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. On the one hand, 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 open nonlinear Bell inequalities from quantum physics are settled.

CONTENTS

1. Introduction

Consider two groups of independent variables, $x_1, \ldots, x_n$ and $\mathtt{m}_{i_1,\ldots,i_n}$ for $i_1, \ldots i_n \in \mathbb{N}_0$. Here, the latter are viewed as formal mixed moments, also denoted $\mathtt{m}_{i_1,\ldots,i_n} = \mathtt{m}(x_1^{i_1} \cdots x_n^{i_n})$, 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 $x_1, \ldots, x_n$ and their formal moments, and their inequalities. A very rudimentary instance of a moment polynomial inequality is that every random variable X has a nonnegative variance,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \mathrm{E}\big((X - \mathrm{E}(X))^2\big) \geq 0;$$

in the language of this paper, we say that the moment polynomial $\mathtt{m}(x_1^2) - \mathtt{m}(x_1)^2$ is nonnegative. Problems involving moment polynomial inequalities and optimization arise in various fields. For example, concentration inequalities in moments provide bounds on deviation of a random variable, and the search for such inequalities has been a flourishing area of probability theory [BBLM05, BP05, MJC⁺14]. In statistics, distributions of structural parameters are partially identified by moment inequalities [LJ10]; similarly, moment inequality constraints can be used as a basis for estimation and inference in partially identified behavioral models in economics [PPHI15] and industrial organization [KPT21]. Another instance of moment polynomial inequalities arises in operator theory [CP10], where they characterize hyponormality of operators. A bounded operator T is k-hyponormal if the $k \times k$ block matrix with commutator entries $[(T^*)^i, T^j]$ is positive semidefinite; this concept is suitable for extending the spectral results on normal operators. Moment optimization is also used to study partial differential equations: in [HIKV23], solving a heat equation with a nonlinear perturbation is formulated as a linear optimization problem on moments, while [Fan22] applies polynomial optimization to verify integral inequalities arising from Lyapunov analysis of fluid flows. Lastly, nonlinear Bell inequalities in quantum information theory are moment inequalities that certify nonlocality in quantum networks [PHBB17, TGB21, TPKLR22].

To systematically approach moment polynomial inequalities, 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 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 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 the 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 to 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 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.

1.1. **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,

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

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

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

A moment polynomial without freely occurring x_j, e.g. $\mathtt{m}(x_1^2 x_2^2) - \mathtt{m}(x_1)^4 + \mathtt{m}(x_1)\mathtt{m}(x_2)\mathtt{m}(x_1 x_2)$, 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.

To study constrained positivity of moment polynomials, let $S_1 \subseteq \mathbb{R}[\underline{x}]$ and $S_2 \subseteq \mathscr{M}$ be collections of constraints. 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 $\mathbf{K}(S_1, S_2)$ 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

$$\left\{ f^2 \mathtt{m}(g^2 s), \ f^2 t \colon s \in S_1 \cup \{1\}, \ t \in S_1 \cup S_2, \ f, g \in \mathscr{M}[\underline{x}] \right\}.$$

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

1.2. **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 (even more, the quadratic module in $\mathbb{R}[\underline{x}]$ generated by S_1 is archimedean in the usual sense). 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 $\mathbf{K}(S_1, S_2) \times K(S_1)$.

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

(i) $f \geq 0$ *on* $\mathbf{K}(S_1, S_2) \times K(S_1)$;
(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 procedure for solving the moment polynomial optimization problem

(1.1) $\qquad$ minimize $f(\mu, \underline{X})$ subject to $\underline{X} \in K(S_1)$ and $\mu \in \mathbf{K}(S_1, S_2)$

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* $\mathbf{K}(S_1, S_2) \times K(S_1)$.

