A Sparse Flat Extension Theorem for Moment Matrices

Monique Laurent and Bernard Mourrain

Abstract. In this note we prove a generalization of the flat extension theorem of Curto and Fialkow [4] for truncated moment matrices. It applies to moment matrices indexed by an arbitrary set of monomials and its border, assuming that this set is connected to 1. When formulated in a basis-free setting, this gives an equivalent result for truncated Hankel operators.

Mathematics Subject Classification (2000). Primary 30E05; Secondary 12D10.

Keywords. Truncated moment problem, moment matrix, Hankel operator, polynomial optimization.

1. Introduction

Throughout this note, $\mathbb{K}$ denotes a field, $\mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1, \ldots, x_n]$ is the ring of multivariate polynomials in n variables $\mathbf{x} = (x_1, \ldots, x_n)$ with coefficients in $\mathbb{K}$, $\mathcal{M}_n = \{\mathbf{x}^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n} \mid \alpha \in \mathbb{N}^n\}$ is the set of monomials in the variables $\mathbf{x}$, and $\mathcal{M}_{n,t}$ (resp., $\mathbb{K}[\mathbf{x}]_t$) is the set of monomials (resp., of polynomials) of degree at most t. The dual basis of $\mathcal{M}_n$ in the dual space $\mathbb{K}[\mathbf{x}]^*$ is denoted as $\mathcal{D}_n = \{\mathbf{d}^\beta \mid \beta \in \mathbb{N}^n\}$. The natural action of $\mathbb{K}[\mathbf{x}]$ on $\mathbb{K}[\mathbf{x}]^*$ is denoted by

$$(p, \Lambda) \in \mathbb{K}[\mathbf{x}] \times \mathbb{K}[\mathbf{x}]^* \mapsto p \cdot \Lambda \in \mathbb{K}[\mathbf{x}]^*$$

where $(p \cdot \Lambda)(q) := \Lambda(pq)$ for $q \in \mathbb{K}[\mathbf{x}]$.

1.1. The moment problem

In this section, we consider $\mathbb{K} = \mathbb{R}$. The moment problem (see e.g. [1, 7]) deals with the characterization of the sequences of moments of measures. Given a probability measure μ on $\mathbb{R}^n$, its moment of order $a = \mathbf{x}^\alpha \in \mathcal{M}_n$ is the quantity $\int \mathbf{x}^\alpha \mu(dx)$. The moment problem concerns the characterization of the sequences $y = (y_a)_{a \in \mathcal{M}_n}$ that are the sequences of moments of some nonnegative measure μ, in which case one says that μ is a representing measure for y, with $y_1 = 1$ if μ is a probability measure. Let $\Lambda \in \mathbb{R}[\mathbf{x}]^*$ denote the linear form on $\mathbb{R}[\mathbf{x}]$ associated to the sequence

y, defined by $\Lambda(p) = \sum_a p_a y_a$ for any polynomial $p = \sum_{a \in \mathcal{M}_n} p_a a \in \mathbb{R}[\mathbf{x}]$. Then, y has a representing measure μ precisely when Λ is given by $\Lambda(p) = \int p(x)\mu(dx)$ for all $p \in \mathbb{R}[\mathbf{x}]$. A well known necessary condition for the existence of a representing measure is the positivity of Λ, i.e. $\Lambda(p^2) \geq 0$ for all $p \in \mathbb{R}[\mathbf{x}]$, which is equivalent to requiring that the matrix $M(y) := (y_{ab})_{a,b \in \mathcal{M}_n}$ be positive semidefinite. As is well known this necessary condition is also sufficient in the univariate case ($n = 1$) (Hamburger's theorem), but it is not sufficient in the multivariate case ($n \geq 2$). However, positivity is sufficient for the existence of a representing measure under some additional assumptions. This is the case, for instance, when the sequence y is bounded [2] or, more generally, exponentially bounded [3]. The next result of Curto and Fialkow [4] shows that this is also the case when the matrix $M(y)$ has finite rank (cf. also [14, 15] for a short proof).

Theorem 1.1. [4] *If $M(y)$ is positive semidefinite and the rank of $M(y)$ is finite, then y has a (unique) representing measure (which is finitely atomic with* rank $M(y)$ *atoms).*

In the univariate case $n = 1$, a matrix of the form $M(y)$ is a Hankel matrix. In the multivariate case, $M(y)$ is known as a *generalized Hankel matrix* (see [17]) or *moment matrix* (see [15]). One can also define *truncated* moment matrices: A matrix M indexed by a subset $\mathcal{C} \subseteq \mathcal{M}_n$ is said to be a *moment matrix* if $M_{a,b} = M_{a',b'}$ for all $a, b, a', b' \in \mathcal{C}$ with $ab = a'b'$. Thus its entries are given by a sequence $y = (y_c)_{c \in \mathcal{C} \cdot \mathcal{C}}$, where $\mathcal{C} \cdot \mathcal{C} := \{ab \mid a, b \in \mathcal{C}\}$, and we can write $M = M_{\mathcal{C}}(y)$. When $\mathcal{C} = \mathcal{M}_{n,t}$, we also write $M = M_t(y)$, where the entries of y are indexed by $\mathcal{M}_{n,2t}$. Such matrices arise naturally in the context of the truncated moment problem, which asks for the existence of a representing measure for a truncated sequence indexed by a subset of monomials. A solution to the truncated moment problem would in fact imply a solution to the moment problem. Indeed, Stochel [19] shows that a sequence $y = (y_a)_{a \in \mathcal{M}_n}$ has a representing measure if and only if the truncated sequence $(y_a)_{a \in \mathcal{M}_{n,t}}$ has a representing measure for all $t \in \mathbb{N}$.

1.2. The flat extension theorem of Curto and Fialkow

Curto and Fialkow studied intensively the truncated moment problem (cf. e.g. [4, 5, 6] and further references therein). In particular, they observed that the notion of *flat extension* of matrices plays a central role in this problem. Given matrices $M_{\mathcal{C}}$ and $M_{\mathcal{B}}$ indexed, respectively, by $\mathcal{C}$ and $\mathcal{B} \subseteq \mathcal{C}$, $M_{\mathcal{C}}$ is said to be a *flat extension* of $M_{\mathcal{B}}$ if $M_{\mathcal{B}}$ coincides with the principal submatrix of $M_{\mathcal{C}}$ indexed by $\mathcal{B}$ and rank $M_{\mathcal{C}}$ = rank $M_{\mathcal{B}}$. Curto and Fialkow [4] show the following result for truncated moment matrices.

