The K-moment problem for continuous linear functionals

Jean-Bernard Lasserre

THE K-MOMENT PROBLEM FOR CONTINUOUS LINEAR FUNCTIONALS

JEAN B. LASSERRE

Abstract. Given a closed (and non necessarily compact) basic semi-algebraic set $\mathbf{K} \subseteq \mathbb{R}^n$, we solve the $\mathbf{K}$-moment problem for continuous linear functionals. Namely, we introduce a weighted ℓ_1-norm $\ell_{\mathbf{w}}$ on $\mathbb{R}[\mathbf{x}]$, and show that the $\ell_{\mathbf{w}}$-closures of the preordering P and quadratic module Q (associated with the generators of $\mathbf{K}$) is the cone $\mathrm{Psd}(\mathbf{K})$ of polynomials nonnegative on $\mathbf{K}$. We also prove that P and Q solve the $\mathbf{K}$-moment problem for $\ell_{\mathbf{w}}$-continuous linear functionals and completely characterize those $\ell_{\mathbf{w}}$-continuous linear functionals nonnegative on P and Q (hence on $\mathrm{Psd}(\mathbf{K})$). When $\mathbf{K}$ has a nonempty interior we also provide in explicit form a canonical $\ell_{\mathbf{w}}$-projection $g_f^{\mathbf{w}}$ for any polynomial f, on the (degree-truncated) preordering or quadratic module. Remarkably, the support of $g_f^{\mathbf{w}}$ is very sparse and does not depend on $\mathbf{K}$! This enables us to provide an explicit Positivstellensatz on $\mathbf{K}$. At last but not least, we provide a simple characterization of polynomials nonnegative on $\mathbf{K}$, which is crucial in proving the above results.

1. Introduction

This paper is concerned with basic closed semi-algebraic sets $\mathbf{K} \subseteq \mathbb{R}^n$ and the *preordering* $P(g)$ and *quadratic module* $Q(g)$ associated with the finite family of polynomials (g_j), $j \in J$, that generate $\mathbf{K}$. In particular, when $\mathbf{K} = \mathbb{R}^n$ then the latter two convex cones coincide with the cone $\Sigma[\mathbf{x}]$ of sums of squares (s.o.s.) of polynomials. The convex cones $P(g)$ and $Q(g)$ (which are subcones of the convex cone $\mathrm{Psd}(\mathbf{K})$ of polynomials nonnegative on $\mathbf{K}$) are of practical importance because on the one hand nonnegative polynomials are ubiquitous but on the other hand, polynomials in $P(g)$ or $Q(g)$ are much easier to handle. For instance, and in contrast with nonnegative polynomials, checking whether a given polynomial is in $P_d(g)\,(\subset P(g))$ or $Q_d(g)\,(\subset Q(g))$ (i.e., with an *a priori* degree bound d on its representation) can be done efficiently by solving a semidefinite program, a powerful technique of convex optimization.

The celebrated $\mathbf{K}$-moment problem is concerned with characterizing all real sequences $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, which can be realized as the moment sequence of some finite Borel measure on $\mathbf{K}$. For such sequences $\mathbf{y}$, the

1991 *Mathematics Subject Classification.* 44A60 13B25 14P10 30C10.

Key words and phrases. Moment problems; real algebraic geometry; positive polynomials; semi-algebraic sets.

linear form $L_{\mathbf{y}}$ associated with $\mathbf{y}$ is called a $\mathbf{K}$-*moment functional* and Haviland's theorem states that $L_{\mathbf{y}}$ is a $\mathbf{K}$-moment functional if and only if $L_{\mathbf{y}}(f) \geq 0$ for all $f \in \mathrm{Psd}(\mathbf{K})$. And so one says that $P(g)$ (resp. $Q(g)$) solves the $\mathbf{K}$-moment problem if the $\mathbf{K}$-moment functionals are those which satisfy $L_{\mathbf{y}}(f) \geq 0$ on $P(g)$ (resp. $Q(g)$). However, this is true if and only if $P(g)$ (resp. $Q(g)$) is dense in $\mathrm{Psd}(\mathbf{K})$ for the finest locally convex topology on $\mathbb{R}[\mathbf{x}]$. When $\mathbf{K}$ is compact, the $\mathbf{K}$-moment problem was completely solved (with $\overline{P(g)} = \mathrm{Psd}(\mathbf{K})$) in Schmüdgen [15] and soon refined to $\overline{Q(g)} = \mathrm{Psd}(\mathbf{K})$ in Putinar [13] for Archimedean quadratic modules $Q(g)$.

Since then, recent contributions have tried to better understand (in even a more general framework) the links between $\mathrm{Psd}(\mathbf{K})$ and closures (and sequential closures) of preorderings and quadratic modules, one important goal being to extend (or provide analogues of) Schmüdgen and Putinar's Positivstellensatzë [15, 13] to cases where $\mathbf{K}$ is not compact. In particular, Scheiderer [14] has shown rather negative results in this direction. For more details on those recent results, the interested is referred to e.g. Powers and Scheiderer [12], Kuhlmann and Marshall [8], Kuhlmann et al. [9], and Cimpric et al. [5].

On the one hand, *all* linear functionals are continuous in the finest locally convex topology, on the other hand, the negative results of Scheiderer [14] suggest that solving the $\mathbf{K}$-moment problem via preorderings or quadratic modules is possible only in specific cases, and so this topology is not the most appropriate in general. So why not rather consider other topologies on $\mathbb{R}[\mathbf{x}]$ and the $\mathbf{K}$-moment problem for linear functionals $L_{\mathbf{y}}$ that are *continuous* for these topologies? This is the point of view taken in Ghasemi et al. [6] where the authors consider certain (weighted) norm-topologies and show that the closure of the cone of sums of squares is $\mathrm{Psd}([-1,1]^n)$ as did Berg [3] for the ℓ_1-norm. Notice that this was also the point of view taken in Schmüdgen [16] for a class of non commutative (enveloping) $*$-algebra; in the latter context [16], the author proves that the cone of *strongly positive* elements is the closure of the (smallest) cone of sums of squares in certain topologies.

Contribution. In view of the negative results in Scheiderer [14], we also consider the above mentioned viewpoint and look at the $\mathbf{K}$-moment problem by using a particular weighted ℓ_1-norm on $\mathbb{R}[\mathbf{x}]$ (denoted $\ell_{\mathbf{w}}$ for a certain sequence $\mathbf{w} : \mathbb{N}^n \to \mathbb{R}_{>0}$) rather than the usual finest locally convex topology. In this framework we solve the $\mathbf{K}$-moment problem for basic closed semi-algebraic sets $\mathbf{K} \subseteq \mathbb{R}^n$ in the following sense. We prove that (a) the $\ell_{\mathbf{w}}$-closure of $P(g)$ and $Q(g)$ is exactly $\mathrm{Psd}(\mathbf{K})$, and (b) $P(g)$ and $Q(g)$ solve the moment problem, i.e., the $\mathbf{K}$-moment ($\ell_{\mathbf{w}}$-continuous) functionals are those $L_{\mathbf{y}}$ that are $\ell_{\mathbf{w}}$-continuous and nonnegative on $Q(g)$ (or $P(g)$). In fact, such linear functionals $L_{\mathbf{y}}$ are characterized by:

- $L_{\mathbf{y}}(h^2 g_j) \geq 0$ for all $h \in \mathbb{R}[\mathbf{x}]$, and every generator g_j of $\mathbf{K}$.

- $\exists M > 0$ such that $|y_\alpha| \leq M\, w_\alpha$ for all $\alpha \in \mathbb{N}^n$.

• Next, when $\mathbf{K}$ has a nonempty interior, there exist $\ell_{\mathbf{w}}$-projections of a polynomial onto $P_d(g)$ and $Q_d(g)$, where $P_d(g)$ (resp. $Q_d(g)$) denotes the subcone of elements of $P(g)$ (resp. of $Q(g)$) which have a degree bound d in their representation. In general these projections are not unique but we provide a *canonical* $\ell_{\mathbf{w}}$-projection g_f for any polynomial f, which takes a remarkably simple and "sparse" form, and particularly when $\ell_{\mathbf{w}}$ is the usual ℓ_1-norm, in which case

$$g_f \;=\; f + \lambda_0 + \sum_{i=1}^{n} \lambda_i\, x_i^{2d}, \tag{1.1}$$

for some nonnegative vector $\lambda \in \mathbb{R}^{n+1}$. In other words, the *support* $\|g_f\|_0$ of g_f does not depend on $\mathbf{K}$ and does not depend on d either! The dependence of g_f on the g_j's that define $\mathbf{K}$ is only through the coefficients (λ_j^*). In addition, the support is very sparse since $\|g_f\|_0 \leq \|f\|_0 + n + 1$. This confirms the property of the ℓ_1-norm with respect to sparsity, already observed in other contexts. Minimizing the ℓ_1-norm aims at finding a solution with small support (where $\|\mathbf{x}\|_0 = \#\{i : x_i \neq 0\}$). Finally, the vector λ in (1.1) is an optimal solution of an explicit semidefinite program, and so can be computed efficiently.

• We also provide a canonical ℓ_1-projection of f onto $\overline{P(g) \cap \mathbb{R}[\mathbf{x}]_{2d}}$ which is again of the form (1.1), and use this result to characterize the sequential closure $P(g)^{\ddagger}$ of $P(g)$ for the finest locally convex topology. Namely,

$$P(g)^{\ddagger} \;=\; \left\{ f \in \mathbb{R}[\mathbf{x}] \;:\; \exists\, d \text{ s.t. } \forall \epsilon > 0,\; f + \epsilon \left(1 + \sum_{i=1}^{n} x_i^{2d}\right) \in P(g) \right\},$$

and the same statement is true for the quadratic module $Q(g)$. This latter result exhibits the particularly simple form $q := (1 + \sum_{i=1}^{n} x_i^{2d})$ for the possible polynomials q in the characterization of $P(g)^{\ddagger}$ (and $Q(g)^{\ddagger}$) provided in e.g. [5, 9]; e.g., in [9] it is stated that one may take the polynomial $(1 + \|\mathbf{x}\|^2)^s$ for some s.

• Thanks to the characterization of canonical $\ell_{\mathbf{w}}$-projections, we finally obtain a Positivstellensatz on $\mathbf{K}$ of the following form: $f \in \mathrm{Psd}(\mathbf{K})$ if and only if for every $\epsilon > 0$, there is some $d \in \mathbb{N}$ such that the polynomial $f + \epsilon(1 + \sum_{i=1}^{n} \sum_{k=1}^{d} x_i^{2k}/(2k)!)$ is in $P(g)$ (or $Q(g)$).

• At last but not least, and crucial for the above results, we prove a result of independent interest, concerned with sequences $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, that have a finite representing Borel measure μ. Namely, we prove that a polynomial f is nonnegative on the support of μ if and only if $\int h^2 f\, d\mu \geq 0$ for all $h \in \mathbb{R}[\mathbf{x}]$.

