CONVEX SETS WITH SEMIDEFINITE REPRESENTATION

JEAN B. LASSERRE

Abstract. We provide a sufficient condition on a class of compact basic semialgebraic sets $\mathbf{K} \subset \mathbb{R}^n$ for their convex hull $\mathrm{co}(\mathbf{K})$ to have a semidefinite representation (SDr). This SDr is explicitly expressed in terms of the polynomials g_j that define $\mathbf{K}$. Examples are provided. We also provide an approximate SDr; that is, for every fixed $\epsilon > 0$, there is a convex set $\mathbf{K}_\epsilon$ such that $\mathrm{co}(\mathbf{K}) \subseteq \mathbf{K}_\epsilon \subseteq \mathrm{co}(\mathbf{K}) + \epsilon\mathbf{B}$ (where $\mathbf{B}$ is the unit ball of $\mathbb{R}^n$), and $\mathbf{K}_\epsilon$ has an explicit SDr in terms of the g_j's. For convex and compact basic semi-algebraic sets $\mathbf{K}$ defined by concave polynomials, we provide a simpler explicit SDr when the nonnegative Lagrangian L_f associated with $\mathbf{K}$ and any linear $f \in \mathbb{R}[X]$ is a sum of squares. We also provide an approximate SDr specific to the convex case.

1. Introduction

An important issue raised in e.g. Ben-Tal and Nemirovski [2], Helton and Vinnikov [4], Parrilo and Sturmfels [14], is to characterize convex sets of $\mathbb{R}^n$ that have a *lifted LMI (Linear Matrix Inequalities)* (or a *semidefinite* representation (SDr)), and called SDr sets in [2] (that is, sets which are semidefinite representable).

Recall that a convex set $\Omega \subset \mathbb{R}^n$ is SDr if there exist integers m, p and real $p \times p$ symmetric matrices $\{A_i\}_{i=0}^n$, $\{B_j\}_{j=1}^m$ such that:

$$\Omega = \{x \in \mathbb{R}^n \,:\, \exists\, y \in \mathbb{R}^m \text{ s.t. } A_0 + \sum_{i=1}^n A_i x_i + \sum_{j=1}^m B_j y_j \succeq 0\,\} \tag{1.1}$$

(where the notation $A \succeq 0$ stands for the matrix A is positive semidefinite). In other words, Ω is the linear projection on $\mathbb{R}^n$ of the convex set

$$\Omega' := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m \,:\, A_0 + \sum_{i=1}^n A_i x_i + \sum_{j=1}^m B_j y_j \succeq 0\,\}\,(\subset \mathbb{R}^{n+m})$$

of the *lifted* space $\mathbb{R}^{n+m}$. The set Ω' is called a *semidefinite* representation (SDr) of Ω and is a lifted LMI because one sometimes needs additional variables $y \in \mathbb{R}^m$ to obtain a description of Ω via appropriate LMIs. For instance:

1991 *Mathematics Subject Classification.* 52A20 52A27 52A41 90C22 90C25 13B25 12D15.

Key words and phrases. Convex sets; semidefinite representation; representation of positive polynomials; sum of squares.

- The intersection of half-spaces, i.e., a polyhedron $\{x \in \mathbb{R}^n : Ax \leq b\}$, is a trivial example of convex sets whose SDr is readily available without lifting. Indeed $Ax \leq b$ is an LMI with diagonal matrices A_i in (1.1).
- The intersection of ellipsoids

$$\Omega := \{x \in \mathbb{R}^n : x^T Q_j x + b^T x + c_j \geq 0,\ j = 1, \ldots, m\}$$

(where $-Q_j \succeq 0$ for all $j =, \ldots, m$) is a SDr set with lifted LMI representation in $\mathbb{R}^{(n+1)(n+2)/2-1}$:

$$\Omega' = \left\{ (x, Y) : \begin{array}{l} \left[\begin{array}{cc} 1 & x^T \\ x & Y \end{array}\right] \succeq 0 \\ \\ \mathrm{trace}\,(Q_j Y) + b^T x + c_j \geq 0, \quad j = 1, \ldots, m. \end{array} \right\}. \tag{1.2}$$

- The epigraph of a univariate convex polynomial is a SDr set.
- Convex sets of $\mathbb{R}^2$ described from genus-zero plane curves are SDr sets; see Parrilo [14].
- *Hyperbolic* cones obtained from 3-variables hyperbolic homogeneous polynomials are SDr sets; see the proof of the Lax conjecture in Lewis et al. [13].

So far, and except for the special cases cited above, little is known. In addition, even if a convex set $\mathbf{K}$ is known to be SDr, there is no systematic procedure to obtain its SDr, i.e., the set of lifted LMIs whose projection describe $\mathbf{K}$. However, Helton and Vinnikov [4] have proved recently that *rigid convexity* is a necessary condition for a set to be SDr (and sufficient for dimension $n = 2$). Chua and Tuncel [3] consider even more general lifted conic representations of convex sets, called lifted G-representations (SDr being a special case) and discuss various geometric properties of convex sets admitting such lifted G-representations, as well as measures of "goodness" for such representations.

In this paper, we consider the convex hull co($\mathbf{K}$) of compact basic semi-algebraic sets $\mathbf{K} \subset \mathbb{R}^n$ of the form

$$\mathbf{K} = \{ x \in \mathbb{R}^n : \quad g_j(x) \geq 0, \quad j = 1, \ldots, m \}, \tag{1.3}$$

for some given polynomials $g_j \in \mathbb{R}[X]$, $j = 1, \ldots, m$. Notice that the class of sets (1.3) is fairly general as $\mathbf{K}$ can be nonconvex (even disconnected), as well as discrete.

Contribution: Our contribution is twofold:

I. We first provide a sufficient condition (and a variant of it) on the defining polynomials $(g_j) \subset \mathbb{R}[X]$ of $\mathbf{K}$ that we call *Schmüdgen's Bounded Degree Representation* (S-BDR) of *affine* polynomials and its *Putinar-Prestel* variant (PP-BDR). A basic compact semi-algebraic set has the S-BDR (resp.

PP-BDR) property if almost all affine polynomials $f \in \mathbb{R}[X]$ positive on $\mathbf{K}$ (hence on $\mathrm{co}(\mathbf{K})$) belong to $P_r(g)$ (resp. $Q_r(g)$), a subset of the preordering $P(g)$ (resp. quadratic module $Q(g)$) generated by the g_j. When $f \in P_r(g)$ or $Q_r(g)$, all elements in the representation of f in the preordering $P(g)$ or in the quadratic module $Q(g)$, have degree at most r.

Recall that when $\mathbf{K}$ is compact then $f > 0$ on $\mathbf{K}$ implies $f \in P(g)$ (or $f \in Q(g)$ if $N - \|X\|^2 \in Q(g)$ for some N), and so the S-BDR (or PP-BDR) property is stronger in that it requires $f \in P_r(g)$ (or $Q_r(g)$). On the other hand, this requirement is only concerned with the class of positive *affine* polynomials.

For instance, this property holds for intersections of halfspaces and ellipsoids, i.e., when the g_j's are affine or quadratic and concave. But we also exhibit some nontrivial non convex compact semi-algebraic sets $\mathbf{K}$ with the PP-BDR property. For instance, we show that when $m = 2$ and the g_j's are quadratic, or when $n = 2$ and the g_j's are quartic, then the PP-BDR property holds generically, and with order $r = 1$ and $r = 2$, respectively.

When the S-BDR or PP-BDR property holds then one can immediately obtain an *explicit* SDr of $\mathrm{co}(\mathbf{K})$, expressed directly in terms of the defining polynomials g_j.

We also obtain an approximate result of the following flavor. For every fixed $\epsilon > 0$, we exhibit a convex set $\mathbf{K}_\epsilon$ such that (a) $\mathrm{co}(\mathbf{K}) \subseteq \mathbf{K}_\epsilon \subseteq \mathrm{co}(\mathbf{K}) + \epsilon\mathbf{B}$ (where $\mathbf{B}$ is the unit ball of $\mathbb{R}^n$), and (b) $\mathbf{K}_\epsilon$ has an *explicit* SDr expressed directly in terms of the polynomials g_j that define $\mathbf{K}$. This result improves significantly upon [8] where we have provided *outer* convex approximations of $\mathrm{co}(\mathbf{K})$, i.e., a monotone nonincreasing sequence of convex sets $\mathbf{K}_r$, with $\mathbf{K}_r \downarrow \mathrm{co}(\mathbf{K})$, and where each $\mathbf{K}_r$ has a SDr. And so, if $x \notin \mathrm{co}(\mathbf{K})$ then $x \notin \mathbf{K}_r$ for all $r \geq r(x)$ for some $r(x)$ that *depends* on x, an undesirable feature.

II. However, for general basic semialgebraic sets $\mathbf{K}$, one cannot expect that the S-BDR (or PP-BDR) property holds (if it ever holds) for *nice* values of the order r. Indeed otherwise one could minimize *any* affine polynomial on $\mathbf{K}$ *efficiently*. Therefore, from a practical point of view, the most interesting case is essentially when $\mathbf{K}$ is convex ... and even more ... when the defining polynomials g_j in (1.3) are concave, because then one may hope for the S-BDR or PP-BDR property to hold for interesting values of r.

