SUMS OF SQUARES CERTIFICATES FOR POLYNOMIAL MOMENT INEQUALITIES

IGOR KLEP, VICTOR MAGRON, AND JURIJ VOLČIČ

Abstract. This paper introduces and develops the algebraic framework of moment polynomials, which are polynomial expressions in commuting variables and their formal mixed moments. Their positivity and optimization over probability measures supported on semialgebraic sets and subject to moment polynomial constraints is investigated. A positive solution to Hilbert's 17th problem for pseudo-moments is given. On the other hand, moment polynomials positive on actual measures are shown to be sums of squares and formal moments of squares up to arbitrarily small perturbation of their coefficients. When only measures supported on a bounded semialgebraic set are considered, a stronger algebraic certificate for moment polynomial positivity is derived. This result gives rise to a converging hierarchy of semidefinite programs for moment polynomial optimization. Finally, as an application, two nonlinear Bell inequalities from quantum physics are settled.

1. Introduction

Let $x_1, \dots, x_n$ be independent variables, and let $\mathtt{m}(x_1^{i_1} \cdots x_n^{i_n})$ denote their formal mixed moments. That is, $\mathtt{m}(x_1^{i_1} \cdots x_n^{i_n})$ are algebraically independent variables, which in the presence of a probability measure μ on $\mathbb{R}^n$ evaluate as $\int x_1^{i_1} \cdots x_n^{i_n} \, \mathrm{d}\mu$. This paper focuses on the class of *moment polynomials*, i.e., polynomials in $x_1, \dots, x_n$ and their formal moments. Problems involving moment polynomial inequalities and optimization arise in various fields, for instance in probability [BBLM05, BP05, MJC+14], statistics [LJ10], operator theory [CP10], economics [PPHI15], industrial organization [KPT21], partial differential equations [FF23, HIKV23], and quantum information theory [PHBB17, TGB21, TPKLR22]. To approach such problems, it is natural to start from the theory built around their moment-free analogs, namely real algebraic geometry [Mar08]. The cornerstone of real algebraic geometry are sums of squares certificates for nonnegative polynomials. Artin's solution of Hilbert's

2020 *Mathematics Subject Classification.* 13J30, 44A60, 60E15, 90C22, 46G12, 47L60, 81-08.

Key words and phrases. Moment polynomial, Positivstellensatz, polynomial optimization, moment problem, semidefinite programming, nonlinear Bell inequality.

IK was supported by the Slovenian Research Agency grants J1-2453, N1-0217, J1-3004 and P1-0222.

VM was supported by the EPOQCS grant funded by the LabEx CIMI (ANR-11-LABX-0040), the FastQI grant funded by the Institut Quantique Occitan, the PHC Proteus grant 46195TA, the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions, grant agreement 813211 (POEMA), by the AI Interdisciplinary Institute ANITI funding, through the French "Investing for the Future PIA3" program under the Grant agreement n° ANR-19-PI3A-0004 as well as by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

JV was supported by the NSF grant DMS-1954709.

17th problem characterizes nonnegative polynomials on $\mathbb{R}^n$ in terms of sums of squares and denominators, and Putinar's Positivstellensatz [Put93] describes polynomials positive on compact semialgebraic sets in $\mathbb{R}^n$. The latter was groundbreakingly applied to polynomial optimization in [Las01], resulting in the so-called Lasserre's hierarchy, based on semidefinite programming. This hierarchy yields a sequence of nondecreasing lower bounds converging to the global infimum of a polynomial over a compact semialgebraic set. Positive polynomials also play a crucial role in functional analysis and measure theory through moment problems [Sch17]. The duality between polynomials positive on a semialgebraic set K and measures supported on K connects sums of squares certificates with necessary conditions for solvability of the moment problem on K. The monographs [Las09, HKL20] present many applications of the moment problem, Lasserre's hierarchy and its variations. More recent developments in this field concern nonlinear expressions in moments, and infinite-dimensional moment problem. In [BRS$^+$22], techniques of tropical geometry are applied to nonnegative polynomials and moment problems, resulting in classification of moment *binomial* inequalities. In the recent work [HIKV23], nonlinear partial differential equations are formulated as moment problems for measures supported on infinite-dimensional vector spaces, and then results about the infinite-dimensional moment problem in nuclear spaces [IKR14, IKKM23] are leveraged to derive converging approximations of solutions of differential equations.

In the noncommutative setting, the Helton-McCullough Positivstellensatz [HM04] leads to similar methods for optimizing eigenvalues of polynomials in matrix or operator variables [BKP16]. The famous Navascués-Pironio-Acín hierarchy [NPA08] yields bounds over the maximal violation levels of linear Bell inequalities, which also relates to the so-called quantum moment problem [DLTW08]. Motivated by the more difficult study of nonlinear Bell inequalities [PHBB17] for correlations in quantum networks [TPKLR22], the three authors have recently proposed two nonlinear extensions to optimization problems over trace [KMV22] and state polynomials [KMVW23], derived from Positivstellensätze for polynomials in noncommuting variables and formal traces or states of their products. In this paper, we let noncommutative real algebraic geometry offer a new perspective on commutative problems involving moment polynomials.

Moment polynomials. This paper investigates positivity and optimization of moment polynomials subject to polynomial relations between the problem variables x_j and their formal mixed moments. For example,

$$f = \mathtt{m}(x_1x_2^3)x_1x_2 - \mathtt{m}(x_1^2)^3x_2^2 + x_2 - \mathtt{m}(x_2)\mathtt{m}(x_1x_2) - 2$$

is a moment polynomial; at a probability measure μ on $\mathbb{R}^2$ with fourth order moments and a pair $(X_1, X_2) \in \mathbb{R}^2$, f evaluates as

$$f\big(\mu,(X_1,X_2)\big) = X_1X_2\int x_1x_2^3\,\mathrm{d}\mu - X_2^2\left(\int x_1^2\,\mathrm{d}\mu\right)^3 + X_2 - \int x_2\,\mathrm{d}\mu\int x_1x_2\,\mathrm{d}\mu - 2.$$

A moment polynomial without freely occurring x_j, e.g. $\mathtt{m}(x_1^2x_2^2)-\mathtt{m}(x_1)^4+\mathtt{m}(x_1)\mathtt{m}(x_2)\mathtt{m}(x_1x_2)$, is called *pure*. The algebra of pure moment polynomials is denoted by $\mathscr{M}$, and the algebra of moment polynomials is denoted by $\mathscr{M}[\underline{x}]$. There is a natural $\mathscr{M}$-linear map $\mathtt{m} : \mathscr{M}[\underline{x}] \to \mathscr{M}$ that corresponds to formal integration. While the algebras $\mathscr{M}$ and $\mathscr{M}[\underline{x}]$ are infinitely generated, we emphasize that the aforementioned infinite-dimensional moment problems [IKR14, IKKM23] are incompatible with our setup. Namely, as we focus on positivity of moment polynomials over measures on finite-dimensional semialgebraic sets, our algebraic certificates of positivity are based on the finitely many underlying variables $x_1, \dots, x_n$.

Let $S_1 \subseteq \mathbb{R}[\underline{x}]$ and $S_2 \subseteq \mathscr{M}$. Let $K(S_1)$ be the set of points $\underline{X} \in \mathbb{R}^n$ such that all polynomials in S_1 are nonnegative at $\underline{X}$. Let $\mathbf{P}(K(S_1))$ be the set of all Borel probability measures supported on $K(S_1)$, and let $\boldsymbol{\mathcal{K}}(S_1, S_2)$ be the set of measures $\mu \in \mathbf{P}(K(S_1))$ such that all pure moment polynomials in S_2 are nonnegative at μ. Adapting a standard notion from real algebra [Mar08], we define the *quadratic module* $\mathrm{QM}(S_1, S_2) \subseteq \mathscr{M}[\underline{x}]$ as the convex hull of

$$\Big\{ f^2 s_1,\ f^2 \mathtt{m}(p^2 s_1),\ f^2 s_2 \colon s_i \in S_i \cup \{1\},\ p \in \mathbb{R}[\underline{x}],\ f \in \mathscr{M}[\underline{x}] \Big\}.$$

Elements of $\mathrm{QM}(S_1, S_2)$ are clearly nonnegative on $\boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$. This paper addresses the converse, and provides certificates for moment polynomial positivity on $\boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$ in terms of $\mathrm{QM}(S_1, S_2)$.

Main results. The first positivity certificate applies to archimedean quadratic modules. Here, $\mathrm{QM}(S_1, S_2)$ is *archimedean* if $N - x_1^2 - \cdots - x_n^2 \in \mathrm{QM}(S_1, S_2)$ for some $N \in \mathbb{N}$. Note that the constrained set $K(S_1)$ is bounded in this instance. Conversely, if $K(S_1)$ is contained in a ball of radius R, we may add $R^2 - x_1^2 - \cdots - x_n^2$ to S_1 to obtain an archimedean quadratic module without shrinking $\mathcal{K}(S_1, S_2) \times K(S_1)$.

Theorem A (Theorem 4.2)**.** *If $\mathrm{QM}(S_1, S_2)$ is archimedean, the following are equivalent for $f \in \mathscr{M}[\underline{x}]$:*

(i) *$f \geq 0$ on $\boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$;*
(ii) *$f + \varepsilon \in \mathrm{QM}(S_1, S_2)$ for every $\varepsilon > 0$.*

Theorem A is proved using results from real algebra and the solution of the moment problem for compactly supported measures. Analogously to Lasserre's hierarchy [Las01] leveraging Putinar's Positivstellensatz [Put93] in polynomial optimization, we utilize Theorem A to derive a moment polynomial optimization procedure based on semidefinite programming.

Theorem B (Corollary 5.2)**.** *Let $\mathrm{QM}(S_1, S_2)$ be archimedean and f a moment polynomial. The Positivstellensatz-induced hierarchy of semidefinite programs produces a nondecreasing sequence converging to the infimum of f on $\boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$.*

We apply Theorem B to moment polynomial optimization problems from quantum information theory. Two nonlinear Bell inequalities proposed in [PHBB17] and [TGB21, TPKLR22] are established. For example, our optimization scheme allows us to solve the following problem:

$$\begin{aligned}
&\sup\ \frac{1}{3} \sum_{i \in \{1,2,3\}} \Big(\mathtt{m}(x_{i+3} x_{i+6}) - \mathtt{m}(x_i x_{i+3}) \Big) - \sum_{\{i,j,k\} = \{1,2,3\}} \mathtt{m}(x_i x_{j+3} x_{k+6}) \\
&\text{subject to} \\
&\mathtt{m}(x_1^{k_1} x_2^{k_2} x_3^{k_3} x_7^{k_4} x_8^{k_5} x_9^{k_6}) = \mathtt{m}(x_1^{k_1} x_2^{k_2} x_3^{k_3}) \mathtt{m}(x_7^{k_4} x_8^{k_5} x_9^{k_6}) \quad \text{for } k_i \in \{0,1\}, \\
&x_j^2 = 1 \text{ and } \mathtt{m}(x_j) = 0 \quad \text{for } j \in \{0, \dots, 9\}, \\
&\mathtt{m}(x_i x_{j+3}) = \mathtt{m}(x_{i+3} x_{j+6}) = 0 \quad \text{for } i,j \in \{1,2,3\},\ i \neq j, \\
&\mathtt{m}(x_i x_{j+3} x_{k+6}) = 0 \quad \text{for } i,j,k \in \{1,2,3\},\ |\{i,j,k\}| \leq 2.
\end{aligned} \tag{1.1}$$

In Subsection 5.2 it is shown that the solution of (1.1) is 4, attained by certain binary variables and the uniform measure on 16 points, which answers a question in [TGB21, TPKLR22].

Next, we address certificates for moment polynomial positivity subject to constraint sets S_1 and S_2 without the archimedean assumption. In particular, we aim to describe everywhere nonnegative moment polynomials (i.e., $S_1 = S_2 = \emptyset$). In this particular case, one might first consider an analog of Hilbert's 17th problem for moment polynomials (H17): if $f \in \mathscr{M}[\underline{x}]$ is nonnegative on $\mathbf{P}(\mathbb{R}^n) \times \mathbb{R}^n$, can we write it as a quotient of sums of products of elements of the form f^2 and $\mathtt{m}(f^2)$ for $f \in \mathscr{M}[\underline{x}]$? It turns out that the answer to this question is negative (cf. Example 3.7). More precisely, the algebraic certificate in (H17) characterizes a strictly smaller class of moment polynomials that are nonnegative under *pseudo-moment evaluations*. A pseudo-moment evaluation is a homomorphism $\varphi : \mathscr{M}[\underline{x}] \to \mathbb{R}$ satisfying $\varphi(\mathtt{m}(p^2)) \geq 0$ for all $p \in \mathbb{R}[\underline{x}]$.

Theorem C (Theorem 3.6)**.** *The following are equivalent for $f \in \mathscr{M}[\underline{x}]$:*

(i) *$\varphi(f) \geq 0$ for every pseudo-moment evaluation φ;*
(ii) *f is a quotient of sums of products of elements of the form h^2 and $\mathtt{m}(h^2)$ for $h \in \mathscr{M}[\underline{x}]$.*

The proof of Theorem C relies on the Krivine-Stengle Positivstellensatz and extensions of positive functionals. The negative answer to (H17) for moment evaluations motivates a search for a different positivity certificate. In [BRS+22], nonnegative moment *binomials* are classified in combinatorial terms. For nonnegative (classical) polynomials in $\mathbb{R}[\underline{x}]$, Lasserre [Las06] showed that they become sums of squares of polynomials after an arbitrary small perturbation of their coefficients. Our second main result generalizes Lasserre's certificate to moment polynomials.

Theorem D (Theorem 6.7)**.** *If S_2 is finite, the following are equivalent for $f \in \mathscr{M}[\underline{x}]$:*

(i) *$f \geq 0$ on $\boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$;*
(ii) *for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that*

$$f + \varepsilon \sum_{j=1}^{n} \sum_{k=0}^{r} \frac{1}{k!} \left(x_j^{2k} + \mathtt{m}(x_j^{2k}) \right) \in \mathrm{QM}(S_1, S_2).$$

The proof of Theorem D uses constructions and techniques from functional analysis, conic programming and duality. When Theorem D is restricted to polynomials in $\mathbb{R}[\underline{x}]$, it improves the approximation result in [LN07], which was established under several additional conditions.