The paper is organized as follows. In Section 2, and after introducing the notation and definitions, we present the intermediate result mentioned above. In Section 3 we show that $P(g)$ and $Q(g)$ solve the $\mathbf{K}$-moment problem for $\ell_{\mathbf{w}}$-continuous linear functionals. In section 4, we provide explicit

expressions for the canonical $\ell_{\mathbf{w}}$-projections onto $P(g)$, $Q(g)$ and their truncated versions. Moreover, we characterize the sequential closures $Q(g)^{\ddagger}$ and $P(g)^{\ddagger}$ and we end up with a Positivstellensatz for $\mathbf{K}$.

2. Notation, definitions and preliminaries

2.1. **Notation and definitions.** The notation $\mathcal{B}$ stands for the Borel σ-field of $\mathbb{R}^n$. Let $\mathbb{R}[\mathbf{x}]$ (resp. $\mathbb{R}[\mathbf{x}]_d$) denote the ring of real polynomials in the variables $\mathbf{x} = (x_1, \ldots, x_n)$ (resp. polynomials of degree at most d), whereas $\Sigma[\mathbf{x}]$ (resp. $\Sigma[\mathbf{x}]_d$) denotes its subset of sums of squares (s.o.s.) polynomials (resp. of s.o.s. of degree at most $2d$). For every $\alpha \in \mathbb{N}^n$ the notation $\mathbf{x}^\alpha$ stands for the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $|\alpha|$ stands for the integer $a := \alpha_1 + \cdots + \alpha_n$, and $\alpha!$ stands for the integer $a!$ For an arbitrary set $\mathbf{S} \subseteq \mathbb{R}^n$, let $\mathrm{Psd}(\mathbf{S})$ denote the convex cone of polynomials that are nonnegative on $\mathbf{S}$.

For every $i \in \mathbb{N}$, let $\mathbb{N}^n_d := \{\beta \in \mathbb{N}^n : \sum_j \beta_j \leq d\}$ whose cardinal is $s(d) = \binom{n+d}{n}$. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is written

$$\mathbf{x} \mapsto f(\mathbf{x}) \;=\; \sum_{\alpha \in \mathbb{N}^n} f_\alpha \, \mathbf{x}^\alpha,$$

and f can be identified with its vector of coefficients $\mathbf{f} = (f_\alpha)$ in the canonical basis $(\mathbf{x}^\alpha)$, $\alpha \in \mathbb{N}^n$. The *support* of $f \in \mathbb{R}[\mathbf{x}]$ is the set $\{\alpha \in \mathbb{N}^n : f_\alpha \neq 0\}$ and let $\|f\|_0 := \mathrm{card}\,\{\alpha : f_\alpha \neq 0\}$. Denote by $\|f\|_1$ the ℓ_1-norm $\sum_\alpha |f_\alpha|$ of the coefficient vector $\mathbf{f}$.

Crucial in the sequel is the use of the following $\ell_{\mathbf{w}}$-norm which is a *weighted* ℓ_1-norm defined from the sequence $\mathbf{w} = (w_\alpha)$, $\alpha \in \mathbb{N}^n$, where $w_\alpha := (2\lceil |\alpha|/2 \rceil)!$. Namely, denote by $\|f\|_{\mathbf{w}}$ the $\ell_{\mathbf{w}}$-norm $\sum_\alpha w_\alpha |f_\alpha|$ of the coefficient vector $\mathbf{f}$; hence the ℓ_1-norm corresponds to the case where $w_\alpha = 1$ for all $\alpha \in \mathbb{N}^n$. Both ℓ_1 and $\ell_{\mathbf{w}}$ also define a norm on $\mathbb{R}[\mathbf{x}]_d$.

Let $\mathcal{S}^p \subset \mathbb{R}^{p \times p}$ denote the space of real $p \times p$ symmetric matrices. For any two matrices $\mathbf{A}, \mathbf{B} \in \mathcal{S}^p$, the notation $\mathbf{A} \succeq 0$ (resp. $\succ 0$) stands for $\mathbf{A}$ is positive semidefinite (resp. positive definite), and the notation $\langle \mathbf{A}, \mathbf{B} \rangle$ stands for trace $\mathbf{AB}$.

Let $\mathbf{v}_d(\mathbf{x}) = (\mathbf{x}^\alpha)$, $\alpha \in \mathbb{N}^n_d$, and let $\mathbf{B}^0_\alpha \in \mathbb{R}^{s(d) \times s(d)}$ be real symmetric matrices such that

$$\mathbf{v}_d(\mathbf{x})\, \mathbf{v}_d(\mathbf{x})^T \;=\; \sum_{\alpha \in \mathbb{N}^n_{2d}} \mathbf{x}^\alpha \, \mathbf{B}^0_\alpha. \tag{2.1}$$

Recall that a polynomial $g \in \mathbb{R}[\mathbf{x}]_{2d}$ is a s.o.s. if and only if there exists a real positive semidefinite matrix $\mathbf{X} \in \mathbb{R}^{s(d) \times s(d)}$ such that

$$g_\alpha \;=\; \langle \mathbf{X}, \mathbf{B}^0_\alpha \rangle, \qquad \forall \alpha \in \mathbb{N}^n_{2d}.$$

Let $g_j \in \mathbb{R}[\mathbf{x}]$, $j = 0, 1, \ldots, m$, with g_0 being the constant polynomial $g_0(\mathbf{x}) = 1$ for all $\mathbf{x}$, and let $\mathbf{K} \subseteq \mathbb{R}^n$ be the basic closed semi algebraic set:

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n \,:\, g_j(\mathbf{x}) \geq 0,\; j = 1, \ldots, m\}, \tag{2.2}$$

For every $J \subseteq \{1,\ldots,m\}$ let $g_J := \prod_{k\in J} g_k$, with the convention $g_\emptyset := 1$, and let $v_J := \lceil (\deg g_J)/2 \rceil$ (with $v_j := v_{\{j\}}$).

Definition 2.1. *With* $\mathbf{K}$ *as in (2.2), let* $P(g), Q(g) \subset \mathbb{R}[\mathbf{x}]$ *and* $P_k(g), Q_k(g) \subset \mathbb{R}[\mathbf{x}]_{2k}$ *be the convex cones:*

$$\begin{aligned}
P(g) &:= \left\{ \sum_{J\subseteq\{1,\ldots,m\}} \sigma_J\, g_J \;:\; \sigma_J \in \Sigma[\mathbf{x}],\; J \subseteq \{1,\ldots,m\} \right\},\\
P_k(g) &:= \left\{ \sum_{J\subseteq\{1,\ldots,m\}} \sigma_J\, g_J \;:\; \sigma_J \in \Sigma[\mathbf{x}]_{k-v_J},\; J \subseteq \{1,\ldots,m\} \right\},\\
Q(g) &:= \left\{ \sum_{j=0}^{m} \sigma_j\, g_j \;:\; \sigma_j \in \Sigma[\mathbf{x}],\; j = 1,\ldots,m \right\},\\
Q_k(g) &:= \left\{ \sum_{j=0}^{m} \sigma_j\, g_j \;:\; \sigma_j \in \Sigma[\mathbf{x}]_{k-v_j},\; j = 1,\ldots,m \right\}.
\end{aligned}$$

The set $P(g)$ (resp. $Q(g)$) is a convex cone called the *preordering* (resp. the *quadratic module*) associated with the g_j's. Obviously, if $h \in P(g)$ (resp. $h \in Q(g)$), the associated s.o.s. weights σ_J's (resp. σ_j's) of its representation provide a certificate of nonnegativity of h on $\mathbf{K}$. The convex cone $P_k(g)$ (resp. $Q_k(g)$) is the subset of elements $h \in P(g)$ (resp. $h \in Q(g)$) with a *degree bound* $2k$ certificate. Observe that $P_k(g) \subset P(g) \cap \mathbb{R}[\mathbf{x}]_{2k}$ and $Q_k(g) \subset Q(g) \cap \mathbb{R}[\mathbf{x}]_{2k}$.

Moment matrix. With a sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, let $L_\mathbf{y} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ be the linear functional

$$f \quad (= \sum_\alpha f_\alpha\, \mathbf{x}^\alpha) \quad \mapsto \quad L_\mathbf{y}(f) \,=\, \sum_\alpha f_\alpha\, y_\alpha, \quad f \in \mathbb{R}[\mathbf{x}].$$

With $d \in \mathbb{N}$, the d-moment matrix associated with $\mathbf{y}$ is the real symmetric matrix $\mathbf{M}_d(\mathbf{y})$ with rows and columns indexed in $\mathbb{N}^n_d$, and defined by:

$$\mathbf{M}_d(\mathbf{y})(\alpha,\beta) \,:=\, L_\mathbf{y}(\mathbf{x}^{\alpha+\beta}) \,=\, y_{\alpha+\beta}, \qquad \forall \alpha,\beta \in \mathbb{N}^n_d. \tag{2.3}$$

Alternatively, $\mathbf{M}_d(\mathbf{y}) = \sum_{\alpha\in\mathbb{N}^n_{2d}} y_\alpha \mathbf{B}^0_\alpha$. It is straightforward to check that

$$\left\{\, L_\mathbf{y}(g^2) \geq 0 \quad \forall g \in \mathbb{R}[\mathbf{x}]_d \,\right\} \quad \Leftrightarrow \quad \mathbf{M}_d(\mathbf{y}) \,\succeq\, 0, \quad d = 0,1,\ldots.$$

A sequence $\mathbf{y} = (y_\alpha)$ has a representing measure if there exists a finite Borel measure μ on $\mathbb{R}^n$, such that $y_\alpha = \int \mathbf{x}^\alpha d\mu$ for every $\alpha \in \mathbb{N}^n$. Moreover, the measure μ is said to be *determinate* if it is the unique such measure. Notice that with the $\ell_\mathbf{w}$-norm on $\mathbb{R}[\mathbf{x}]$ is associated a dual norm $\|\cdot\|^*_\mathbf{w}$ on the dual space $\mathbb{R}[\mathbf{x}]^*$ of $\ell_\mathbf{w}$-continuous linear functionals on $\mathbb{R}[\mathbf{x}]$, by $\|L_\mathbf{y}\|^*_\mathbf{w} = \sup\{|y_\alpha|/w_\alpha : \alpha \in \mathbb{N}^n\}$.