Theorem 1.2 (The flat extension theorem [4]**).** *For a sequence $y = (y_a)_{a \in \mathcal{M}_{n,2t}}$, if $M_t(y)$ is a flat extension of $M_{t-1}(y)$, then there exists a (unique) sequence $\tilde{y} = (\tilde{y}_a)_{a \in \mathcal{M}_n}$ for which $M(\tilde{y})$ is a flat extension of $M_t(y)$.*

The flat extension theorem combined with Theorem 1.1 directly implies the following sufficient condition for existence of a representing measure.

Corollary 1.3. *For a sequence $y = (y_a)_{a\in\mathcal{M}_{n,2t}}$, if $M_t(y)$ is positive semidefinite and $M_t(y)$ is a flat extension of $M_{t-1}(y)$, then y has a representing measure.*

Curto and Fialkow [5] show moreover that the flat extension condition is in some sense necessary and sufficient for the existence of a representing measure. More precisely, they show that a sequence $y = (y_a)_{a\in\mathcal{M}_{n,2t}}$ has a representing measure if and only if it can be extended to a sequence $y' = (y'_a)_{a\in\mathcal{M}_{n,2t+2k+2}}$ (for some $k \geq 0$) for which $M_{t+k+1}(y')$ is a flat extension of $M_{t+k}(y')$.

The proof of Theorem 1.2 relies on a "truncated ideal like" property of the kernel of flat moment matrices (see (2.2) below). This permits to set up a linear system of equations in order to construct the flat extension $M_{t+1}(\tilde{y})$ of $M_t(y)$ (and then iteratively the infinite flat extension $M(\tilde{y})$). See also [15] for an exposition of this proof. Schweighofer [18] proposes an alternative proof which is less technical and relies on properties of Gröbner bases. We propose in this note another simple alternative proof, which applies more generally to truncated moment matrices indexed by (suitable) general monomial sets (see Theorem 1.4).

1.3. A generalized flat extension theorem

We need some definitions to state our extension of Theorem 1.2. For $\mathcal{C} \subseteq \mathcal{M}_n$,

$$\mathcal{C}^+ := \mathcal{C} \cup \bigcup_{i=1}^{n} x_i\mathcal{C} = \{m, x_1m, \ldots, x_nm \mid m \in \mathcal{C}\} \quad \text{and} \quad \partial\mathcal{C} := \mathcal{C}^+ \setminus \mathcal{C}$$

are called, respectively, the *closure* and the *border* of $\mathcal{C}$. The set $\mathcal{C} \subseteq \mathcal{M}_n$ is said to be *connected to 1* if $1 \in \mathcal{C}$ and every monomial $m \in \mathcal{C} \setminus \{1\}$ can be written as $m = x_{i_1} \cdots x_{i_k}$ with $x_{i_1}, x_{i_1}x_{i_2}, \ldots, x_{i_1} \cdots x_{i_k} \in \mathcal{C}$. For instance, $\mathcal{C}$ is connected to 1 if $\mathcal{C}$ is closed under taking divisions. For example, $\{1, x_2, x_1x_2\}$ is connected to 1 but $\{1, x_1x_2\}$ is not. We now state our main result.

Theorem 1.4. *Consider a sequence $y = (y_a)_{a\in\mathcal{C}^+\cdot\mathcal{C}^+}$, where $\mathcal{C} \subseteq \mathcal{M}_n$ is finite and connected to 1. If $M_{\mathcal{C}^+}(y)$ is a flat extension of $M_{\mathcal{C}}(y)$, then there exists a (unique) sequence $\tilde{y} = (\tilde{y})_{a\in\mathcal{M}_n}$ for which $M(\tilde{y})$ is a flat extension of $M_{\mathcal{C}^+}(y)$.*

The proof is delayed till Section 2. Note that Theorem 1.2 follows directly from Theorem 1.4 applied to the case $\mathcal{C} = \mathcal{M}_{n,t-1}$. Thus our result can be seen as a *sparse* version of Theorem 1.2, which applies to a more general monomial set $\mathcal{C}$, not necessarily the full set of monomials up to a given degree. We now give an example showing that the assumption that $\mathcal{C}$ is connected to 1 cannot be omitted.

Example. For $n = 1$, consider the set $\mathcal{C} = \{1, x^3\}$, which is not connected to 1, with $\partial\mathcal{C} = \{x, x^4\}$. Consider the sequence $y \in \mathbb{R}^{\mathcal{C}^+\cdot\mathcal{C}^+}$ defined by $y_1 = y_x = y_{x^2} = 1$, $y_{x^3} = y_{x^4} = y_{x^5} = a$ and $y_{x^6} = y_{x^7} = y_{x^8} = b$, where a, b are scalars with $b \neq a^2$. Then, $\operatorname{rank} M_{\mathcal{C}^+}(y) = \operatorname{rank} M_{\mathcal{C}}(y) = 2$. If there is a flat extension $M(\tilde{y})$ of $M_{\mathcal{C}^+}(y)$,

then its principal submatrix indexed by $\mathcal{C}^+ \cup \{x^2\}$ has the form:

$$M_{\mathcal{C}^+\cup\{x^2\}}(\tilde{y}) = \begin{array}{c} \\ 1 \\ x^3 \\ x \\ x^4 \\ x^2 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & x^3 & x & x^4 & x^2 \end{array} \\ \begin{pmatrix} 1 & a & 1 & a & 1 \\ a & b & a & b & a \\ 1 & a & 1 & a & a \\ a & b & a & b & b \\ 1 & a & a & b & a \end{pmatrix} \end{array}$$

However, $1-x \in \ker M_{\mathcal{C}^+}(y)$ implies $x-x^2 \in \ker M_{\mathcal{C}^+\cup\{x^2\}}(\tilde{y})$ (see (2.2)) and thus $1 = a = b$, contradicting our choice $b \neq a^2$. Hence no flat extension exists.