2. Preliminaries

We start by introducing the notation and terminology pertaining to moment polynomials and their evaluations. Let $\mathbb{R}[\underline{x}] = \mathbb{R}[x_1, \dots, x_n]$ be the polynomial ring in n variables. Consider the polynomial ring in countably many variables $\mathscr{M} = \mathbb{R}[\mathtt{m}_{i_1,\dots,i_n} : i_j \in \mathbb{N}_0]$ where $\mathtt{m}_{0,\dots,0} := 1$, and denote $\mathscr{M}[\underline{x}] = \mathscr{M} \otimes \mathbb{R}[\underline{x}]$. Elements of $\mathscr{M}[\underline{x}]$ and $\mathscr{M}$ are called *moment polynomials* and *pure moment polynomials*, respectively. There is a canonical unital $\mathscr{M}$-linear map $\mathtt{m} : \mathscr{M}[\underline{x}] \to \mathscr{M}$ determined by $\mathtt{m}(x_1^{i_1} \cdots x_n^{i_n}) = \mathtt{m}_{i_1,\dots,i_n}$. In terms of polynomial functions on vector spaces, one may (for $V = \mathbb{R}^n$) write $\mathbb{R}[\underline{x}] = \mathbb{R}[V]$, $\mathscr{M} = \mathbb{R}[\mathbb{R}[V]]$, $\mathscr{M}[\underline{x}] = \mathbb{R}[\mathbb{R}[V] \times V]$.

Recalling a standard notion from real algebra [Mar08, Section 2.1], a subset M of a commutative unital ring A is a *quadratic module* if $1 \in M$, $M + M \subseteq M$, and $a^2 M \subseteq M$ for $a \in A$. Given $S_1 \subseteq \mathbb{R}[\underline{x}]$ and $S_2 \subseteq \mathscr{M}$ let $\operatorname{qm}(S_1, S_2) \subseteq \mathscr{M}$ be the quadratic module in $\mathscr{M}$ generated by

$$\{\mathtt{m}(p^2 s) \colon p \in \mathbb{R}[\underline{x}], s \in \{1\} \cup S_1\} \cup S_2,$$

and let $\operatorname{QM}(S_1, S_2) \subseteq \mathscr{M}[\underline{x}]$ be the quadratic module in $\mathscr{M}[\underline{x}]$ generated by

$$S_1 \cup \{\mathtt{m}(p^2 s) \colon p \in \mathbb{R}[\underline{x}], s \in \{1\} \cup S_1\} \cup S_2.$$

More concretely, $\operatorname{qm}(S_1, S_2)$ is the convex hull of

$$q^2 \mathtt{m}(p^2 s_1), \quad q^2 s_2 \tag{2.1}$$

for $s_i \in \{1\} \cup S_i$, $p \in \mathbb{R}[\underline{x}]$ and $q \in \mathscr{M}$, and $\operatorname{QM}(S_1, S_2)$ is the convex hull of

$$f^2 s_1, \quad f^2 \mathtt{m}(p^2 s_1), \quad f^2 s_2 \tag{2.2}$$

for $s_i \in S_i \cup \{1\}$, $p \in \mathbb{R}[\underline{x}]$ and $f \in \mathscr{M}[\underline{x}]$. Also, let $M(S_1) \subset \mathbb{R}[\underline{x}]$ denote the quadratic module in $\mathbb{R}[\underline{x}]$ generated by S_1.

Remark 2.1. Observe that $\operatorname{qm}(S_1, S_2) \subseteq \operatorname{QM}(S_1, S_2) \cap \mathscr{M} \subseteq \mathtt{m}(\operatorname{QM}(S_1, S_2))$, and these inclusions are strict in general (the first one because of term cancellations, and the second one because of terms of the form $\mathtt{m}(f^2)\mathtt{m}(p^2)$). Therefore for a pure moment polynomial, membership in $\operatorname{qm}(S_1, S_2)$ is a stronger property than membership in $\operatorname{QM}(S_1, S_2)$. Thus when stating our results for moment polynomials and $\operatorname{QM}(S_1, S_2)$, we also state refinements for pure moment polynomials and $\operatorname{qm}(S_1, S_2)$, and the proofs are analogous. The reason for persisting with $\operatorname{qm}(S_1, S_2)$ is that it leads to smaller optimization problems than $\mathtt{m}(\operatorname{QM}(S_1, S_2))$.

There is a natural notion of a degree deg on $\mathscr{M}[\underline{x}]$ satisfying

$$\deg x_j = 1 \qquad \text{and} \qquad \deg \mathtt{m}_{i_1,\dots,i_n} = i_1 + \cdots + i_n.$$

For $r \in \mathbb{N}$ let $\mathbb{R}[\underline{x}]_r, \mathscr{M}_r, \mathscr{M}[\underline{x}]_r$ be the finite-dimensional subspaces of $\mathbb{R}[\underline{x}], \mathscr{M}, \mathscr{M}[\underline{x}]$ of elements of degree at most r. Also, let $\operatorname{qm}(S_1, S_2)_{2r} \subseteq \mathscr{M}_{2r}$ and $\operatorname{QM}(S_1, S_2)_{2r} \subseteq \mathscr{M}[\underline{x}]_{2r}$ be the convex hulls of elements in $\mathscr{M}[\underline{x}]_{2r}$ of the form (2.1) and (2.2), respectively.

The following lemma identifies certain non-obvious elements of $\operatorname{qm}(\emptyset, \emptyset)$ that are required later.

Lemma 2.2 (Symbolic univariate Hölder's inequality)**.** *Let $n = 1$. Then $\mathtt{m}_{2k} - \mathtt{m}_1^{2k} \in \operatorname{qm}(\emptyset, \emptyset)$ for all $k \in \mathbb{N}$.*

Proof. For $k \in \mathbb{N}$ and $\ell = \lceil \log_2 k \rceil$ let $a_0, \dots, a_{\ell-1}$ be recursively defined as $a_0 = k$ and $a_{i+1} = \lceil \frac{a_i}{2} \rceil$. Denote $\mathrm{r}(a_i) = 0$ if a_i is even and $\mathrm{r}(a_i) = 1$ if a_i is odd. Observe that $a_i + \mathrm{r}(a_i) = 2a_{i+1}$ for $i < \ell - 1$, and $a_{\ell-1} = 2$ if $\ell > 1$. We claim that

$$\mathtt{m}_{2k} - \mathtt{m}_1^{2k} = k\, \mathtt{m}\left(\left(\mathtt{m}_1^k - x_1 \mathtt{m}_1^{k-1}\right)^2\right) + \sum_{i=0}^{\ell-1} 2^i\, \mathtt{m}\left(\left(x_1^{\mathrm{r}(a_i)} \mathtt{m}_1^{k-\mathrm{r}(a_i)} - x_1^{a_i} \mathtt{m}_1^{k-a_i}\right)^2\right). \tag{2.3}$$

Indeed, the right-hand side of (2.3) expands as

$$
\begin{aligned}
&k\left(\mathtt{m}_2\mathtt{m}_1^{2(k-1)}-\mathtt{m}_1^{2k}\right)+\sum_{i=0}^{\ell-1}2^i\left(\mathtt{m}_{2\mathrm{r}(a_i)}\mathtt{m}_1^{2(k-\mathrm{r}(a_i))}-2\mathtt{m}_{a_i+\mathrm{r}(a_i)}\mathtt{m}_1^{2k-(a_i+\mathrm{r}(a_i))}+\mathtt{m}_{2a_i}y_1^{2(k-a_i)}\right)\\
&=k\left(\mathtt{m}_2\mathtt{m}_1^{2(k-1)}-\mathtt{m}_1^{2k}\right)+\sum_{i=0}^{\ell-1}2^i\mathtt{m}_{2\mathrm{r}(a_i)}\mathtt{m}_1^{2(k-\mathrm{r}(a_i))}+\mathtt{m}_{2a_0}\mathtt{m}_1^{2(k-a_0)}-2^\ell\mathtt{m}_{a_{\ell-1}+\mathrm{r}(a_{\ell-1})}\mathtt{m}_1^{2k-(a_{\ell-1}+\mathrm{r}(a_{\ell-1}))}\\
&=k\left(\mathtt{m}_2\mathtt{m}_1^{2(k-1)}-\mathtt{m}_1^{2k}\right)+\mathtt{m}_{2k}-2^\ell\mathtt{m}_2\mathtt{m}_1^{2(k-1)}+\sum_{i=0}^{\ell-1}2^i\mathtt{m}_{2\mathrm{r}(a_i)}\mathtt{m}_1^{2(k-\mathrm{r}(a_i))}\\
&=\mathtt{m}_{2k}-\mathtt{m}_1^{2k}-\left((k-1)\mathtt{m}_1^{2k}+(2^\ell-k)\mathtt{m}_2\mathtt{m}_1^{2(k-1)}\right)+\sum_{i=0}^{\ell-1}2^i\mathtt{m}_{2\mathrm{r}(a_i)}\mathtt{m}_1^{2(k-\mathrm{r}(a_i))}\\
&=\mathtt{m}_{2k}-\mathtt{m}_1^{2k}
\end{aligned}
$$

because $2^\ell-k=\sum_{i=0}^{\ell-1}2^i\mathrm{r}(a_i)$ and $k-1=\sum_{i=0}^{\ell-1}2^i(1-\mathrm{r}(a_i))$. $\square$

2.1. **Moment evaluations of moment polynomials.** There are two natural (and closely related) types of evaluations of moment polynomials. For a closed (but not necessarily bounded) set $K\subseteq\mathbb{R}^n$ let $\mathbf{P}(K)$ denote the set of Borel probability measures μ on $\mathbb{R}^n$ that are supported on K and admit all marginal moments (that is, $\int x_j^{2k}\,\mathrm{d}\mu<\infty$ for all $j=1,\dots,n$ and all $k\in\mathbb{N}$). By Hölder's inequality, it follows that such measures admit all mixed moments. Note that a Borel probability measure on $\mathbb{R}^n$ is always a Radon measure by [Par05, Theorem II.3.2]. Each pair $(\mu,\underline{X})\in\mathbf{P}(\mathbb{R}^n)\times\mathbb{R}^n$ gives rise to the homomorphism

$$\mathscr{M}[\underline{x}]\to\mathbb{R},\quad f\mapsto f(\mu,\underline{X}) \tag{2.4}$$

determined by

$$\mathtt{m}_{i_1,\dots,i_n}\mapsto\int x_1^{i_1}\cdots x_n^{i_n}\,\mathrm{d}\mu,\qquad x_j\mapsto X_j.$$

Such homomorphisms are called *moment evaluations*.

Let $(\mathcal{P},\Sigma,\pi)$ be a probability space. For $p\in\mathbb{N}$ let $\mathcal{L}^p(\mathcal{P},\Sigma,\pi)$ be the space of real-valued random variables F on $\mathcal{P}$ such that $\int|F|^p\,\mathrm{d}\pi<\infty$. There is a partial order $\succeq$ on $\mathcal{L}^p(\mathcal{P},\Sigma,\pi)$, given as $f\succeq g$ if $f\geq g$ almost everywhere. Consider the ring $\mathcal{L}^\omega(\pi):=\bigcap_{p=1}^\infty\mathcal{L}^p(\mathcal{P},\Sigma,\pi)$ introduced in [Are46]. For $\underline{F}=(F_1,\dots,F_n)\in\mathcal{L}^\omega(\pi)^n$, all the mixed moments of $\underline{F}$ exist, and we can define the homomorphism

$$\mathscr{M}[\underline{x}]\to\mathcal{L}^\omega(\pi),\quad f\mapsto f[\pi,\underline{F}] \tag{2.5}$$

determined by

$$\mathtt{m}_{i_1,\dots,i_n}\mapsto\int F_1^{i_1}\cdots F_n^{i_n}\,\mathrm{d}\pi,\qquad x_j\mapsto F_j.$$

The restrictions of homomorphisms (2.4) and (2.5) to $\mathscr{M}$ coincide (when we view $\mathbb{R}^n$ as a probability space with the Borel sigma algebra). Observe that the homomorphism (2.5) intertwines $\mathtt{m}:\mathscr{M}[\underline{x}]\to\mathscr{M}$ and integration with respect to π. In contrast, the homomorphism (2.4) does not satisfy such an intertwining property.

For $S_1\subseteq\mathbb{R}[\underline{x}]$ and $S_2\subseteq\mathscr{M}$ let

$$K(S_1)=\{\underline{X}\in\mathbb{R}^n\colon p(\underline{X})\geq 0\text{ for all }p\in S_1\}\subseteq\mathbb{R}^n,$$

$$\boldsymbol{\mathcal{K}}(S_1,S_2)=\{\mu\in\mathbf{P}(K(S_1))\colon s(\mu)\geq 0 \text{ for all } s\in S_2\}\subseteq\mathbf{P}(K(S_1)).$$

The following proposition indicates that evaluations (2.4) and (2.5) are essentially equivalent from the perspective of moment polynomial positivity.

Proposition 2.3. *Let $S_1\subseteq\mathbb{R}[\underline{x}]$ and $S_2\subseteq\mathscr{M}$. The following are equivalent for $f\in\mathscr{M}[\underline{x}]$:*

(i) *$f(\mu,\underline{X})\geq 0$ for all $(\mu,\underline{X})\in\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$;*

(ii) *$f[\pi,\underline{F}]\succeq 0$ for every probability measure π and a random variable $\underline{F}\in\mathcal{L}^{\omega}(\pi)^n$ with values in $K(S_1)^n$ such that $s[\pi,\underline{F}]\geq 0$ for all $s\in S_2$.*

Proof. (i)⇒(ii): Suppose a probability space $(\mathcal{P},\Sigma,\pi)$ and a random variable $\underline{F}$ on $\mathcal{P}$ with values in $K(S_1)^n$ satisfy $s[\pi,\underline{F}]\geq 0$ for all $s\in S_2$. Let μ be the pushforward of π induced by $\underline{F}$. Then for all $P\in\mathcal{P}$ we have $(\mu,\underline{F}(P))\in\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$ and $f(\mu,\underline{F}(P))=f[\pi,\underline{F}](P)$. Thus (i) implies (ii). Conversely, (i) is a special case of (ii) (where the probability space is $K(S_1)$ endowed with the σ-algebra of Borel sets and the measure μ, and the coordinate functions are considered as random variables), so (ii) implies (i). □

In the rest of the paper, we mostly deal with evaluations of the first type. The following statement is a straightforward consequence of definitions.