So, our second contribution is concerned with the case of compact convex basic semialgebraic sets $\mathbf{K}$ defined by *concave* polynomials. We first show that the PP-BDR property holds for $\mathbf{K}$ whenever the Lagrangian L_f associated with $\mathbf{K}$ and an arbitrary *linear* $f \in \mathbb{R}[X]$ is a sum of squares (s.o.s.) (by construction it is already nonnegative). In this case, $\mathbf{K}$ has a natural SDr based on the Karush-Kuhn-Tucker optimality conditions. This makes an interesting connection between convexity and s.o.s. Finally, we also provide an *approximate* SDr of $\mathbf{K}$, specific to the convex case.

2. I. Semidefinite representation of $\mathrm{co}(\mathbf{K})$

2.1. **Notation and definitions.** For a real symmetric matrix A the notation $A \succeq 0$ (resp. $A \succ 0$) stands for A is positive semidefinite (resp. positive definite). Let $\mathbb{R}[X]$ be the ring of real polynomials in the variables $X = (X_1, \dots, X_n)$ and let $\Sigma^2 \subset \mathbb{R}[X]$ be its subset of sums of squares (s.o.s.) (whereas Σ^2_d is that of degree at most $2d$). For $x \in \mathbb{R}^n$, let $\|x\|$ denote its euclidean norm.

With $d \in \mathbb{N}$, let $s(d) := \binom{n+d}{n}$, and let $u(X) \in \mathbb{R}^{s(d)}$ be the column vector

$$u_d(X) = (1, X_1, \dots, X_n, X_1^2, X_1X_2, \dots, X_n^d)^T,$$

whose components form the usual canonical basis of the vector space $\mathbb{R}[X]_d$ (of dimension $s(d)$) of real polynomials of degree at most d.

Given a infinite sequence $y := \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ indexed in the canonical basis $u_\infty(X)$, let $L_y : \mathbb{R}[X] \to \mathbb{R}$ be the linear mapping

$$f \in \mathbb{R}[X] \; (= \sum_{\alpha\in\mathbb{N}^n} f_\alpha \, X^\alpha) \;\longmapsto\; L_y(f) \;:=\; \sum_{\alpha\in\mathbb{N}^n} f_\alpha \, y_\alpha, \tag{2.1}$$

and let $\mathbf{f} = \{f_\alpha\} \in \mathbb{R}^{s(d)}$ be the vector of coefficients of $f \in \mathbb{R}[X]_d$ in the basis $u_d(X)$.

Moment matrix. Let $M_d(y)$ be the $s(d) \times s(d)$ real matrix with rows and columns indexed in the basis $u_d(X)$, and defined by:

$$M_d(y)(\alpha, \beta) \;=\; y_{\alpha+\beta}, \qquad \alpha, \beta \in \mathbb{N}^n, \quad |\alpha|, |\beta| \leq d, \tag{2.2}$$

where for every $\alpha \in \mathbb{N}^n$, the notation $|\alpha|$ stands for $\sum_{i=1}^n \alpha_i$.

Equivalently, $M_d(y) = L_y(u_d(X)u_d(X)^T)$, meaning that L_y is applied *entrywise* to the polynomial matrix $u_d(X)u_d(X)^T$. The matrix $M_d(y)$ is called the *moment* matrix associated with the sequence y; see e.g. [9]. If y has a *representing measure* μ_y (i.e., if $y_\alpha = \int X^\alpha d\mu_y$ for every $\alpha \in \mathbb{N}^n$) then, one has

$$\langle \mathbf{f}, M_d(y)\mathbf{f} \rangle \;=\; \int f^2 \, d\mu_y \;\geq\; 0, \qquad \forall f \in \mathbb{R}[X]_d, \tag{2.3}$$

so that $M_d(y) \succeq 0$.

Localizing matrix. Similarly, given $y = \{y_\alpha\}$ and $\theta \in \mathbb{R}[X]$, let $M_d(\theta y)$ be the $s(d) \times s(d)$ matrix defined by:

$$M_d(\theta y) \;:=\; L_y(\theta(X)u_d(X)u_d(X)^T), \tag{2.4}$$

i.e., L_y is applied entrywise to the polynomial matrix $\theta(X)u_d(X)u_d(X)^T$. The matrix $M_d(\theta y)$ is called the *localizing* matrix associated with the sequence y and the polynomial θ (see again [9]). Notice that the localizing matrix with respect to the constant polynomial $\theta \equiv 1$ is the moment matrix $M_d(y)$ in (2.2).

If y has a representing measure μ_y with support contained in the level set $\{x \in \mathbb{R}^n : \theta(x) \geq 0\}$ (where $\theta \in \mathbb{R}[X]$), then

$$\langle \mathbf{f}, M_d(\theta y)\mathbf{f}\rangle \;=\; \int f^2\theta\, d\mu_y \;\geq\; 0 \qquad \forall\, f \in \mathbb{R}[X]_d, \tag{2.5}$$

so that $M_d(\theta y) \succeq 0$.

2.2. **Semidefinite representation of** co($\mathbf{K}$)**.** Let $\mathbf{K} \subset \mathbb{R}^n$ be the basic closed semi-algebraic set defined in (1.3) for some polynomials $g_j \in \mathbb{R}[X]$, $j = 1, \ldots, m$.

For every $J \subseteq \{1, \ldots, m\}$, let $g_J := \prod_{j\in J} g_j$, with the convention $g_\emptyset \equiv 1$, and let $P(g) \subset \mathbb{R}[X]$ be the *preordering* generated by the g_j's, i.e.,

$$P(g) \;:=\; \{ \sum_{J\subseteq\{1,\ldots,m\}} \sigma_J\, g_J \;:\quad \sigma_J \in \Sigma^2\},$$

and given $r \in \mathbb{N}$, define $P_r(g) \subset P(g)$ to be the set

$$P_r(g) \;:=\; \{ \sum_{J\subseteq\{1,\ldots,m\}} \sigma_J\, g_J \;:\quad \sigma_J \in \Sigma^2,\, \deg \sigma_J + \deg\, g_J \leq 2r\}. \tag{2.6}$$

Similarly, let $Q(g) \subset \mathbb{R}[X]$ be the *quadratic module* generated by the g_j's, i.e.,

$$Q(g) \;:=\; \{ \sum_{j=0}^{m} \sigma_j\, g_j \;:\quad \sigma_j \in \Sigma^2\}$$

(with the convention $g_0 \equiv 1$), and given $r \in \mathbb{N}$, define $Q_r(g) \subset Q(g)$ to be the set

$$Q_r(g) \;:=\; \{ \sum_{j=0}^{m} \sigma_j\, g_j \;:\quad \sigma_j \in \Sigma^2,\, \deg \sigma_j + \deg\, g_j \leq 2r\}. \tag{2.7}$$

Definition 1 (Semi Definite representation (SDr))**.** A convex set $\mathbf{\Omega} \subset \mathbb{R}^n$ has a SDr (or is a SDr set) if it has the form

$$\mathbf{\Omega} \;=\; \{\, x \in \mathbb{R}^n \;:\quad \exists\, y \in \mathbb{R}^m \text{ s.t.} \qquad A_0 + \sum_{i=1}^{n} A_i x_i + \sum_{k=1}^{p} B_k y_k \;\succeq\; 0\,\} \tag{2.8}$$

for some integer p and real symmetric matrices $\{A_i\}$ and $\{B_k\}$.

For an affine polynomial $X \mapsto f_0 + \sum_{i=1}^n f_i X_i$, let $(f_0, \mathbf{f}) \in \mathbb{R} \times \mathbb{R}^n$ be its vector of coefficients.

Definition 2 (Schmüdgen's Bounded Degree Representation of affine polynomials)**.** Given a compact set $\mathbf{K} \subset \mathbb{R}^n$ defined as in (1.3), one says that *Schmüdgen's Bounded Degree Representation* (S-BDR) of affine polynomials holds for $\mathbf{K}$ if there exists $r \in \mathbb{N}$ such that

$$[\; f \text{ affine and positive on } \mathbf{K}\;] \quad \Rightarrow \quad f \in P_r(g), \tag{2.9}$$

except perhaps on a set of vectors $\mathbf{f}$ in $\mathbb{R}^n$ with Lebesgue measure zero. Call r its order.

Definition 3 (Putinar-Prestel's Bounded Degree Representation of affine polynomials)**.** Given a compact set $\mathbf{K} \subset \mathbb{R}^n$ defined as in (1.3), one says that *Putinar-Prestel's Bounded Degree Representation* (PP-BDR) of affine polynomials holds for $\mathbf{K}$ if there exists $r \in \mathbb{N}$ such that

$$(2.10) \qquad [\, f \text{ affine and positive on } \mathbf{K}\,] \quad \Rightarrow \quad f \in Q_r(g),$$

except perhaps on a set of vectors $\mathbf{f}$ in $\mathbb{R}^n$ with Lebesgue measure zero. Call r its order.

Remark 1. (a) Observe that if $\mathbf{K}$ is compact, by Schmüdgen's Positivstellensatz [17]

$$[\, f \in \mathbb{R}[X] \;\text{ and }\; f \text{ positive on } \mathbf{K}\,] \quad \Rightarrow \quad f \in P_r(g),$$

for some $r(f)$. The S-BDR property states that $r(f) < r$ for almost all affine $f \in \mathbb{R}[X]$, positive on $\mathbf{K}$.

(b) If for some N, the polynomial $N - \|X\|^2$ is in $Q(g)$, then by Putinar's Positivstellensatz [16]

$$[\, f \in \mathbb{R}[X] \;\text{ and }\; f \text{ positive on } \mathbf{K}\,] \quad \Rightarrow \quad f \in Q_r(g),$$

for some $r(f)$; see also Jacobi and Prestel [7]. The PP-BDR property states that $r(f) < r$ for almost all affine $f \in \mathbb{R}[X]$, positive on $\mathbf{K}$.

