}^5_\Lambda$ *is 19 and* $\min \Lambda(f)$ *=0. The minimums are equal, hence we compute the kernel of* $\overline{H}^4_\Lambda$, *which is generated by 6 polynomials. We compute the border basis and we obtain* $B = \{1, x, y, x^2, xy, y^2, x^2y, xy^2, x^2y^2\}$. *There exists an element of* B *with degree* ≥ 4, *so we go to the next degree.*
- *We compute a generic form* Λ *for* f *in degree* 6, *the size of the Hankel matrix* $\overline{H}^6_\Lambda$ *is 22 and* $\min \Lambda(f)$ *=0. The minimums are equal, hence we compute the kernel of* $\overline{H}^5_\Lambda$, *which is generated by 11 polynomials. We compute the border basis and obtain the basis* $B = \{1, x, y, x^2, xy, y^2, x^2y, xy^2\}$. *All the elements of* B *have degree* < 5.
- *We compute a generic form for* f *with this border basis, the size of the Hankel matrix* $\overline{H}^5_\Lambda$ *is 8,* $\min \Lambda(f)$ *=0 and* $\ker \overline{H}^5_\Lambda = \{0\}$.

After fourth iteration the algorithm stops and we obtain

(1) $I_{min} = (x^3 - x, y^3 - y, x^2y^2 - x^2 - y^2 + 1)$.
(2) *The basis* $B = \{1, x, y, x^2, xy, y^2, x^2y, xy^2\}$.
(3) *The points which minimize f,* $\{(x = 1, y = 1), (x = 1, y = -1), (x = -1, y = 1), (x = -1, y = -1), (x = 1, y = 0), (x = -1, y = 0), (x = 0, y = 1), (x = 0, y = -1)\}$.

The complete process of resolution took $0.315s$ *on which* $0.134s$ *is spent computing SDP.*

Example 6.3. *We consider the polynomial,*

$$f(x, y) = -12x^3 + 3xy^2 + 4y^3 - 16x^2y + 48x^2 - 12y^2$$

We compute its gradient ideal, $I_{grad}(f) = (-36x^2 + 3y^2 - 32xy + 96x, 6xy + 12y^2 - 16x^2 - 24y)$

- *In the first iteration the degree is 2, the size of the Hankel matrix* $\overline{H}^2_\Lambda$ *is 4, min* $\Lambda(f) = -18.6$. *The minimum differs from the previous minimum* $-\infty$. *Hence, a new iteration is needed.*
- *In the second iteration the degree is 3, the size of the Hankel matrix* $\overline{H}^3_\Lambda$ *is 4 and* $\min \Lambda(f) = -18.6$. *The minimums are equal hence we compute the kernel of* $\overline{H}^2_\Lambda$,

which is generated by 3 polynomials. We compute the border basis and we obtain $B = \{1\}$. *All the elements of* B *have degree* < 2.

- *We compute a generic form for* f *with this border basis, the size of the Hankel matrix* $\overline{H}^2_\Lambda$ *is 1,* $\min \Lambda(f) = -18.6$ *and* $\ker \overline{H}^2_\Lambda = \{0\}$.

After second iteration the algorithm stops and we obtain

(1) $I_{min} = (x + 0.43636, y - 2.32727)$.
(2) *The basis* $B = \{1\}$.
(3) *The points which minimize f,* $\{(x = -0.43636, y = 2.32727)\}$.

The complete process of resolution took $0.427s$ *on which* $0.130s$ *were spent computing SDP.*

Example 6.4. *We consider the Leep and Starr polynomial,*

$$f(x, y) = 16 + x^2y^4 + 2x^2y^3 - 4x^3y^3 + 4xy^2 + 20x^2y^2 + 8x^3y^2 + 6x^4y^2 + 8xy - 16x^2y$$

that is positive on $\mathbb{R}^2$ *but cannot be written as sum of squares in* $\mathbb{R}[x, y]$. *We compute its gradient ideal,*
$I_{grad}(f) = (2xy^4 + 4xy^3 - 12x^2y^3 + 4y^2 + 40xy^2 + 24x^2y^2 + 24x^3y^2 + 8y - 32xy, 4x^2y^3 + 6x^2y^2 - 12x^3y^2 + 8xy + 40x^2y + 16x^3y + 12x^4y + 8x - 16x^2)$