Proposition 2.4. *Let $S_1\subseteq\mathbb{R}[\underline{x}]$ and $S_2\subseteq\mathscr{M}$. If $f\in\operatorname{QM}(S_1,S_2)$ then $f\geq 0$ on $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$.*

We conclude this section with a renowned quadrature result, which is also relevant for evaluations of moment polynomials, and is utilized in several subsequent proofs.

Proposition 2.5 (Tchakaloff's theorem [Put97, Theorem 2])**.** *Let $S_1\subseteq\mathbb{R}[\underline{x}]$. For every $\mu\in\mathbf{P}(K(S_1))$ and $d\in\mathbb{N}$ there exists $\nu\in\mathbf{P}(K(S_1))$ with $|\operatorname{supp}\nu|\leq\binom{n+d}{d}$ such that $\mathtt{m}_{i_1,\dots,i_n}(\nu)=\mathtt{m}_{i_1,\dots,i_n}(\mu)$ for all $i_1+\cdots+i_n\leq d$.*

Thus if $S_2\subset\mathscr{M}$ is finite, $f\in\mathscr{M}[\underline{x}]$ and $d=\max\{\deg f,\deg s\colon s\in S_2\}$, then $f\geq 0$ on $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$ if and only if $f\geq 0$ on $\left\{\nu\in\boldsymbol{\mathcal{K}}(S_1,S_2)\colon|\operatorname{supp}\nu|\leq\binom{n+d}{d}\right\}\times K(S_1)$.

Remark 2.6. Proposition 2.5 in principle allows to reformulate a moment polynomial optimization problem as a classical polynomial optimization problem in the following way. Let $S_1\subset\mathbb{R}[\underline{x}]$ and $S_2\subset\mathscr{M}$ be finite, and $f\in\mathscr{M}[\underline{x}]$. Denote $d=\max\{\deg f,\deg s\colon s\in S_2\}$ and $D=\binom{n+d}{d}$. By Proposition 2.5, the infimum of f on $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$ is equal to

$$\begin{aligned}
&\inf_{\underline{X},\underline{Y}_1,\dots,\underline{Y}_D,\alpha_1,\dots,\alpha_D} \; f\left(\sum_{i=1}^{D}\alpha_i\delta_{\underline{Y}_i},\underline{X}\right)\\
&\text{subject to}\\
&s\left(\sum_{i=1}^{D}\alpha_i\delta_{\underline{Y}_i},\underline{X}\right)\geq 0\quad\text{for } s\in S_2,\\
&\underline{X},\underline{Y}_1,\dots,\underline{Y}_D\in K(S_1),\\
&\alpha_1,\dots,\alpha_D\geq 0,\ \sum_{i=1}^{D}\alpha_i=1.
\end{aligned}\tag{2.6}$$

Here, $\delta_{\underline{Y}_i}$ denotes the Dirac delta measure concentrated at $\underline{Y}_i \in \mathbb{R}^n$. While (2.6) minimizes a polynomial function subject to polynomial constraints, and can be thus approached with standard methods of polynomial optimization, it has $n + Dn + D = (\binom{n+d}{d} + 1)(n+1) - 1$ variables. This number quickly rises beyond the capabilities of solvers for global nonlinear optimization (see Subsection 5.1 for a concrete example). Furthermore, the problem (2.6) does not fully utilize the structure of moment polynomials; e.g. $\underline{Y}_i$ can be permuted, α_i always appears jointly with $\underline{Y}_i$, and so on. Moment polynomial optimization procedure bypassing these issues is developed in Section 5 below.

3. Pseudo-moments and Hilbert's 17th problem for moment polynomials

In this section we consider pseudo-moment evaluations of moment polynomials. We give a solution to a natural version of Hilbert's 17th problem for pseudo-moment evaluations (Theorem 3.6). In particular, since positivity on pseudo-moments is stricter than positivity on moments, our solution implies that moment polynomials with nonnegative moment evaluations are not necessarily rational consequences of f^2 and $\mathtt{m}(f^2)$ for $f \in \mathscr{M}[\underline{x}]$.

Let $[\underline{x}]_d$ denote all monomials in $\mathbb{R}[\underline{x}]$ of degree at most d, ordered degree-lexicographically according to $x_1 > \cdots > x_n$. For $d \in \mathbb{N}$ let $H_d = (uv)_{u,v\in[\underline{x}]_d}$ be the *symbolic Hankel matrix* over $\mathbb{R}[\underline{x}]$ of order d. For any map α on $\mathbb{R}[\underline{x}]$, let $\alpha(H_d)$ denote the matrix obtained by applying α entry-wise to H_d.

Lemma 3.1. *Let $\phi : \mathscr{M}[\underline{x}] \to \mathbb{R}$ be a homomorphism of $\mathbb{R}$-algebras. The following are equivalent:*

- (i) $\phi(\mathtt{m}(f^2)) \geq 0$ *for all* $f \in \mathscr{M}[\underline{x}]$*;*
- (ii) $\phi(\mathtt{m}(p^2)) \geq 0$ *for all* $p \in \mathbb{R}[\underline{x}]$*;*
- (iii) $(\phi \circ \mathtt{m})(H_d)$ *is positive semidefinite for all* $d \in \mathbb{N}$*.*

Proof. (ii)⇔(iii) and (i)⇒(ii) are clear. Suppose (ii) holds. Let $f \in \mathscr{M}[\underline{x}]$, and write it as $f = \sum_i q_i p_i$ for $q_i \in \mathscr{M}$ and $p_i \in \mathbb{R}[\underline{x}]$. Then

$$\phi\left(\mathtt{m}(f^2)\right) = \sum_{i,j} \phi(q_i)\phi(q_j)\phi\left(\mathtt{m}(p_ip_j)\right) = \phi\left(\mathtt{m}\left(\Big(\sum_i \phi(q_i)p_i\Big)^2\right)\right) \geq 0,$$

so (i) holds. □

Homomorphisms satisfying the equivalent conditions in Lemma 3.1 are called *pseudo-moment evaluations* of moment polynomials. Note that a *pseudo-moment evaluation* $\phi : \mathscr{M}[\underline{x}] \to \mathbb{R}$ is uniquely determined by $\phi(x_j)$ for $j = 1, \dots, n$, and a unital linear functional $L : \mathbb{R}[\underline{x}] \to \mathbb{R}$ given by $L(p) = \phi(\mathtt{m}(p))$ and satisfying $L(p^2) \geq 0$ for all $p \in \mathbb{R}[\underline{x}]$.

Remark 3.2. There is a certain nuance in Lemma 3.1. Namely, the implication (ii)⇒(i) fails in general for homomorphisms $\mathscr{M}[\underline{x}] \to R$ where R is a closed real field containing $\mathbb{R}$, even when $n = 1$. Indeed, by [KPV21, Example 2.6] there exist a real closed field R and a homomorphism $\phi : \mathscr{M} \to R$ such that $\phi(\mathtt{m}(p^2)) \geq 0$ for all $p \in \mathbb{R}[\underline{x}]$, and $\phi(\mathtt{m}_2 - \mathtt{m}_1^2) < 0$, even though $\mathtt{m}_2 - \mathtt{m}_1^2 = \det \mathtt{m}(H_2) = \mathtt{m}((x_1 - \mathtt{m}_1)^2) \in \mathrm{qm}(\emptyset, \emptyset)$.

Pseudo-moment evaluations form a strictly larger class than moment evaluations. The pure moment polynomial $\mathtt{m}_{4,2}\mathtt{m}_{2,4} - \mathtt{m}_{2,2}^3$ is nonnegative under all moment evaluations, but not under all pseudo-moment evaluations by [BRS+22, Example 4.15] (see Example 3.7 for an alternative argument). Theorem 3.6 below gives a sums of squares certificate with denominators for moment polynomials that are nonnegative under all pseudo-moment evaluations. For a sums of squares certificate with perturbations for moment polynomials that are nonnegative under all moment evaluations, see Theorem 6.7.

The proof of Theorem 3.6 requires some additional terminology from real algebra [Mar08]. A *preordering* P in a commutative unital ring A is a quadratic module closed under multiplication. For $d \in \mathbb{N}$ let $\mathscr{M}_d \subset \mathscr{M}$ be the polynomial ring generated by $\{\mathtt{m}(u) \colon u \in [\underline{x}]_d\}$, and let $\mathscr{M}_d[\underline{x}] = \mathscr{M}_d \otimes \mathbb{R}[\underline{x}]$. Let P_d denote the preordering in $\mathscr{M}_{2d}[\underline{x}]$ generated by the principal minors of $\mathtt{m}(H_d)$, and let Ω denote the preordering in $\mathscr{M}$ generated by $\mathrm{QM}(\emptyset, \emptyset)$.

Lemma 3.3. *Every principal minor of H_d is a quotient of elements in Ω.*

Proof. Straightforward adaptation of [KMVW23, Proposition 4.2]. ☐

Lemma 3.4. *If a unital functional $L : \mathbb{R}[\underline{x}]_{2d} \to \mathbb{R}$ satisfies $L(p^2) > 0$ for $p \in \mathbb{R}[\underline{x}]_d \setminus \{0\}$, then it extends to a unital functional $\widetilde{L} : \mathbb{R}[\underline{x}] \to \mathbb{R}$ satisfying $\widetilde{L}(p^2) > 0$ for $p \in \mathbb{R}[\underline{x}] \setminus \{0\}$.*

Proof. For $\alpha > 0$ consider the linear functional $L_\alpha : \mathbb{R}[\underline{x}]_{2d+2} \to \mathbb{R}$ defined on monomials $u \in [\underline{x}]_{2d+2}$ as follows:

$$L(u) = \begin{cases} L(u) & \text{if } u \in [\underline{x}]_{2d}, \\ 0 & \text{if } u \in [\underline{x}]_{2d+1} \setminus [\underline{x}]_{2d}, \\ \alpha \int_{[0,1]^n} u \, \mathrm{d}x_1 \cdots \mathrm{d}x_n & \text{if } u \in [\underline{x}]_{2d+2} \setminus [\underline{x}]_{2d+1}. \end{cases}$$

Applying L_α entry-wise to H_{d+1} results in

$$L_\alpha(H_{d+1}) = \begin{pmatrix} L(H_d) & B^* \\ B & \alpha K \end{pmatrix},$$

where $L(H_d)$ and K are positive definite matrices. Since $L(H_d)$ is invertible, $L(H_{d+1})$ is positive definite if and only if $\alpha K - B^* L(H_d)^{-1} B$ is positive definite. This is indeed the case for a sufficiently large $\alpha > 0$. Thus we showed that L extends to a functional on $\mathbb{R}[\underline{x}]_{2d+2}$ that is positive on $\mathtt{m}(p^2)$ for $p \in \mathbb{R}[\underline{x}]_{d+1} \setminus \{0\}$. Continuing inductively in this fashion, we obtain $\widetilde{L} : \mathbb{R}[\underline{x}] \to \mathbb{R}$ that extends L and satisfies $\widetilde{L}(p^2) > 0$ for $p \in \mathbb{R}[\underline{x}] \setminus \{0\}$. ☐

Lemma 3.5. *For every $\varepsilon > 0$, $d \in \mathbb{N}$, and a unital linear functional $L : \mathbb{R}[\underline{x}]_{2d} \to \mathbb{R}$ satisfying $L(p^2) \geq 0$ for $p \in \mathbb{R}[\underline{x}]_d$, there exists a unital linear functional $\widetilde{L} : \mathbb{R}[\underline{x}]_{2d} \to \mathbb{R}$ satisfying $\widetilde{L}(p^2) > 0$ for $p \in \mathbb{R}[\underline{x}]_d \setminus \{0\}$, and $|\widetilde{L}(u) - L(u)| < \varepsilon$ for $u \in \mathbb{R}[\underline{x}]_{2d}$.*

Proof. Note that the unital functional $L_0 : \mathbb{R}[\underline{x}] \to \mathbb{R}$ given by $L_0(p) = \int_{[0,1]^n} p \, \mathrm{d}x_1 \cdots \mathrm{d}x_n$ satisfies $L_0(p^2) > 0$ for $p \in \mathbb{R}[\underline{x}] \setminus \{0\}$. Then $\widetilde{L} = (1-\delta)L + \delta L_0$ for a sufficiently small $\delta > 0$ has the desired properties. ☐

Theorem 3.6. *Let $f \in \mathscr{M}[\underline{x}]$. Then all pseudo-moment evaluations of f are nonnegative if and only if f is a quotient of sums of products of elements in $\mathrm{QM}(\emptyset, \emptyset)$.*

Proof. ($\Rightarrow$): Let $d = \deg f$. Assume f is not a quotient of sums of products of elements in $\mathrm{QM}(\emptyset,\emptyset)$. Then f is not a quotient of elements in P_d by Lemma 3.3. By the Krivine-Stengle Positivstellensatz [Mar08, Theorem 2.2.1] there is a homomorphism $\phi : \mathscr{M}_d[\underline{x}] \to \mathbb{R}$ such that $\phi(f) < 0$ and $(\phi \circ \mathtt{m})(H_d)$ is positive semidefinite. Note that ϕ is determined by $\phi(x_j)$ for $j = 1, \dots, n$ and the linear functional $L : \mathbb{R}[\underline{x}]_d \to \mathbb{R}$ given by $L(p) = \phi(\mathtt{m}(p))$. By Lemma 3.5 we can slightly perturb L, so that $L(\mathtt{m}(p^2)) > 0$ for $p \in \mathbb{R}[\underline{x}]_d \setminus \{0\}$, and still $\phi(f) < 0$. By Lemma 3.4, L extends to $\widetilde{L} : \mathbb{R}[\underline{x}] \to \mathbb{R}$ such that $\widetilde{L}(p^2) \geq 0$ for $p \in \mathbb{R}[\underline{x}]$. Define a homomorphism $\widetilde{\phi} : \mathscr{M}[\underline{x}] \to \mathbb{R}$ determined by $\widetilde{\phi}(p) = \phi(p)$ and $\widetilde{\phi}(\mathtt{m}(p)) = \widetilde{L}(p)$ for $p \in \mathbb{R}[\underline{x}]$. Then $\widetilde{\phi}$ is a pseudo-moment evaluation by Lemma 3.1, and $\widetilde{\phi}(f) < 0$.