(c) Finally, the PP-BDR property implies the S-BDR property.

For every $J \subseteq \{1, \ldots, m\}$ let $r_J := \lceil \deg g_J/2 \rceil$.

Theorem 2. *Let $\mathbf{K} \subset \mathbb{R}^n$ be compact and defined as in (1.3).*

(a) *If the S-BDR property holds for $\mathbf{K}$ with order r, then* $\mathrm{co}(\mathbf{K})$ *is a SDr set with SDr*

(2.11)
$$\left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^{s(2r)} \;:\; \left\{ \begin{array}{lcll} M_{r-r_J}(g_J\, y) & \succeq & 0, & J \subseteq \{1, \ldots, m\} \\ L_y(X_i) & = & x_i, & i = 1, \ldots, n \\ y_0 & = & 1 & \end{array} \right. \right\}.$$

(b) *If the PP-BDR property holds for $\mathbf{K}$ with order r, then* $\mathrm{co}(\mathbf{K})$ *is a SDr set with SDr*

(2.12)
$$\left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^{s(2r)} \;:\; \left\{ \begin{array}{lcll} M_{r-r_j}(g_j\, y) & \succeq & 0, & j = 0, 1, \ldots, m \\ L_y(X_i) & = & x_i, & i = 1, \ldots, n \\ y_0 & = & 1 & \end{array} \right. \right\}.$$

Proof. We only prove (a) as (b) is proved in exactly the same manner. Let $\Omega \subset \mathbb{R}^n \times \mathbb{R}^{s(r)}$ be the set defined in (2.11). We have to show that $(\exists\, y : (x,y) \in \Omega) \Leftrightarrow x \in \mathrm{co}(\mathbf{K})$.

1. $x \in \mathrm{co}(\mathbf{K}) \Rightarrow (x,y) \in \Omega$ for some $y \in \mathbb{R}^{s(r)}$. Observe that by the definition of $\mathrm{co}(\mathbf{K})$,

$$x \in \mathrm{co}(\mathbf{K}) \quad \Leftrightarrow \quad x = \int X \, d\mu$$

for some probability measure μ supported on $\mathbf{K}$. Let $y = (y_\alpha) \in \mathbb{R}^{s(r)}$ be the sequence of moments of μ up to order $2r$, i.e.

$$y_\alpha \,=\, L_y(X^\alpha) \,=\, \int X^\alpha \, d\mu, \quad \alpha \in \mathbb{N}^n;\, |\alpha| \,\leq\, 2r.$$

The sequence y is well defined because μ has compact support; in particular $y_0 = 1$. From the definition of μ one has $L_y(X_i) = \int X_i d\mu = x_i$. In addition, as μ is supported on $\mathbf{K}$, one has $M_{r-r_J}(g_J\, y) \succeq 0$ for all subsets $J \subseteq \{1,\ldots,m\}$ (just take $\theta := g_J$ in (2.5)). And so $(x,y) \in \Omega$.

2. $\exists\, y:\ (x,y) \in \Omega \Rightarrow x \in \mathrm{co}(\mathbf{K})$. We prove it by contradiction.

Let $x \notin \mathrm{co}(\mathbf{K})$ and assume that there exists $y \in \mathbb{R}^{s(r)}$ such that $(x,y) \in \Omega$. As $\mathrm{co}(\mathbf{K})$ is convex and compact, by the Hahn-Banach separation theorem, there exists $(f_0, \mathbf{f}) \in \mathbb{R} \times \mathbb{R}^n$ such that

$$\langle \mathbf{f}, x\rangle \,<\, f_0 \quad \text{and} \quad \langle \mathbf{f}, z\rangle \,>\, f_0 \quad \forall\, z \in \mathrm{co}(\mathbf{K}), \tag{2.13}$$

and so the affine polynomial $f \in \mathbb{R}[X]$, $X \mapsto f(X) := -f_0 + \sum_i f_i X_i$ is positive on $\mathbf{K}$.

By the PP-BDR property of $\mathbf{K}$ with order r, one has $f \in P_r(g)$ or $\tilde{f} \in P_r(g)$ for some affine $\tilde{f} \in \mathbb{R}[X]$ with coefficient vector $(\tilde{\mathbf{f}}, \tilde{f}_0)$ such that $\|\tilde{\mathbf{f}}\| = 1$ and some $\epsilon > 0$ such that $\|(\tilde{\mathbf{f}}, \tilde{f}_0) - (\mathbf{f}, f_0)\| < \epsilon$, with $\epsilon > 0$ as small as desired. Therefore, one may choose ϵ sufficiently small to ensure that $(\tilde{\mathbf{f}}, \tilde{f}_0)$ also satisfies (2.13) and so, one may rename $\tilde{f}$ as f and safely assume that $f \in P_r(g)$. Hence,

$$f(X) \,=\, \sum_{J \subseteq \{1,\ldots,m\}} \sigma_J\, g_J, \quad \sigma_J \in \Sigma^2;\, \deg \sigma_J + \deg g_J \leq 2r. \tag{2.14}$$

Observe that as σ_J is s.o.s. and $\deg \sigma_J + \deg\, g_J \leq 2r$, one has

$$L_y(\sigma_J\, g_J) \,\geq\, 0 \qquad \forall\, J \subseteq \{1,\ldots,m\}. \tag{2.15}$$

Applying the linear functional L_y to the polynomial f in (2.14) yields the contradiction

$$\begin{aligned} 0 \,>\, \mathbf{f}^T x - f_0 \;&=\; -f_0 y_0 + \sum_{i=1}^n L_y(f_i X_i) \quad [\text{as } y_0 = 1,\, L_y(X_i) = x_i \,\forall\, i] \\ &=\; L_y(f(X)) \,=\, \sum_{J \subseteq \{1,\ldots,m\}} L_y(g_J\, \sigma_J) \quad [\text{ by (2.14) }] \\ &\geq\; 0 \quad [\text{ by (2.15) }]. \end{aligned}$$

This proves that there is no y such that $(x,y) \in \Omega$, the desired result. □

Notice that in Theorem 2, the SDr (2.11) and (2.12) of $\mathrm{co}(\mathbf{K})$ are given *explicitly* in terms of the data g_j's that define $\mathbf{K}$.

2.3. **Examples of convex K.** We have already seen that the intersection $\mathbf{K}$ of half-spaces and/or ellipsoids is a SDr set. But we here show that the PP-BDR property holds for such sets $\mathbf{K}$, and also for the intersection of level sets of quartic polynomials in two variables. Of course, one already knows how to build up a SDr for $\mathbf{K}$ at least in the first two cases. But this is to illustrate that the domain of application of Theorem 2 is not empty and not trivial.

Example 1. Let us start with $\mathbf{K}$ being a convex polytope defined by linear inequalities, i.e., $g_j \in \mathbb{R}[X]$ is affine in X for all $j = 1, \ldots, m$. Hence $\mathrm{co}(\mathbf{K}) \equiv \mathbf{K}$ and this description of $\mathbf{K}$ by the g_j's is already a SDr; it is even a linear system. We briefly prove that the PP-BDR property holds for $\mathbf{K}$ with order 0. Let $f \in \mathbb{R}[X]$ be affine with coefficient vector $(f_0, \mathbf{f}) \in \mathbb{R} \times \mathbb{R}^n$, and write

$$\begin{aligned} g_j(X) &= g_{j0} + \sum_{i=1}^n g_{ji} X_i, \qquad j = 1, \ldots, m \\ f(X) &= f_0 + \sum_{i=1}^n f_i\, X_i. \end{aligned}$$

Next, let $G \in \mathbb{R}^{m \times n}$ be the matrix $G(j,i) = g_{ji}$, $j = 1, \ldots, m$, $i = 1, \ldots, n$, and $\mathbf{g} = (g_{j0}) \in \mathbb{R}^m$. If f is nonnegative on $\mathbf{K}$ then by Farkas lemma $\mathbf{f} = \lambda^T G$ and $f_0 \geq \lambda^T \mathbf{g}$, for some nonnegative vector $\lambda \in \mathbb{R}^m$. Therefore $f(X) = u + \sum_{j=1}^m \lambda_j g_j(X)$, for some nonnegative scalar u, which proves that $f \in Q_1(g)$. That is, the PP-BDR property holds for $\mathbf{K}$ with order $r = 1$.

Example 2. Let $g_j \in \mathbb{R}[X]$ be concave and quadratic, for all $j = 1, \ldots, m$. Then $\mathbf{K}$ is convex and it is well-known that $\mathbf{K}$ is a SDr set. Let $f \in \mathbb{R}[X]$ be affine with coefficient vector $(f_0, \mathbf{f}) \in \mathbb{R} \times \mathbb{R}^n$, and nonnegative on $\mathbf{K}$, so that $f^* := \min_{x \in \mathbf{K}} f(x) \geq 0$. Assume that $\mathbf{K}$ is compact with nonempty interior. Convexity along with Slater's condition[1] imply that the KKT optimality conditions hold at any global minimizer $x^* \in \mathbf{K}$, i.e.,

$$\mathbf{f} - \sum_{j=1}^m \lambda_j \, \nabla g_j(x^*) \;=\; 0\,; \quad \lambda_j g_j(x^*) \;=\; 0,\; j = 1, \ldots, m,$$

for some nonnegative Lagrange multipliers $\lambda \in \mathbb{R}^m_+$. Then x^* is also a global minimizer of the (convex) quadratic Lagrangian $L_f := f - \sum_{j=1}^m \lambda_j \, g_j$ on the whole $\mathbb{R}^n$. Therefore, $L_f - f^* \geq 0$ on $\mathbb{R}^n$ and being quadratic, $L_f - f^* \in \Sigma^2$. Hence

$$f \;=\; f^* + (L_f - f^*) + \sum_{j=1} \lambda_j g_j,$$

[1]Slater's condition states that there exists $x_0 \in \mathbf{K}$ such that $g_j(x_0) > 0$ for every $j = 1, \ldots, m$. If Slater's condition holds and f is convex and differentiable, then the Karush-Kuhn-Tucker (KKT) optimality conditions hold at any minimizer $x^* \in \mathbf{K}$ of the convex optimization problem: $\min_x \{f(x) \,:\, x \in \mathbf{K}\}$

that is, $f \in Q_1(g)$ (as $f^* \geq 0$). And so the PP-BDR property holds for $\mathbf{K}$ with order $r = 1$, and $\mathbf{K}$ has the SDr (2.12). Writing $g_j(x) = x^T Q_j x + b^T x + c_j$ for some positive semidefinite matrix $-Q_j \succeq 0$, $j = 1, \ldots, m$, the SDr (2.12) is nothing less than (1.2) already encountered in the introduction.