We apply Theorem B to moment polynomial optimization problems of the form (1.1) arising from quantum information theory. Two nonlinear Bell inequalities conjectured in [PHBB17] and [TGB21, TPKLR22] are established. For example, our optimization scheme allows us to solve the following problem:

$$\sup \frac{1}{3} \sum_{i \in \{1,2,3\}} \left(\mathtt{m}(x_{i+3}x_{i+6}) - \mathtt{m}(x_i x_{i+3}) \right) - \sum_{\{i,j,k\}=\{1,2,3\}} \mathtt{m}(x_i x_{j+3} x_{k+6})$$

subject to

$$\begin{aligned}
(1.2) \qquad & \mathtt{m}(x_1^{k_1} x_2^{k_2} x_3^{k_3} x_7^{k_4} x_8^{k_5} x_9^{k_6}) = \mathtt{m}(x_1^{k_1} x_2^{k_2} x_3^{k_3}) \mathtt{m}(x_7^{k_4} x_8^{k_5} x_9^{k_6}) \quad \text{for } k_i \in \{0,1\}, \\
& x_j^2 = 1 \text{ and } \mathtt{m}(x_j) = 0 \quad \text{for } j \in \{0, \ldots, 9\}, \\
& \mathtt{m}(x_i x_{j+3}) = \mathtt{m}(x_{i+3} x_{j+6}) = 0 \quad \text{for } i, j \in \{1,2,3\}, \ i \neq j, \\
& \mathtt{m}(x_i x_{j+3} x_{k+6}) = 0 \quad \text{for } i, j, k \in \{1,2,3\}, \ |\{i,j,k\}| \leq 2.
\end{aligned}$$

In Subsection 5.2 it is shown that the solution of (1.2) is 4, attained by certain binary variables and the uniform measure on 16 points, thereby answering 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 3.8). More precisely, the algebraic certificate in (H17) characterizes the 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 3.7). *The following statements 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 $h \in \mathscr{M}[\underline{x}]$.*

In other words, (ii) means that f is a quotient of two elements from the preordering in $\mathscr{M}[\underline{x}]$ generated by formal moments of squares. 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 arbitrarily small perturbation of their coefficients. Our second main result generalizes Lasserre's certificate to moment polynomials.

Theorem D (Theorem 6.9). *If S_2 is finite, the following statements are equivalent for $f \in \mathscr{M}[\underline{x}]$:*

(i) *$f \geq 0$ on $\mathbf{K}(S_1, S_2) \times K(S_1)$;*

(ii) *for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that*

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

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 restrictive assumptions.

2. Preliminaries

We start by introducing the notation and terminology pertaining to moment polynomials and their evaluations, and establishing some basic facts that are used throughout the paper. Let $\mathbb{R}[\underline{x}] = \mathbb{R}[x_1, \ldots, x_n]$ be the polynomial ring in n variables. Consider the polynomial ring in countably many variables $\mathscr{M} = \mathbb{R}[\mathtt{m}_{i_1,\ldots,i_n} : i_1, \ldots, i_n \in \mathbb{N}_0]$ where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $\mathtt{m}_{0,\ldots,0} := 1$. Furthermore, 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 $\mathtt{m}(x_1^{i_1} \cdots x_n^{i_n}) = \mathtt{m}_{i_1,\ldots,i_n}$.

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 $\mathrm{qm}(S_1, S_2) \subseteq \mathscr{M}$ be the quadratic module in $\mathscr{M}$ generated by

$$\{\mathtt{m}(f^2 s) \colon s \in \{1\} \cup S_1, f \in \mathscr{M}[\underline{x}]\} \cup S_2,$$

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

$$S_1 \cup \{\mathtt{m}(f^2 s) \colon s \in \{1\} \cup S_1, f \in \mathscr{M}[\underline{x}]\} \cup S_2.$$

Alternatively, $\mathrm{qm}(S_1, S_2)$ consists of sums of elements of the form

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

for $s_1 \in \{1\} \cup S_1$, $s_2 \in S_2$, $f \in \mathscr{M}[\underline{x}]$ and $q \in \mathscr{M}$. Indeed, note that sums of (2.1) are preserved under multiplication by squares in $\mathscr{M}$ because $q^2 \mathtt{m}(f^2 s) = \mathtt{m}((qf)^2 s)$ for $q \in \mathscr{M}$ and $f, s \in \mathscr{M}[\underline{x}]$ (that is, $\mathtt{m}$ is $\mathscr{M}$-linear). Similarly, $\mathrm{QM}(S_1, S_2)$ consists of sums of elements of the form