($\Leftarrow$): Let $f = \frac{g}{h}$ where $g, h \neq 0$ are sums of products of elements in $\mathrm{QM}(\emptyset,\emptyset)$. Suppose $f(\mu, \underline{X}) < 0$ for some $(\mu, \underline{X}) \in \mathbf{P}(\mathbb{R}^n) \times \mathbb{R}^n$. Note that $g(\nu, \underline{Y}), h(\nu, \underline{Y}) \geq 0$ for all $(\nu, \underline{Y}) \in \mathbf{P}(\mathbb{R}^n) \times \mathbb{R}^n$. Since $h \neq 0$, there exists $(\mu', \underline{X}') \in \mathbf{P}(\mathbb{R}^n) \times \mathbb{R}^n$ such that $h(\mu, \underline{X}) > 0$. For $\varepsilon \in [0, 1]$ let $\underline{X}_\varepsilon = (1-\varepsilon)\underline{X} + \varepsilon \underline{X}'$ and $\mu_\varepsilon = (1-\varepsilon)\mu + \varepsilon\mu'$. Then there exists a sufficiently small $\varepsilon > 0$ so that $f(\mu_\varepsilon, \underline{X}_\varepsilon) < 0$ and $h(\mu_\varepsilon, \underline{X}_\varepsilon) > 0$. Then

$$0 > h(\mu_\varepsilon, \underline{X}_\varepsilon) f(\mu_\varepsilon, \underline{X}_\varepsilon) = g(\mu_\varepsilon, \underline{X}_\varepsilon) \geq 0,$$

a contradiction. $\square$

Example 3.7. Let $f = \mathtt{m}_{4,2}\mathtt{m}_{2,4} - \mathtt{m}_{2,2}^3$. All moment evaluations of f are nonnegative by Hölder's inequality. On the other hand, consider the functional $L : \mathbb{R}[\underline{x}]_6 \to \mathbb{R}$ given on the Hankel matrix H_3 as

$$L(H_3) = \begin{pmatrix} 1 & 0 & 0 & 5 & 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 & 26 & 0 & 2 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 2 & 0 & 563 \\ 5 & 0 & 0 & 26 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 2 & 0 & 563 & 0 & 0 & 0 & 0 \\ 0 & 26 & 0 & 0 & 0 & 0 & 587 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 563 & 0 & 0 & 0 & 0 & 1 & 0 & 319642 \end{pmatrix} \tag{3.1}$$

Note that the right-hand side of (3.1) is positive definite, and $L(x_1^4x_2^2)L(x_1^2x_2^4) - L(x_1^2x_2^2)^3 = 1 - 2^3 = -7$. By Lemma 3.4, L extends to $\widetilde{L} : \mathbb{R}[\underline{x}] \to \mathbb{R}$ such that $\widetilde{L}(p^2) \geq 0$ for $p \in \mathbb{R}[\underline{x}]$. Therefore $\phi(f) < 0$ for the pseudo-moment evaluation ϕ determined by $\widetilde{L}$ (and any evaluation on x_1, x_2), so f is neither a quotient of sums of products of elements in $\mathrm{QM}(\emptyset,\emptyset)$, nor in $\mathrm{qm}(\emptyset,\emptyset)$, by Theorem 3.6.

4. Archimedean Positivstellensatz for moment polynomials

The main result of this section, Theorem 4.2, describes moment polynomials that are positive subject to constraints on measures with a given compact support. Recall [Mar08, Section 5.2] that a quadratic module A in a commutative unital ring A is *archimedean* if for every $a \in A$ there exists $N \in \mathbb{N}$ such that $N \pm a \in M$. Equivalently, a quadratic module M in $\mathbb{R}[\underline{x}]$ is archimedean if and only if there is an $N \in \mathbb{N}$ such that $N - x_1^2 - \cdots - x_n^2 \in M$ [Mar08, Corollary 5.2.4].

Lemma 4.1. *Let $S \subseteq \mathbb{R}[\underline{x}]$. If $M(S) \subseteq \mathbb{R}[\underline{x}]$ is archimedean, then $\operatorname{qm}(S, \emptyset) \subseteq \mathscr{M}$ and $\operatorname{QM}(S, \emptyset) \subseteq \mathscr{M}[\underline{x}]$ are archimedean.*

Proof. Let $(i_1, \dots, i_n) \in \mathbb{N}_0^n$ be arbitrary. Since $M(S)$, there exists $N > 0$ such that $N \pm x_1^{i_1} \cdots x_n^{i_n}$ is a convex combination of some $p^2 s$ for $p \in \mathbb{R}[\underline{x}]$ and $s \in S \cup \{1\}$. Therefore $N \pm \mathtt{m}_{i_1,\dots,i_n} \in \operatorname{qm}(S, \emptyset)$. Therefore $\operatorname{qm}(S, \emptyset)$ and $\operatorname{QM}(S, \emptyset)$ are archimedean by [Mar08, Proposition 5.2.3]. □

Theorem 4.2 (Archimedean Positivstellensatz)**.** *Let $S_1 \subseteq \mathbb{R}[\underline{x}]$ and $S_2 \subseteq \mathscr{M}$, and suppose $M(S_1)$ is archimedean in $\mathbb{R}[\underline{x}]$. The following are equivalent for $f \in \mathscr{M}[\underline{x}]$:*

(i) *$f \geq 0$ on $\boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$;*
(ii) *$f + \varepsilon \in \operatorname{QM}(S_1, S_2)$ for all $\varepsilon > 0$.*

The following are equivalent for $f \in \mathscr{M}$:

(i') *$f \geq 0$ on $\boldsymbol{\mathcal{K}}(S_1, S_2)$;*
(ii') *$f + \varepsilon \in \operatorname{qm}(S_1, S_2)$ for all $\varepsilon > 0$.*

Proof. We only prove the first equivalence (the proof of second one is analogous). The implication (ii)⇒(i) is straightforward. Now suppose (ii) is false. By the Kadison-Dubois representation theorem [Mar08, Theorem 5.4.4] there exists a homomorphism $\varphi : \mathscr{M}[\underline{x}] \to \mathbb{R}$ such that $\varphi(f) < 0$ and $\varphi(\operatorname{QM}(S_1, S_2)) = \mathbb{R}_{\geq 0}$. Then $\underline{X} := (\varphi(x_1), \dots, \varphi(x_n)) \in K(S_1)$. Consider the unital functional $L : \mathbb{R}[\underline{x}] \to \mathbb{R}$ given by $L(p) = \varphi(\mathtt{m}(p))$. Then L is nonnegative on $M(S_1)$, so by the solution of the moment problem on compact sets [Sch17, Theorem 12.36 (ii)] there is $\mu \in \mathbf{P}(K(S_1))$ such that $L(p) = \int p \, \mathrm{d}\mu$ for all $p \in \mathbb{R}[\underline{x}]$. By the construction, $\mu \in \boldsymbol{\mathcal{K}}(S_1, S_2)$. Therefore $(\mu, \underline{X}) \in \boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$ and $f(\mu, \underline{X}) = \varphi(f) < 0$. □

Remark 4.3. The equivalence (i')⇔(ii') in Theorem 4.2 also follows from [KMVW23, Theorem 5.5] on state polynomials and their evaluations on constrained tuples of bounded operators and states. Indeed, the class of admissible constraints in [KMVW23, Theorem 5.5] is large enough to allow for commutators, and thus one can consider positivity of state polynomials on commuting bounded operators subject to archimedean constraints. The second part of Theorem 4.2 can be then obtained using the spectral theorem for tuples of commuting bounded operators [Sch12, Theorem 5.23].

However, note that the results of [KMVW23] carry implications only for pure moment polynomials, but not for general moment polynomials, and are not applicable to the first part of Theorem 4.2.

Corollary 4.4. *Let $S_1 \subset \mathbb{R}[\underline{x}]$ and $S_2 \subset \mathscr{M}$, and suppose $M(S_1)$ is archimedean in $\mathbb{R}[\underline{x}]$. If $f \in \mathscr{M}[\underline{x}]$ is strictly positive on $\boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$, then $f \in \operatorname{QM}(S_1, S_2)$.*

Proof. Since $K(S_1)$ is compact, the set of Borel probability measures supported on $K(S_1)$ is also compact by [Par05, Theorem II.6.4], and is equal to $\mathbf{P}(K(S_1))$ (the existence of all marginal moments for Borel measures on a compact subset of $\mathbb{R}^n$ is automatic). Therefore $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$ is compact, so there is $\varepsilon>0$ such that $f-\varepsilon\geq 0$ on $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$. Then $f\in \mathrm{QM}(S_1,S_2)$ by Theorem 4.2. □

Corollary 4.5. *Let $S_1\subset\mathbb{R}[\underline{x}]$ and $S_2\subset\mathscr{M}$, and suppose $K(S_1)\subset\mathbb{R}^n$ is bounded. Then the following are equivalent for $f\in\mathscr{M}[\underline{x}]$:*

(i) *$f\geq 0$ on $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$;*
(ii) *$f+\varepsilon\in\mathrm{QM}(\widetilde{S}_1,S_2)$ for all $\varepsilon>0$, where $\widetilde{S}_1$ is the set of all square-free products of elements in S_1.*

Proof. If $K(S_1)$ is bounded, then $M(\widetilde{S}_1)$ is archimedean in $\mathbb{R}[\underline{x}]$ by [Mar08, Corollary 6.1.2]. The rest then follows from Theorem 4.2. □

5. Moment polynomial optimization and examples

Theorem 4.2 can be applied to design a converging hierarchy of semidefinite programs (SDPs) for moment polynomial optimization. For the sake of simplicity, we first focus on pure moment polynomial objective functions, and then indicate the necessary changes for general moment polynomial objective functions.

Let $S_1\subset\mathbb{R}[\underline{x}]$ and $S_2\subset\mathscr{M}$ be finite, and $r\in\mathbb{N}$. Recall that $\mathrm{qm}(S_1,S_2)_{2r}$ is the convex hull of

$$q_1^2\mathtt{m}(p^2s_1),\ q_2^2s_2:\quad \begin{array}{l} s_i\in\{1\}\cup S_i,\ p\in\mathbb{R}[\underline{x}],\ q_i\in\mathscr{M},\\ \deg s_1+2(\deg q_1+\deg p),\ \deg s_2+2\deg q_2\leq 2r.\end{array}$$

Membership in $\mathrm{qm}(S_1,S_2)_{2r}$ is a feasibility linear conic program, but not a semidefinite program because of the terms $q_1^2\mathtt{m}(p^2s_1)=\mathtt{m}((q_1p)^2s_1)$. Thus we consider a larger set $\widetilde{\mathrm{qm}}(S_1,S_2)$, which is the convex hull of

$$\mathtt{m}(f^2s_1),\ q^2s_2:\quad \begin{array}{l} s_i\in\{1\}\cup S_i,\ f\in\mathscr{M}[\underline{x}],\ q\in\mathscr{M},\\ \deg s_1+2\deg f,\ \deg s_2+2\deg q\leq 2r.\end{array}$$

It is easy to see that membership in $\widetilde{\mathrm{qm}}(S_1,S_2)_{2r}$ can be certified by an SDP; indeed, its members can be represented as

$$\sum_{s\in\{1\}\cup S_1}\sum_{v_1,v_2}G^{(s)}_{v_1,v_2}\cdot\mathtt{m}(v_1v_2s)+\sum_{t\in S_2}\sum_{u_1,u_2}H^{(t)}_{u_1,u_2}\cdot u_1u_2t,$$

where v_i are monomials in $\mathscr{M}[\underline{x}]_{r-\frac{\deg s}{2}}$, u_i are monomials in $\mathscr{M}_{r-\frac{\deg t}{2}}$, and $G^{(s)}$, $H^{(t)}$ are positive semidefinite matrices of dimensions $\dim\mathscr{M}[\underline{x}]_{r-\frac{\deg s}{2}}$ and $\dim\mathscr{M}_{r-\frac{\deg t}{2}}$, respectively.

For $f\in\mathscr{M}$ and $r\geq\frac{\deg f}{2}$ consider the sequence of SDPs

$$f_r=\sup\{\alpha\in\mathbb{R}:\ f-\alpha\in\widetilde{\mathrm{qm}}(S_1,S_2)_{2r}\}. \tag{5.1}$$

Corollary 5.1. *Let $S_1\subseteq\mathbb{R}[\underline{x}]$, $S_2\subseteq\mathscr{M}$, $f\in\mathscr{M}$, and suppose $M(S_1)$ is archimedean in $\mathbb{R}[\underline{x}]$. Then the sequence $\{f_r\}_{r\geq\frac{\deg f}{2}}$ arising from the SDP hierarchy (5.1) converges monotonically to $f_*:=\inf_{\mu\in\boldsymbol{\mathcal{K}}(S_1,S_2)}f(\mu)$ from below.*

Proof. The sequence $\{f_r\}_r$ is increasing since $\widetilde{\mathrm{qm}}(S_1,S_2)_{2r} \subset \widetilde{\mathrm{qm}}(S_1,S_2)_{2(r+1)}$. Also, $f_r \le f_*$ by the definition of $\widetilde{\mathrm{qm}}(S_1,S_2)_{2r}$. Let $\varepsilon > 0$ be arbitrary. By Theorem 4.2 there exists $r \in \mathbb{N}$ such that

$$f - f_* + \varepsilon \in \mathrm{qm}(S_1,S_2)_{2r} \subseteq \widetilde{\mathrm{qm}}(S_1,S_2)_{2r}.$$

Therefore $\lim_{r\to\infty} f_r = f_*$. □

Corollary 5.1 is also a specialization of [KMVW23, Corollary 6.1] from the state polynomial setup. Let us note a few further consequences of [KMVW23, Section 6] without proofs:

(1) If $N - x_1^2 - \cdots - x_n^2$ for some $N > 0$ is a conic combination of $S_1 \cup \{\ell^2 \colon \ell \in \mathbb{R}[\underline{x}]_1\}$, then there is no duality gap between SDP (5.1) and its dual,

$$f_r = \inf\left\{L(f) \colon L \in \mathscr{M}_{2r}^\vee,\ L(1) = 1,\ L(\widetilde{\mathrm{qm}}(S_1,S_2)_{2r}) = \mathbb{R}_{\ge 0}\right\}.$$

(2) If the solution of the dual of (5.1) certain rank conditions, then the SDP hierarchy (5.1) stops, and one can extract a concrete finitely supported optimizer for f_*.