Example 3. Let $n = 2$ with g_j concave and $\deg g_j = 2$ or 4, for all $j = 1, \ldots, m$, so that $\mathbf{K}$ is convex. Assume $\mathbf{K}$ is compact with nonempty interior. It is known that in general $\mathbf{K}$ is not representable by a LMI in the variables x_1 and x_2 only;. For instance take $m = 1$ and $g_1(X) = 1 - X_1^4 - X_2^4$. The rigid convexity condition of Helton and Vinnikov [4] is violated, but on the other hand, $\mathbf{K}$ is known to be SDr.

Let $f \in \mathbb{R}[X]$ be affine and nonnegative on $\mathbf{K}$ with global minimum $f^* \geq 0$ on $\mathbf{K}$. Again, convexity along with Slater's condition implies that the KKT optimality conditions hold at any global minimizer $x^* \in \mathbf{K}$. And so there exist nonnegative Lagrange multipliers $\lambda \in \mathbb{R}^m_+$ such that the (convex) Lagrangian $L_f := f - \sum_{j=1}^m \lambda_j g_j$ also has optimal value f^* and, in addition, $x^* \in \mathbf{K}$ is a global minimizer of L_f on $\mathbb{R}^2$. Therefore, the polynomial $L_f - f^*$ being nonnegative on $\mathbb{R}^2$ and being quadratic or quartic in 2 variables, is s.o.s. That is $L_f - f^* = \sigma$ for some $\sigma \in \Sigma^2$ and $\deg \sigma \leq 4$. But then

$$f \,=\, f^* + (L_f - f^*) + \sum_{j=1}^m \lambda_j\, g_j \,\in\, Q_2(g)$$

because as $f^* \geq 0$, $f^* + (L_f - f^*) \in \Sigma^2$. That is, the PP-BDR property holds for $\mathbf{K}$ with order $r = 2$. Hence, $\mathbf{K}$ has the SDr (2.12).

2.4. **Examples with nonconvex K.**

Example 4. Let $m = 2$ with

$$g_i(X) \,=\, X^T A_i X + c_i\,, \qquad i = 1, 2, \tag{2.16}$$

for some real symmetric matrices A_i, and vector $\mathbf{c} = (c_1, c_2) \in \mathbb{R}^2$.

Given a linear polynomial $f \in \mathbb{R}[X]$ with coefficient vector $\mathbf{f} = (f_i)_{i=1}^n \in \mathbb{R}^n$, consider the SDP

$$\mathbf{Q}: \quad \min_y \{\, L_y(f) \;:\quad M_1(y) \succeq 0;\; L_y(g_i) \geq 0;\; i = 1, 2;\; y_0 = 1 \,\} \tag{2.17}$$

with optimal value denoted $\inf \mathbf{Q}$ ($\min \mathbf{Q}$ if the infimum is achieved at some y^*), and with dual

$$\mathbf{Q}^*: \quad \max_{\lambda,\gamma,\sigma} \{\, \gamma \,:\, f - \gamma = \sigma + \lambda_1 g_1 + \lambda_2 g_2;\; \lambda_1,\, \lambda_2 \geq 0;\; \sigma \in \Sigma_2^2 \,\}$$

where Σ_2^2 is the set of s.o.s. of degree 2. Let $A_\lambda := \lambda_1 A_1 + \lambda_2 A_2$ and introduce the matrix

$$H(\lambda, \gamma) := \left[\begin{array}{c|c} -\gamma - \langle \lambda, \mathbf{c} \rangle & \mathbf{f}^T/2 \\ \hline \mathbf{f}/2 & -A_\lambda \end{array} \right]. \tag{2.18}$$

Then $\mathbf{Q}^*$ has the equivalent form

$$\mathbf{Q}^* : \quad \max_{\lambda\geq 0,\gamma} \{ \gamma : \quad H(\lambda,\gamma) \succeq 0 \}, \tag{2.19}$$

with optimal value denoted $\sup \mathbf{Q}^*$ ($\max \mathbf{Q}^*$ if the sup is achieved). Obviously

$$\inf \mathbf{Q} \leq f^* := \min_x \{ f(x) : \quad x \in \mathbf{K} \},$$

and $\min \mathbf{Q} = f^*$ holds if for instance $M_1(y^*)$ is rank one at some optimal solution y^*. Indeed, in this case, $y^* = (1, x^*, (x_1^*)^2, x_1^* x_2^*, \ldots, (x_n^*)^2)$, which implies $L_{y^*}(f) = f(x^*)$ and $L_{y^*}(g_i) = g_i(x^*) \geq 0$, $i = 1, 2$.

Theorem 3. *Let $\mathbf{K} \subset \mathbb{R}^n$ be defined as in (1.3) with $m = 2$ and g_j as in (2.16), and let $\mathbf{Q}$ be as in (2.17). Assume that $\mathbf{K}$ is compact with nonempty interior and*

$$\lambda_1 A_1 + \lambda_2 A_2 \prec 0 \tag{2.20}$$

for some $\lambda = (\lambda_1, \lambda_2) \geq 0$. Then for generic $\mathbf{f} \in \mathbb{R}^n$:

(a) $\min \mathbf{Q} = f^*$

(b) *The PP-BDR property holds for* $\text{co}(\mathbf{K})$ *with order* $r = 1$, *and so* $\text{co}(\mathbf{K})$ *has the SDr (2.12), i.e.,*

$$\{ M_1(y) \succeq 0; L_y(g_j) \geq 0,\ j = 1, 2;\ L_y(X_i) = x_i,\ i = 1, \ldots, n;\ y_0 = 1 \}.$$

Proof. (a) Slater's condition holds for $\mathbf{Q}$ and $\mathbf{Q}^*$. Indeed as $\mathbf{K}$ has nonempty interior, let μ be the uniform probability measure on $\mathbf{K}$, with (well-defined) sequence of moment $y = (y_\alpha)$ (hence with $y_0 = 1$). It satisfies $M_1(y) \succ 0$ and $L_y(g_1) > 0$ as well as $L_y(g_2) > 0$.

Next, in view of (2.20), one may find $\lambda_1, \lambda_2 > 0$ and $\gamma \in \mathbb{R}$ such that $A_\lambda \prec 0$ and $H(\lambda, \gamma) \succ 0$. With $X \mapsto \sigma(X) := (1, X)^T H(\lambda, \gamma) \begin{bmatrix} 1 \\ X \end{bmatrix} \in \Sigma_2^2$, one obtains a strictly feasible solution $(\gamma, \lambda, \sigma)$ of $\mathbf{Q}^*$. As the value of both primal and dual strictly feasible solutions are finite, it follows that there is no duality gap, i.e., $\min \mathbf{Q} = \max \mathbf{Q}^*$, and both $\mathbf{Q}$ and $\mathbf{Q}^*$ are solvable.

Next, zero-duality gap yields complementarity[2] at optimal solutions y^* and $(\gamma^*, \lambda^*, \sigma^*)$ of $\mathbf{Q}$ and $\mathbf{Q}^*$, i.e., $\text{trace}(M_1(y^*) H(\lambda^*, \gamma^*)) = 0$. Therefore $H(\lambda^*, \gamma^*)$ must be singular. Notice that $H(\lambda, \gamma) \succeq 0$ implies that

$$-2 A_\lambda\, u = \mathbf{f} \tag{2.21}$$

for some $u \in \mathbb{R}^n$ and $\gamma + \lambda_1 c_1 + \lambda_2 c_2 \leq u^T A_\lambda u$.

We next prove that generically (i.e., except perhaps for a set of vectors $\{\mathbf{f}\} \subset \mathbb{R}^n$ with zero Lebesgue measure) $A_{\lambda^*} \prec 0$, and so $\text{rank}\, H(\lambda^*, \gamma^*) = n - 1$. Indeed, consider the set of $\lambda \in \mathbb{R}^2_+$ with $\lambda_1 \lambda_2 \neq 0$, such that A_λ is singular. Equivalently, after scaling by $\rho := \lambda_1 + \lambda_2 > 0$, and letting $\alpha := \lambda_1 / (\lambda_1 + \lambda_2)$, the set of $\alpha \in [0, 1]$ such that the determinant of the real symmetric matrix $B := A_2 + \alpha(A_1 - A_2)$ vanishes. Such an α must be a root in $[0, 1]$ of the characteristic polynomial of B, which has at most n

[2] See for instance Alizadeh et al. [1] or Pataki and Tuncel [15]

solutions (α_k). So A_λ is singular only on the (at most n) rays $(\lambda_1^k, \lambda_2^k) = \rho(\alpha_k, 1-\alpha_k)$, with $\rho \geq 0$ and $\alpha_k \in [0,1]$. For each α_k, the image space of $A_{\lambda^k} = \rho(\alpha_k A_1 + (1-\alpha_k)A_2)$ is at most $(n-1)$-dimensional, and (2.21) holds if and only if

$$\mathbf{f} \,=\, -2\rho(\alpha_k A_1 + (1-\alpha_k)A_2)\, u,$$

for some $u \in \mathbb{R}^n$, i.e., if and only if

$$v_j^T \mathbf{f} \,=\, 0, \quad j = 1, \ldots, p \tag{2.22}$$

where $(v_j)_{j=1}^p$ is a basis of $\mathrm{Ker}\,(\alpha_k A_1 + (1-\alpha_k)A_2)$.