Localizing matrix. With $\mathbf{y}$ as above, $J \subseteq \{1,\ldots,m\}$, and $g_J \in \mathbb{R}[\mathbf{x}]$ (with $g_J(\mathbf{x}) = \sum_\gamma g_{J\gamma}\,\mathbf{x}^\gamma$), the *localizing* matrix of order d associated with $\mathbf{y}$ and g_J is the real symmetric matrix $\mathbf{M}_d(g_J\,\mathbf{y})$ with rows and columns indexed by $\mathbb{N}^n_d$, and whose entry (α,β) is just

(2.4)
$$\mathbf{M}_d(\mathbf{y})(g_J\,\mathbf{y})(\alpha,\beta) \;:=\; L_{\mathbf{y}}(g_J(\mathbf{x})\mathbf{x}^{\alpha+\beta}) \;=\; \sum_\gamma g_{J\gamma}\,y_{\alpha+\beta+\gamma}, \quad \forall \alpha,\beta \in \mathbb{N}^n_d.$$

If $\mathbf{B}^J_\alpha \in \mathcal{S}^{s(d)}$ is defined by:

$$g_J(\mathbf{x})\,\mathbf{v}_d(\mathbf{x})\,\mathbf{v}_d(\mathbf{x})^T \;=\; \sum_{\alpha\in\mathbb{N}^n_{2d+\deg g_J}} \mathbf{B}^J_\alpha\,\mathbf{x}^\alpha, \qquad \forall \mathbf{x}\in\mathbb{R}^n, \tag{2.5}$$

then $\mathbf{M}_d(g_J\,\mathbf{y}) = \sum_{\alpha\in\mathbb{N}^n_{2d+\deg g_J}} y_\alpha\,\mathbf{B}^J_\alpha$. Alternatively, $\mathbf{M}_d(g_J\,\mathbf{y}) = \mathbf{M}_d(\mathbf{z})$ where $\mathbf{z} = (z_\alpha)$, $\alpha\in\mathbb{N}^n$, with $z_\alpha = L_{\mathbf{y}}(g_J\,\mathbf{x}^\alpha)$.

Multivariate Carleman's condition. Let $\mathbf{y} = (y_\alpha)$, $\alpha\in\mathbb{N}^n$, be such that $\mathbf{M}_d(\mathbf{y}) \succeq 0$ for all $d\in\mathbb{N}$. If for every $i = 1,\ldots,n$,

$$\sum_{k=1}^\infty L_{\mathbf{y}}(x_i^{2k})^{-1/2k} \;=\; \infty, \tag{2.6}$$

then $\mathbf{y}$ has a finite representing Borel measure μ on $\mathbb{R}^n$, which in addition, is determinate; see e.g. Berg [3].

Closures. For a set $A\subset\mathbb{R}[\mathbf{x}]$ we denote by $\overline{A}$ the closure of A for the finest locally convex topology on $\mathbb{R}[\mathbf{x}]$ (treated as a real vector space). With this topology, every finite-dimensional subspace of $\mathbb{R}[\mathbf{x}]$ inherits the euclidean topology, so that $\overline{A}$ also denotes the usual euclidean closure of a subset $A\subset\mathbb{R}[\mathbf{x}]_d$. Following Cimpric et al. [5] and Kuhlmann et al. [9], we also denote by $A^\ddagger$ the set of all elements of $\mathbb{R}[\mathbf{x}]$ which are expressible as the limit of some sequence of elements of A, and so $A^\ddagger$ is called the *sequential* closure of A, and clearly $A\subseteq A^\ddagger\subseteq\overline{A}$. If $A\subset\mathbb{R}[\mathbf{x}]$ is a convex cone

$$A^\ddagger \;=\; \{f\in\mathbb{R}[\mathbf{x}] \;:\; \exists\, q\in\mathbb{R}[\mathbf{x}] \text{ s.t. } f+\epsilon\,q\in A, \quad \forall\epsilon>0\,\},$$

and in fact, q can be chosen to be in A. Moreover, if A has nonempty interior (equivalently, has an algebraic interior) then $A^\ddagger = \overline{A}$.

Semidefinite programming. A semidefinite program is a convex (more precisely convex conic) optimization problem of the form

$$\inf_{\mathbf{X}}\{\,\langle\mathbf{C},\mathbf{X}\rangle \;:\; \mathcal{A}\,\mathbf{X} \;=\mathbf{b};\; \mathbf{X}\succeq 0\,\},$$

for some real symmetric matrices $\mathbf{C},\mathbf{X}\in\mathcal{S}^p$, vector $\mathbf{b}\in\mathbb{R}^m$, and some linear mapping $\mathcal{A}:\mathcal{S}^p\to\mathbb{R}^m$. Semidefinite programming is a powerful technique of convex optimization, ubiquitous in many areas. A semidefinite program can be solved efficiently and even in time polynomial in the input size of the problem, for fixed arbitrary precision. For more details the interested reader is referred to e.g. [17].

2.2. **A preliminary result of independent interest.** Recall that in a complete separable metric space $\mathbf{X}$, the support of a finite Borel measure μ (denoted $\operatorname{supp}\mu$) is the unique smallest closed set $\mathbf{K} \subseteq \mathbf{X}$ such that $\mu(\mathbf{X} \setminus \mathbf{K}) = 0$.

Theorem 2.2. *Let $f \in \mathbb{R}[\mathbf{x}]$ and μ be a finite Borel measure with all moments $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, finite and such that for some $M > 0$ and all $k \in \mathbb{N}$ and all $i = 1, \ldots, n$, $L_{\mathbf{y}}(x_i^{2k}) \leq (2k)!\, M$. Then:*

$$
\begin{aligned}
(2.7) \qquad f \geq 0 \text{ on } \operatorname{supp}\mu \quad &\Longleftrightarrow \quad \int h^2\, f\, d\mu \,\geq\, 0 \quad \forall h \in \mathbb{R}[\mathbf{x}] \\
&\Longleftrightarrow \quad \mathbf{M}_d(f\,\mathbf{y}) \,\succeq\, 0, \qquad \forall d = 0, 1, \ldots
\end{aligned}
$$

Proof. The implication $\Rightarrow$ is clear. For the reverse implication, consider the signed Borel measure $\nu(B) := \int_B f d\mu$, for all Borel sets $B \in \mathcal{B}$, and let $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, be its sequence of moments. By Lemma 5.1, the sequence $\mathbf{z}$ satisfies Carleman's condition (2.6). Next, recalling that $\mathbf{M}_k(f\,\mathbf{y}) = \mathbf{M}_k(\mathbf{z})$ for every $k \in \mathbb{N}$,

$$
\left(\int h^2 f\, d\mu \,\geq\, 0 \quad \forall h \in \mathbb{R}[\mathbf{x}] \right) \Longleftrightarrow \quad \mathbf{M}_k(\mathbf{z}) \,\succeq\, 0, \forall\, k \in \mathbb{N}.
$$

This combined with the fact that $\mathbf{z}$ satisfies Carleman's condition yields that $\mathbf{z}$ is the moment sequence of a finite Borel measure ψ on $\mathbb{R}^n$, which in addition, is determinate. Therefore,

$$
(2.8) \qquad z_\alpha \,=\, \int \mathbf{x}^\alpha \underbrace{f(\mathbf{x}) d\mu(\mathbf{x})}_{d\nu(\mathbf{x})} \,=\, \int \mathbf{x}^\alpha\, d\psi(\mathbf{x}), \qquad \forall \alpha \in \mathbb{N}^n.
$$

Let $\Gamma+ := \{\mathbf{x} : f(\mathbf{x}) \geq 0\}$, $\Gamma^- := \{\mathbf{x} : f(\mathbf{x}) < 0\}$ and let $\mu = \mu^+ + \mu^-$ with $\mu^+(B) = \mu(B \cap \Gamma^+)$, $\mu^-(B) = \mu(B \cap \Gamma^-)$, for all $B \in \mathcal{B}$. Similarly, let $\nu = \nu^+ - \nu^-$ with ν^+ and ν^- being the positive measures defined by $\nu^+(B) = \int_B f d\mu^+$ and $\nu^-(B) = -\int_B f d\mu^-$, for all $B \in \mathcal{B}$. Since $\mu^+, \mu^- \leq \mu$, one has

$$
\int x_i^{2k} d\mu^+(\mathbf{x}) \,\leq\, \int x_i^{2k} d\mu(\mathbf{x}) \quad \text{and} \quad \int x_i^{2k} d\mu^-(\mathbf{x}) \,\leq\, \int x_i^{2k} d\mu(\mathbf{x}),
$$

for all $i = 1, \ldots, n$ and all $k \in \mathbb{N}$. Therefore, again by Lemma 5.1, both ν^+ and ν^- satisfy Carleman's condition (2.6) so that both are determinate. On the other hand, (2.8) can be rewritten,

$$
\int \mathbf{x}^\alpha\, d\nu^+(\mathbf{x}) \,=\, \int \mathbf{x}^\alpha\, d\nu^-(\mathbf{x}) + \int \mathbf{x}^\alpha\, d\psi(\mathbf{x}), \qquad \forall \alpha \in \mathbb{N}^n,
$$

and so $\nu^+ = \nu^- + \psi$ because ν^+ and $\nu^- + \psi$ are determinate. But then $0 = \nu^+(\Gamma^-) \geq \nu^-(\Gamma^-)$ implies that $\nu^- = 0$, i.e., $\nu = \nu^+ = \psi$, and so the signed Borel measure ν is in fact a positive measure. This in turn implies that $f \geq 0$ for all $\mathbf{x} \in \operatorname{supp}\mu \setminus G$, where $G \subset \operatorname{supp}\mu$ and $\mu(G) = 0$. Notice that by minimality of the support , $\overline{\operatorname{supp}\mu \setminus G} = \operatorname{supp}\mu$. Hence let $\mathbf{x} \in \operatorname{supp}\mu$ be fixed, arbitrary. As $\overline{\operatorname{supp}\mu \setminus G} = \operatorname{supp}\mu$, there is sequence

$(\mathbf{x}_\ell) \subset \operatorname{supp}\mu \setminus G$ such that $\mathbf{x}_\ell \to \mathbf{x}$ as $\ell \to \infty$, and $f(\mathbf{x}_\ell) \geq 0$ for all ℓ. But then continuity of f yields that $f(\mathbf{x}) \geq 0$. □

Interestingly, as we next see, Theorem 2.2 yields alternative characterizations of the cone $\mathrm{Psd}(\mathbf{S})$ for an arbitrary closed set $\mathbf{S} \subset \mathbb{R}^n$.

For a finite Borel measure μ (with all moments finite) and a polynomial $f \in \mathbb{R}[\mathbf{x}]$, let μ_f be the finite signed Borel measure defined by $\mu_f(B) := \int_B f d\mu$ for all $B \in \mathcal{B}$. Let $\Theta_\mu := \{\mu_\sigma : \sigma \in \Sigma[\mathbf{x}]\}$, i.e., Θ_μ is the set of finite Borel measures absolutely continuous with respect to μ, and whose density (or Radon Nikodym derivative) is a sum of squares of polynomials.