- *In the first iteration the degree is 3, the size of the Hankel matrix* $\overline{H}^3_\Lambda$ *is 10, min* $\Lambda(f) = -5.4$, *there is duality gap we try with degree 4 but we change the basis we try again with degree 3.*
- *In the second iteration the degree is 3, the size of the Hankel matrix* $\overline{H}^3_\Lambda$ *is 10, min* $\Lambda(f) = 0.6$, *there is no duality gap. As the minimum differs from the minimum in degree* 2, *a new iteration step is performed.*
- *In the third iteration the degree is 4, the size of the Hankel matrix* $\overline{H}^4_\Lambda$ *is 11 (with the reduction) and min* $\Lambda(f) = 0.6$. *The minimums are equal, hence we compute the kernel of* $\overline{H}^3_\Lambda$, *which is generated by 9 polynomials. We compute the border basis and obtain the basis* $B = \{1\}$. *All the elements of* B *have degree* < 3.
- *We compute a generic form for* f *with this border basis, the size of the Hankel matrix* $\overline{H}^3_\Lambda$ *is 1,* $\min \Lambda(f) = 0.6$ *and* $\ker \overline{H}^3_\Lambda = \{0\}$.

After the third iteration the algorithm stops and we obtain

(1) $I_{min} = (x + 3.3884, y - 0.14347)$.
(2) *The basis* $B = \{1\}$.
(3) *The points which minimize f,* $\{(x = -3.3884, y = 0.14347)\}$.

The complete process of resolution took $0.417s$ *on which* $0.130s$ *were spent computing SDP.*

The experiments suggest that due to the small size of the matrices $\overline{H}^i_\Lambda$ most of the resolution time is spent during a classical border basis computation, we all the more emphasize this, as we used a file interface for communicating with `sdpa` which is rather slow.

References

[1] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry.* Springer, 1998.
[2] S. Boyd and L. Vandenbergh. *Convex Optimization.* Cambridge University Press., New York, 2004.
[3] D.A. Cox, J.B. Little, and D.B. O'Shea. *Using Algebraic Geometry.* Springer, 1998.
[4] 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.
[5] 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.
[6] M. Elkadi and B. Mourrain. *Introduction à la résolution des systèmes d'équations algébriques*, volume 59 of *Mathématiques et Applications.* Springer-Verlag, 2007.

[7] S. Graillat and Ph. Trébuchet. A new algorithm for computing certified numerical approximations of the roots of a zero-dimensional system. In *ISSAC2009*, pages 167–173, July 2009.

[8] 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.

[9] 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.

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

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

[12] 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.

[13] 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.

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

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

[16] 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.

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

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

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

[20] 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.

[21] B. Mourrain and P. Trébuchet. Generalized normal forms and polynomials system solving. In M. Kauers, editor, *ISSAC: Proceedings of the ACM SIGSAM International Symposium on Symbolic and Algebraic Computation*, pages 253–260, 2005.

[22] B. Mourrain and Ph. Trébuchet. Stable normal forms for polynomial system solving. *Theoretical Computer Science*, 409(2):229–240, 2008.

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

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

[27] 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.

[28] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.

[29] 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.

[30] M. Schweighofer. Global optimization of polynomials using gradient tentacles and sums of squares. *SIAM Journal on Optimization*, 17(3):920–942, 2006.

[31] N.Z. Shor. Class of global minimum bounds of polynomial functions. *Cybernetics*, 23:731–734, 1987.

[32] Ph. Trébuchet and B. Mourrain. Border basis representation of a general quotient algebra. In Joris van der Hoeven, editor, *ISSAC 2012*, pages 265–272, July 2012.

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

Philippe Trebuchet: APR, LIP6, 4 place Jussieu 75005 Paris, France