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

for $s_1 \in \{1\} \cup S_1$, $s_2 \in S_2$ and $f, g \in \mathscr{M}[\underline{x}]$. Also, let $\mathrm{QM}(S_1) \subset \mathbb{R}[\underline{x}]$ denote the quadratic module in $\mathbb{R}[\underline{x}]$ generated by S_1.

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

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

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

and $\deg(fg) = \deg(f) + \deg(g)$ for $f, g \in \mathscr{M}[\underline{x}]$. 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 $\mathrm{qm}(S_1, S_2)_{2r} \subseteq \mathscr{M}_{2r}$ and $\mathrm{QM}(S_1, S_2)_{2r} \subseteq \mathscr{M}[\underline{x}]_{2r}$ be sums 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 $\mathrm{qm}(\emptyset, \emptyset)$ that are required later. It can be viewed as a formal special case of Hölder's inequality [Rud87, Theorem 3.5].

Lemma 2.2 (Symbolic Hölder's inequality). *Let $k \in \mathbb{N}_0$.*

    (1) *If $n = 1$ then $\mathtt{m}_{2k} - \mathtt{m}_1^{2k} \in \mathrm{qm}(\emptyset, \emptyset)$.*
    (2) *If $n \in \mathbb{N}$ and $i_1, \dots, i_n \in \mathbb{N}_0$ then $\mathtt{m}_{2ki_1, \dots, 2ki_n} - \mathtt{m}_{i_1, \dots, i_n}^{2k} \in \mathrm{qm}(\emptyset, \emptyset)$.*

Proof. (1) The case $k = 0$ holds because $\mathtt{m}_0 - \mathtt{m}_1^0 = 0$, and the case $k = 1$ holds because

$$\mathrm{qm}(\emptyset, \emptyset) \ni \mathtt{m}((\mathtt{m}_1 - x_1)^2) = \mathtt{m}(\mathtt{m}_1^2 - 2\mathtt{m}_1 x_1 + x_1^2) = \mathtt{m}_1^2 - 2\mathtt{m}_1^2 + \mathtt{m}_2 = \mathtt{m}_2 - \mathtt{m}_1^2.$$

For $k \geq 2$, we claim that

$$(2.3) \qquad \mathtt{m}_{2k} - \mathtt{m}_1^{2k} = \sum_{j=1}^{k-1} j \mathtt{m}\left(\left(x_1^{k-j+1} \mathtt{m}_1^{j-1} - x_1^{k-j-1} \mathtt{m}_1^{j+1} \right)^2 \right) + k \mathtt{m}\left(\left(x_1 \mathtt{m}_1^{k-1} - \mathtt{m}_1^k \right)^2 \right),$$

which implies (1). To prove (2.3), denote $A = \sum_{j=1}^{k-1} \mathtt{m}_{2(k-j)} \mathtt{m}_1^{2j}$ and $B = \sum_{j=1}^{k-1} j \mathtt{m}_{2(k-j)} \mathtt{m}_1^{2j}$. The right-hand side of (2.3) expands as

$$\sum_{j=1}^{k-1} j \left(\mathtt{m}_{2(k-j+1)} \mathtt{m}_1^{2(j-1)} + \mathtt{m}_{2(k-j-1)} \mathtt{m}_1^{2(j+1)} - 2\mathtt{m}_{2(k-j)} \mathtt{m}_1^{2j} \right) + k \left(\mathtt{m}_2 \mathtt{m}_1^{2(k-1)} - \mathtt{m}_1^{2k} \right)$$