Let $\Sigma[\mathbf{x}]^* \subset \mathbb{R}[\mathbf{x}]^*$ be the dual cone of the cone of $\Sigma[\mathbf{x}]$, i.e., the set of linear forms $L_\mathbf{y}$ on $\mathbb{R}[\mathbf{x}]$ such that $L_\mathbf{y}(h^2) \geq 0$ for all $h \in \mathbb{R}[\mathbf{x}]$, and similarly, let $\Theta_\mu^* \subset \mathbb{R}[\mathbf{x}]$ be the dual cone of Θ_μ, i.e., $\Theta_\mu^* := \{h \in \mathbb{R}[\mathbf{x}] : \int h d\nu \geq 0, \ \forall \nu \in \Theta_\mu\}$.

Corollary 2.3. *Let $\mathbf{S} \subseteq \mathbb{R}^n$ be an arbitrary closed set and let μ be any finite Borel measure such that $\operatorname{supp}\mu = \mathbf{S}$ and $\mathbf{x} \mapsto \exp(|x_i|)$ is μ-integrable for all $i = 1, \ldots, n$. Then with $f \in \mathbb{R}[\mathbf{x}]$,*

$$f \in \mathrm{Psd}(\mathbf{S}) \iff \mu_f \in \Sigma[\mathbf{x}]^* \tag{2.9}$$

$$\mathrm{Psd}(\mathbf{S}) = \Theta_\mu^*. \tag{2.10}$$

Proof. Let $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, be the moment sequence of μ. Observe that for every $i = 1, \ldots, n$, and all $k \in \mathbb{N}$,

$$L_\mathbf{y}\left(\frac{x_i^{2k}}{(2k)!}\right) \leq \int_\mathbf{S} \exp(|x_i|) d\mu(\mathbf{x}) \leq M,$$

for some $M > 0$. Moreover, $\operatorname{supp}\mu = \mathbf{S}$ and so by Theorem 2.2, $f \geq 0$ on $\mathbf{S}$ if and only if $\int_\mathbf{S} h^2 f d\mu \geq 0$ for all $h \in \mathbb{R}[\mathbf{x}]$. Equivalently, if and only if $\int_\mathbf{S} \sigma d\mu_f \geq 0$ for all $\sigma \in \Sigma[\mathbf{x}]$ (which yields (2.9)), or if and only if $\int_\mathbf{S} f d\mu_\sigma \geq 0$ for all $\sigma \in \Sigma[\mathbf{x}]$, which yields (2.10). □

3. The $\mathbf{K}$-moment problem for $\ell_\mathbf{w}$-continuous linear functionals

We first show that $Q(g)$ (and $P(g)$) solve the $\mathbf{K}$-moment problem for $\ell_\mathbf{w}$-continuous linear functionals.

When equipped with the $\ell_\mathbf{w}$-norm, we may and will identify $\mathbb{R}[\mathbf{x}]$ as the subspace of sequences with finite support, in the Banach space of real infinite sequences $\mathbf{f} = (f_\alpha)$, $\alpha \in \mathbb{N}^n$, that are $\mathbf{w}$-summable, i.e., such that $\sum_\alpha w_\alpha |f_\alpha| < +\infty$.

Proposition 3.1. *The dual of $(\mathbb{R}[\mathbf{x}], \|\cdot\|_\mathbf{w})$ is the space $(\mathbb{R}[\mathbf{x}]^*, \|\cdot\|_\mathbf{w}^*)$ of linear functionals $L_\mathbf{y}$ associated with the sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, which satisfy $\|L_\mathbf{y}\|_\mathbf{w}^* < \infty$, where $\|L_\mathbf{y}\|_\mathbf{w}^* := \sup\{|y_\alpha|/w_\alpha : \alpha \in \mathbb{N}^n\}$.*

Proof. If $L_\mathbf{y} \in \mathbb{R}[\mathbf{x}]^*$ satisfies $\|L_\mathbf{y}\|_\mathbf{w}^* < \infty$, then

$$|L_\mathbf{y}(f)| \leq \sum_\alpha |f_\alpha| \, w_\alpha \frac{y_\alpha}{w_\alpha} \leq \|f\|_\mathbf{w} \, \|L_\mathbf{y}\|_\mathbf{w}^*,$$

and so $L_{\mathbf{y}}$ is bounded, hence $\ell_{\mathbf{w}}$-continuous. Conversely, if $L_{\mathbf{y}}$ is $\ell_{\mathbf{w}}$-continuous, then consider the sequence of polynomials $(f_\alpha) \subset \mathbb{R}[\mathbf{x}]$ with $\mathbf{x} \mapsto f_\alpha(\mathbf{x}) := \mathbf{x}^\alpha / w_\alpha$ for every $\alpha \in \mathbb{N}^n$. Then $\|f_\alpha\|_{\mathbf{w}} = 1$ for every $\alpha \in \mathbb{N}^n$, and $L_{\mathbf{y}}(f_\alpha) = y_\alpha / w_\alpha$ for all $\alpha \in \mathbb{N}^n$. And so, if $L_{\mathbf{y}}$ is $\ell_{\mathbf{w}}$-continuous then $\sup\{|y_\alpha|/w_\alpha : \alpha \in \mathbb{N}^n\} < +\infty$. □

Theorem 3.2. *Let $\mathbf{K} \subseteq \mathbb{R}^n$ be as in (2.2) and let $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}$, be a given real sequence such that $L_{\mathbf{y}}$ is $\ell_{\mathbf{w}}$-continuous. Then $\mathbf{y}$ has a finite representing Borel measure μ on $\mathbf{K}$ if and only if $L_{\mathbf{y}}$ is nonnegative on $Q(g)$. Equivalently if and only if:*
(3.1)

$$L_{\mathbf{y}}(h^2 g_j) \geq 0 \quad \forall\, h \in \mathbb{R}[\mathbf{x}],\ j = 0, \ldots, m; \quad \sup_{\alpha \in \mathbb{N}^n} \frac{|y_\alpha|}{w_\alpha} \leq M \text{ for some } M.$$

Proof. The necessity is clear. Indeed, if $\mathbf{y}$ has a representing measure μ on $\mathbf{K}$ then $L_{\mathbf{y}}(h^2 g_j) = \int_{\mathbf{K}} h^2 g_j d\mu \geq 0$ for all $h \in \mathbb{R}[\mathbf{x}]$ and all $j = 0, 1, \ldots, m$; and so $L_{\mathbf{y}}(f) \geq 0$ for all $f \in Q(g)$. Moreover, $\|L_{\mathbf{y}}\|^*_{\mathbf{w}} < \infty$ because $L_{\mathbf{y}}$ is $\ell_{\mathbf{w}}$-continuous; hence (3.1) holds.

Sufficiency. Suppose that $L_{\mathbf{y}}$ is a non trivial $\ell_{\mathbf{w}}$-continuous linear functional, nonnegative on $Q(g)$, i.e., suppose that (3.1) holds. In particular, $L_{\mathbf{y}}(x_i^{2k}) \leq M(2k)!$ for every $k \in \mathbb{N}$ and every $i = 1, \ldots, n$. Therefore, $\mathbf{y}$ satisfies Carleman's condition (2.6) and since $\mathbf{M}_k(\mathbf{y}) \succeq 0$ for all $k \in \mathbb{N}$, $\mathbf{y}$ has a representing finite Borel measure μ on $\mathbb{R}^n$, which in addition, is determinate. Next, using $L_{\mathbf{y}}(h^2 g_j) \geq 0$ for all $h \in \mathbb{R}[\mathbf{x}]$, and invoking Theorem 2.2, one may conclude that $g_j \geq 0$ on $\operatorname{supp}\mu$, for every $j = 1, \ldots, m$. Hence $\operatorname{supp}\mu \subseteq \mathbf{K}$. □

Theorem 3.2 states that $Q(g)$ solves the **K**-moment problem for $\ell_{\mathbf{w}}$-continuous functionals. Of course, Theorem 3.2 is also true if one replaces the quadratic module $Q(g)$ with the preordering $P(g)$.

The $\ell_{\mathbf{w}}$-closure of $Q(g)$ and $P(g)$. Observe that $\operatorname{Psd}(\mathbf{K})$ is $\ell_{\mathbf{w}}$-closed. To see this, notice that with every every $\mathbf{x} \in \mathbf{K}$ is associated the Dirac measure $\delta_{\mathbf{x}}$, whose associated sequence $\mathbf{y} = (\mathbf{x}^\alpha)$, $\alpha \in \mathbb{N}^n$, is such that $L_{\mathbf{y}}$ is $\ell_{\mathbf{w}}$-continuous. Indeed, let $a := \max_i |\mathbf{x}_i|$ so that $|\mathbf{x}^\alpha| \leq a^{|\alpha|}$, and let $M := \exp(a)$.

$$M^{-1}\,|y_\alpha| \,=\, \exp(a)^{-1}|\mathbf{x}^\alpha| \,\leq\, \exp(a)^{-1}\, a^{|\alpha|} \,<\, \alpha! \,\leq\, (2\lceil|\alpha|/2\rceil)! \,=\, w_\alpha,$$

and so $\|L_{\mathbf{y}}\|^*_{\mathbf{w}} < M$, i..e., $L_{\mathbf{y}}$ is $\ell_{\mathbf{w}}$-continuous. Therefore, let $(f_n) \subset \operatorname{Psd}(\mathbf{K})$ be such that $\|f_n - f\|_{\mathbf{w}} \to 0$ as $n \to \infty$. As $L_{\mathbf{y}}$ is $\ell_{\mathbf{w}}$-continuous one must have $0 \leq \lim_{n\to\infty} L_{\mathbf{y}}(f_n) = L_{\mathbf{y}}(f) = f(\mathbf{x})$. As $\mathbf{x} \in \mathbf{K}$ was arbitrary, $f \in \operatorname{Psd}(\mathbf{K})$.

Theorem 3.3. *Let $\mathbf{K} \subseteq \mathbb{R}^n$ be as in (2.2) and recall that* $\operatorname{Psd}(\mathbf{K}) := \{f \in \mathbb{R}[\mathbf{x}] : f \geq 0$ *on* $\mathbf{K}\}$. *Then* $\operatorname{cl}_{\mathbf{w}}(P(g)) = \operatorname{cl}_{\mathbf{w}}(Q(g)) = \operatorname{Psd}(\mathbf{K})$.