1.4. Basis-free reformulation

Here we reformulate our result in a basis-free setting. Moment matrices correspond indeed to choosing the monomial basis $\mathcal{M}_n$ in the polynomial ring $\mathbb{K}[\mathbf{x}]$ and its dual basis $\mathcal{D}_n$ in the dual space $\mathbb{K}[\mathbf{x}]^*$. Given $\Lambda \in \mathbb{K}[\mathbf{x}]^*$, the operator

$$\begin{array}{rccc} H_\Lambda : & \mathbb{K}[\mathbf{x}] & \to & \mathbb{K}[\mathbf{x}]^* \\ & p & \mapsto & p \cdot \Lambda \end{array}$$

is known as a *Hankel operator*. Its matrix with respect to the bases $\mathcal{M}_n$ and $\mathcal{D}_n$ is precisely the moment matrix $(\Lambda(x^{\alpha+\beta}))_{\alpha,\beta\in\mathcal{M}_n} = M(y)$ of the sequence $y = (\Lambda(a))_{a\in\mathcal{M}_n}$. The kernel of H_Λ,

$$\ker\, H_\Lambda = \{p \in \mathbb{K}[\mathbf{x}] \mid \Lambda(pq) = 0 \ \forall q \in \mathbb{K}[\mathbf{x}]\},$$

is an ideal in $\mathbb{K}[\mathbf{x}]$. Moreover, when $\mathbb{K} = \mathbb{R}$ and Λ is positive, i.e. when $\Lambda(p^2) \geq 0$ for all $p \in \mathbb{R}[\mathbf{x}]$, $\ker H_\Lambda$ is a real radical ideal [14]. Theorem 1.1 means that $\Lambda \in \mathbb{R}[\mathbf{x}]^*$ is positive with $\operatorname{rank} H_\Lambda < \infty$ if and only if there exists a nonnegative finite atomic measure μ for which $\Lambda(p) = \int p(x)\mu(dx)$ for all $p \in \mathbb{R}[\mathbf{x}]$.

Truncated Hankel operators can be analogously defined. Given $\mathcal{C} \subseteq \mathcal{M}_n$ and $\Lambda \in (\mathrm{Span}(\mathcal{C}^+ \cdot \mathcal{C}^+))^*$, the corresponding Hankel operator is

$$\begin{array}{rccc} H_\Lambda^{\mathcal{C}^+} : & \mathrm{Span}(\mathcal{C}^+) & \to & \mathrm{Span}(\mathcal{C}^+)^* \\ & p & \mapsto & p \cdot \Lambda \end{array}$$

and its restriction to $\mathrm{Span}(\mathcal{C})$ is $H_\Lambda^{\mathcal{C}} : \mathrm{Span}(\mathcal{C}) \to \mathrm{Span}(\mathcal{C})^*$. We have the following mappings:

$$\frac{\mathrm{Span}(\mathcal{C})}{\ker\, H_\Lambda^{\mathcal{C}}} \xleftarrow{\sigma_1} \frac{\mathrm{Span}(\mathcal{C})}{\ker\, H_\Lambda^{\mathcal{C}^+} \cap \mathrm{Span}(\mathcal{C})} \xrightarrow{\sigma_2} \frac{\mathrm{Span}(\mathcal{C}^+)}{\ker\, H_\Lambda^{\mathcal{C}^+}} \tag{1.1}$$

where σ_1 is onto and σ_2 is one-to-one, so that

$$\dim\, \frac{\mathrm{Span}(\mathcal{C})}{\ker\, H_\Lambda^{\mathcal{C}}} \ \leq\ \dim\, \frac{\mathrm{Span}(\mathcal{C})}{\ker\, H_\Lambda^{\mathcal{C}^+} \cap \mathrm{Span}(\mathcal{C})} \ \leq\ \dim\, \frac{\mathrm{Span}(\mathcal{C}^+)}{\ker\, H_\Lambda^{\mathcal{C}^+}}. \tag{1.2}$$

Thus, $\operatorname{rank} H_\Lambda^{\mathcal{C}^+} = \operatorname{rank} H_\Lambda^{\mathcal{C}}$ (in which case we also say that $H_\Lambda^{\mathcal{C}^+}$ is a flat extension of $H_\Lambda^{\mathcal{C}}$) if and only if equality holds throughout in (1.2), i.e. both σ_1 and σ_2 in (1.1) are isomorphisms or, equivalently, if

$$\mathrm{Span}(\mathcal{C}^+) = \mathrm{Span}(\mathcal{C}) + \ker H_\Lambda^{\mathcal{C}^+} \quad \text{and} \quad \ker\, H_\Lambda^{\mathcal{C}} = \ker\, H_\Lambda^{\mathcal{C}^+} \cap \mathrm{Span}(\mathcal{C}).$$

Theorem 1.4 can be reformulated as follows.

Theorem 1.5. *Let* $\Lambda \in (\mathrm{Span}(\mathcal{C}^+ \cdot \mathcal{C}^+))^*$*, where* $\mathcal{C} \subseteq \mathcal{M}_n$ *is finite and connected to 1, and assume that* $\mathrm{rank}\ H_\Lambda^{\mathcal{C}^+} = \mathrm{rank}\ H_\Lambda^{\mathcal{C}}$*. Then there exists (a unique)* $\tilde{\Lambda} \in \mathbb{R}[\mathbf{x}]^*$ *for which* $H_{\tilde{\Lambda}}$ *is a flat extension of* $H_\Lambda^{\mathcal{C}^+}$*, i.e.* $\tilde{\Lambda}$ *coincides with* Λ *on* $\mathrm{Span}(\mathcal{C}^+ \cdot \mathcal{C}^+)$ *and* $\mathrm{rank}\ H_{\tilde{\Lambda}} = \mathrm{rank}\ H_\Lambda^{\mathcal{C}^+}$*.*

1.5. Border bases and commuting multiplication operators

We recall here a result of [16] about border bases of polynomial ideals that we exploit to prove our flat extension theorem. Let $\mathcal{B} := \{b_1, \ldots, b_N\}$ be a finite set of monomials. Assume that, for each border monomial $x_i b_j \in \partial\mathcal{B}$, we are given a polynomial of the form

$$g^{(ij)} := x_i b_j - \sum_{h=1}^{N} a_h^{(ij)} b_h \quad \text{where} \quad a_h^{(ij)} \in \mathbb{K}.$$