If $p \geq 1$ then there is no solution in general, except perhaps on a set $\{\mathbf{f}\}_k \subset \mathbb{R}^n$ of zero Lebesgue measure. Therefore, as the set $\cup_k \{\mathbf{f}\}_k$ has zero Lebesgue measure, $A_{\lambda^*} \prec 0$ at an optimal solution $\lambda^* > 0$, for generic $\mathbf{f} \in \mathbb{R}^n$. Similar arguments are also valid if $\lambda_1 = 0$ or $\lambda_2 = 0$, as $\mathbf{f}$ must belong to the image space of A_1 or A_2.

And so, $H(\lambda^*, \gamma^*)$ has only one zero eigenvalue, which by complementarity, implies that $M_1(y^*)$ is rank-one. This in turn implies the desired result $\min \mathbf{Q} = f^*$.

(b) From $\min \mathbf{Q} = \max \mathbf{Q}^* = f^*$, for generic $\mathbf{f} \in \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^2$

$$f - f^* \,=\, \sigma^* + \lambda_1^* g_1 + \lambda_2^* g_2,$$

for some $\lambda^* \in \mathbb{R}^2_+$ and some $\sigma^* \in \Sigma^2$ of degree 2, that is, $f - f^* \in Q_1(g)$. In other words, the PP-BDR property holds for $\mathbf{K}$ with order $r = 1$, and so, $\mathrm{co}(\mathbf{K})$ has the SDr (2.12) which is the same as that of Theorem 3(b). $\square$

Figures 1, 2 and 3 respectively, display three examples of sets $\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3 \subset \mathbb{R}^2$ that have the PP-BDR property with order $r = 1$. In all cases $g_1(X) = 1 - X_1^2 - X_2^2$, and

$$\begin{array}{rcll} g_2(X) & = & (X_1-1)^2 + X_2^2 - 1 & [\text{ for } \mathbf{K}_1 \text{ }] \\ & = & 1/8 - X_1 X_2 & [\text{ for } \mathbf{K}_2 \text{ }] \\ & = & X_1 X_2 - 1/8 & [\text{ for } \mathbf{K}_3 \text{ }]. \end{array}$$

Notice that $\mathbf{K}_3$ is not even connected, and that for $\mathbf{K}_1$, one even has a linear term X_1 in the polynomial g_2.

Remark 4. Theorem 3 illustrates the fact that the PP-BDR property is specific to the representation of affine polynomials. Indeed if $f \in \mathbb{R}[X]$ is now an arbitrary quadratic polynomial $X \mapsto f(X) = X^T A_0 X + \mathbf{f}^T x + f_0$, then in general (and except in some special cases as those treated in [21]) $f - f^* \notin Q_1(g)$ even for generic data $A_0, \mathbf{f}$. See for instance some complexity results in quadratic optimization in Ye and Zhang [21].

Example 5. With $\mathbf{K}$ as in (1.3), let $\widehat{\mathbf{K}} := \mathbf{K} \cap \{-1,1\}^n$. The results in Lasserre [9, 10] show that $\widehat{\mathbf{K}}$ has the PP-BDR property with order $r = n + \max_j \lceil \deg g_j/2 \rceil$. Hence $\mathrm{co}(\widehat{\mathbf{K}})$ has the SDr (2.12) with the additional constraints $y_\alpha = y_{\alpha \bmod 2}$ for all α. In this case, the PP-BDR property is not useful for practical purposes because r depends on n, and the corresponding SDP has 2^n variables y_α.

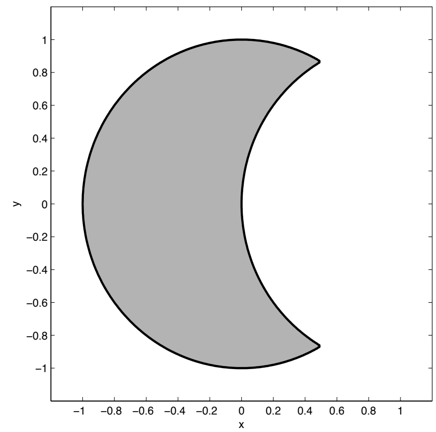

FIGURE 1. $\mathbf{K}_1 : \; g_2(X) = (X_1 - 1)^2 + X_2 - 1$

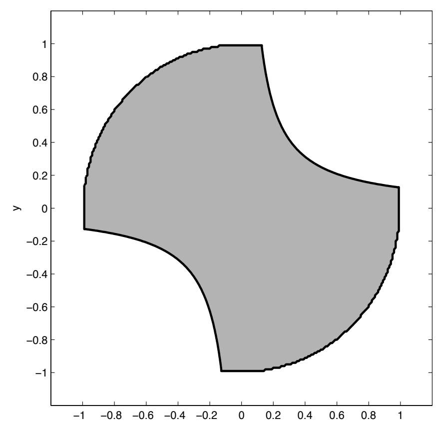

FIGURE 2. $\mathbf{K}_2 : \; g_2(X) = 1/8 - X_1 X_2$

2.5. **An approximate SDr set.** With $\mathbf{B} := \{ \, x \in \mathbb{R}^n \; : \; \|x\| \leq 1\}$ and given a compact set $\Omega \subset \mathbb{R}^n$ and $\rho > 0$, let

$$\Omega + \rho\,\mathbf{B} \;=\; \{ \, x \in \mathbb{R}^n \mid \inf_{y\in\Omega} \, \|x-y\| \leq \rho \, \}.$$

In this section we prove that given any $\epsilon > 0$, there is a convex SDr set in sandwich between $\mathrm{co}(\mathbf{K})$ and $\mathrm{co}(\mathbf{K}) + \epsilon\mathbf{B}$ and with an explicit SDr in terms of the g_j's that define $\mathbf{K}$. For this purpose we use a result of Prestel (later refined by Schweighofer [18]) on a degree bound in Schmüdgen's Positivstellensatz (and similarly a result of Nie and Schweighofer [19] on a degree bound in Putinar's Positivstellensatz).

We first need the following intermediate result.

Lemma 5. *(a) Let $\Omega \subset \mathbb{R}^n$ be a compact convex set and let $\epsilon > 0$ be fixed. If $x \notin \Omega + \epsilon\,\mathbf{B}$ then there exists a linear $f \in \mathbb{R}[X]$ whose coefficient vector*

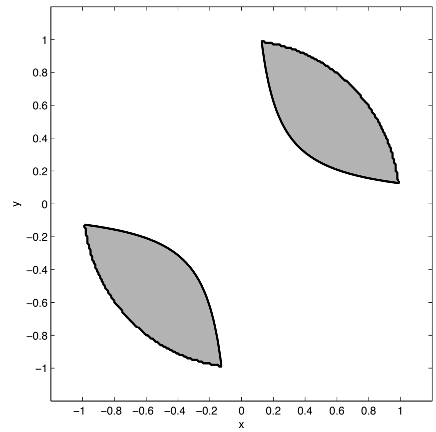

FIGURE 3. $\mathbf{K}_2 : \; g_2(X) = X_1X_2 - 1/8$

$\mathbf{f} \in \mathbb{R}^n$ *satisfies* $\|\mathbf{f}\| = 1$, *and a scalar* f^* *such that*

$$f(z) \geq f^* \quad \forall\, z \in \Omega \quad \textit{and} \quad f(x) \,<\, f^* - \epsilon. \tag{2.23}$$

In addition, $|f^*| \leq \tau_\Omega := \max\{\, \|x\| \,:\, x \in \Omega\}$.

(b) For any compact set $\mathbf{K} \subset \mathbb{R}^n$, *and any* $\mathbf{f} \in \mathbb{R}^n$ *with* $\|\mathbf{f}\| = 1$, *let* $f^* := \min_{x\in\mathbf{K}} \mathbf{f}^T x$, *and let* $\tau_\mathbf{K} := \max\{\|x\| \,:\, x \in \mathbf{K}\}$. *Then*

$$\min_{x\in\mathrm{co}(\mathbf{K})} \mathbf{f}^T x \,=\, f^* \quad \textit{and} \quad |f^*| \,\leq\, \tau_\mathbf{K}. \tag{2.24}$$

Proof. (a) With $x \notin \Omega + \epsilon\mathbf{B}$, let $x^* \in \Omega$ be its projection on Ω (well defined because Ω is compact and convex). Let $\mathbf{f} := (x^* - x)/\|x - x^*\|$ so that $\|\mathbf{f}\| = 1$, and let $f^* := \mathbf{f}^T x^*$, so that $|f^*| \leq \|\mathbf{f}\| \max\{\|x\| \,:\, x \in \Omega\} = \tau_\Omega$.