Proof. As $\operatorname{Psd}(\mathbf{K})$ is $\ell_{\mathbf{w}}$-closed and $Q(g) \subset \operatorname{Psd}(\mathbf{K})$, $\operatorname{cl}_{\mathbf{w}}(Q(g)) \subseteq \operatorname{Psd}(\mathbf{K})$, and so we only have to prove the reverse inclusion. Let $f \notin \operatorname{cl}_{\mathbf{w}}(Q(g))$. Since $Q(g)$ is a convex cone, by the Hahn-Banach separation theorem there

exists an $\ell_{\mathbf{w}}$-continuous linear functional $L_{\mathbf{y}}$ that strictly separates f from $\mathrm{cl}_{\mathbf{w}}(Q(g))$. That is, there exists $\mathbf{y} \in \mathbb{N}^n$ such that $L_{\mathbf{y}}$ is $\ell_{\mathbf{w}}$-continuous, $L_{\mathbf{y}}(f) < 0$ and $L_{\mathbf{y}}(h) \geq 0$ for all $h \in \mathrm{cl}_{\mathbf{w}}(Q(g))$. By Theorem 3.2, such a $\mathbf{y}$ has a representing finite Borel measure μ on $\mathbf{K}$, and so $L_{\mathbf{y}}(f) = \int_{\mathbf{K}} f d\mu < 0$ yields that necessarily $f(\mathbf{x}_0) < 0$ for some $\mathbf{x}_0 \in \mathbf{K}$. Hence $\mathrm{Psd}(\mathbf{K}) \subseteq \mathrm{cl}_{\mathbf{w}}(Q(g))$, which in turn yields the desired result. □

For instance, from Berg [3], the ℓ_1-closure of $\Sigma[\mathbf{x}]$ is $\mathrm{Psd}([-1,1]^n)$. On the other hand, its $\ell_{\mathbf{w}}$-closure is now $\mathrm{Psd}(\mathbb{R}^n)$, which is what we really want.

It is worth mentioning that Theorem 3.2 and 3.3 can be extended to any preordering or quadratic module (finitely generated or not). The proof is the same. Any closed set $\mathbf{K} \subset \mathbb{R}^n$ may be represented as the non-negativity set of such a quadratic module (taking generators of the form $\mathbf{x} \mapsto g(\mathbf{x}) := \sum_{i=1}^n (x_i - a_i) - r^2$, for suitable $\mathbf{a} \in \mathbb{R}^n$ and $r > 0$). However, for reasons that become obvious in the next section, the focus of the present paper is on the finitely generated case.

4. Canonical $\ell_{\mathbf{w}}$-projections onto $P_d(g)$ and $Q_d(g)$

As we next see, the $\ell_{\mathbf{w}}$- and ℓ_1-norm have the nice feature that one may find particular (canonical) projections onto various truncations of $P(g)$ and $Q(g)$ with a particularly simple expression. For $\ell_{\mathbf{w}}$- and ℓ_1- projections to be well-defined we assume that $\mathbf{K}$ has a nonempty interior.

We first provide an explicit form of *canonical* ℓ_1- and $\ell_{\mathbf{w}}$-projection of any given polynomial f onto $P_d(g)$ and $Q_d(g)$ respectively, and analyze their limit as $d \to \infty$. Then we will consider the projections onto $P_d(g) \cap \mathbb{R}[\mathbf{x}]_s$ for fixed $s, d \in \mathbb{N}$, which, letting $s \to \infty$, will permit to obtain the projection of f onto $\overline{P(g) \cap \mathbb{R}[\mathbf{x}]_d}$, and so to also characterize the sequential closure $P(g)^{\ddagger}$.

As in the previous section, for a polynomial in $\mathbb{R}[\mathbf{x}]_t$ we use indifferently the notation h for both the polynomial and its vector of coefficients $\mathbf{h} \in \mathbb{R}^{s(t)}$. The context will make clear which one of the two is concerned.

Let $\mathbf{K} \subseteq \mathbb{R}^n$ be as in (2.2) and consider the following optimization problem:

$$p_{d\mathbf{w}} := \inf_h \{ \, \|f - h\|_{\mathbf{w}} \; : \; h \in P_d(g) \, \}. \tag{4.1}$$

That is, one searches for the best $\ell_{\mathbf{w}}$-approximation of f by an element h^* of $P_d(g)$, or equivalently, an $\ell_{\mathbf{w}}$-projection of f onto $P_d(g)$. In general, such a best approximation $h^* \in P_d(g)$ is not unique[1] but we provide a *canonical* solution with a very simple expression. Of course, and even though (4.1) is well defined for an arbitrary $f \in \mathbb{R}[\mathbf{x}]$, such a problem is of particular interest when f is nonnegative on $\mathbf{K}$ but not necessarily an element of $P(g)$.

[1]The following example was kindly provided by an anonymous referee: Let $n = 2, d = 1$, $\mathbf{x} \mapsto f(\mathbf{x}) := -2x_1x_2$, and let $C \subset \mathbb{R}[\mathbf{x}]_{2d}$ be the cone of sums of squares of linear polynomials. Then any polynomial $\mathbf{x} \mapsto p_\lambda(\mathbf{x}) := \lambda(x_1 - x_2)^2$ with $\lambda \in [0,1]$, is an $\ell_{\mathbf{w}}$-projection (and an ℓ_1-projection) of f onto C and $\|f - p_\lambda\|_{\mathbf{w}} = 2$.

Theorem 4.1. *Let $\mathbf{K} \subseteq \mathbb{R}^n$ in (2.2) be with nonempty interior. Let $f \in \mathbb{R}[\mathbf{x}]$ and let $2d \geq \deg f$. There is an $\ell_{\mathbf{w}}$-projection of f onto $P_d(g)$ which is a polynomial $g_f^{P\mathbf{w}} \in \mathbb{R}[\mathbf{x}]_{2d}$ of the form:*

$$\mathbf{x} \mapsto g_f^{P\mathbf{w}}(\mathbf{x}) := f(\mathbf{x}) + \left(\lambda_0^{P\mathbf{w}} + \sum_{i=1}^{n} \sum_{k=1}^{d} \lambda_{ik}^{P\mathbf{w}} \frac{x_i^{2k}}{(2k)!} \right), \tag{4.2}$$

where the nonnegative vector $\lambda^{P\mathbf{w}} \in \mathbb{R}^{nd+1}$ is an optimal solution of the semidefinite program:

$$\inf_{\lambda \geq 0} \left\{ \lambda_0 + \sum_{i=1}^{n} \sum_{k=1}^{d} \lambda_{ik} \ : \ f + \lambda_0 + \sum_{i=1}^{n} \sum_{k=1}^{d} \lambda_i \frac{x_{ik}^{2k}}{(2k)!} \in P_d(g) \right\}, \tag{4.3}$$

and $p_{d\mathbf{w}} = \|f - g_f^{P\mathbf{w}}\|_{\mathbf{w}} = \lambda_0^{P\mathbf{w}} + \sum_{i=1}^{n} \sum_{k=1}^{d} \lambda_{ik}^{P\mathbf{w}}$.

Proof. Consider f as an element of $\mathbb{R}[\mathbf{x}]_{2d}$ by setting $f_\alpha = 0$ whenever $|\alpha| > \deg f$ (where $|\alpha| = \sum \alpha_i$), and rewrite (4.1) as the semidefinite program:

$$\begin{array}{rll} p_{d\mathbf{w}} := \inf\limits_{\lambda, \mathbf{X}_J, h} & \displaystyle\sum_{\alpha \in \mathbb{N}^n_{2d}} w_\alpha \lambda_\alpha & \\ \text{s.t.} & \lambda_\alpha + h_\alpha \geq f_\alpha, & \forall \alpha \in \mathbb{N}^n_{2d} \\ & \lambda_\alpha - h_\alpha \geq -f_\alpha, & \forall \alpha \in \mathbb{N}^n_{2d} \\ & h_\alpha - \displaystyle\sum_{J \subseteq \{1,\ldots,m\}} \langle \mathbf{X}_J, \mathbf{B}_\alpha^J \rangle = 0, & \forall \alpha \in \mathbb{N}^n_{2d} \\ & \lambda \geq 0;\ h \in \mathbb{R}[\mathbf{x}]_{2d};\ \mathbf{X}_J \succeq 0,\ \forall J \subseteq \{1,\ldots,m\}. & \end{array} \tag{4.4}$$

The dual semidefinite program of (4.4) reads:

$$\left\{ \begin{array}{rll} \sup\limits_{\mathbf{u},\mathbf{v} \geq 0, \mathbf{y}} & \displaystyle\sum_{\alpha \in \mathbb{N}^n_d} f_\alpha (u_\alpha - v_\alpha) & \\ \text{s.t.} & u_\alpha + v_\alpha \ \leq \ w_\alpha & \forall \alpha \in \mathbb{N}^n_{2d} \\ & u_\alpha - v_\alpha + y_\alpha \ = \ 0 & \forall \alpha \in \mathbb{N}^n_{2d}, \\ & \mathbf{M}_d(g_J\, \mathbf{y}) \ \succeq \ 0, & \forall J \subseteq \{1,\ldots,m\}, \end{array} \right. \tag{4.5}$$

or, equivalently,

$$\left\{ \begin{array}{ll} \sup\limits_{\mathbf{y}} & -L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_d(g_J\, \mathbf{y}) \ \succeq \ 0, \qquad \forall J \subseteq \{1,\ldots,m\} \\ & |y_\alpha| \ \leq \ w_\alpha, \quad \forall \alpha \in \mathbb{N}^n_{2d}. \end{array} \right. \tag{4.6}$$

The semidefinite program (4.6) has an optimal solution $\mathbf{y}^*$ because the feasible set is nonempty and compact. In addition, let $\mathbf{y} = (y_\alpha)$ be the moment sequence of the finite Borel measure $\mu(B) = \int_{\mathbf{K} \cap B} \mathrm{e}^{-\|\mathbf{x}\|^2} d\mathbf{x}$, for all $B \in \mathcal{B}$, scaled so that $|y_\alpha| < w_\alpha$ for all $\alpha \in \mathbb{N}^n_{2d}$. Then $(\mathbf{y}, \mathbf{u}, \mathbf{v})$ with $\mathbf{u} = -\min[\mathbf{y}, 0]$

and $\mathbf{v} = \max[\mathbf{y}, 0]$, is strictly feasible in (4.5). Indeed, as $\mathbf{K}$ has nonempty interior, then necessarily $\mathbf{M}_d(g_J\,\mathbf{y}) \succ 0$ for all $J \subseteq \{1,\ldots,m\}$, and so Slater's condition[2] holds for (4.5). Therefore, by a standard duality result in convex optimization, there is no duality gap between (4.4) and (4.5) (or (4.6)), and (4.4) has an optimal solution $(\lambda^*, (\mathbf{X}_j^*), g_f^P)$. Hence $p_{d\mathbf{w}} = -L_{\mathbf{y}^*}(f)$ for any optimal solution $\mathbf{y}^*$ of (4.6).