The set

$$F := \{g^{(ij)} \mid i = 1, \ldots, n,\ j = 1, \ldots, N \ \text{ with } x_i b_j \in \partial\mathcal{B}\} \tag{1.3}$$

is known as a *border prebasis* [9] or a *rewriting family* for $\mathcal{B}$ [16]. When the set $\mathcal{B}$ contains the constant monomial 1, one can easily verify that $\mathcal{B}$ is a generating set for the quotient space $\mathbb{K}[\mathbf{x}]/(F)$, where (F) is the ideal generated by the set F. When $\mathcal{B}$ is connected to 1, Theorem 1.6 below characterizes the case when $\mathcal{B}$ is a basis of $\mathbb{K}[\mathbf{x}]/(F)$, in which case F is said to be a *border basis* of the ideal (F). For this, for each $i = 1, \ldots, n$, consider the linear operator:

$$\begin{array}{rccl} \chi_i : & \mathrm{Span}(\mathcal{B}) & \to & \mathrm{Span}(\mathcal{B}) \\ & b_j & \mapsto & \chi_i(b_j) = \left\{ \begin{array}{ll} x_i b_j & \text{if } x_i b_j \in \mathcal{B}, \\ \sum_{h=1}^{N} a_h^{(ij)} b_h & \text{if } x_i b_j \in \partial\mathcal{B} \end{array} \right. \end{array} \tag{1.4}$$

extended to $\mathrm{Span}(\mathcal{B})$ by linearity. When $\mathcal{B}$ is a basis of $\mathbb{K}[\mathbf{x}]/(F)$, χ_i corresponds to the "multiplication operator by x_i" from $\mathbb{K}[\mathbf{x}]/(F)$ to $\mathbb{K}[\mathbf{x}]/(F)$ and thus the operators $\chi_1, \ldots, \chi_n$ commute pairwise. The next result of [16] shows that the converse implication holds when $\mathcal{B}$ is connected to 1; this was also proved later in [9] when $\mathcal{B}$ is closed under taking divisions.

Theorem 1.6. [16] *Let* $\mathcal{B} \subseteq \mathcal{M}_n$ *be a finite set of monomials which is connected to 1, let* F *be a rewriting family for* $\mathcal{B}$ *as in* (1.3)*, and let* $\chi_1, \ldots, \chi_n$ *be defined as in* (1.4)*. The set* $\mathcal{B}$ *is a basis of the quotient space* $\mathbb{K}[\mathbf{x}]/(F)$ *if and only if the operators* $\chi_1, \ldots, \chi_n$ *commute pairwise.*

The proof of our sparse flat extansion theorem is an adaptation of this result to kernels of Hankel operators, where we omit the assumption that B is connected to 1.

1.6. Contents of the paper

Section 2 contains the proof of our generalized flat extension theorem and we mention some applications in Section 3. In particular, we observe that Theorem 1.2 is an 'easy' instance of our flat extension theorem (since one can prove existence of a basis connected to 1). We also point out the relevance of the flat extension theorem to polynomial optimization and to the problem of computing real roots to systems of polynomial equations.

2. Proof of the flat extension theorem

We give here the proof of Theorem 1.5 (equivalently, of Theorem 1.4). We will often use the following simple observations, which follow directly from the assumption that rank $H_\Lambda^{\mathcal{C}^+} = \text{rank } H_\Lambda^{\mathcal{C}}$: For all $p \in \text{Span}(\mathcal{C}^+)$,

$$p \in \ker H_\Lambda^{\mathcal{C}^+} \overset{\text{def.}}{\iff} \Lambda(ap) = 0 \ \forall a \in \mathcal{C}^+ \iff \Lambda(ap) = 0 \ \forall a \in \mathcal{C}, \tag{2.1}$$

$$p \in \ker H_\Lambda^{\mathcal{C}^+} \ \text{ and } \ x_i p \in \text{Span}(\mathcal{C}^+) \implies x_i p \in \ker H_\Lambda^{\mathcal{C}^+}. \tag{2.2}$$

Our objective is to construct a linear form $\tilde{\Lambda} \in \mathbb{K}[\mathbf{x}]^*$ whose Hankel operator $H_{\tilde{\Lambda}}$ is a flat extension of $H_\Lambda^{\mathcal{C}^+}$.

Let $\mathcal{B} \subseteq \mathcal{C}$ for which rank $H_\Lambda^{\mathcal{C}^+} = \text{rank } H_\Lambda^{\mathcal{B}} = |\mathcal{B}|$. Note that we can assume that $1 \in \mathcal{B}$. Indeed, if no such B exists containing 1, then $\Lambda(p) = 0 \ \ \forall p \in \text{Span}(\mathcal{C}^+)$ and one can easily verify that this implies that Λ is identically zero, in which case the theorem trivially holds.

From the assumption: rank $H_\Lambda^{\mathcal{C}^+} = \text{rank } H_\Lambda^{\mathcal{B}} = |\mathcal{B}|$, we have the direct sum decomposition: $\text{Span}(\mathcal{C}^+) = \text{Span}(\mathcal{B}) \oplus \ker H_\Lambda^{\mathcal{C}^+}$, and thus

$$\forall p \in \text{Span}(\mathcal{C}^+) \quad \exists!\, \pi(p) \in \text{Span}(\mathcal{B}) \ \text{ such that } \ f(p) := p - \pi(p) \in \ker H_\Lambda^{\mathcal{C}^+}. \tag{2.3}$$

Then the set

$$F := \{f(m) = m - \pi(m) \mid m \in \partial\mathcal{B}\}$$

is a rewriting family for $\mathcal{B}$ and, for $i = 1, \ldots, n$, the linear operator χ_i in (1.4) maps $p \in \text{Span}(\mathcal{B})$ to $\chi_i(p) = \pi(x_i p) \in \text{Span}(\mathcal{B})$. We show that $\chi_1, \ldots, \chi_n$ commute pairwise. Set $K := \ker H_\Lambda^{\mathcal{C}^+}$.

Lemma 2.1. $\chi_i \circ \chi_j = \chi_j \circ \chi_i$.