(3) While the sizes of SDPs (5.1) and their duals grow quickly in concrete applications, one can mitigate this by employing sparsity [MW23] and symmetry reductions.

To apply semidefinite programming to optimization of general moment polynomials, one needs to replace $\mathrm{QM}(S_1,S_2)_{2r}$ with a larger cone $\widetilde{\mathrm{QM}}(S_1,S_2)_{2r}$. Similarly as in the case of $\mathrm{qm}(S_1,S_2)_{2r}$, the problematic elements in $\mathrm{QM}(S_1,S_2)_{2r}$ are conic combinations of $f^2\mathtt{m}(p^2s)$ for $s \in S_1 \cup \{1\}$, $p \in \mathbb{R}[\underline{x}]$ and $f \in \mathscr{M}[\underline{x}]$ with $\deg s + 2(\deg p + \deg f) \le 2r$. To obtain $\widetilde{\mathrm{QM}}(S_1,S_2)_{2r}$, we replace them by

$$\sum_{(u_1,v_1),(u_2,v_2)} G_{(u_1,v_1),(u_2,v_2)} \cdot \mathtt{m}(u_1u_2s)v_1v_2,$$

where (u_i,v_i) are pairs of monomials $u_i \in \mathbb{R}[\underline{x}]$ and $v_i \in \mathscr{M}[\underline{x}]$ with $\deg s + 2(\deg u_i + \deg v_i) \le 2r$, and G is a positive semidefinite matrix of dimension $\sum_{i+j\le r-\frac{\deg s}{2}}(\dim \mathbb{R}[\underline{x}]_i + \dim \mathscr{M}[\underline{x}]_j)$. Given $f \in \mathscr{M}[\underline{x}]$, the optimization problems

$$f_r = \sup\{\alpha \colon f - \alpha \in \widetilde{\mathrm{QM}}(S_1,S_2)_{2r}\} \tag{5.2}$$

are then SDPs, and the following analog of Corollary 5.1 holds.

Corollary 5.2. *Let $S_1 \subseteq \mathbb{R}[\underline{x}]$, $S_2 \subseteq \mathscr{M}$, $f \in \mathscr{M}[\underline{x}]$, and suppose $M(S_1)$ is archimedean in $\mathbb{R}[\underline{x}]$. Then the sequence $\{f_r\}_{r \ge \frac{\deg f}{2}}$ arising from the SDP hierarchy (5.2) converges monotonically to $f_* := \inf_{(\mu,\underline{X}) \in \boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)} f(\mu,\underline{X})$ from below.*

Next, we demonstrate the above SDP hierarchy method on two optimization problems arising from nonlinear Bell inequalities in quantum physics. Section 5.1 confirms a covariance Bell inequality proposed in [PHBB17], and Section 5.1 rectifies a bilocal Bell inequality proposed in [TGB21].

5.1. **Covariance Bell inequality.** Let X_j, Y_j for $j = 1,2,3$ be binary random variables (valued in $\{-1,1\}$) on some probability space $(\mathcal{P},\Sigma,\pi)$, and consider the expression

$$\begin{aligned}\mathrm{cov}_{3322}(\underline{X},\underline{Y}) :=\ & \mathrm{cov}(X_1,Y_1) + \mathrm{cov}(X_1,Y_2) + \mathrm{cov}(X_1,Y_3)\\ & + \mathrm{cov}(X_2,Y_1) + \mathrm{cov}(X_2,Y_2) - \mathrm{cov}(X_2,Y_3)\end{aligned}$$

$$+\operatorname{cov}(X_3,Y_1)-\operatorname{cov}(X_3,Y_2)$$

where $\operatorname{cov}(X,Y)=\int XY\,\mathrm{d}\pi-\int X\,\mathrm{d}\pi\cdot\int Y\,\mathrm{d}\pi$. In [PHBB17], the authors ask what is the largest possible value of $\operatorname{cov}_{3322}$. They provide concrete examples of probability spaces (on a three-element set) and binary random variables where $\operatorname{cov}_{3322}$ attains the value 4.5. In the quest for proving that $\operatorname{cov}_{3322}\le 4.5$ for all binary random variables, they propose a reduction to solving a certain number of linear systems. Nonetheless, for establishing this particular inequality, they estimate that more than 10^{14} linear systems would have to be solved, thus rendering this particular approach infeasible. As an alternative, they suggest maximizing $\operatorname{cov}_{3322}$ via classical polynomial optimization similarly as in Remark 2.6. However, the corresponding polynomial problem has too many variables for global optimization tools to apply. Thus they use numerical nonlinear optimization to look for local maxima of $\operatorname{cov}_{3322}$ from numerous starting points, which lends confidence to their conjecture that $\operatorname{cov}_{3322}\le 4.5$.

Let

$$\begin{aligned} f=&\mathtt{m}_{100100}-\mathtt{m}_{100000}\,\mathtt{m}_{000100}+\mathtt{m}_{100010}-\mathtt{m}_{100000}\,\mathtt{m}_{000010}+\mathtt{m}_{100001}-\mathtt{m}_{100000}\,\mathtt{m}_{000001}\\ &+\mathtt{m}_{010100}-\mathtt{m}_{010000}\,\mathtt{m}_{000100}+\mathtt{m}_{010010}-\mathtt{m}_{010000}\,\mathtt{m}_{000010}-\mathtt{m}_{010001}+\mathtt{m}_{010000}\,\mathtt{m}_{000001}\\ &+\mathtt{m}_{001100}-\mathtt{m}_{001000}\,\mathtt{m}_{000100}-\mathtt{m}_{001010}+\mathtt{m}_{001000}\,\mathtt{m}_{000010}. \end{aligned}$$

The question of [PHBB17] is equivalent to the moment polynomial optimization problem

$$f_*=\sup\ f\quad \text{subject to } x_j^2=1 \text{ for } j=1,\dots,6.$$

By Corollary 5.1 we have $f_r\searrow f_*$ for

$$f_r=\inf\{\alpha\colon \alpha-f\in\widetilde{\mathrm{qm}}(S,\emptyset)\}$$

and $S=\{\pm(1-x_j^2)\colon j=1,\dots,6\}$. When constructing SDPs for f_r, we encode the relations of S as substitution rules, to reduce the size of the SDPs. For $r=2$, the resulting SDP has 4146 indeterminates and the semidefinite constraint of size 100×100, and yields $f_2=4.5$. Therefore we have $f_*=4.5$.

5.2. **Bilocal Bell inequality.** In [TGB21, TPKLR22], the authors ask about the largest value of

$$\frac{1}{3}\sum_{i\in\{1,2,3\}}\Big(E(B_iC_i)-E(A_iB_i)\Big)-\sum_{\{i,j,k\}=\{1,2,3\}}E(A_iB_jC_k) \tag{5.3}$$

where A_i,B_i,C_i for $i=1,2,3$ are binary random variables on a probability space $(\mathcal{P},\Sigma,\pi)$ satisfying bilocality constraints

$$E(A_1^{k_1}A_2^{k_2}A_3^{k_3}C_1^{k_4}C_2^{k_5}C_3^{k_6})=E(A_1^{k_1}A_2^{k_2}A_3^{k_3})E(C_1^{k_4}C_2^{k_5}C_3^{k_6}) \tag{5.4}$$

for all $k_i\in\{0,1\}$, and additional vanishing constraints

$$\begin{aligned} &E(A_i)=E(B_i)=E(C_i)=0\quad \text{for } i\in\{1,2,3\},\\ &E(A_iB_j)=E(B_iC_j)=0\quad \text{for } i\ne j,\\ &E(A_iB_jC_k)=0\quad \text{for } |\{i,j,k\}|\le 2. \end{aligned} \tag{5.5}$$

Here, $E(X)=\int X\,\mathrm{d}\pi$. In [TGB21] it is shown that the largest value of (5.3) for bilocal models with the tetrahedral symmetry is 3. Furthermore, [TGB21, TPKLR22] suggest that (5.3) can be at most 3 in general, and support this claim with numerical methods that search for local maxima. However, as shown below, this claim is false; the largest value of (5.3) subject to (5.4) and (5.5) is 4.

Consider the moment polynomial optimization problem

$$\begin{aligned}
&\sup\ \frac{1}{3}\sum_{i\in\{1,2,3\}}\Big(\mathtt{m}(x_{i+3}x_{i+6})-\mathtt{m}(x_ix_{i+3})\Big)-\sum_{\{i,j,k\}=\{1,2,3\}}\mathtt{m}(x_ix_{j+3}x_{k+6})\\
&\text{subject to}\\
&\mathtt{m}(x_1^{k_1}x_2^{k_2}x_3^{k_3}x_7^{k_4}x_8^{k_5}x_9^{k_6})=\mathtt{m}(x_1^{k_1}x_2^{k_2}x_3^{k_3})\mathtt{m}(x_7^{k_4}x_8^{k_5}x_9^{k_6})\quad\text{for } k_i\in\{0,1\},\\
&x_j^2=1 \text{ and } \mathtt{m}(x_j)=0\quad\text{for } j\in\{0,\dots,9\},\\
&\mathtt{m}(x_ix_{j+3})=\mathtt{m}(x_{i+3}x_{j+6})=0\quad\text{for } i,j\in\{1,2,3\},\ i\neq j,\\
&\mathtt{m}(x_ix_{j+3}x_{k+6})=0\quad\text{for } i,j,k\in\{1,2,3\},\ |\{i,j,k\}|\leq 2.
\end{aligned}\tag{5.6}$$

Corollary 5.1 provides a converging sequence of upper bounds for the solution of (5.6). For $r=3$, one obtains the upper bound 4 by solving an SDP with 31017 indeterminates and the semidefinite constraint of size 263, or more practically, by solving its dual with 4549 indeterminates and the semidefinite constraint of size 325. Therefore (5.3) subject to (5.4) and (5.5) is at most 4. Next, we show that the value 4 is indeed attained. Denote

$$\eta_0=(1\ 1\ 1\ 1),\quad \eta_1=(1\ 1\ \text{-1}\ \text{-1}),\quad \eta_2=(1\ \text{-1}\ 1\ \text{-1}),\quad \eta_3=(1\ \text{-1}\ \text{-1}\ 1),$$

and let $e_i\in\mathbb{R}^4$ be the i^{th} standard unit vector. Endow $\{1,2,3,4\}^2$ with the uniform probability distribution, and consider the following binary random variables on it:

$$A_i=\eta_0\otimes\eta_i,\qquad B_i=\left(\eta_0\otimes\eta_0-2\sum_{k=1}^{4}e_k\otimes e_k\right)\cdot\eta_i\otimes\eta_0,\qquad C_i=\eta_i\otimes\eta_0,$$

for $i\in\{1,2,3\}$. Here, we identified the algebra of random variables on $\{1,2,3,4\}^2$ with $\mathbb{R}^4\otimes\mathbb{R}^4$. The bilocality constraints (5.4) are satisfied because of the tensor structure of A_i,C_i (and the uniform distribution on a product is the product of uniform distributions), and the vanishing constraints (5.5) follow by direct calculation. Finally, (5.3) evaluates to 4 for this ensemble of binary random variables.

6. Lasserre's Störungspositivstellensatz for moment polynomials

In this section we show that moment polynomials nonnegative on $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$ belong to the quadratic module $\mathrm{QM}(S_1,S_2)$ *up to an arbitrarily small perturbation* of their coefficients (Theorem 6.7). This is achieved through the analysis of a sequence of conic optimization problems and their duals, and the resolution of an infinite-dimensional moment problem. Finally, a corollary for polynomial positivity on semialgebraic sets is given (Corollary 6.13).

For $r\in\mathbb{N}$ let

$$\Omega_r=\sum_{j=1}^{n}\sum_{k=0}^{r}\frac{x_j^{2k}}{k!}.$$

If $\underline{X}\in\mathbb{R}^n$, then $n\leq\Omega_r(\underline{X})\leq\sum_{j=1}^n\exp(X_j^2)$ is uniformly bounded for all $r\in\mathbb{N}$. Similarly, if $\nu\in\mathbf{P}(\mathbb{R}^n)$ is finitely supported, i.e., $\nu=\sum_{i=1}^{\ell}\alpha_i\delta_{\underline{X}_i}$ is a convex combination of the Dirac delta measures $\delta_{\underline{X}_i}$ concentrated at $\underline{X}_i \in \mathbb{R}^n$, then

$$n \le \mathtt{m}(\Omega_r)(\nu) \le \sum_{i=1}^{\ell} \alpha_i \sum_{j=1}^{n} \exp(X_{ij}^2)$$

is uniformly bounded for all $r \in \mathbb{N}$.

To $S_1 \subseteq \mathbb{R}[\underline{x}]$, $S_2 \subseteq \mathscr{M}$ and $f \in \mathscr{M}[\underline{x}]$ we assign a pair of optimization problems for every $r \ge \frac{\deg f}{2}$ and $M > 0$:

$$Q_{r,M}: \begin{cases} \sup\limits_{z\in\mathbb{R}} & z \\ s.t. & f - z \in \mathrm{QM}(S_1,S_2)_{2r} + \mathbb{R}_{\ge 0}\big(M - \Omega_r - \mathtt{m}(\Omega_r)\big)\,; \end{cases} \tag{6.1}$$

$$Q_{r,M}^{\vee}: \begin{cases} \inf\limits_{L\in\mathscr{M}[\underline{x}]_{2r}^{\vee}} & L(f) \\ s.t. & L(1) = 1\,, \\ & L\big(M - \Omega_r - \mathtt{m}(\Omega_r)\big) \ge 0\,, \\ & L(g) \ge 0 \quad \text{for all } g \in \mathrm{QM}(S_1,S_2)_{2r}\,. \end{cases} \tag{6.2}$$

In the following two lemmas and their proofs we abbreviate $\mathcal{C}_{r,M} = \mathrm{QM}(S_1,S_2)_{2r} + \mathbb{R}_{\ge 0}(M - \Omega_r - \mathtt{m}(\Omega_r))$.