Next, recall that with $J := \emptyset$, $\mathbf{M}_d(g_\emptyset\,\mathbf{y}) = \mathbf{M}_d(\mathbf{y})$. Moreover, $\mathbf{M}_d(\mathbf{y}) \succeq 0$ implies $\mathbf{M}_k(\mathbf{y}) \succeq 0$ for all $k \le d$. By [11, Lemma 1], $\mathbf{M}_k(\mathbf{y}) \succeq 0$ implies that $|y_\alpha| \le \max[L_{\mathbf{y}}(1), \max_i L_{\mathbf{y}}(x_i^{2k})]$, for every $\alpha \in \mathbb{N}^n_{2k}$, and all $k \le d$. Therefore, (4.6) has the equivalent formulation

$$
(4.7)\quad \begin{cases} p_d = -\inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f)) & & \\ \text{s.t.} & \mathbf{M}_d(g_J\,\mathbf{y}) & \succeq 0, & \forall J \subseteq \{1,\ldots,m\} \\ & L_{\mathbf{y}}(1) & \le 1 & \\ & L_{\mathbf{y}}(x_i^{2k}) & \le (2k)!, & k = 1,\ldots,d;\ i = 1,\ldots,n. \end{cases}
$$

Indeed, any feasible solution of (4.7) satisfies

$$|y_\alpha| \le \max[L_{\mathbf{y}}(1), \max_i L_{\mathbf{y}}(x_i^{2k})] \le (2k)! = w_\alpha,$$

for every α with $|\alpha| = 2k$ and $2k-1$, $k = 1,\ldots,d$. The dual of (4.7) is exactly the semidefinite program (4.3). Again Slater's condition holds for (4.7) and it has an optimal solution $\mathbf{y}^*$. Therefore (4.3) also has an optimal solution $\lambda^{P\mathbf{w}} \in \mathbb{R}^{nd+1}_+$ with $p_{d\mathbf{w}} = \lambda_0^{P\mathbf{w}} + \sum_{i=1}^n \sum_{k=1}^d \lambda_{ik}^{P\mathbf{w}}$, which is the desired result. □

The polynomial $g_f^{P\mathbf{w}} \in P_d(g)$ in (4.2) is what we call the canonical $\ell_{\mathbf{w}}$-projection of f onto $P_d(g)$.

Of course, all statements in Theorem 4.1 remain true if one replaces $P_d(g)$ with $Q_d(g)$. Moreover, if $w_\alpha = 1$ for all α (and so $\ell_{\mathbf{w}}$ is now the usual ℓ_1-norm) the polynomial $g_f^{P\mathbf{w}}$ in (4.2) simplifies and is of the form:

$$
(4.8)\qquad \mathbf{x} \mapsto g_f^P(\mathbf{x}) := f(\mathbf{x}) + (\lambda_0^P + \sum_{i=1}^n \lambda_i^P\, x_i^{2d}),
$$

for some nonnegative vector $\lambda^P \in \mathbb{R}^{n+1}$. If $\mathbf{K} = \mathbb{R}^n$ then g_f^P is the canonical ℓ_1-projection of f onto the cone of s.o.s. polynomials., as illustrated in the following example.

Example 1. Let $n = 2$ and $\mathbf{K} = \mathbb{R}^2$, in which case $P_d(g) = Q_d(g) = \Sigma[\mathbf{x}]_d$ for all $d \in \mathbb{N}$. Consider the Motzkin-like polynomial[3] $\mathbf{x} \mapsto f(\mathbf{x}) =$

[2]Slater's condition holds for the conic optimization problem $\mathbf{P}$: $\inf_{\mathbf{x}}\{\mathbf{c}'\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b};\ \mathbf{x} \in \mathbf{K}\}$, where $\mathbf{K} \subset \mathbb{R}^n$ is a convex cone and $\mathbf{A} \in \mathbb{R}^{p\times n}$, if there exists a feasible solution $\mathbf{x}_0 \in \operatorname{int}\mathbf{K}$. In this case, there is no duality gap between $\mathbf{P}$ and its dual $\mathbf{P}^*$: $\sup_{\mathbf{z}}\{\mathbf{b}'\mathbf{z} : \mathbf{c} - \mathbf{A}'\mathbf{z} \in \mathbf{K}^*\}$. In addition, if the optimal value is bounded then $\mathbf{P}^*$ has an optimal solution.

[3]Computation was made by running the GloptiPoly software [7] dedicated to solving the generalized problem of moments.

$x_1^2x_2^2(x_1^2+x_2^2-1)+1/27$ of degree 6, which is nonnegative but not a s.o.s., and with a global minimum $f^*=0$ attained at 4 global minimizers $\mathbf{x}^*=(\pm(1/3)^{1/2},\pm(1/3)^{1/2})$. The results are displayed in Table 1 for $d=3,4,5$.

d	λ^*	p_d
3	$\approx 10^{-3}\,(5.445, 5.367, 5.367)$	$\approx 1.6\,10^{-2}$
4	$\approx 10^{-4}\,(2.4, 9.36, 9.36)$	$\approx 2.\,10^{-3}$
5	$\approx 10^{-5}\,(0.04, 4.34, 4.34)$	$\approx 8.\,10^{-5}$

TABLE 1. Best ℓ_1-approximation for the Motzkin polynomial.

4.1. **Canonical $\ell_{\mathbf{w}}$-projection onto $P(g)_t\cap\mathbb{R}[\mathbf{x}]_{2d}$ and $Q(g)_t\cap\mathbb{R}[\mathbf{x}]_{2d}$.** We now consider $\ell_{\mathbf{w}}$-projection of f onto $P_t(g)\cap\mathbb{R}[\mathbf{x}]_{2d}$ for given integers $d,t\in\mathbb{N}$, i.e.,

$$p_t^{d\mathbf{w}}:=\inf_g\{\,\|f-g\|_{\mathbf{w}}\;:\;g\in P_t(g)\,\cap\,\mathbb{R}[\mathbf{x}]_{2d}\,\}. \tag{4.9}$$

In other words, we are interested in searching for a polynomial of degree $2d$ in $P_t(g)$ which is the closest to f for the $\ell_{\mathbf{w}}$-norm. For instance, if $2d=\deg f$, one wishes to find an $\ell_{\mathbf{w}}$-projection onto $P_t(g)$ of same degree as f. One may also consider the analogue problem with the quadratic module, i.e., an $\ell_{\mathbf{w}}$-projection onto $Q_t(g)\cap\mathbb{R}[\mathbf{x}]_{2d}$. We also analyze the limit as $t\to\infty$.

Theorem 4.2. *Let $\mathbf{K}\subseteq\mathbb{R}^n$ in (2.2) be with nonempty interior and let $d\in\mathbb{N}$. Let $f\in\mathbb{R}[\mathbf{x}]$ and let $2t\geq\max[2d,\deg f]$. There is an $\ell_{\mathbf{w}}$-projection of f onto $P_t(g)\cap\mathbb{R}[\mathbf{x}]_{2d}$ which is a polynomial $g_f^{P\mathbf{w}}\in\mathbb{R}[\mathbf{x}]_{2d}$ of the form:*

$$\mathbf{x}\mapsto g_f^{P\mathbf{w}}(\mathbf{x})\;=\;f(\mathbf{x})+\left(\lambda_0^{P\mathbf{w}}+\sum_{i=1}^n\sum_{k=1}^d\lambda_{ik}^{P\mathbf{w}}\,\frac{x_i^{2d}}{(2k)!}\right), \tag{4.10}$$

where the nonnegative vector $\lambda^{P\mathbf{w}}\in\mathbb{R}^{nd+1}$ is an optimal solution of the semidefinite program:

$$p_t^{d\mathbf{w}}\;=\;\inf_{\lambda\geq0}\left\{\lambda_0+\sum_{i=1}^n\sum_{k=1}^d\lambda_{ik}\;:\;f+\lambda_0+\sum_{i=1}^n\sum_{k=1}^d\lambda_{ik}\,\frac{x_i^{2d}}{(2k)!}\in P_t(g)\right\}, \tag{4.11}$$

and $p_t^{d\mathbf{w}}=\|f-g^PP\mathbf{w}_f\|_{\mathbf{w}}=\lambda_0^{P\mathbf{w}}+\sum_{i=1}^n\sum_{k=1}^d\lambda_{ik}^{P\mathbf{w}}$.

Proof. The proof is almost a verbatim copy of that of Theorem 4.1 with a slight difference. For instance, (4.4) is now replaced with

$$
\begin{aligned}
p_t^{d\mathbf{w}} := \inf_{\lambda,\mathbf{X}_J,h} \quad & \sum_{\alpha\in\mathbb{N}^n_{2d}} w_\alpha\,\lambda_\alpha \\
\text{s.t.} \quad & \lambda_\alpha + h_\alpha \geq f_\alpha, \quad \forall \alpha \in \mathbb{N}^n_{2d} \\
& \lambda_\alpha - h_\alpha \geq -f_\alpha, \quad \forall \alpha \in \mathbb{N}^n_{2d} \\
& h_\alpha - \sum_{J\subseteq\{1,\ldots,m\}}^{m} \langle \mathbf{X}_J, \mathbf{B}^J_\alpha\rangle = 0, \quad \forall \alpha \in \mathbb{N}^n_{2t} \\
& \lambda \geq 0;\ h \in \mathbb{R}[\mathbf{x}]_{2d};\ \mathbf{X}_J \succeq 0,\ \forall J \subseteq \{1,\ldots,m\}.
\end{aligned} \tag{4.12}
$$

(where $h_\alpha = 0$ whenever $|\alpha| > 2d$) and the dual (4.6) now reads

$$
\begin{cases}
\sup\limits_{\mathbf{y}} & -L_{\mathbf{y}}(f) \\
\text{s.t.} & \mathbf{M}_t(g_J\,\mathbf{y}) \,\succeq\, 0, \qquad \forall J \subseteq \{1,\ldots,m\} \\
& |y_\alpha| \,\leq\, w_\alpha, \quad \forall \alpha \in \mathbb{N}^n_{2d}.
\end{cases} \tag{4.13}
$$

Again with exactly the same arguments, (4.12) has a feasible solution and (4.13) has a strictly feasible solution $\mathbf{y}$, and so Slater's condition holds for (4.13), which in turn implies that there is no duality gap between (4.12) and (4.13), and (4.12) has an optimal solution $(\lambda, (\mathbf{X}_J), g_f^{P\mathbf{w}})$. However (and this is the only difference with the proof of Theorem 4.1) now one cannot guarantee any more that (4.13) has an optimal solution because $|y_\alpha| \leq w_\alpha$ only for $\alpha \in \mathbb{N}^n_{2d}$ and not for all $\alpha \in \mathbb{N}^n_{2t}$. □

As before, we call $g_f^{P\mathbf{w}}$ in (4.10) the canonical $\ell_{\mathbf{w}}$-projection of f onto $P_t(g) \cap \mathbb{R}[\mathbf{x}]_{2d}$. Of course, an analogue of Theorem 4.2 (with obvious *ad hoc* adjustments) holds for the canonical $\ell_{\mathbf{w}}$-projection $g_f^{Q\mathbf{w}}$ onto $Q_t(g) \cap \mathbb{R}[\mathbf{x}]_{2d}$. Also and again, if $w_\alpha = 1$ for all $\alpha \in \mathbb{N}^n_{2d}$, then the polynomial $g_f^{P\mathbf{w}}$ in (4.10) simplifies to the form in (4.8).