Proof. Let $m \in \mathcal{B}$. Write $\pi(x_i m) := \sum_{b\in\mathcal{B}} \lambda_b^i b$ ($\lambda_b^i \in \mathbb{R}$). We have:

$$\begin{aligned}\chi_j \circ \chi_i(m) &= \chi_j(\textstyle\sum_{b\in\mathcal{B}} \lambda_b^i b) = \sum_{b\in\mathcal{B}} \lambda_b^i \chi_j(b) = \sum_{b\in\mathcal{B}} \lambda_b^i (x_j b - f(x_j b)) \\ &= x_j(\textstyle\sum_{b\in\mathcal{B}} \lambda_b^i b) - \sum_{b\in\mathcal{B}} \lambda_b^i f(x_j b) = x_j(x_i m - f(x_i m)) - \sum_{b\in\mathcal{B}} \lambda_b^i f(x_j b).\end{aligned}$$

Therefore,

$$p := \chi_j \circ \chi_i(m) - \chi_i \circ \chi_j(m) = \underbrace{x_i f(x_j m) - x_j f(x_i m)}_{p_1} + \underbrace{\sum_{b\in\mathcal{B}} \lambda_b^j f(x_j b) - \lambda_b^i f(x_i b)}_{p_2}.$$

We show that $p_1 \in K$. Indeed, $\forall a \in \mathcal{C}$, $\Lambda(ap_1) = \Lambda(ax_i f(x_j m) - ax_j(x_i m)) = 0$ since $ax_i, ax_j \in \mathcal{C}^+$ and $f(x_i m),\ f(x_j m) \in K$; by (2.1), this shows that $p_1 \in K$. As $p_2 \in K$ too, this implies $p \in K$ and thus $p = 0$, because $p \in \mathrm{Span}(\mathcal{B})$. □

Our objective now is to show that $\mathcal{B}$ is a basis of $\mathbb{K}[\mathbf{x}]/(F)$ and that, if $\tilde{\pi}$ denotes the projection from $\mathbb{K}[\mathbf{x}]$ onto $\mathrm{Span}(\mathcal{B})$ along (F), then the operator $\tilde{\Lambda}$ defined by $\tilde{\Lambda}(p) = \Lambda(\tilde{\pi}(p))$ for $p \in \mathbb{K}[\mathbf{x}]$, defines the desired flat extension of Λ. Note that when $\mathcal{B}$ is connected to 1, Theorem 1.6 implies directly that $\mathcal{B}$ is a basis of $\mathbb{K}[\mathbf{x}]/(F)$. As we do not assume $\mathcal{B}$ connected to 1, we cannot apply Theorem 1.6, but our arguments below are inspired from its proof. In particular, we construct the projection $\tilde{\pi}$ via the mapping φ from (2.4) below.

As the χ_i's commute, the operator $f(\chi) := f(\chi_1, \ldots, \chi_n)$ is well defined for any polynomial $f \in \mathbb{K}[\mathbf{x}]$. Then $\mathbb{K}[\mathbf{x}]$ acts on $\mathrm{Span}(\mathcal{B})$ by

$$(f, p) \in \mathbb{K}[\mathbf{x}] \times \mathrm{Span}(\mathcal{B}) \mapsto f(\chi)(p) \in \mathrm{Span}(\mathcal{B}).$$

Recall that $1 \in B$. The mapping

$$\begin{array}{rccc} \varphi: & \mathbb{K}[\mathbf{x}] & \to & \mathrm{Span}(\mathcal{B}) \\ & f & \mapsto & f(\chi)(1) \end{array} \tag{2.4}$$

is a homomorphism and, by the following property,

$$\varphi(fg) = f(\chi)(g(\chi)(1)) = f(\chi)(\varphi(g))\ \forall f, g \in \mathbb{K}[\mathbf{x}], \tag{2.5}$$

$\ker \varphi$ is an ideal in $\mathbb{K}[\mathbf{x}]$. We now prove that φ coincide on $\mathrm{Span}(\mathcal{C}^+)$ with the projection π on $\mathrm{Span}(\mathcal{B})$ along $K = \ker H_\Lambda^{\mathcal{C}^+}$.

Lemma 2.2. *For any element $m \in \mathcal{C}^+$, $\varphi(m) = \pi(m)$.*

Proof. We use induction on the degree of m. If $m = 1$, we have $\varphi(1) = \pi(1) = 1$ since $1 \in \mathcal{B}$. Let $m \neq 1 \in \mathcal{C}^+$. As $\mathcal{C}$ is connected to 1, m is of the form $m = x_i m_1$ for some $m_1 \in \mathcal{C}^+$. By the induction assumption, we have $\varphi(m_1) = \pi(m_1)$. Then,

$$\varphi(m) = \varphi(x_i m_1) = \chi_i(\varphi(m_1)) = \chi_i(\pi(m_1)) = x_i \pi(m_1) - \kappa,$$

with $\kappa \in F \subseteq K$. But we also have

$$m = x_i\, m_1 = x_i(\pi(m_1) + m_1 - \pi(m_1)) = x_i \pi(m_1) + x_i\, \kappa_1$$

where $\kappa_1 = m_1 - \pi(m_1) \in K$. We deduce that

$$m = \varphi(m) + \kappa + x_i\, \kappa_1 = \varphi(m) + \kappa_2$$

with $\kappa_2 = \kappa + x_i\, \kappa_1 \in K^+ \cap \mathrm{Span}(\mathcal{C}^+)$. As $\kappa_1 \in K$ and $x_i \kappa_1 \in \mathrm{Span}(\mathcal{C}^+)$, we deduce using (2.1) that $x_i \kappa_1 \in K$, thus implying $\kappa_2 \in K$. As $\varphi(m) \in \mathrm{Span}(\mathcal{B})$, it coincides with the projection of m on $\mathrm{Span}(\mathcal{B})$ along K. □

This implies directly:

$$\varphi(b) = b, \quad \varphi(x_i b) = \chi_i(b) \quad \forall b \in \mathcal{B} \quad \forall i = 1, \ldots, n, \tag{2.6}$$

$$\Lambda(pq) = \Lambda(p\ \varphi(q)) = \Lambda(\varphi(p)\varphi(q)) \quad \forall p, q \in \mathrm{Span}(\mathcal{C}^+). \tag{2.7}$$

Lemma 2.3. *For all $p, q \in \mathrm{Span}(\mathcal{C}^+)$, $\Lambda(pq) = \Lambda(\varphi(pq))$.*