Lemma 6.1. *For all $r \in \mathbb{N}$ and $M > 0$, the closure of the cone $\mathscr{M}[\underline{x}]_r \cap \mathcal{C}_{2^r,M}$ in $\mathscr{M}[\underline{x}]_r$ is contained in*

$$\{g \in \mathscr{M}[\underline{x}]_r \colon g + \varepsilon \in \mathcal{C}_{2^r,M} \text{ for all } \varepsilon > 0\}\,. \tag{6.3}$$

Proof. Let us fix $r \in \mathbb{N}$, $M > 0$ and let us endow the finite-dimensional space $\mathscr{M}[\underline{x}]_r$ with some norm $\|\cdot\|$. The identities $\pm 2uv = (u \pm v)^2 - u^2 - v^2$ and $\pm 2\mathtt{m}(uv) = \mathtt{m}((u\pm v)^2) - \mathtt{m}(u^2) - \mathtt{m}(v^2)$ for $u, v \in \mathscr{M}[\underline{x}]$ imply that there exists $A > 0$ (dependent on r and M) such that

$$g + A\|g\| \in \mathrm{QM}(\emptyset,\emptyset)_{2^r} + \mathbb{R}_{\ge 0}\big(M - \Omega_{2^{r-1}} - \mathtt{m}(\Omega_{2^{r-1}})\big)$$

for all $g \in \mathscr{M}[\underline{x}]_r$. Define $F : \mathscr{M}[\underline{x}]_r \to [-\infty, \infty]$ as

$$F(g) = \sup\{z \in \mathbb{R} \colon g - z \in \mathcal{C}_{2^r,M}\}\,.$$

This function satisfies the following properties:

(i) $F(g) \ge -A\|g\|$ for all $g \in \mathscr{M}[\underline{x}]_r$;
(ii) $F(g_1 + g_2) \ge F(g_1) + F(g_2)$ for all $g_i \in \mathscr{M}[\underline{x}]_r$;
(iii) $F(g) \ge 0$ if and only if g belongs to (6.3).

Now suppose $(g_i)_i$ is a sequence in $\mathscr{M}[\underline{x}]_r \cap \mathcal{C}_{2^r,M}$ that converges to $g \in \mathscr{M}[\underline{x}]_r$. Then

$$F(g) \ge F(g_i) + F(g - g_i) \ge -A\|g - g_i\|$$

for all i, and so $F(g) \ge 0$. Therefore g belongs to (6.3). □

Remark 6.2. Note that $\mathcal{C}_{r,M}$ is not closed in general. Indeed, following [PS01, Remark 2.8] let $S_1 = \{-x_1^2\}$; then $x_1 + \varepsilon \in \mathrm{QM}(S_1, \emptyset) = \mathcal{C}_{1,1}$ for all $\varepsilon > 0$ but $x_1 \notin \mathcal{C}_{1,1}$.

Lemma 6.3. *For all r and M, the optimization problem (6.1) is a linear conic problem, and (6.2) is its dual. Sequences $(\sup Q_{r,M})_r$ and $(\inf Q^\vee_{r,M})_r$ are increasing. If S_2 is finite, $\boldsymbol{\mathcal{K}}(S_1,S_2)\neq\emptyset$ and M is large enough, then*

$$\lim_{r\to\infty}\sup Q_{r,M}=\lim_{r\to\infty}\inf Q^\vee_{r,M}. \tag{6.4}$$

Proof. The first part of the claim follows by inspection. The sequence $(\sup Q_{r,M})_r$ is increasing since $\mathcal{C}_{r,M}\subseteq\mathcal{C}_{r+1,M}$. If L is feasible for $Q^\vee_{r,M}$, then its restriction is feasible for $Q^\vee_{r',M}$ for $r'<r$; hence the sequence $(\inf Q^\vee_{r,M})_{r\geq d}$ is increasing. By weak duality [Bar02, Theorem IV.6.2] we have $\sup Q_{r,M}\leq\inf Q^\vee_{r,M}$. By Proposition 2.5 there exists $(\nu,\underline{X})\in\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$ with a finitely supported ν. Let $M\geq\sup_r(\Omega_r(\underline{X})+\mathtt{m}(\Omega_r)(\nu))$ (note that the right hand side is finite). Then $L\in\mathscr{M}[\underline{x}]^\vee_{2r}$ defined by $L(p)=p(\nu,\underline{X})$ is clearly feasible for (6.2), whence $\inf Q^\vee_{r,M}<\infty$ for all $r\geq\frac{\deg f}{2}$. Note that $L(f-\inf Q^\vee_{r,M})\geq 0$ for all $L\in\mathcal{C}^\vee_{r,M}$. This implies that $f-\inf Q^\vee_{r,M}$ is in $\mathcal{C}^{\vee\vee}_{r,M}$, which is the closure of $\mathcal{C}_{r,M}$. Therefore $f-\inf Q^\vee_{r,M}+\varepsilon\in\mathcal{C}_{2^r,M}$ for all $\varepsilon>0$ by Lemma 6.1, so $\inf Q^\vee_{r,M}\leq\sup Q_{2^r,M}$. Thus (6.4) follows from $\sup Q_{r,M}\leq\inf Q^\vee_{r,M}\leq\sup Q_{2^r,M}$. □

The following proposition resolves the unbounded moment problem for positive functionals on moment polynomials (cf. [IKR14, AJK15] for related infinite-dimensional moment problems), in the spirit of Nussbaum's theorem [Sch17, Theorem 14.25] on functionals satisfying the multivariate Carleman condition.

Proposition 6.4. *Let $S_1\subseteq\mathbb{R}[\underline{x}]$, $S_2\subseteq\mathscr{M}$, $f\in\mathscr{M}[\underline{x}]$ and $M\in\mathbb{R}_{>0}$. Suppose $L\in\mathscr{M}[\underline{x}]^\vee$ satisfies*

(a) $L(1)=1$,
(b) $L(\mathrm{QM}(S_1,S_2))=\mathbb{R}_{\geq 0}$,
(c) $|L(w)|\leq(\deg w)!M^{\deg w}$ *for all monomials* $w\in\mathscr{M}[\underline{x}]$.

Then there exists $(\mu,\underline{X})\in\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$ such that $L(f)=f(\mu,\underline{X})$.

Proof. Denote $\alpha=L(f)$, and endow $\mathscr{M}[\underline{x}]$ with the finest locally convex topology. Let $\mathcal{C}$ be the set of all $L'\in\mathscr{M}[\underline{x}]^\vee$ that satisfy (a), (b), (c) and $L'(f)=\alpha$. Then $\mathcal{C}$ is a nonempty convex set, and Tychonoff's theorem [Wil70, Theorem 17.8] implies that $\mathcal{C}$ is weak-* compact in $\mathscr{M}[\underline{x}]^\vee$. By the Krein-Milman theorem [Bar02, Theorem III.4.1] we may therefore assume that the functional L is an extreme point of $\mathcal{C}$.

On $\mathscr{M}[\underline{x}]$ we define a semi-inner product $\langle p,q\rangle=L(pq)$. Let $\mathcal{N}=\{p\in\mathscr{M}[\underline{x}]\colon L(p^2)=0\}$. By the Cauchy-Schwarz inequality for semi-inner products, $\mathcal{N}$ is an ideal of $\mathscr{M}[\underline{x}]$. Let $\mathcal{H}$ be the completion of the inner product space $\mathscr{M}[\underline{x}]/\mathcal{N}$. Multiplication with generators x_j and $\mathtt{m}_{i_1,\dots,i_n}$ in $\mathscr{M}[\underline{x}]$ induces symmetric unbounded operators X_j and $Y_{i_1,\dots,i_n}$ on $\mathcal{H}$ with a dense domain $\mathscr{M}[\underline{x}]/\mathcal{N}$. Moreover, the elements of $\mathscr{M}[\underline{x}]/\mathcal{N}$ are analytic vectors for X_j and $Y_{i_1,\dots,i_n}$ according to [Sch12, Definition 7.1] by (c). By [Sch12, Theorem 7.18], the closures $\overline{X_j}$ and $\overline{Y_{i_1,\dots,i_n}}$ are strongly commuting self-adjoint operators. Since real polynomials in strongly commuting self-adjoint operators are again self-adjoint,

$$\varphi(x_j)=\overline{X_j},\quad \varphi(\mathtt{m}_{i_1,\dots,i_n})=\overline{Y_{i_1,\dots,i_n}}$$

defines an integrable representation φ of $\mathscr{M}[\underline{x}]$ on $\mathcal{H}$ according to [Sch90, Definition 9.1.1]. Note that φ admits a cyclic unit vector $1 \in \mathscr{M}[\underline{x}]/\mathcal{N}$, and $L(p) = \langle \varphi(p)1, 1\rangle$.

Let $P \in \mathcal{B}(\mathcal{H})$ be a projection that strongly commutes with $\varphi(\mathscr{M}[\underline{x}])$. Suppose $P \notin \{0, I\}$. Therefore $P1 \neq 0$ and $(I-P)1 \neq 0$ since 1 is a cyclic vector for φ. Consider the following $L_1, L_2 \in \mathscr{M}[\underline{x}]^\vee$:

$$L_1(p) = \frac{\langle \varphi(p)P1, P1\rangle}{\|P1\|^2}, \qquad L_2(p) = \frac{\langle \varphi(p)(I-P)1, (I-P)1\rangle}{\|(I-P)1\|^2}.$$

Then $L_1, L_2 \in \mathcal{C}$ and L is a convex combination of L_1 and L_2. By the extremal property of L it follows $L = L_1 = L_2$. Then

$$\langle \varphi(p)1, 1\rangle = \frac{\langle \varphi(p)P1, P1\rangle}{\|P1\|^2} = \left\langle \varphi(p)1, \frac{1}{\|P1\|^2}P1\right\rangle$$

for every $p \in \mathscr{M}[\underline{x}]$ implies $1 = \frac{1}{\|P1\|^2}P1$ because 1 is a cyclic vector. Therefore $P1 = 1$ because P is a projection, which contradicts $(I-P)1 \neq 0$.

Hence there are no nontrivial projections in the strong commutant of $\varphi(\mathscr{M}[\underline{x}])$. Therefore φ is irreducible by the unbounded analog of von Neumann's theorem [Sch90, Lemma 8.3.5], and $\mathcal{H}$ is one-dimensional by the unbounded analog of Schur's lemma [Sch90, Corollary 9.1.11]. Therefore φ is a homomorphism from $\mathscr{M}[\underline{x}]$ to $\mathbb{R}$. In particular, $L(p) = \langle \varphi(p)1, 1\rangle = \varphi(p)$ for all $p \in \mathscr{M}[\underline{x}]$. Consider the functional $\widetilde{L} : \mathbb{R}[\underline{x}] \to \mathbb{R}$ given by $\widetilde{L}(p) = L(\mathtt{m}(p))$; note that $\widetilde{L}(M(S_1)) = \mathbb{R}_{\geq 0}$ by (b), and $\widetilde{L}$ satisfies the Carleman condition $|\widetilde{L}(x_j^k)| \leq k! M^k$ for all $j = 1, \dots, n$ and $k \in \mathbb{N}$. By a refined version of Nussbaum's theorem [Sch17, Theorem 14.25] applied to $\widetilde{L}$, there is $\mu \in \boldsymbol{\mathcal{K}}(S_1, S_2)$ such that $L(\mathtt{m}_{i_1,\dots,i_n}) = \widetilde{L}(x_1^{i_1}\cdots x_n^{i_n}) = \int x_1^{i_1}\cdots x_n^{i_n}\,\mathrm{d}\mu$ for all $(i_1, \dots, i_n) \in \mathbb{N}_0^n$. Let $\underline{X} = (\varphi(x_1), \dots, \varphi(x_n))$; then (b) implies $\underline{X} \in K(S_1)$. Lastly, $L(f) = f(\mu, \underline{X})$. □

Lemma 6.5. *Let $r \in \mathbb{N}$ and $M \geq 1$. Suppose $L \in \mathscr{M}[\underline{x}]_{2r}^\vee$ satisfies $L(\mathrm{QM}(\emptyset, \emptyset)_{2r}) = \mathbb{R}_{\geq 0}$ and*

$$L\left(x_j^{2k}\right), L\left(\mathtt{m}(x_j^{2k})\right) \leq k!\, M \tag{6.5}$$

for $j = 1, \dots, n$ and $k = 1, \dots, r$. Then

$$|L(w)| \leq \sqrt{(\deg w)!\, M} \tag{6.6}$$

for all monomials w in $\mathscr{M}[\underline{x}]_r$.

Proof. By applying [Las06, Lemma 6.2] to the moment matrix $(L(\mathtt{m}(\alpha\beta)))_{\alpha,\beta}$ indexed by $x_1^{i_1}\cdots x_n^{i_n}$ for $i_1 + \cdots + i_n \leq k$, one obtains

$$L(\mathtt{m}_{2i_1,\dots,2i_n}) \leq (i_1 + \cdots + i_n)!\, M \tag{6.7}$$

for $i_1 + \cdots + i_n \leq r$. Next,

$$L(\mathtt{m}_{i_1,\dots,i_n}^{2k}) \leq L(\mathtt{m}_{2ki_1,\dots,2ki_n}) \tag{6.8}$$

for $k(i_1+\cdots+i_n) \le r$. Indeed, $\mathtt{m}_{2k} - \mathtt{m}_1^{2k} \in \mathrm{qm}(\emptyset,\emptyset)$ by Lemma 2.2, and after applying the $\mathtt{m}$-intertwining homomorphism $x_1 \mapsto x_1^{i_1}\cdots x_1^{i_n}$ to it, we obtain $\mathtt{m}_{2ki_1,\dots,2ki_n} - \mathtt{m}_{i_1,\dots,i_n}^{2k} \in \mathrm{qm}(\emptyset,\emptyset)$, which then implies (6.8). By (6.7), (6.8) and [Las06, Lemma 6.2],

$$L(w^2) \le (\deg w)!\, M \tag{6.9}$$

for all monomials $w \in \mathscr{M}[\underline{x}]_r$. Finally, (6.6) follows from $L(w^2)-L(w)^2 = L((w-L(w))^2) \ge 0$ for $\deg w \le r$. □

Lemma 6.6. *Suppose S_2 is finite, $\boldsymbol{\mathcal{K}}(S_1,S_2) \ne \emptyset$ and $f \in \mathscr{M}[\underline{x}]$ is bounded below on $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$; denote $f_* := \inf_{(\mu,\underline{X})\in\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_2)} f(\mu,\underline{X}) > -\infty$. For large enough $M>0$, (6.2) is feasible for $2r \ge \deg f$, and $\inf Q^\vee_{r,M} \nearrow f_M$ as $r\to\infty$ for some $f_M \ge f_*$.*

Proof. Feasibility of (6.1) follow by the same argument as in the proof of Lemma 6.3. Let L be feasible for $Q^\vee_{r,M}$. Observe that for $k \le r$, the values of $L(x_j^{2k})$ and $L(\mathtt{m}(x_j^{2k}))$ are bounded by $k!\,M$. Let $d\in\mathbb{N}$; for $r \ge d$, Lemma 6.5 implies

$$|L(w)| \le \sqrt{d!\,M} =: c_d \tag{6.10}$$

for all monomials w in $\mathscr{M}[\underline{x}]_d$. In particular, $L(f)$ is uniformly bounded for large enough r. Hence $(\inf Q^\vee_{r,M})_r$ is an increasing function bounded from above, whence $\inf Q^\vee_{r,M} \nearrow f_M$ as $r\to\infty$, for some f_M. It remains to show $f_M \ge f_*$.