Then with $f \in \mathbb{R}[X]$ being the linear polynomial with coefficient vector $\mathbf{f}$, one has $f(z) \geq f^*$ for all $z \in \Omega$ because

$$f(z) \,=\, \mathbf{f}^T z \,=\, \mathbf{f}^T x^* + \mathbf{f}^T(z - x^*) \,=\, f^* + \langle \vec{xx^*}, \vec{x^*z}\rangle \geq f^*$$

(since $\langle \vec{xx^*}, \vec{x^*z}\rangle \geq 0$), and

$$f(x) - f^* \,=\, \mathbf{f}^T(x - x^*) \,=\, -\|x - x^*\| < -\,\epsilon.$$

(b) Indeed, $f^* = \min\limits_{x\in\mathbf{K}} \mathbf{f}^T x = \min\limits_{x\in\mathrm{co}(\mathbf{K})} \mathbf{f}^T x$. Moreover, $|\mathbf{f}^T x| \leq \|\mathbf{f}\|\cdot\|x\| \leq \tau_\mathbf{K}$ for all $x \in \mathbf{K}$. $\square$

Then we have the following result.

Theorem 6. *Let* $\mathbf{K} \subset \mathbb{R}^n$ *be a compact set as defined in (1.3).*

(a) *For every fixed $\epsilon > 0$ there is a integer $r_\epsilon \in \mathbb{N}$ such that the SDr set $\mathbf{K}_\epsilon$ defined by*

$$\mathbf{K}_\epsilon := \left\{ x \in \mathbb{R}^n \;:\; \left\{ \begin{array}{lll} \exists\, y \in \mathbb{R}^{s(2r_\epsilon)} : & & \\ M_{r_\epsilon - r_J}(g_J\, y) & \succeq & 0, \quad J \subseteq \{1,\ldots,m\} \\ L_y(X_i) & = & x_i, \quad i = 1,\ldots,n \\ y_0 & = & 1 \end{array} \right. \right\} \tag{2.25}$$

satisfies $\mathrm{co}(\mathbf{K}) \subseteq \mathbf{K}_\epsilon \subset \mathbf{K} + \epsilon\mathbf{B}$.

(b) *Assume that the polynomial $N - \|X\|^2$ is in the quadratic module $Q(g)$. Then for every fixed $\epsilon > 0$ there is an integer $r_\epsilon \in \mathbb{N}$ such that the SDr set $\mathbf{K}_\epsilon$ defined by*

$$\mathbf{K}_\epsilon := \left\{ x \in \mathbb{R}^n \;:\; \left\{ \begin{array}{lll} \exists\, y \in \mathbb{R}^{s(2r_\epsilon)} : & & \\ M_{r_\epsilon - r_j}(g_j\, y) & \succeq & 0, \quad j = 0,\ldots,m \\ L_y(X_i) & = & x_i, \quad i = 1,\ldots,n \\ y_0 & = & 1 \end{array} \right. \right\} \tag{2.26}$$

satisfies $\mathrm{co}(\mathbf{K}) \subseteq \mathbf{K}_\epsilon \subset \mathbf{K} + \epsilon\mathbf{B}$.

In both cases (a) and (b), bounds on r_ϵ are available.

Proof. (a) That $\mathrm{co}(\mathbf{K}) \subseteq \mathbf{K}_\epsilon$ is straightforward and as in the proof of Theorem 2. Next, let $x \not\in \mathrm{co}(\mathbf{K}) + \epsilon\mathbf{B}$ be fixed. Then by Lemma 5 (with $\Omega := \mathrm{co}(\mathbf{K})$) there exists $\mathbf{f} \in \mathbb{R}^n$ and $f^* := \min_{x \in \mathrm{co}(\mathbf{K})} \mathbf{f}^T x$ such that (2.23) holds. In addition, $\|\mathbf{f}\| = 1$ and $|f^*| \leq \tau_{\mathbf{K}}$.

Let $f \in \mathbb{R}[X]$ be the affine polynomial with coefficient vector $(\mathbf{f}, -f^*) \in \mathbb{R}^n \times \mathbb{R}$ so that $f + \epsilon \geq \epsilon > 0$ on $\mathbf{K}$. By Schmüdgen Positivstellensatz [17], $f + \epsilon \in P(g)$. Even more, $f + \epsilon \in P_{r_\epsilon}(g)$ for some integer $r_\epsilon \in \mathbb{N}$ that does *not* depend on the precise value of f but *only* on its degree (here 1) and norm (here $\|\mathbf{f}\| = 1$ and $|f^*| \leq \tau_{\mathbf{K}}$); see Schweighofer [18]. So let $\mathbf{K}_\epsilon$ be the SDr set defined in (2.25) with this r_ϵ. If $x \in \mathbf{K}_\epsilon$, we obtain the contradiction

$$\begin{aligned} 0 \;>\; & \mathbf{f}^T x - f^* + \epsilon \\ & = (\epsilon - f^*) y_0 + \sum_{i=1}^n L_y(f_i X_i) \quad [\text{as } y_0 = 1 \text{ and } L_y(X_i) = x_i \quad \forall i] \\ & = L_y(f(X) + \epsilon) \;=\; \sum_{J \subseteq \{1,\ldots,m\}} L_y(g_J\, \sigma_J) \quad [\text{ as } f + \epsilon \in P_{r_\epsilon}(g) \,] \\ & \geq 0 \quad [\text{ by (2.15) }]. \end{aligned}$$

Hence $\mathbf{K}_\epsilon \subset \mathrm{co}(\mathbf{K}) + \epsilon\mathbf{B}$, the desired result.

(b) The proof is very similar except that now we invoke Putinar Positivstellenstaz [16] and Nie and Schweighofer [19] to replace $P_{r_\epsilon}(g)$ with $Q_{r_\epsilon}(g)$. Finally, bounds on r_ϵ can be found for both cases (a) and (b) in [18] and [19] respectively. □

Hence, no matter if $\mathrm{co}(\mathbf{K})$ is SDr, for every $\epsilon > 0$, there is always a SDr set $\mathbf{K}_\epsilon$ in sandwich between $\mathrm{co}(\mathbf{K})$ and $\mathrm{co}(\mathbf{K}) + \epsilon\mathbf{B}$. In addition, the SDr of $\mathbf{K}_\epsilon$ is explicit in terms of the polynomials (g_j) that define $\mathbf{K}$. This is a significant improvement upon the *outer* convex approximations $\Delta_r \downarrow \mathrm{co}(\mathbf{K})$ of [8], where each Δ_r has a SDr. Indeed in [8], if $x \notin \mathrm{co}(\mathbf{K})$ then $x \notin \Delta_r$ for all $r \geq r(x)$ for some $r(x)$ that *depends* on x, an undesirable feature.

3. II. SDr for Compact convex basic semialgebraic sets

In this section, $\mathbf{K} \subset \mathbb{R}^n$ defined in (1.3) is compact and convex, and we assume that one knows a scalar $\tau_{\mathbf{K}}$ such that:

$$x \in \mathbf{K} \quad \Rightarrow \quad \|x\| \leq \tau_{\mathbf{K}}. \tag{3.1}$$

Lemma 7. *Let $\mathbf{K} \subset \mathbb{R}^n$ be as in (1.3), and assume that the g_j's that define $\mathbf{K}$ are all concave and Slater's condition holds. Given $f \in \mathbb{R}[X]$, let $f^* := \min_{x \in \mathbf{K}} f(x)$.*

For every linear $f \in \mathbb{R}[X]$ with $\|\mathbf{f}\| = 1$, there exists $\lambda(f) \in \mathbb{R}^m_+$ such that

$$X \mapsto L_f(X) \quad := \quad f(X) - f^* - \sum_{j=1}^m \lambda_j(f)\, g_j(X) \;\geq\; 0 \quad \text{on } \mathbb{R}^n \tag{3.2}$$

$$|f^*| \quad \leq \quad \tau_{\mathbf{K}}; \quad \lambda_j(f) \;\leq\; M_{\mathbf{K}}, \quad j = 1, \ldots, m, \tag{3.3}$$

where $M_{\mathbf{K}}$ is independent of f.

Proof. As the g_j's are concave, $\mathbf{K}$ is compact and convex. In addition, as Slater's condition holds and f is convex, there exist nonnegative Lagrange multipliers $\lambda(f) \in \mathbb{R}^m_+$ such that

$$\nabla f(x^*) = \sum_{j=1}^m \lambda_j(f) \nabla g_j(x^*); \quad \lambda_j(f) g_j(x^*) \;=\; 0, \; \forall\, j = 1, \ldots, m,$$

where $x^* \in \mathbf{K}$ is a (global) minimizer of f on $\mathbf{K}$. Therefore the Lagrangian L_f defined in (3.2) is convex, with f^* as its global minimum on $\mathbb{R}^n$ and x^* as global minimizer. Recall that Slater's condition states that $g_j(x_0) > 0$, $j = 1, \ldots, m$, for some x_0. And so, from

$$L_f(x_0) \;=\; f(x_0) - f^* - \sum_{j=1}^m \lambda_j(f)\, g_j(x_0) \;\geq\; 0,$$

we deduce that for every $j = 1, \ldots, m$,

$$0 \;\leq\; \lambda_j(f) \;\leq\; \frac{f(x_0) - f^*}{g_j(x_0)} \;\leq\; \frac{2\tau_{\mathbf{K}}}{g_j(x_0)} \;\leq\; \frac{2\tau_{\mathbf{K}}}{\displaystyle\min_{j=1,\ldots,m} g_j(x_0)} \;=:\; M_{\mathbf{K}},$$

where we have used $\|\mathbf{f}\| = 1$. Therefore (3.3) holds and $M_{\mathbf{K}}$ above is independent of f. $\square$

Theorem 8. *Let $\mathbf{K} \subset \mathbb{R}^n$ be compact and defined as in (1.3). Assume that the g_j's that define $\mathbf{K}$ are all concave and Slater's condition holds. Given a linear polynomial $f \in \mathbb{R}[X]$, let L_f be the Lagrangian defined in (3.2).*

If L_f is s.o.s. for every linear $f \in \mathbb{R}[X]$, then the PP-BDR property holds for $\mathbf{K}$ with order $r = \max_{j=1,\ldots,m} \lceil \deg g_j/2 \rceil$, and $\mathbf{K}$ is a SDr set. In addition, the convex set

$$(3.4) \quad \mathbf{\Omega} := \left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^{s(2r)} \; : \; \left[\begin{array}{lll} M_r(y) & \succeq 0 & \\ L_y(g_j) & \geq 0, & j = 1, \ldots, m \\ L_y(X_i) & = x_i, & i = 1, \ldots, n \\ y_0 & = 1 & \end{array} \right. \right\}$$

is a SDr of $\mathbf{K}$.