Proof. We first show by induction on the degree of $m \in \mathcal{C}^+$ that

$$\Lambda(mb) = \Lambda(\varphi(mb)) \quad \forall b \in \mathcal{B}. \tag{2.8}$$

The result is obvious if $m = 1$. Else, as $\mathcal{C}^+$ is connected to 1, we can write $m = x_i m_1$ where $m_1 \in \mathcal{C}^+$. Using first (2.7) and then (2.6), we find:

$$\Lambda(mb) = \Lambda(m_1 x_i b) = \Lambda(m_1 \varphi(x_i b)) = \Lambda(m_1 \chi_i(b)).$$

Next, using first the induction assumption and then (2.5), (2.6), we find:

$$\Lambda(m_1\chi_i(b)) = \Lambda(\varphi(m_1\chi_i(b))) = \Lambda(m_1(\chi)(\chi_i(b))) = \Lambda(m(\chi)(b)) = \Lambda(\varphi(mb)),$$

thus showing (2.8). We can now conclude the proof of the lemma. Let $p, q \in \mathrm{Span}(\mathcal{C}^+)$. Then, using successively (2.7), (2.8), (2.5), (2.6), $\Lambda(pq)$ is equal to

$$\Lambda(p\ \varphi(q)) = \Lambda(\varphi(p\varphi(q))) = \Lambda(p(\chi)(\varphi(\varphi(q)))) = \Lambda(p(\chi)(\varphi(q))) = \Lambda(\varphi(pq)). \quad \square$$

We can now conclude the proof of Theorem 1.5. Let $\tilde{\Lambda}$ be the linear operator on $\mathbb{K}[\mathbf{x}]$ defined by

$$\tilde{\Lambda}(p) := \Lambda(\varphi(p)) \quad \text{for } p \in \mathbb{K}[\mathbf{x}].$$

We show that $H_{\tilde{\Lambda}}$ is the unique flat extension of $H_\Lambda^{\mathcal{C}^+}$.

First, $H_{\tilde{\Lambda}}$ is an extension of $H_\Lambda^{\mathcal{C}^+}$ since, for all $p, q \in \mathrm{Span}(\mathcal{C}^+)$, $\tilde{\Lambda}(pq) = \Lambda(\varphi(pq)) = \Lambda(pq)$ (by Lemma 2.3).

Next, we have $K = \ker H_\Lambda^{\mathcal{C}^+} \subseteq \ker H_{\tilde{\Lambda}}$. Indeed, let $\kappa \in K$. By Lemma 2.2, $\varphi(\kappa) = \pi(\kappa) = 0$. Thus for any $p \in \mathbb{K}[\mathbf{x}]$, we have $\tilde{\Lambda}(p\,\kappa) = \Lambda(\varphi(p\,\kappa)) = \Lambda(p(\chi)(\varphi(\kappa))) = 0$, which shows that $\kappa \in \ker H_{\tilde{\Lambda}}$.

As F is a rewritting family for $\mathcal{B}$ and $\mathcal{B}$ contains 1, $\mathcal{B}$ is a generating set of $\mathbb{K}[\mathbf{x}]/(F)$ and thus $\dim \mathbb{K}[\mathbf{x}]/(F) \leq |\mathcal{B}|$. Set $\mathcal{A}_{\tilde{\Lambda}} := \mathbb{K}[\mathbf{x}]/\ker H_{\tilde{\Lambda}}$. Then, as $F \subseteq K \subseteq \ker H_{\tilde{\Lambda}}$, we have $\dim \mathcal{A}_{\tilde{\Lambda}} \leq \dim \mathbb{K}[\mathbf{x}]/(F) \leq |\mathcal{B}|$. On the other hand, $\dim \mathcal{A}_{\tilde{\Lambda}} = \mathrm{rank} H_{\tilde{\Lambda}} \geq \mathrm{rank} H_\Lambda^{\mathcal{B}} = \mathrm{rank} H_\Lambda^{\mathcal{B}} = |\mathcal{B}|$. Therefore, $\dim \mathcal{A}_{\tilde{\Lambda}} = \mathrm{rank} H_{\tilde{\Lambda}} = |\mathcal{B}|$, $\ker H_{\tilde{\Lambda}} = (K)$, $H_{\tilde{\Lambda}}$ is a flat extension of $H_\Lambda^{\mathcal{C}^+}$, and we have the direct sum: $\mathbb{K}[\mathbf{x}] = \mathrm{Span}(\mathcal{B}) \oplus \ker H_{\tilde{\Lambda}}$. Moreover, $\varphi(p)$ is the projection of $p \in \mathbb{K}[\mathbf{x}]$ on $\mathrm{Span}(\mathcal{B})$ along $\ker H_{\tilde{\Lambda}}$. Indeed, $\varphi(p) \in \mathrm{Span}(\mathcal{B})$ and $p - \varphi(p) \in \ker H_{\tilde{\Lambda}}$ for any $p \in \mathbb{K}[\mathbf{x}]$ since, for any $q \in \mathbb{K}[\mathbf{x}]$,

$$\begin{aligned} &\tilde{\Lambda}(pq) = \Lambda(\varphi(pq)) = \Lambda(p(\chi)(\varphi(q))), \\ &\tilde{\Lambda}(p\varphi(q)) = \Lambda(\varphi(p\varphi(q))) = \Lambda(p(\chi)(\varphi(q))) = \tilde{\Lambda}(pq). \end{aligned}$$

Finally, if $\Lambda' \in \mathbb{K}[\mathbf{x}]^*$ is another linear form whose Hankel operator $H_{\Lambda'}$ is a flat extension of $H_\Lambda^{\mathcal{C}^+}$, then $\ker H_{\tilde{\Lambda}} = (K) \subseteq \ker H_{\Lambda'}$. This implies that for all $p \in \mathbb{K}[\mathbf{x}]$, $\Lambda'(p) = \Lambda'(\varphi(p)) = \Lambda(\varphi(p)) = \tilde{\Lambda}(p)$. This shows the unicity of the flat extension of $H_\Lambda^{\mathcal{C}^+}$, which concludes the proof of Theorem 1.5.