Let ℓ^∞ be the space of bounded functions on monomials in $\mathscr{M}[\underline{x}]$. For every $r\in\mathbb{N}$ let $L^{(r)}$ be an optimizer of (6.1), and let $s_r \in \ell^\infty$ be given as $s_r(w) = \frac{1}{c_{\deg w}} L(w)$ for monomials w with $\deg w \le 2r$, and $s_r(w)=0$ for all other monomials w. Note that for every monomial w, $s_r(w)$ is bounded by 1 for all sufficiently large r. By the Banach-Alaoglu theorem [Bar02, Theorem III.2.9], $(s_r)_r$ has an accumulation point ℓ^∞ with respect to the weak-* topology. Hence there is $s\in\ell^\infty$ and a subsequence $(s_{r_k})_k$ converging to s. Define

$$L : \mathscr{M}[\underline{x}] \to \mathbb{R}, \qquad L(w) = c_{\deg w}\cdot s(w).$$

Then $L^{(r_k)}|_{\mathscr{M}[\underline{x}]_d} \to L|_{\mathscr{M}[\underline{x}]_d}$ as $k\to\infty$, for every $d\in\mathbb{N}$. In particular, L is a unital linear functional, $L(f)=f_M$, and $L(\mathrm{QM}(S_1,S_2)) = \mathbb{R}_{\ge 0}$.

Let $d\in\mathbb{N}$ be arbitrary. Then for every $r_k \ge d$,

$$|L^{(r_k)}(w)| \le \sqrt{d!\,M}$$

for all monomials $w \in \mathscr{M}[\underline{x}]_d$ by (6.10). Consequently

$$|L(w)| \le \sqrt{(\deg w)!\,M}$$

for all monomials $w\in\mathscr{M}[\underline{x}]$. Therefore $L(f)\ge f_*$ by Proposition 6.4. □

Theorem 6.7 (Störungspositivstellensatz)**.** *Let $S_1 \subseteq \mathbb{R}[\underline{x}]$ and $S_2 \subseteq \mathscr{M}$, and suppose S_2 is finite. The following are equivalent for $f\in\mathscr{M}[\underline{x}]$:*

(i) *$f \ge 0$ on $\boldsymbol{\mathcal{K}}(S_1,S_2)\times K(S_1)$;*
(ii) *for every $\varepsilon>0$ there exists $r\in\mathbb{N}$ such that $f+\varepsilon(\Omega_r+\mathtt{m}(\Omega_r)) \in \mathrm{QM}(S_1,S_2)$.*

The following are equivalent for $f\in\mathscr{M}$:

(i') *$f \geq 0$ on $\boldsymbol{\mathcal{K}}(S_1, S_2)$;*

(ii') *for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that $f + \varepsilon \mathtt{m}(\Omega_r) \in \mathrm{qm}(S_1, S_2)$.*

Proof. (ii)⇒(i) Let $\underline{X} \in K(S_1)$ be arbitrary, and let ν be a finitely supported measure in $\boldsymbol{\mathcal{K}}(S_1, S_2)$. There is $0 < M < \infty$ such that

$$\Omega_r(\underline{X}) + \mathtt{m}(\Omega_r)(\nu) \leq M$$

for all $r \in \mathbb{N}$. Then for every $\varepsilon > 0$ one has $f(\nu, \underline{X}) \geq -\varepsilon M$, and so $f(\nu, \underline{X}) \geq 0$. Since $\underline{X}$ and ν were arbitrary, and finitely supported measures in $\boldsymbol{\mathcal{K}}(S_1, S_2)$ interpolate any measure in $\boldsymbol{\mathcal{K}}(S_1, S_2)$ up to moments of any fixed order by Proposition 2.5, it follows that $f(\mu, \underline{X}) \geq 0$ for all $(\mu, \underline{X}) \in \boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$.

(i)⇒(ii) We divide the proof into two main cases (a) and (b), according to whether $\boldsymbol{\mathcal{K}}(S_1, S_2)$ is empty or not.

Case (a): assume $\boldsymbol{\mathcal{K}}(S_1, S_2) \neq \emptyset$ and denote $f_* = \inf_{(\mu, \underline{X}) \in \boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)} f(\mu, \underline{X})$. We further divide this case in two sub-cases.

First suppose $f_* > 0$. By Proposition 2.5 there exists $(\nu, \underline{X}) \in \boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$, with a finitely supported ν. Denote $M_0 := \sup_r (\Omega_r(\underline{X}) + \mathtt{m}(\Omega_r)(\nu)) < \infty$, and let $M > \max\{\frac{1}{f_*}, M_0\}$ be arbitrary. By Lemmas 6.3 and 6.6 there exists $r_M > 0$ such that $\sup Q_{r_M, M} > f_* - \frac{1}{M}$. That is, there are $z_M \geq f_* - \frac{1}{M}$, $\lambda_M \geq 0$ and $q_M \in \mathrm{QM}(S_1, S_2)_{2r_M}$ such that

$$f - z_M = q_M + \lambda_M \big(M - \Omega_{r_M} - \mathtt{m}(\Omega_{r_M})\big). \tag{6.11}$$

Evaluating (6.11) at $(\nu, \underline{X}) \in \boldsymbol{\mathcal{K}}(S_1, S_2) \times K(S_1)$ gives

$$\begin{aligned} f(\nu, \underline{X}) - f_* + \frac{1}{M} &\geq f(\nu, \underline{X}) - z_M \\ &= q_M(\nu, \underline{X}) + \lambda_M \big(M - \Omega_{r_M}(\underline{X}) - \mathtt{m}(\Omega_{r_M})(\nu)\big) \\ &\geq \lambda_M (M - M_0), \end{aligned}$$

and therefore

$$\lambda_M \leq \frac{f(\nu, \underline{X}) - f_* + \frac{1}{M}}{M - M_0}. \tag{6.12}$$

The right-hand side of (6.12) goes to 0 as $M \to \infty$. By (6.11),

$$f + \lambda_M \big(\Omega_{r_M} + \mathtt{m}(\Omega_{r_M})\big) = z_M + q_M + \lambda_M M \in \mathrm{QM}(S_1, S_2)_{2r_M},$$

and $\lambda_M \to 0$ as $M \to \infty$. Therefore (ii) holds.

Now suppose $f_* = 0$, and let $\varepsilon > 0$ be arbitrary. By applying (i)⇒(ii) to $f + n\varepsilon$ and $\frac{\varepsilon}{2} > 0$, there exists $r \in \mathbb{N}$ such that $(f + n\varepsilon) + \frac{\varepsilon}{2}(\Omega_r + \mathtt{m}(\Omega_r)) \in \mathrm{QM}(S_1, S_2)_{2r}$. But the latter equals $f + \varepsilon(\Omega_r + \mathtt{m}(\Omega_r)) - \frac{\varepsilon}{2}(\Omega_r - n + \mathtt{m}(\Omega_r) - n)$, so $f + \varepsilon(\Omega_r + \mathtt{m}(\Omega_r)) \in \mathrm{QM}(S_1, S_2)_{2r}$.

Case (b): assume $\boldsymbol{\mathcal{K}}(S_1, S_2) = \emptyset$, and let $f \in \mathbb{R}[\underline{x}]$ and $\varepsilon > 0$ be arbitrary. Let x_{n+1} be an auxiliary variable, and consider $S_1' = x_{n+1} \cdot (\{1\} \cup S_1) \subset \mathbb{R}[x_1, \dots, x_{n+1}]$ and $S_2' = \mathtt{m}_{1,0,\dots,0}^2 \cdot S_2 \subset \mathscr{M}$. Then $K(S_1')$ contains $\mathbb{R}^n \times \{0\}$, $\boldsymbol{\mathcal{K}}(S_1', S_2')$ contains all $\mu \in \mathbf{P}(\mathbb{R}^n \times \{0\})$

such that $\int x_1\,\mathrm{d}\mu = 0$ (and is thus nonempty), and $x_{n+1}f \geq 0$ on $\boldsymbol{\mathcal{K}}(S_1', S_2')$. By the case (a) of the proof above, there exists $r \in \mathbb{N}$ such that

$$x_{n+1}f + \frac{\varepsilon}{n+2e}\left(\Omega_r + \mathtt{m}(\Omega_r) + \sum_{k=0}^{r}\frac{1}{k!}\left(x_{n+1}^{2k} + \mathtt{m}(x_{n+1}^{2k})\right)\right) \in \mathrm{QM}(S_1', S_2'). \tag{6.13}$$

Consider the homomorphism ξ, from moment polynomials generated by $x_1, \dots, x_{n+1}$ to moment polynomials generated by $x_1, \dots, x_n$, that is determined by

$$\xi(x_j) = x_j \text{ for } j \leq n, \quad \xi(x_{n+1}) = 1, \quad \xi(\mathtt{m}_{i_1,\dots,i_{n+1}}) = \mathtt{m}_{i_1,\dots,i_n}.$$

Note that ξ intertwines with $\mathtt{m}$. Applying ξ to (6.13) thus gives

$$f + \frac{\varepsilon}{n+2e}\left(\Omega_r + \mathtt{m}(\Omega_r) + \sum_{k=0}^{r}\frac{2}{k!}\right) \in \mathrm{QM}(S_1, S_2') \subseteq \mathrm{QM}(S_1, S_2),$$

and therefore $f + \varepsilon\Omega_r \in \mathrm{QM}(S_1, S_2)$.

(i')$\Leftrightarrow$(i') The proof is analogous to (i)$\Leftrightarrow$(ii), and utilizes the straightforward counterparts of Lemmas 6.3, 6.5, 6.6 and Proposition 6.4 for $\mathrm{qm}(S_1, S_2)$. $\square$

Remark 6.8. In Theorem 6.7, one can replace $\mathrm{qm}(S_1, S_2)$ and $\mathrm{QM}(S_1, S_2)$ with larger cones $\widetilde{\mathrm{qm}}(S_1, S_2)$ and $\widetilde{\mathrm{QM}}(S_1, S_2)$ from Section 5, respectively. While the resulting statement is slightly weaker than Theorem 6.7, it has an advantage that checking membership in $\widetilde{\mathrm{qm}}(S_1, S_2)$ or $\widetilde{\mathrm{QM}}(S_1, S_2)$ can be done via SDP.

Remark 6.9. As it is evident from the proof of Theorem 6.7, the sequence of polynomials Ω_r can be replaced by

$$\sum_{j=1}^{n}\sum_{k=0}^{r}\frac{x_j^{2k}}{c_k} \qquad \text{for } r \in \mathbb{N}, \tag{6.14}$$

where $c_k > 0$ are such that $(c_k)_k$ has super-exponential growth (to ensure point-wise convergence of (6.14), which is used for (ii)$\Rightarrow$(i) of Theorem 6.7 and for feasibility of (6.1) and (6.2)) and $(k^{-k}c_k)_k$ has at most exponential growth (which is needed for applying Proposition 6.4).

Example 6.10. The implications (ii)$\Rightarrow$(i) and (ii')$\Rightarrow$(i') of Theorem 6.7 fail in general when S_2 is not finite. Let $n = 1$, $f = -1$ and

$$S_2 = \{\mathtt{m}_{2i} - (4i+1)!\colon i \in \mathbb{N}\}.$$

Since the $2i^{\text{th}}$ moment of $\mu = e^{-\sqrt{|t|}}\,\mathrm{d}t$ is $(4i+1)!$, we have $\boldsymbol{\mathcal{K}}(\emptyset, S_2) \neq \emptyset$ and therefore (i') and (i) are false. Now let $\varepsilon > 0$ be arbitrary; then there exists $r \in \mathbb{N}$ such that $r! \leq \varepsilon(4r+1)!$, and so

$$-1 + \varepsilon\mathtt{m}(\Omega_r) = \frac{\varepsilon}{r!}\left(\mathtt{m}_{2r} - \frac{r!}{\varepsilon}\right) + \varepsilon\mathtt{m}(\Omega_{r-1}) \in \mathrm{qm}(\emptyset, S_2).$$

Thus (ii) and (ii') are true.

Example 6.11. Let $f = \mathtt{m}_{2,0}\mathtt{m}_{0,2}$. Since f is a product of two elements in $\mathrm{qm}(\emptyset, \emptyset)$, it is nonnegative on $\mathbf{P}(\mathbb{R}^2)$; on the other hand, f does not belong to $\mathrm{qm}(\emptyset, \emptyset)$. For a fixed $r \in \mathbb{N}$, searching for the smallest $\varepsilon(r) > 0$ such that $f + \varepsilon(r)\mathtt{m}(\Omega_r) \in \widetilde{\mathrm{qm}}(\emptyset, \emptyset)$ can be formulated as an SDP. For small values of r one obtains $\varepsilon(2) = 0.5$, $\varepsilon(3) = 0.10447$, $\varepsilon(4) = 0.02202$.