Proof. Let $x \in \mathbf{K}$ and let $y = (x^\alpha) \in \mathbb{R}^{s(r)}$. Then $M_r(y) \succeq 0$ and $L_y(g_j) = g_j(x) \geq 0$ for all $j = 1, \ldots, m$. Therefore, $(x, y) \in \mathbf{\Omega}$.

Conversely, let $x \notin \mathbf{K}$, and suppose that $(x, y) \in \mathbf{\Omega}$ for some $y \in \mathbb{R}^{s(r)}$. As $x \notin \mathbf{K}$ there exists $(f^*, \mathbf{f}) \in \mathbb{R} \times \mathbb{R}^n$ with $\|\mathbf{f}\| = 1$ such that $\mathbf{f}^T z \geq f^*$ for all $z \in \mathbf{K}$ and $\mathbf{f}^T x < f^*$. Actually, $f^* = \min_{x \in \mathbf{K}} f(x)$ where $f \in \mathbb{R}[X]$ is linear with vector of coefficients $\mathbf{f}$. Let L_f be as in (3.2). If L_f is s.o.s. then $f - f^* = \sigma + \sum_{j=1}^m \lambda_j(f)\, g_j$ for some s.o.s. polynomial $\sigma \in \mathbb{R}[X]$ of degree at most $2r$. Therefore, one obtains the contradiction

$$\begin{aligned} 0 \; &> \; \mathbf{f}^T x - f^* \; = \; f(x) - f^* \\ &\geq \; L_y(f - f^*) = L_y(\sigma + \sum_{j=1}^m \lambda_j g_j) \\ &\geq \; 0 \qquad [\text{ as } (x, y) \in \mathbf{\Omega} \text{ }]. \end{aligned}$$

□

Remark 9. Interestingly, Theorem 8 has a rephrasing in terms of the support function $\mathbf{f} \mapsto \sigma_{\mathbf{K}}(\mathbf{f})$ of $\mathbf{K}$, defined by:

$$\mathbf{f} \mapsto \sigma_{\mathbf{K}}(\mathbf{f}) \; := \; \sup \{ \, \langle \mathbf{f}, x \rangle \; : \quad x \in \mathbf{K} \, \}.$$

For more details on the support function and its properties, the interested reader is referred to e.g. Hiriart-Urruty and Lemarechal [6, Chapter V].

For every linear polynomial $f \in \mathbb{R}[X]$, let $\mathbf{f} \in \mathbb{R}^n$ be its vector of coefficients. Then observe that in Theorem 8, and with $r = \max_{j=1,\ldots,m} \lceil \deg g_j/2 \rceil$, the statement "*$L_f$ is s.o.s. for every linear $f \in \mathbb{R}[X]$*" can be replaced with the new statement "*$f + \sigma_{\mathbf{K}}(-\mathbf{f}) \in Q_r(g)$ for every linear $f \in \mathbb{R}[X]$*".

The SDr (3.4) of $\mathbf{K}$ is very natural as it is based on the Karush-Kuhn-Tucker optimality conditions. Existence of such a SDr reduces to the real algebraic problem of checking whether the Lagrangian L_f is s.o.s. for every (in fact, almost all) *linear* $f \in \mathbb{R}[X]$. Examples 2 and 3 in §2.3 provide such instances of sets $\mathbf{K}$ with the PP-BDR property and with SDr (3.4).

Hence, an important issue to find sufficient conditions to ensure that the Lagrangian L_f is s.o.s., and if possible, conditions that can be checked directly from the data g_j. For instance, in Lasserre [12] one finds sets of sufficient conditions on the coefficients of a polynomial f to ensure it is s.o.s. Also, after the present paper was written, Helton and Nie [5] have provided several sufficient conditions for the Lagrangian L_f to be s.o.s. In particular, if the Hessian $-\nabla^2 g_j(X)$ can be written $P_j(X)P_j(X)^T$ for some (not necessarily square) matrix $P_j(X)$ (i.e. $-\nabla^2 g_j(X)$ is a sum of squares), $j = 1,\dots,m$, then L_f is s.o.s.

Example 6. Consider the class of convex sets $\mathbf{K} \subset \mathbb{R}^n$ with $g_j \in \mathbb{R}[X]$ concave and of the form

$$g_j(X) = -\sum_{i=1}^n g_{ji}\, X_i^{2d} + h_j(X) + g_{j0}, \quad j = 1,\dots,m, \tag{3.5}$$

with $(g_{ji}) \subset \mathbb{R}_+$, and $h_j \in \mathbb{R}[X]$ linear for every $j = 1,\dots,m$. Then $-\nabla^2 g_j(X)$ is the diagonal matrix with diagonal elements $(g_{ii}X_i^{2d-2})$ and so can be written as $P_j(X)P_j(X)^T$ with $P_j(X) = (\nabla^2 g_j(X))^{1/2}$. Therefore, by Theorem 8, $\mathbf{K}$ has the SDr (3.4).

Taking $\mathbf{K} := \{x \in \mathbb{R}^n \ : \ \|x\|_d \,(:= (\sum_{i=1}^n x_i^{2d})^{1/2d}) \ \leq 1\}$ as a particular case of Example 6, one may thus conclude that the d-Euclidean ball is SDr, for all $d \geq 1$.

Approximate SDr. When one does not know whether the Lagrangian L_f is s.o.s. for all (in fact, almost all) linear $f \in \mathbb{R}[X]$, we next provide an approximation result. Namely we provide a semidefinite representation $\mathbf{\Omega_r}$ for an arbitrarily close convex approximation $\mathbf{K}_r$ of $\mathbf{K}$. This approximation is in the spirit of that of §2.5, but specific to the convex case. We first need the following crucial auxiliary results.

Lemma 10. *Let $\mathbf{K} \subset \mathbb{R}^n$ be as in (1.3), $\tau_{\mathbf{K}}$ as in (3.1), and assume that the g_j's that define $\mathbf{K}$ are all concave and Slater's condition holds. Let $X \mapsto \Theta_r(X) := \sum_{i=1}^n \left(\frac{X_i}{\tau_{\mathbf{K}}}\right)^{2r}$.*

Then for every $\epsilon > 0$ there exists $r(\epsilon)$ such that for every linear $f \in \mathbb{R}[X]$ with $\|\mathbf{f}\| = 1$ and L_f as in (3.2),

$$L_f + \epsilon\,(1 + \Theta_r) \quad \textit{is s.o.s.} \quad \forall\, r \geq r(\epsilon). \tag{3.6}$$

Equivalently, $f - f^* + \epsilon\,(1 + \Theta_r) \in Q_r(g)$.

Proof. By Lemma 7, $L_f \geq 0$ and observe that the coefficients of the polynomial L_f are all uniformly bounded in $\mathbf{f}$ whenever $\|\mathbf{f}\| = 1$. Indeed,

$$0 \leq \lambda_j(f) \leq M_{\mathbf{K}} \quad \forall\, j = 1,\dots,m; \qquad |f^*| \leq \tau_{\mathbf{K}},$$

with $\tau_{\mathbf{K}}$ as in (3.1) and $M_{\mathbf{K}}$ as in Lemma 7. Hence, in view of the definition (3.2) of the polynomial L_f, its coefficients $(L_f)_\alpha$ are all bounded, uniformly in $\mathbf{f}$.