$$= \sum_{j=1}^{k-1} j \mathtt{m}_{2(k-j+1)} \mathtt{m}_1^{2(j-1)} + \sum_{j=1}^{k-1} j \mathtt{m}_{2(k-j-1)} \mathtt{m}_1^{2(j+1)} - 2 \sum_{j=1}^{k-1} j \mathtt{m}_{2(k-j)} \mathtt{m}_1^{2j} + k \left(\mathtt{m}_2 \mathtt{m}_1^{2(k-1)} - \mathtt{m}_1^{2k} \right)$$

$$= \left(B + A + \mathtt{m}_{2k} - k \mathtt{m}_2 \mathtt{m}_1^{2(k-1)} \right) + \left(B - A + (k-1) \mathtt{m}_{2k} \right) - 2B + k \mathtt{m}_2 \mathtt{m}_1^{2(k-1)} - k \mathtt{m}_1^{2k}$$

$$= \mathtt{m}_{2k} - \mathtt{m}_1^{2k},$$

as desired.

(2) Given $i_1, \ldots, i_n \in \mathbb{N}_0$, consider the homomorphism $\Theta : \mathbb{R}[x_1, \mathtt{m}_k \colon k \in \mathbb{N}] \to \mathscr{M}[x]$ determined by $\Theta(x_1) = x_1^{i_1} \cdots x_n^{i_n}$ and $\Theta(\mathtt{m}_k) = \mathtt{m}_{ki_1, \ldots, ki_n}$. Then Θ intertwines with the map $\mathtt{m}$ (i.e., $\Theta \circ \mathtt{m} = \mathtt{m} \circ \Theta$), so applying it to $\mathtt{m}_{2k} - \mathtt{m}_1^{2k} \in \mathrm{qm}(\emptyset, \emptyset)$ results in $\mathtt{m}_{2ki_1, \ldots, 2ki_n} - \mathtt{m}_{i_1, \ldots, i_n}^{2k} \in \mathrm{qm}(\emptyset, \emptyset)$. $\square$

Remark 2.3. Not all cases of Hölder's inequality admit a symbolic interpretation as in Lemma 2.2. For example, $\mathtt{m}_{11} \mathtt{m}_{22} - \mathtt{m}_{12}^2 \notin \mathrm{qm}(\emptyset, \emptyset)$. Indeed, if $f_1, \ldots, f_\ell \in \mathscr{M}[x]$ and uv with $u \in \mathscr{M}$ and $v \in [x]$ is a monomial in one of the f_j with maximal $\deg v$, then $\mathtt{m}(\sum_j f_j^2)$ contains the monomial $u^2 \mathtt{m}(v^2)$ with a positive coefficient, and consequently $\mathtt{m}(\sum_j f_j^2) \neq \mathtt{m}_{11} \mathtt{m}_{22} - \mathtt{m}_{12}^2$.

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, $\int x_j^{2k} \, d\mu < \infty$ for all $j = 1, \ldots, n$ 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

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

determined by

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

Such homomorphisms are called *moment evaluations*.

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

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

determined by

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

The bracket notation in (2.5) is ad hoc, and intended simply as a distinction from (2.4). 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 σ-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

$$K(S_1) = \{\underline{X} \in \mathbb{R}^n \colon p(\underline{X}) \geq 0 \text{ for all } p \in S_1\} \subseteq \mathbb{R}^n,$$
$$\mathbf{K}(S_1, S_2) = \{\mu \in \mathbf{P}(K(S_1)) \colon s(\mu) \geq 0 \text{ for all } s \in S_2\} \subseteq \mathbf{P}(K(S_1)).$$

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

Proposition 2.4. *Let $S_1 \subseteq \mathbb{R}[\underline{x}]$ and $S_2 \subseteq \mathscr{M}$. The following statements are equivalent for $f \in \mathscr{M}[\underline{x}]$:*

(i) *$f(\mu, \underline{X}) \geq 0$ for all $(\mu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_1)$;*
(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)$\Rightarrow$(ii): Suppose a probability space $(\mathcal{X}, \Sigma, \pi)$ and a random variable $\underline{F}$ on $\mathcal{X}$ 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 $P \in \mathcal{X}$ we have $(\mu, \underline{F}(P)) \in \mathbf{K}(S_1, S_2) \times K(S_1)$ and $f(\mu, \underline{F}(P)) = f[\pi, \underline{F}](P)$. This shows that (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). $\qquad\square$