3. Applications

3.1. Application to the flat extension theorem of Curto and Fialkow

Theorem 1.2 is in some sense an 'easy' instance of Theorem 1.4. Indeed, under its assumptions, one can show existence of a maximum rank principal submatrix of $M_{t-1}(y)$ indexed by a monomial set $\mathcal{B}$ connected to 1 which, as noted in the proof of Theorem 1.4, permits to apply Theorem 1.6.

Proposition 3.1. *Let* $\Lambda \in (\mathrm{Span}(\mathcal{C}^+ \cdot \mathcal{C}^+))^*$, *where* $\mathcal{C} := \mathcal{M}_{n,t-1}$. *If* $\mathrm{rank}\ H_\Lambda^{\mathcal{C}^+} = \mathrm{rank}\ H_\Lambda^{\mathcal{C}}$, *then there exists* $\mathcal{B} \subseteq \mathcal{C}$ *closed under taking divisions (and thus connected to 1) for which* $\mathrm{rank}\ H_\Lambda^{\mathcal{C}^+} = \mathrm{rank}\ H_\Lambda^{\mathcal{B}} = |\mathcal{B}|$.

Proof. Let $M = (\Lambda(ab))_{a,b\in\mathcal{C}^+}$ denote the matrix of $H_\Lambda^{\mathcal{C}^+}$ in the canonical bases. Consider a total degree monomial ordering $\preceq$ of $\mathcal{C}$ and let $\mathcal{B} \subseteq \mathcal{C}$ index a maximum linearly independent set of columns of M which is constructed by the greedy algorithm using the ordering $\preceq$. One can easily verify that $\mathcal{B}$ is closed under taking divisions (cf. [13]). □

The following example shows that, even if $\mathcal{C}$ is connected to 1, there may not always exist a base $\mathcal{B}$ connected to 1 for $H_\Lambda^{\mathcal{C}}$ (which justifies our generalisation of Theorem 1.6 to kernels of Hankel operators).

Example. For $n = 2$, let $\mathcal{C} = \{1, x_1, x_1x_2\}$ with $\partial\mathcal{C} = \{x_2, x_1x_2^2, x_1^2, x_1^2x_2\}$, and let $\Lambda \in (\mathrm{Span}(\mathcal{C}^+ \cdot \mathcal{C}^+))^*$ be defined by $\Lambda(x_1^ix_2^j) = 1$ if $j = 0, 1$, and $\Lambda(x_1^ix_2^j) = a$ if $j = 2, 3, 4$, except $\Lambda(x_1^2x_2^4) = a^2$, where a is a scalar with $a \neq 1$. The associated moment matrix has the form

$$\begin{array}{c} \\ 1 \\ x_1 \\ x_1x_2 \\ x_1^2 \\ x_1^2x_2 \\ x_2 \\ x_1x_2^2 \end{array} \begin{array}{c} \begin{array}{ccccccc} 1 & x_1 & x_1x_2 & x_1^2 & x_1^2x_2 & x_2 & x_1x_2^2 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & a \\ 1 & 1 & 1 & 1 & 1 & 1 & a \\ 1 & 1 & a & 1 & a & a & a \\ 1 & 1 & 1 & 1 & 1 & 1 & a \\ 1 & 1 & a & 1 & a & a & a \\ 1 & 1 & a & 1 & a & a & a \\ a & a & a & a & a & a & a^2 \end{pmatrix} \end{array}$$

and $\mathrm{rank}\ H_\Lambda^{\mathcal{C}} = \mathrm{rank}\ H_\Lambda^{\mathcal{C}^+} = 2$. As $1 - x_1 \in \ker\ H_\Lambda^{\mathcal{C}^+}$, the only sets indexing a column base for $H_\Lambda^{\mathcal{C}}$ are $\mathcal{B} = \{1, x_1x_2\}$ and $\{x_1, x_1x_2\}$, thus not connected to 1.

Combining Theorem 1.5 with Theorem 1.1 we obtain the following extension of Corollary 1.3.

Theorem 3.2. *Let* $\Lambda \in (\mathrm{Span}(\mathcal{C}^+ \cdot \mathcal{C}^+))^*$, *where* $\mathcal{C} \subseteq \mathcal{M}_n$ *is finite and connected to 1. Assume that* Λ *is positive and that* $\mathrm{rank}\ H_\Lambda^{\mathcal{C}^+} = \mathrm{rank}\ H_\Lambda^{\mathcal{C}}$. *Then the sequence* $y = (\Lambda(a))_{a\in\mathcal{C}^+\cdot\mathcal{C}^+}$ *has a representing measure.*

3.2. Application to polynomial optimization

We point out here the relevance of the flat extension theorems to polynomial optimization and to the problem of computing the real roots to polynomial equations. In this section, we take again $\mathbb{K} = \mathbb{R}$.

The truncated moment problem has recently attracted a lot of attention also within the optimization community, since it can be used to formulate semidefinite programming relaxations to polynomial optimization problems (see [11]). Moreover the flat extension theorem of Curto and Fialkow permits to detect optimality of the relaxations and to extract global optimizers to the original optimization problem (see [8]). Here is a brief sketch; see e.g. [15] and references therein for details.

Suppose we want to compute the infimum p^* of a polynomial p over a semi-algebraic set K defined by the polynomial inequalities $g_1 \geq 0, \ldots, g_m \geq 0$. For any integer $t \geq \deg(p)/2$ and such that $t \geq d_j := \lceil \deg(g_j)/2 \rceil$, consider the program:

$$p_t^* := \inf \ \Lambda(p) \ \text{ s.t. } \Lambda \in (\mathbb{R}[\mathbf{x}]_{2t})^*, \ \Lambda(1) = 1, \ \Lambda \succeq 0, \ g_j \cdot \Lambda \succeq 0 \ (\forall j \leq m). \quad (3.1)$$

Here, $\Lambda \succeq 0$ means that Λ is positive (i.e., $\Lambda(p^2) \geq 0$ for all $p \in \mathbb{R}[\mathbf{x}]_t$) and the localizing conditions $g_j \cdot \Lambda \succeq 0$ (i.e. $\Lambda(g_j p^2) \geq 0$ for all $p \in \mathbb{R}[\mathbf{x}]_{t-d_j}$) aim to restrict the search for a representing measure suported by the set K (cf. [6, 11]). Using moment matrices, the program (3.1) can be formulated as an instance of semidefinite programming for which efficient algorithms exist (see e.g. [20, 21]). We have: $p_t^* \leq p^*$, with equality if $H_\Lambda^{\mathcal{M}_{n,t}}$ is a flat extension of $H_\Lambda^{\mathcal{M}_{n,t-d}}$ for an optimum solution Λ to (3.1) ($d := \max_j d_j$). In that case, the atoms of the representing measure (which exists by Corollary 1.3) are global minimizers of p over the semi-algebraic set K and they can be computed from Λ [8]. Moreover, they are *all* the global minimizers when $H_\Lambda^{\mathcal{M}_{n,t}}$ has the maximum possible rank among all optimum solutions to the semidefinite program (3.1).