Next, $L_f \geq 0$ implies that L_f is nonnegative on the box $[-\tau_{\mathbf{K}}, \tau_{\mathbf{K}}]^n$. Therefore (3.6) follows from Lasserre and Netzer [11, §3.3], where it was proved that the degree $r(\epsilon)$ does *not* depend on the precise value of the coefficients of L_f but only on ϵ, the dimension n, the degree of L_f and the *size* of its coefficients. Here, whenever $\mathbf{f}$ varies, the degree of L_f takes *finitely* many values (depending on which Lagrange multipliers λ_j are zero), and its coefficients are uniformly bounded. $\square$

Next, in view of (3.1) and with $\Theta_r \in \mathbb{R}[X]$ as in Lemma 10,

$$\Theta_r(x) \,\leq\, 1 \qquad \forall\, x \in \mathbf{K}, \quad \forall\, r \in \mathbb{N}. \tag{3.7}$$

Theorem 11. *Let $\mathbf{K} \subset \mathbb{R}^n$ as in (1.3) be compact, with $\tau_{\mathbf{K}}$ as in (3.1). Assume that the g_j's that define $\mathbf{K}$ are all concave and Slater's condition holds. With $r \in \mathbb{N}$, $r \geq \lceil \deg g_j/2 \rceil$, $j = 1, \ldots, m$, let $\Theta_r(X) = \sum_{i=1}^n \left(\frac{X_i}{\tau_{\mathbf{K}}}\right)^{2r}$, and let $\mathbf{K}_r \subset \mathbb{R}^n$ be the convex set:*

$$\mathbf{K}_r \,:=\, \left\{ x \in \mathbb{R}^n \,:\, \left[\begin{array}{lll} \exists\, y \in \mathbb{R}^{s(2r)} & \text{s.t.} & \\ M_r(y) & \succeq 0 & \\ L_y(g_j) & \geq 0, & j = 1, \ldots, m \\ L_y(\Theta_r) & \leq 1 & \\ L_y(X_i) & = x_i, & i = 1, \ldots, n \\ y_0 & = 1 & \end{array}\right. \right\}. \tag{3.8}$$

Then for every $\epsilon > 0$, there exists $r \in \mathbb{N}$ such that

$$\mathbf{K} \,\subseteq\, \mathbf{K_r} \,\subseteq\, \mathbf{K} + \epsilon\,\mathbf{B}, \tag{3.9}$$

and the convex set

$$\mathbf{\Omega_r} \,:=\, \left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^{s(2r)} \,:\, \left[\begin{array}{lll} M_r(y) & \succeq 0 & \\ L_y(g_j) & \geq 0, & j = 1, \ldots, m \\ L_y(\Theta_r) & \leq 1 & \\ L_y(X_i) & = x_i, & i = 1, \ldots, n \\ y_0 & = 1 & \end{array}\right. \right\} \tag{3.10}$$

is a SDr of $\mathbf{K}_r$.

Proof. Let $x \in \mathbf{K}$. Then the vector $y = (x^\alpha) \in \mathbb{R}^{s(r)}$ satisfies the constraints described in (3.8), so that $\mathbf{K} \subseteq \mathbf{K}_r$ for all $r \geq \lceil \deg g_j/2 \rceil$, $j = 1, \ldots, m$.

To prove $\mathbf{K}_r \subseteq \mathbf{K} + \epsilon\mathbf{B}$, we proceed by contradiction. With $\epsilon > 0$ fixed, let $x \not\in \mathbf{K} + \epsilon\,\mathbf{B}$ be fixed but arbitrary, and with $r(\epsilon/2)$ as in Lemma 10, let $r \geq r(\epsilon/2)$ be fixed arbitrary. Let $f \in \mathbb{R}[X]$ be as in Lemma 5 so that $f(x) - f^* < -\,\epsilon$. Next, with L_f being the Lagrangian associated with f, by Lemma 10,

$$L_f + \frac{\epsilon}{2}\,(1 + \Theta_r) \,=\, \sigma, \tag{3.11}$$

for some s.o.s. polynomial $\sigma \in \mathbb{R}[X]$ of degree $2r$. Equivalently,

$$f - f^* + \frac{\epsilon}{2}(1 + \Theta_r) \,=\, \sigma + \sum_{j=1}^m \lambda_j(f)\, g_j.$$

Now, suppose that $x \in \mathbf{K}_r$. There exists $y \in \mathbb{R}^{s(r)}$ such that $(x, y) \in \mathbf{\Omega_r}$. In particular, $L_y(\Theta_r) \leq 1$, $L_y(g_j) \geq 0$, $j = 1, \ldots, m$, and $L_y(\sigma) \geq 0$ for every $\sigma \in \Sigma_r^2$ (because $M_r(y) \succeq 0$). And so, we obtain the contradiction

$$\begin{aligned} 0 > \mathbf{f}^T x - f^* + \epsilon &= f(x) - f^* + \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &\geq L_y(f - f^*) + \frac{\epsilon}{2} L_y(1 + \Theta_r) \\ &= L_y(\sigma) + \sum_{j=1}^{m} \lambda_j(f) L_y(g_j) \qquad [\text{ by } (3.11)\ ] \\ &\geq 0 \qquad [\text{ as } (x, y) \in \mathbf{\Omega_r}\ ]. \end{aligned}$$

Therefore $x \notin \mathbf{K}_r$. As $x \notin \mathbf{K} + \epsilon\mathbf{B}$ was arbitrary, this implies $\mathbf{K}_r \subseteq \mathbf{K} + \epsilon\,\mathbf{B}$. Finally, that $\mathbf{\Omega_r}$ in (3.10) is a SDr of $\mathbf{K}_r$, follows from the definition (3.8) of $\mathbf{K}_r$. □

The SDr $\mathbf{\Omega_r}$ of the convex set $\mathbf{K}_r$ in Theorem 11 resembles the SDr $\mathbf{\Omega}$ of $\mathbf{K}$ in Theorem 8. The only difference is the index r which is larger than $\max_j \lceil \deg g_j/2 \rceil$, and the additional constraint $L_y(\Theta_r) \leq 1$.

Hence, it is worth noticing that if $\mathbf{K}$ does not admit the SDr $\mathbf{\Omega}$ of Theorem 8, one still obtains a SDr $\mathbf{\Omega_r}$ of an *arbitrarily close* convex approximation $\mathbf{K}_r$ of $\mathbf{K}$, explicit in terms of the concave polynomials (g_j) that define $\mathbf{K}$.

4. Conclusion

We have considered the class of compact basic semialgebraic sets $\mathbf{K} \subset \mathbb{R}^n$, and have provided sufficient conditions for its convex hull $\mathrm{co}(\mathbf{K})$ to have a SDr expressed directly in terms of the polynomials that define $\mathbf{K}$. When $\mathbf{K}$ is convex and defined by concave polynomials, we have shown that if for every linear polynomial $f \in \mathbb{R}[X]$, the associated (nonnegative) Lagrangian L_f is s.o.s., then $\mathbf{K}$ has a simpler specific SDr. Finally, we have also provided a SDr of an arbitrarily close approximation $\mathbf{K}_\epsilon$ of $\mathrm{co}(\mathbf{K})$ (and of $\mathbf{K}$ in the convex case). An interesting issue of further investigation is to provide concrete conditions on the concave polynomials g_j's, to ensure that the Lagrangian L_f is s.o.s. The work in [5] provides some interesting results in this direction.

Acknowledgement: This work was supported by french ANR-grant NT05$-$3 $-$ 41612.

References

[1] F. Alizadeh, J-P. Haeberly, M. Overton. Complementarity and nondegeneracy in semidefinite programming, Math. Programming **77** (1997), 111–128.
[2] A. Ben-Tal, A. Nemirovski. *Lectures on Modern Convex Optimization*, SIAM, Philadelphia, 2001.
[3] C. Beng Chua, L. Tuncel. Invariance and efficiency of convex representations, Math. Programming **111** (2008), 113–140.

[4] J.W. Helton, V. Vinnikov. Linear matrix inequality representation of sets, Comm. Pure Appl. Math., to appear. arXiv:math.OC/0306180.

[5] J.W. Helton, J. Nie. Semidefinite representation of convex sets, Technical report, Mathematics Dept., University of California at San Diego, USA, 2007. arXiv:0705.4068.

[6] J-B. Hiriart-Urruty, C. Lemarechal. *Convex Analysis and Minimization Algorithms I*, Springer-Verlag, Berlin, 1993.

[7] T. Jacobi, A. Prestel. Distinguished representations of strictly positive polynomials, J. Reine. Angew. Math. **532** (2001), 223–235.

[8] R. Laraki, J.B. Lasserre. Computing uniform convex approximations for convex envelopes and convex hulls, J. of Convex Analysis, to appear.

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

[10] J.B. Lasserre. An explicit equivalent positive semidefinite program for nonlinear 0-1 programs, SIAM J. Optim. **12** (2002), 756–769.

[11] J.B. Lasserre, T. Netzer. SOS approximation of nonnegative polynomial via simple high degree perturbations, Math. Z. **256** (2006), 99–112.

[12] J.B. Lasserre. Conditions for a real polynomial to be sum of squares, Archiv der Mathematik, to appear. doi: 10.1007/s00013-007-2251-y Available at http://arxiv.org/abs/math.AG/0612358

[13] A.S. Lewis, P. Parrilo, M.V. Ramana. The Lax conjecture is true, Proc. Am. Math. Soc., **133** (2005), 2495-2499.

[14] P. Parrilo, Exact semidefinite representations for genus zero curves, Talk at the Banff workshop "Positive Polynomials and Optimization", Banff, Canada, October 8-12th 2006.

[15] G. Pataki, L. Tuncel. On the generic properties of convex optimization problems in conic form, Math. Programming **89** (2001), 449–457.

[16] M. Putinar. Positive polynomials on compact semi-algebraic sets, Indiana Univ. Math. J. **42** (1993), 969–984.

[17] K. Schmüdgen. The $\mathbb{K}$-moment problem for compact semi-algebraic sets, Math. Ann. **289** (1991), 203–206.

[18] M. Schweighofer. On the complexity of Schmdgen's Positivstellensatz, J. Complexity **20** (2004), pp. 529-543.

[19] Jiawang Nie, M. Schweighofer. On the complexity of Putinar's Positivstellensatz, J. Complexity **23** (2007), pp. 135–150.

[20] L. Vandenberghe, S. Boyd. Semidefinite programming, SIAM Review **38** (1996), pp. 49-95.

[21] Y. Ye, S. Zhang, New results on quadratic minimization, SIAM J. Optim. **14** (2003), 245–267.

LAAS-CNRS and Institute of Mathematics, LAAS, 7 avenue du Colonel Roche, 31077 Toulouse cedex 4, France

E-mail address: lasserre@laas.fr