We next analyze the behavior of $g_f^{P\mathbf{w}}$ as $t \to \infty$ to obtain the canonical $\ell_{\mathbf{w}}$-projection of f onto $\overline{P(g) \cap \mathbb{R}[\mathbf{x}]_{2d}}$. Recall that

$$
\overline{P(g) \cap \mathbb{R}[\mathbf{x}]_{2d}} \,=\, \overline{\left(\bigcup_{t\geq 0} P_t(g)\right) \cap \mathbb{R}[\mathbf{x}]_{2d}} \,=\, \overline{\bigcup_{t\geq 0} P^d_t(g)}.
$$

Corollary 4.3. *Let $\mathbf{K} \subseteq \mathbb{R}^n$ be as in (2.2) and with nonempty interior, $f \in \mathbb{R}[\mathbf{x}]_{2d}$, and let $g_f^{P\mathbf{w}}(t) \in P_t(g) \cap \mathbb{R}[\mathbf{x}]_{2d}$ be an optimal solution of (4.9). Then there is an $\ell_{\mathbf{w}}$-projection of f onto $\overline{P(g) \cap \mathbb{R}[\mathbf{x}]_{2d}}$ which is a polynomial $g_f^{\mathbf{w}} \in \mathbb{R}[\mathbf{x}]_{2d}$ of the form*

$$
\mathbf{x} \,\mapsto\, g_f^{\mathbf{w}}(\mathbf{x}) \,=\, f(\mathbf{x}) + \left(\lambda_0^* + \sum_{i=1}^{n}\sum_{k=1}^{d} \lambda^*_{ik}\, x_i^{2d}\right), \tag{4.14}
$$

for some nonnegative vector $\lambda^* \in \mathbb{R}^{dn+1}$. In particular, if $\mathbf{K}$ is compact and $f \geq 0$ on $\mathbf{K}$ then $\lambda^* = 0$ and $g_f^{\mathbf{w}} = f$.

Proof. Let $(p_t^{d\mathbf{w}})$, $t \in \mathbb{N}$, be the sequence of optimal values of (4.9), which is nonnegative and monotone non increasing. Therefore it converges to some nonnegative value $p^{d\mathbf{w}} \geq 0$. For every $t \in \mathbb{N}$, (4.9) has an optimal solution $g_f^{P\mathbf{w}}(t)$ of the form

$$(4.15)\quad \mathbf{x} \mapsto g_f^{P\mathbf{w}}(t)(\mathbf{x}) \,=\, f(\mathbf{x}) + \lambda_0^{P\mathbf{w}}(t) + \sum_{i=1}^{n}\sum_{k=1}^{d} \lambda_{ik}^{P\mathbf{w}}(t)\frac{x_i^{2k}}{2k!}, \qquad \forall \mathbf{x} \in \mathbb{R}^n,$$

with $\lambda^{P\mathbf{w}}(t) \geq 0$ and $\sum_{i,k} \lambda_{ik}^{P\mathbf{w}}(t) = p_t^{d\mathbf{w}} \leq p_{t_0}^{d\mathbf{w}}$. Hence there is a subsequence (t_j), $j \in \mathbb{N}$, and some nonnegative vector $\lambda^* \in \mathbb{R}_+^{nd+1}$ such that $\lambda^{P\mathbf{w}}(t_j) \to \lambda^*$ as $j \to \infty$. In addition,

$$p^{d\mathbf{w}} \,=\, \lim_{j\to\infty} p_{t_j}^{d\mathbf{w}} \,=\, \lim_{j\to\infty} \lambda_0^{P\mathbf{w}}(t_j) + \sum_{i=1}^{n}\sum_{k=1}^{d} \lambda_{ik}^{P\mathbf{w}}(t_j) \,=\, \lambda_0^* + \sum_{i=1}^{n}\sum_{k=1}^{d} \lambda_{ik}^*.$$

Hence $g_f^{P\mathbf{w}}(t_j) \to g_f^{\mathbf{w}} \in \mathbb{R}[\mathbf{x}]_{2d}$ as $j \to \infty$ where $g_f^{\mathbf{w}}$ is as in (4.14). And of course, as $\|g_f^{P\mathbf{w}}(t) - g_f^{\mathbf{w}}\|_{\mathbf{w}} \to 0$, $g_f^{\mathbf{w}}$ is in the closure $\overline{P(g) \cap \mathbb{R}[\mathbf{x}]_{2d}}$ of $P(g) \cap \mathbb{R}[\mathbf{x}]_{2d}$.

Next, suppose that there exists $h \in \overline{P(g) \cap \mathbb{R}[\mathbf{x}]_{2d}}$ with $\|f - h\|_{\mathbf{w}} < \|f - g_f^{\mathbf{w}}\|_{\mathbf{w}}$. Then there exists a sequence $(h_t) \subset \mathbb{R}[\mathbf{x}]_{2d}$, $t \in \mathbb{N}$, with $h_t \in P_t^d(g)$ such that $\|h_t - h\|_{\mathbf{w}} \to 0$ as $t \to \infty$. But then

$$\|f - h\|_{\mathbf{w}} \,=\, \|f - h_t + h_t - h\|_{\mathbf{w}} \,\geq\, \underbrace{\|f - h_t\|_{\mathbf{w}}}_{\geq \|f - g_f^{P\mathbf{w}}(t)\|_{\mathbf{w}}} - \underbrace{\|h_t - h\|_{\mathbf{w}}}_{\to 0 \text{ as } t\to\infty}, \qquad \forall\, t,$$

and so taking limit as $t \to \infty$ yields the contradiction

$$\|f - g_f^{\mathbf{w}}\|_{\mathbf{w}} \,>\, \|f - h\|_{\mathbf{w}} \,\geq\, \lim_{t\to\infty} \|f - g_f^{P\mathbf{w}}(t)\|_{\mathbf{w}} \,=\, \|f - g_f^{\mathbf{w}}\|_{\mathbf{w}}.$$

Therefore $\|f - g_f^{\mathbf{w}}\|_{\mathbf{w}} = \min_h \{\|f - h\|_{\mathbf{w}} \,:\, h \in \overline{P(g) \cap \mathbb{R}[\mathbf{x}]_{2d}}\}$.

The last statement (when **K** is compact) follows from Schmüdgen's Positivstellensatz [15] which implies that if f is nonnegative on **K** then $f + \epsilon \in P(g)$ for every $\epsilon > 0$. □

Of course there is an analogue of Corollary 4.3 with $Q(g)$ in lieu of $P(g)$. The only change is concerned with the last statement where one needs that $Q(g)$ is Archimedean. And also, if $w_\alpha = 1$ for every $\alpha \in \mathbb{N}^n$, then $g_f^{\mathbf{w}}$ in (4.14) simplifies to the form (4.8).

4.2. **The sequential closures of $P(g)$ and $Q(g)$.** Recall that for any convex cone $A \subset \mathbb{R}[\mathbf{x}]$

$$(4.16)\qquad A^{\ddagger} \,=\, \{f \in \mathbb{R}[\mathbf{x}] \,:\, \exists\, q \,\in \mathbb{R}[\mathbf{x}] \text{ s.t. } f + \epsilon\, q \in A, \quad \forall \epsilon > 0\,\}.$$

We have seen that $P(g) \subset P(g)^{\ddagger} \subseteq \overline{P(g)}$, where $\overline{A}$ denotes the closure of A for the finest locally convex topology. Therefore, it is of particular interest

to describe $P(g)^\ddagger$, which the goal of this section. We know that

$$P(g)^\ddagger = \bigcup_{d\in\mathbb{N}} \overline{P(g)\cap\mathbb{R}[\mathbf{x}]_d}, \tag{4.17}$$

and if for instance $P(g)$ has an algebraic interior then $\overline{P(g)} = P(g)^\ddagger$. (See e.g. Kuhlmann and Marshall [9, Prop. 1.4] and Cimpric et al. [5, Prop. 1.3].)

Notice that $Q(g)^\ddagger = \bigcup_{d\in\mathbb{N}} \overline{Q(g)\cap\mathbb{R}[\mathbf{x}]_d}$ ([9]) and by [5, Prop. 1.3], we also have $\overline{Q(g)} = Q(g)^\ddagger$ if $Q(g)$ is archimedean.

Theorem 4.4. *Assume that $\mathbf{K}\subseteq\mathbb{R}^n$ in (2.2) has a nonempty interior and let $f\in\mathbb{R}[\mathbf{x}]$. Then:*

(a) *$f\in P(g)^\ddagger$ if and only if there is some $d\in\mathbb{N}$ such that for every $\epsilon>0$, the polynomial*

$$\mathbf{x}\mapsto f(\mathbf{x}) + \epsilon\left(1+\sum_{i=1}^n x_i^{2d}\right) \quad \textit{is in } P(g). \tag{4.18}$$

(b) *The same statement as (a) also holds with $Q(g)$ instead of $P(g)$.*

In other words, $q\in\mathbb{R}[\mathbf{x}]$ in (4.16) for $A=P(g)$ can be taken as $\mathbf{x}\mapsto(1+\sum_{i=1}^n x_i^{2d})$ independently of $\mathbf{K}$. The dependence of q on f is through the power d only.

Proof. (a) From (4.17) $f\in P(g)^\ddagger$ if and only if $f\in\overline{P(g)\cap\mathbb{R}[\mathbf{x}]_{2d}}$ for some $d\in\mathbb{N}$. Next, by Corollary 4.3, $f\in\overline{P(g)\cap\mathbb{R}[\mathbf{x}]_{2d}}$ if and only if the polynomial $g_f^{\mathbf{w}}\in\mathbb{R}[\mathbf{x}]_{2d}$ defined in (4.14) (and which simplifies to (4.8) when $w_\alpha=1$ for all $\alpha\in\mathbb{N}^n$) is identical to f. But then this implies that the polynomial $g_f^{P\mathbf{w}}(t)\in P_t(g)\cap\mathbb{R}[\mathbf{x}]_{2d}$ in (4.15) is such that $\sum_{i=0}^n\lambda_i^{P\mathbf{w}}(t)\to 0$ as $t\to\infty$ (recall that $w_\alpha=1$ for all $\alpha\in\mathbb{N}^n$). Let $\lambda(t):=\max_i[\lambda_i^{P\mathbf{w}}(t)]$ so that $\lambda(t)\to 0$ as $t\to\infty$, and the polynomial

$$\mathbf{x}\mapsto f(\mathbf{x}) + \lambda(t)\left(1+\sum_{i=1}^n x_i^{2d}\right)$$

belongs to $P_t(g)\cap\mathbb{R}[\mathbf{x}]_{2d}$ because

$$f+\lambda(t)\left(1+\sum_{i=1}^n x_i^{2d}\right) = g_f^{P\mathbf{w}}(t) + \underbrace{\lambda(t)-\lambda_0^P(t)}_{\geq 0} + \sum_{i=1}^n \underbrace{(\lambda(t)-\lambda_f^P(t))}_{\geq 0} x_i^{2d}.$$

Therefore, for every $\epsilon>0$, choosing t_ϵ such that $\lambda(t_\epsilon)\leq\epsilon$ ensures that the polynomial $f+\epsilon(1+\sum_{i=1}^n x_i^{2d})$ is in $P_{t_\epsilon}(g)\cap\mathbb{R}[\mathbf{x}]_{2d}$, which implies the desired result in (a).