As shown in [12], the truncated moment problem also yields an algorithmic approach to the problem of computing the real roots to polynomial equations $g_1 = 0, \ldots, g_m = 0$ (assuming their number is finite). Indeed, this amounts to finding all global minimizers to a constant polynomial, say $p = 0$, over the real variety $K := \{x \in \mathbb{R}^n \mid g_j(x) = 0 \ \forall j = 1, \ldots, m\}$. Consider the semidefinite program (3.1) where the localizing conditions now read $g_j \cdot \Lambda = 0 \ \forall j$. For t large enough, the program (3.1) has a maximum rank solution which is a flat extension and thus, as noted above, all points of K can be computed from this solution. See [12] for details.

A concern in this type of approach is the size of the matrices appearing in the semidefinite program (3.1). In order to improve the practical applicability of this approach, it is crucial to derive semidefinite programs involving matrices of moderate sizes. For this one may want to consider moment matrices indexed by *sparse* sets of monomials instead of the full degree levels $\mathcal{M}_{n,t}$. This is where our new sparse flat extension theorem may become very useful. It will be used, in particular, in [10].

The approach in [12] also permits to find the real radical of the ideal generated by the polynomials $g_1, \ldots, g_m$. Indeed, if $\Lambda \in (\mathbb{R}[\mathbf{x}])^*$ is positive, then the kernel of its Hankel operator H_Λ is a real radical ideal [14] and, under the conditions of Theorem 3.2, $\ker H_\Lambda^{\mathcal{C}^+}$ generates a real radical ideal. These facts explain the relevance of moment matrices and Hankel operators to the problem of finding the real radical of a polynomial ideal. For instance, this permits to weaken the assumptions in Proposition 4.1 of [12] and to strengthen its conclusions; more precisely, we do not need to assume the commutativity of the operators χ_i's (as this holds automatically, by Lemma 2.1) and we can claim that the returned ideal is real radical (by the above argument).

References

[1] N.I. Akhiezer. *The Classical Moment Problem.* Hafner, New York, 1965.

[2] C. Berg, J.P.R. Christensen, and P. Ressel. Positive definite functions on Abelian semigroups. *Mathematische Annalen*, 223:253–272, 1976.

[3] C. Berg and P.H. Maserick. Exponentially bounded positive definite functions. *Illinois Journal of Mathematics*, 28:162–179, 1984.

[4] R.E. Curto and L.A. Fialkow. Solution of the truncated complex moment problem for flat data. *Memoirs of the American Mathematical Society,* vol. 119, Amer. Math. Soc., Providence, RI, 1996.

[5] R.E. Curto and L.A. Fialkow. Flat extensions of positive moment matrices: recursively generated relations. *Memoirs of the American Mathematical Society,* vol. 648, Amer. Math. Soc., Providence, RI, 1998.

[6] R.E. Curto and L.A. Fialkow. The truncated complex K-moment problem. *Trans. Amer. Math. Soc.*, 352:28252855, 2000.

[7] B. Fuglede. The multidimensional moment problem. *Expositiones Mathematicae*, 1:47–65, 1983.

[8] D. Henrion and J.B. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. In *Positive Polynomials in Control*, D. Henrion and A. Garulli (eds.), Lecture Notes on Control and Information Sciences, 312:293–310, Springer, Berlin, 2005.

[9] A. Kehrein, M. Kreuzer, and L. Robbiano. An algebraist's view on border bases. In *Solving Polynomial Equations - Foundations, Algorithms and Applications*, A. Dickenstein and I.Z. Emiris (eds.), pages 169–202. Springer, 2005.

[10] J.B. Lasserre, M. Laurent, B. Mourrain, P. Rostalski, and P. Trebuchet. Moment matrices, border bases and real radical ideals. In preparation, 2009.

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

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

[13] J.B. Lasserre, M. Laurent, and P. Rostalski. A prolongation-projection algorithm for computing the finite real variety of an ideal. arXiv:0806.3874v1, June 2008.

[14] M. Laurent. Revisiting two theorems of Curto and Fialkow on moment matrices. *Proceedings of the American Mathematical Society* 133(10):2965–2976, 2005.

[15] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry,* Vol. 149 of IMA Volumes in Mathematics and its Applications, M. Putinar and S. Sullivant (eds.), Springer, pages 157-270, 2009.

[16] B. Mourrain. A new criterion for normal form algorithms. In *Proceedings of the 13th International Symposium on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes,* H. Imai, S. Lin, and A. Poli (eds.), vol. 1719 of *Lecture Notes In Computer Science*, pages 430–443. Springer Verlag, 1999.

[17] B. Mourrain and V.Y. Pan. Multivariate Polynomials, Duality, and Structured Matrices. *J. Complexity*, 16(1):110–180, 2000.

[18] M. Schweighofer. A Gröbner basis proof of the flat extension theorem for moment matrices. arXiv:0801.4243v1, January 2008.

[19] J. Stochel. Solving the truncated moment problem solves the moment problem. *Glasgow Journal of Mathematics*, 43:335–341, 2001.

[20] L. Vandenberghe and S. Boyd. Semidefinite Programming. *SIAM Review* 38(1):49–95, 1996.

[21] H. Wolkowicz, R. Saigal, L. Vandeberghe (eds.). *Handbook of Semidefinite Programming*, Kluwer, 2000.

Monique Laurent
Centrum Wiskunde & Informatica (CWI)
Kruislaan 413
1098 SJ Amsterdam
The Netherlands
e-mail: M.Laurent@cwi.nl

Bernard Mourrain
GALAAD, INRIA Méditerranée
BP 93, 06902 Sophia Antipolis
France
e-mail: mourrain@sophia.inria.fr