We have introduced the evaluations of the second type mainly to indicate the scope of our framework, and they appear only in applications in Sections 5.1 and 5.2; in the rest of the paper, we deal with evaluations of the first type for the sake of simplicity (and this is sufficient by Proposition 2.4). The following statement is a straightforward consequence of definitions, and represents the easy part of the algebra-geometry correspondence for moment polynomial positivity that is pursued in this paper. Namely, elements of the moment quadratic module are nonnegative on the corresponding moment semialgebraic set.

Proposition 2.5. *Let $S_1 \subseteq \mathbb{R}[\underline{x}]$ and $S_2 \subseteq \mathscr{M}$. If $f \in \mathrm{QM}(S_1, S_2)$ then $f \geq 0$ on $\mathbf{K}(S_1, S_2) \times K(S_1)$.*

Let us comment on the restriction to probability measures in this paper. While moment polynomials can be evaluated on more general measures, the algebraic certificates of this paper involve constant term perturbations, or quotients of moment polynomials. The fact that such certificates indeed guarantee positivity is rooted in the integral of the constant 1 being a fixed positive number; or, from an algebraic perspective, that $\mathtt{m} \colon \mathscr{M}[\underline{x}] \to \mathscr{M}$ is $\mathscr{M}$-linear. In general, our results can be modified to moment evaluations on measures μ with $\int \mathrm{d}\mu = t$ for a fixed $t \in \mathbb{R}_{>0}$, by simply renormalizing the measures to probability measures. For this reason we restrict ourselves to probability measures for the sake of simplicity.

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 2.6 (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 $\mathrm{m}_{i_1,\dots,i_n}(\nu) = \mathrm{m}_{i_1,\dots,i_n}(\mu)$ for all $i_1 + \cdots + i_n \leq d$.*

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

Remark 2.7. Proposition 2.6 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 $S_2 \subset \mathcal{M}$ be finite, and $f \in \mathcal{M}[\underline{x}]$. Denote $d = \max\{\deg f, \deg s \colon s \in S_2\}$ and $D = \binom{n+d}{d}$. By Proposition 2.6, the infimum of f on $\mathbf{K}(S_1, S_2) \times K(S_1)$ is equal to

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

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 consequently approached with standard methods of polynomial optimization, it has $n + Dn + D = \left(\binom{n+d}{d} + 1\right)(n+1) - 1$ 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. A moment polynomial optimization procedure bypassing these issues is developed in Section 5 below.

2.2. Comparison with existing work.
Now that moment polynomials and their evaluations have been rigorously introduced, let us compare them to related constructions in the literature.

2.2.1. Linear moment optimization.
A special case of the moment polynomial optimization problem 1.1 has already been well-studied. Let $p \in \mathbb{R}[\underline{x}]$, $S_1 \subset \mathbb{R}[\underline{x}]$ and $S_2 \subset \{\mathrm{m}(q) \colon q \in \mathbb{R}[\underline{x}]\} \subset \mathcal{M}$. Then

$$(2.7) \qquad \text{minimize } \int p \, \mathrm{d}\mu \text{ subject to } \mu \in \mathbf{K}(S_1, S_2)$$

can be viewed as a linear moment polynomial optimization problem, since both its objective function $\mathrm{m}(p)$ and its pure moment polynomial constraints S_2 are linear in the moment symbols m. Approaches to (2.7) are typically based on solutions of the classical moment problem [Sch17]. The optimization problem (2.7) has numerous applications, especially in polynomial optimization; its linear nature in particular allows for more refined results than the ones presented in this paper. We refer to [Las09, HKL20] for a comprehensive source of this special case of moment polynomial optimization.