The proof of (b) is omitted as it follows similar arguments. Indeed, it was already mentioned that Theorem 4.2 and Corollary 4.3 have obvious analogues for the quadratic module $Q(g)$. □

Of course in (4.18) one may replace $1+\sum_{i=1}^n x_i^{2d}$ with $1+\sum_{i=1}^n \sum_{k=1}^d \frac{x_i^{2k}}{(2k)!}$. In fact, and as pointed out by an anonymous referee, in (a) one may also replace $1+\sum_{i=1}^n x_i^{2d}$ with any point in the interior of $P_d(g)$ in $\mathbb{R}[\mathbf{x}]_{2d}$ and in (b) with any point in the interior of $Q_d(g)$ in $\mathbb{R}[\mathbf{x}]_{2d}$. The fact that $1+\sum_{i=1}^n x_i^{2d}$ is an interior point of $\Sigma[\mathbf{x}]_d$ in $\mathbb{R}[\mathbf{x}]_{2d}$ (and hence also an interior point of $P_d(g)$ and $Q_d(g)$ in $\mathbb{R}[\mathbf{x}]_{2d}$) seems to be well-known. For instance, it can be deduced from [11, Theorem 3].

4.3. **A Positivstellensatz for non compact K.** As we know how to project with the $\ell_{\mathbf{w}}$-norm, we are now able to obtain the following Positivstellensatz on **K**.

Corollary 4.5. *Let $\mathbf{K} \subseteq \mathbb{R}^n$ in (2.2) be nonempty interior. Then $f \geq 0$ on $\mathbf{K}$ if and only if for every $\epsilon > 0$ there exists $d \in \mathbb{N}$ such that*

$$\mathbf{x} \mapsto f(\mathbf{x}) + \epsilon \left(1 + \sum_{i=1}^n \sum_{k=1}^d \frac{x_i^{2k}}{(2k)!}\right) \in P(g). \tag{4.19}$$

Proof. The *only if part:* Recall that $P(g) = \bigcup_{d\geq 0} P_d(g)$, and from Theorem 3.3, $\mathrm{Psd}(\mathbf{K}) = \mathrm{cl}_{\mathbf{w}}(P(g))$. Let $g_f^{P\mathbf{w}}(d) \in \mathbb{R}[\mathbf{x}]_{2d}$ be the canonical $\ell_{\mathbf{w}}$-projection of f onto $P_d(g)$ given in (4.2), where $p_{d\mathbf{w}} = \lambda_0^{P\mathbf{w}} + \sum_{i=1}^n \sum_{k=1}^d \lambda_{ik}^{P\mathbf{w}}$. As $f \in \mathrm{cl}_{\mathbf{w}}(P(g))$, we necessarily have $\lim_{d\to\infty} p_{d\mathbf{w}} = 0$, because $\|f - g_f^{P\mathbf{w}}(d)\|_{\mathbf{w}} \to 0$. Hence given $\epsilon > 0$, let d be such that $\max_{i,k} \lambda_{ik}^{P\mathbf{w}} \leq \epsilon$. Then

$$f + \epsilon \left(1 + \sum_{i=1}^n \sum_{k=1}^d \frac{x_i^{2k}}{(2k)!}\right) = g_f^{P\mathbf{w}}(d) + \underbrace{(\epsilon - \lambda_0^{P\mathbf{w}}) + \sum_{i=1}^n \sum_{k=1}^d (\epsilon - \lambda_{ik}^{P\mathbf{w}}) \frac{x_i^{2k}}{(2k)!}}_{\in \Sigma[\mathbf{x}]},$$

and so $f + \epsilon(1 + \sum_{i=1}^n \sum_{k=1}^d \frac{x_i^{2k}}{(2k)!}) \in P(g)$.

The *if part.* Let $q_d \in \mathbb{R}[\mathbf{x}]$ be the polynomial in (4.19), and let $\mathbf{x} \in \mathbf{K}$ be fixed, arbitrary. Then $0 \leq q_d(\mathbf{x}) \leq f(\mathbf{x}) + \epsilon \sum_{i=1}^n \exp|x_i|$. Therefore, letting $\epsilon \to 0$ yields $f(\mathbf{x}) \geq 0$. □

5. Appendix

Lemma 5.1. *Let μ a finite Borel measure whose sequence of moments $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, is such that for all $i = 1, \ldots, n$, and all $k \in \mathbb{N}$, $L_{\mathbf{y}}(\mathbf{x}_i^{2k}) \leq (2k!)M$ for some M. Let $f \in \mathbb{R}[\mathbf{x}]$ be such that $L_{\mathbf{y}}(\mathbf{x}_i^{2t} f) \geq 0$ for all $i = 1, \ldots, n$, and all $t \in \mathbb{N}$. Then the sequence $\mathbf{z}^f = (z_\alpha^f)$, $\alpha \in \mathbb{N}^n$, where $z_\alpha^f = L_{\mathbf{z}^f}(\mathbf{x}^\alpha) := L_{\mathbf{y}}(\mathbf{x}^\alpha f)$ for all $\alpha \in \mathbb{N}^n$, satisfies Carleman's condition (2.6).*

Proof. Let $1 \leq i \leq n$ be fixed arbitrary, and let $2s \geq \deg f$. Observe that whenever $|\alpha| \leq k$, $|\mathbf{x}|^\alpha \leq |\mathbf{x}_j|^k$ on the subset $W_j := \{\mathbf{x} \in \mathbb{R}^n \setminus [-1,1]^n$:

$|x_j| = \max_i |x_i|\}$. And so, $|f(\mathbf{x})| \le \|f\|_1 \, |\mathbf{x}_j|^{2s}$ for all $\mathbf{x} \in W_j$. Hence,

$$\begin{aligned} L_{\mathbf{z}^f}(x_i^{2k}) &= \int f(\mathbf{x})\, x_i^{2k} d\mu(\mathbf{x}) \\ &\le \int_{[-1,1]^n} |f(\mathbf{x})|\, x_i^{2k} d\mu(\mathbf{x}) + \|f\|_1 \sum_{j=1}^n \int_{W_j} x_j^{2(k+s)} d\mu(\mathbf{x}) \\ &\le M\|f\|_1 + Mn\|f\|_1\,(2(k+s))! \;\le\; 2Mn\|f\|_1\,(2(k+s))!, \end{aligned}$$

and so we have

$$\begin{aligned} L_{\mathbf{z}^f}(x_i^{2k})^{-1/2k} &\ge (2Mn\|f\|_1)^{-1/2k} \left(((2(k+s))!)^{-1/2(k+s)}\right)^{(k+s)/k} \\ &\ge \frac{1}{2}\left(((2(k+s))!)^{-1/2(k+s)}\right)^{(k+s)/k} \\ &\ge \frac{1}{2}\left(\frac{1}{2(k+s)}\right)^{(k+s)/k}, \end{aligned}$$

where $k \ge k_0$ is sufficiently large so that $(2Mn\|f\|_1)^{-1/2k} \ge 1/2$. Therefore,

$$\sum_{k=1}^{\infty} L_{\mathbf{z}^f}(x_i^{2k})^{-1/2k} \ge \frac{1}{2}\sum_{k=k_0}^{\infty}\left(\frac{1}{2(k+s)}\right)^{(k+s)/k} = +\infty.$$

where the last equality follows from $(\frac{1}{2(k+s)})^{(k+s)/k} = (\frac{1}{2(k+s)})(\frac{1}{2(k+s)})^{s/k}$ and $(\frac{1}{2(k+s)})^{s/k} \ge 1/2$ whenever k is sufficiently large, say $k \ge k_1$. Hence the sequence $\mathbf{z}^f$ satisfies Carleman's condition (2.6). □

Acknowledgement. The author wishes to thank an anonymous referee for several comments and suggestions that helped writing the final version.

References

[1] R. Ash. *Real Analysis and Probability*, Academic Press, Inc., Boston (1972).
[2] C. Berg, J.P.R. Christensen and P. Ressel, Positive definite functions on Abelian semigroups. Math. Ann. **223**, 253–274 (1976)
[3] C. Berg, The multidimensional moment problem and semigroups. Proc. Symp. Appl. Math. **37**, 110–124 (1987).
[4] G. Blekherman, There are significantly more nonnegative polynomials than sums of squares, Isr. J. Math. **153**, 355-380 (2006)
[5] J. Cimpric, M. Marshall, T. Netzer, Closures of quadratic modules, Israel J. Math., to appear.
[6] M. Ghasemi, S. Kuhlmann, E. Samei, The moment problem for continuous positive semidefinite linear functionals, `arXiv:1010.279v3`, November 2010.
[7] D. Henrion, J.B. Lasserre and J. Lofberg, Gloptipoly 3: moments, optimization and semidefinite programming, Optim. Methods and Software **24**, 761–779 (2009)
[8] S. Kuhlmann, M. Marshall, Positivity sums of squares and the multidimensional moment problem, Trans. Amer. Math. Soc. **354**, 4285–4301 (2002)
[9] S. Kuhlmann, M. Marshall, N. Schwartz, Positivity sums of squares and the multidimensional moment problem II, Adv. Geom. **5**, 583–606 (2005)
[10] J.B. Lasserre and T. Netzer, SOS approximations of nonnegative polynomials via simple high degree perturbations, Math. Z. **256**, 99–112 (2006)

[11] J.B. Lasserre, Sufficient conditions for a real polynomial to be a sum of squares, Arch. Math. **89**, 390–398 (2007)
[12] V. powers, C. Scheiderer, The moment problem for non-compact semialgebraic sets, Adv. Geom. **1**, 71–88 (2001)
[13] M. Putinar, Positive polynomials on compact sets, Ind. Univ. Math. J. **42**, 969–984 (1993).
[14] C. Scheiderer, Positivity and sums of squares: A guide to recent results. In: *Emerging Applications of Algebraic Geometry* (M. Putinar, S. Sullivant, eds.), IMA Volumes Math. Appl. **149**, Springer, 2009, pp. 271-324.
[15] K. Schmüdgen, The K-moment problem for compact semi-algebraic sets, Math. Ann. **289**, 203–206 (1991).
[16] K. Schmüdgen, Positive cones in enveloping algebras, Rep. Math. Physics **14**, 385–404 (1978).
[17] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49–95 (1996)

LAAS-CNRS and Institute of Mathematics, University of Toulouse, LAAS, 7 avenue du Colonel Roche, 31077 Toulouse Cédex 4,France
E-mail address: lasserre@laas.fr