Example 6.12. Let $f = \mathtt{m}_{1,2}\mathtt{m}_{2,1} - \mathtt{m}_{1,1}^3$ [BRS+22, Example 1.1]. By the results of [BRS+22], f is nonnegative on $\mathbf{P}(\mathbb{R}^2_{\geq 0})$, but $\phi(f) < 0$ for some homomorphism $\phi : \mathscr{M} \to \mathbb{R}$ such that $\phi(\mathrm{qm}(\{x_1, x_2\}, \emptyset)) = \mathbb{R}_{\geq 0}$ (that is, the inequality $f \geq 0$ is not valid for all pseudo-moments on the nonnegative orthant). Nevertheless, for every $\varepsilon > 0$, Theorem 6.7 guarantees an $r \in \mathbb{N}$ such that $f + \varepsilon\mathtt{m}(\Omega_r) \in \mathrm{qm}(\{x_1, x_2\}, \emptyset)$. Alternatively, since f is homogeneous with respect to the degree on $\mathscr{M}$, its nonnegativity on $\mathbf{P}(\mathbb{R}^2_{\geq 0})$ is equivalent to nonnegativity on $\mathbf{P}([0, 1]^2)$. By Theorem 4.2, $f + \varepsilon \in \mathrm{qm}(\{x_1, 1 - x_1, x_2, 1 - x_2\}, \emptyset)$ for every $\varepsilon > 0$.

6.1. **Polynomial positivity on arbitrary semialgebraic sets.** Theorem 6.7 also carries implications for (classical, non-moment) polynomials.

Corollary 6.13. *Let $S \subseteq \mathbb{R}[\underline{x}]$. Then the following are equivalent for $f \in \mathbb{R}[\underline{x}]$:*

(i) *$f \geq 0$ on $K(S)$;*

(ii) *for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that $f + \varepsilon\Omega_r \in M(S)$.*

Proof. The homomorphism $\zeta : \mathscr{M}[\underline{x}] \to \mathbb{R}[\underline{x}]$ determined by $\zeta|_{\mathbb{R}[\underline{x}]} = \mathrm{id}_{\mathbb{R}[\underline{x}]}$ and $\zeta(\mathtt{m}_{i_1,\dots,i_n}) = 0$ (for i_j not all zero) maps $\mathrm{QM}(S, \emptyset)$ into $M(S)$. Applying ζ to the conclusions of Theorem 4.2 for f and $\mathrm{QM}(S, \emptyset)$ gives the desired statement. □

Remark 6.14. Corollary 6.13 is a strengthening of [LN07, Corollary 3.7]; the latter assumes that S is finite and has the strong moment property, the interior of $K(S)$ is nonempty, and requires preorderings instead of quadratic modules.

Corollary 6.13 thus characterizes polynomial positivity on arbitrary basic closed semialgebraic sets, but differently from the renowned Krivine-Stengle Positivstellensatz [Mar08, Theorem 2.2.1]; while the latter certificate involves preorderings and denominators, the former involves quadratic modules and coefficient perturbations.

Example 6.15. Let us record one of the simplest cases to which [LN07, Corollary 3.7] does not apply. Clearly, $x_1x_2 \geq 0$ on $K(\{x_1, x_2\})$. By Corollary 6.13, for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that $x_1x_2 + \varepsilon\Omega_r \in M(\{x_1, x_2\})$.

For a fixed $r \in \mathbb{N}$, one can find the smallest $\varepsilon(r) > 0$ such that $f + \varepsilon(r)\mathtt{m}(\Omega_r) \in \widetilde{\mathrm{qm}}(\emptyset, \emptyset)$ by solving an SDP. For $r = 2, \dots, 8$ the values of $\varepsilon(r)$ are

$$0.5,\ 0.012428,\ 0.002016,\ 0.000580, 0.000238,\ 0.000117,\ 0.000065,\ 0.000032.$$

Example 6.16. Remark 6.9 shows that Corollary 6.13 allows one for certain modifications of polynomials Ω_r. One might contemplate whether only the constant term and the leading terms are essential; this is indeed true in certain cases [KSV22, Example 7.9]. However, the following example shows this is not true in general.

Let $n = 1$, $f = -2$ and $S = \{-1 + \frac{x_1^{2k}}{k!-1} \colon k \geq 2\}$. Then $K(S) = \emptyset$ and $f \geq 0$ on $K(S)$. We claim that $-2 + (1 + \frac{x_1^{2r}}{r!}) \notin M(S)$ for every $r \in \mathbb{N}$. Indeed, suppose

$$-1 + \frac{x_1^{2r}}{r!} = \sigma_1 + \sum_{k=2}^{\ell} \sigma_k \cdot \left(-1 + \frac{x_1^{2k}}{k!-1}\right) \tag{6.15}$$

where $\ell \geq 2$ and $\sigma_k \in \mathbb{R}[\underline{x}]$ are sums of squares. Note that $\ell \leq r$. Let $X = \sqrt[2\ell]{\ell! - 1}$. Then the right-hand side of (6.15) is nonnegative at X, while the left-hand side of (6.15) is negative at X, a contradiction.

On the other hand, if $\varepsilon > 0$ is arbitrary and $r \geq \frac{2}{\varepsilon} - 1$, then

$$-1 + \varepsilon\Omega_r = \left(-1 + \varepsilon\left(1 + \sum_{k=2}^{r} \frac{k!-1}{k!}\right)\right) + \varepsilon x_1^2 + \varepsilon \sum_{k=2}^{r} \frac{k!-1}{k!}\left(-1 + \frac{x_1^{2k}}{k!-1}\right) \in M(S),$$

as anticipated by Corollary 6.13.

We conclude the section with a modified Lasserre's SDP hierarchy, applicable to arbitrary semialgebraic sets. Let $S \subseteq \mathbb{R}[\underline{x}]$, $f \in \mathbb{R}[\underline{x}]$ and $\varepsilon > 0$. Let $f_* = \inf_{\underline{X} \in K(S)} f(\underline{X})$. For $r \geq \frac{\deg f}{2}$ consider the SDP

$$f_r^{(\varepsilon)} = \sup\left\{z \in \mathbb{R} \colon f - z + \varepsilon\Omega_r \in M(S)_{2r}\right\}.$$

Corollary 6.17. *Let S, f, ε be as above. Then $(f_r^{(\varepsilon)})_r$ is an increasing sequence, and*

$$f_* \leq \lim_{r\to\infty} f_r^{(\varepsilon)} \leq \inf_{\underline{X}\in K(S)} \left(f(\underline{X}) + \varepsilon(\exp(X_1^2) + \cdots + \exp(X_n^2)\right). \tag{6.16}$$

In particular,

$$\lim_{\varepsilon\downarrow 0} \lim_{r\to\infty} f_r^{(\varepsilon)} = f_*. \tag{6.17}$$

Proof. The first inequality in (6.16) holds by Corollary 6.13, and the second inequality in (6.16) is straightforward. Lastly, (6.17) follows from $\lim_{\varepsilon\downarrow 0} \inf_K (f + \varepsilon g) = \inf_K f$ for any nonnegative function g on $\mathbb{R}^n$. ☐

References

[AJK15] Daniel Alpay, Palle E. T. Jorgensen, and David P. Kimsey. Moment problems in an infinite number of variables. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 18(04):1550024, 2015. 17

[Are46] Richard Arens. The space L^ω and convex topological rings. *Bull. Amer. Math. Soc.*, 52:931–935, 1946. 6

[Bar02] Alexander Barvinok. *A course in convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002. 17, 19

[BBLM05] Stéphane Boucheron, Olivier Bousquet, Gábor Lugosi, and Pascal Massart. Moment inequalities for functions of independent random variables. *The Annals of Probability*, 33(2):514 – 560, 2005. 1

[BKP16] Sabine Burgdorf, Igor Klep, and Janez Povh. *Optimization of polynomials in non-commuting variables*. SpringerBriefs in Mathematics. Springer, 2016. 2

[BP05] Dimitris Bertsimas and Ioana Popescu. Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal on Optimization*, 15(3):780–804, 2005. 1

[BRS+22] Grigoriy Blekherman, Felipe Rincón, Rainer Sinn, Cynthia Vinzant, and Josephine Yu. Moments, sums of squares, and tropicalization, 2022. 2, 4, 9, 22

[CP10] Raúl Curto and Mihai Putinar. Polynomially hyponormal operators. In *A Glimpse at Hilbert Space Operators: Paul R. Halmos in Memoriam*, pages 195–207. Springer Basel, 2010. 1

[DLTW08] Andrew C. Doherty, Yeong-Cherng Liang, Ben Toner, and Stephanie Wehner. The quantum moment problem and bounds on entangled multi-prover games. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 199–210. IEEE, 2008. 2

[FF23] Giovanni Fantuzzi and Federico Fuentes. Global minimization of polynomial integral functionals. *preprint arXiv:2305.18801*, 2023. 1

[HIKV23] Didier Henrion, Maria Infusino, Salma Kuhlmann, and Victor Vinnikov. Infinite-dimensional moment-SOS hierarchy for nonlinear partial differential equations. *preprint arXiv:2305.18768*, 2023. 1, 2

[HKL20] Didier Henrion, Milan Korda, and Jean-Bernard Lasserre. *Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*, volume 4. World Scientific, 2020. 2

[HM04] J. William Helton and Scott A. McCullough. A Positivstellensatz for non-commutative polynomials. *Trans. Amer. Math. Soc.*, 356(9):3721–3737, 2004. 2

[IKKM23] Maria Infusino, Salma Kuhlmann, Tobias Kuna, and Patrick Michalski. Moment problem for algebras generated by a nuclear space. *preprint arXiv:2301.12949*, 2023. 2

[IKR14] Maria Infusino, Tobias Kuna, and Aldo Rota. The full infinite dimensional moment problem on semi-algebraic sets of generalized functions. *Journal of Functional Analysis*, 267(5):1382–1418, 2014. 2, 17

[KMV22] Igor Klep, Victor Magron, and Jurij Volčič. Optimization over trace polynomials. *Ann. Henri Poincaré*, 23(1):67–100, 2022. 2

[KMVW23] Igor Klep, Victor Magron, Jurij Volčič, and Jie Wang. State polynomials: positivity, optimization and nonlinear bell inequalities. *arXiv preprint arXiv:2301.12513*, 2023. 2, 9, 11, 13

[KPT21] Brendan Kline, Ariel Pakes, and Elie Tamer. Moment inequalities and partial identification in industrial organization. In Kate Ho, Ali Hortaçsu, and Alessandro Lizzeri, editors, *Handbook of Industrial Organization*, volume 4, pages 345–431. Elsevier, 2021. 1

[KPV21] Igor Klep, James E. Pascoe, and Jurij Volčič. Positive univariate trace polynomials. *J. Algebra*, 579:303–317, 2021. 8

[KSV22] Igor Klep, Claus Scheiderer, and Jurij Volčič. Globally trace-positive noncommutative polynomials and the unbounded tracial moment problem. *Math. Ann. to appear,* `doi.org/10.1007/s00208-022-02495-5`, 2022. 22

[Las01] Jean-Bernard Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2000/01. 2, 3

[Las06] Jean-Bernard Lasserre. A sum of squares approximation of nonnegative polynomials. *SIAM Journal on Optimization*, 16(3):751–765, 2006. 4, 18, 19

[Las09] Jean-Bernard Lasserre. *Moments, Positive Polynomials and Their Applications.* Imperial College Press optimization series. Imperial College Press, 2009. 2

[LJ10] Yuan Liao and Wenxin Jiang. Bayesian analysis in moment inequality models. *The Annals of Statistics*, 38(1):275–316, 2010. 1

[LN07] Jean-Bernard Lasserre and Tim Netzer. Sos approximations of nonnegative polynomials via simple high degree perturbations. *Math. Z.*, 256:1432–1823, 2007. 4, 22

[Mar08] Murray Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008. 1, 3, 5, 9, 10, 11, 12, 22

[MJC+14] Lester Mackey, Michael I. Jordan, Richard Y. Chen, Brendan Farrell, and Joel A. Tropp. Matrix concentration inequalities via the method of exchangeable pairs. *The Annals of Probability*, 42(3):906 – 945, 2014. 1

[MW23] Victor Magron and Jie Wang. *Sparse polynomial optimization: theory and practice*. Series on Optimization and Its Applications, World Scientific Press, 2023. Forthcoming. 13

[NPA08] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. Phys.*, 10(7):073013, 2008. 2

[Par05] Kalyanapuram Rangachari Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Soc., Providence, RI, 2005. 6, 12

[PHBB17] Victor Pozsgay, Flavien Hirsch, Cyril Branciard, and Nicolas Brunner. Covariance Bell inequalities. *Phys. Rev. A*, 96(6):062128, 13, 2017. 1, 2, 3, 13, 14

[PPHI15] Ariel Pakes, Jack Porter, Kate Ho, and Joy Ishii. Moment inequalities and their application. *Econometrica*, 83(1):315–334, 2015. 1

[PS01] Victoria Powers and Claus Scheiderer. The moment problem for non-compact semialgebraic sets. *Adv. Geom.*, 1(1):71–88, 2001. 16

[Put93] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993. 2, 3

[Put97] Mihai Putinar. A note on Tchakaloff's theorem. *Proc. Amer. Math. Soc.*, 125(8):2409–2414, 1997. 7

[Sch90] Konrad Schmüdgen. *Unbounded Operator Algebras and Representation Theory*, volume 37 of *Operator Theory: Advances and Applications*. Birkhäuser Basel, 1990. 18

[Sch12] Konrad Schmüdgen. *Unbounded self-adjoint operators on Hilbert space*, volume 265 of *Graduate Texts in Mathematics*. Springer Dordrecht, 2012. 11, 17

[Sch17] Konrad Schmüdgen. *The moment problem*, volume 277 of *Graduate Texts in Mathematics*. Springer Cham, 2017. 2, 11, 17, 18

[TGB21] Armin Tavakoli, Nicolas Gisin, and Cyril Branciard. Bilocal Bell inequalities violated by the quantum elegant joint measurement. *Phys. Rev. Lett.*, 126(22):Paper No. 220401, 2021. 1, 3, 4, 13, 14

[TPKLR22] Armin Tavakoli, Alejandro Pozas-Kerstjens, Ming-Xing Luo, and Marc-Olivier Renou. Bell nonlocality in networks. *Rep. Progr. Phys.*, 85(5):Paper No. 056001, 41, 2022. 1, 2, 3, 4, 14

[Wil70] Stephen Willard. *General topology*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970. 17

Igor Klep: Faculty of Mathematics and Physics, Department of Mathematics, University of Ljubljana & Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia
Email address: `igor.klep@fmf.uni-lj.si`

Victor Magron: LAAS-CNRS & Institute of Mathematics from Toulouse, France
Email address: `vmagron@laas.fr`

Jurij Volčič: Department of Mathematics, Drexel University, Pennsylvania
Email address: `jurij.volcic@drexel.edu`

NOT FOR PUBLICATION

Contents