2.2.2. *State and trace polynomials.* Noncommutative analogs of the moment polynomial positivity framework have been introduced in [KMV22] and [KMVW23]. While moment polynomials are polynomial expressions in commuting variables $x_1, \dots, x_n$ and formal symbols of their products $\mathtt{m}(x_1^{i_1} \cdots x_n^{i_n})$, *state polynomials* of [KMVW23] are polynomial expressions in *noncommuting* variables $x_1, \dots, x_n$ and formal symbols of their products $\varsigma(x_{j_1} \cdots x_{j_\ell})$, called *formal states*. While moment polynomials are evaluated on pairs of a point $\underline{X} \in \mathbb{R}^n$ and a measure μ on $\mathbb{R}^n$, state polynomials are evaluated on pairs of a tuple of self-adjoint operators $\underline{X} \in \mathcal{B}(\mathcal{H})^n$ and a state λ on $\mathcal{B}(\mathcal{H})$ (a state is a unital positive linear functional $\mathcal{B}(\mathcal{H}) \to \mathbb{C}$). If one is only interested in evaluations of state polynomials on tracial states (a state $\lambda : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ is tracial if $\lambda(AB) = \lambda(BA)$ for all $A, B \in \mathcal{B}(\mathcal{H})$), then one can impose additional tracial relations on the symbol ς, namely $\varsigma(w_1 w_2) = \varsigma(w_2 w_1)$ for all products w_1, w_2 of $x_1, \dots, x_n$, and in this manner one obtains *trace polynomials* [KMV22], which originate in invariant theory [Pro76]. There is a hierarchy among moment, trace and state polynomial frameworks. Let S be a collection of state polynomial inequalities. First, positivity of trace polynomials subject to constraints S can be viewed as positivity of state polynomials subject to constraints S together with tracial constraints, namely $S \cup \{\pm(\varsigma(w_1 w_2) - \varsigma(w_2 w_1)) : w_1, w_2\}$. Second, positivity of moment polynomials subject to constraints S can be viewed as positivity of trace polynomials subject to constraints S together with commutativity constraints, namely $S \cup \{\pm(x_i x_j - x_j x_i) : i, j\}$; this equivalence is exact because states on jointly commuting (bounded) operators correspond to (compactly supported) measures by the spectral theorem [Sch12, Theorem 5.23]. All three frameworks are applied in quantum information theory. Classical shared randomness, maximal quantum entanglement, and general quantum entanglement can be investigated with moment, trace and state polynomials, respectively.

With the above hierarchy in mind, it is reasonable to ask why these three frameworks are investigated separately. The answer is that the scope and the strength of positivity certificates vary considerably among the frameworks. Let us demonstrate this with the main results presented in the introduction. When positivity without constraints is considered, the state analog of Hilbert's 17th problem admits a positive resolution, while trace and moment analogs of Hilbert's 17th problem do not. Nevertheless, Theorem C exactly identifies to what extent such a resolution fails for moment polynomials; namely, a moment polynomial admits a sum-of-squares certificate with denominators if and only if it is nonnegative under all pseudo-moment evaluations. Next, when positivity on bounded domains is considered, Theorem A provides a Positivstellensatz for arbitrary moment polynomials; in contrast, within the other two frameworks, such a statement only holds for *pure* variants of state and trace polynomials (polynomial expressions in states or traces, without freely appearing noncommuting variables $x_1, \dots, x_n$). Finally, Theorem D establishes a perturbative Positivstellensatz for moment polynomial positivity subject to arbitrary (not necessarily bounded) moment semialgebraic sets; on the contrary, no such general result exists for state or trace polynomials. The underlying principle of these differences is that the commutative theory (real algebraic geometry, measure theory) provides more sophisticated techniques than the noncommutative theory (free real algebraic geometry, operator algebras).

2.2.3. *Related advances on polynomial inequalities in moments.* Let us highlight three other works that can be compared to the moment polynomial framework of this paper.

In [BRS$^+$22], positivity of pure moment *binomials* (i.e., polynomials with only two additive terms) is analyzed using tropicalization, resulting in a combinatorial description of nonnegative pure moment binomials subject to binomial inequalities. The certificates of [BRS$^+$22] are more definitive than the ones of this paper (easier to utilize, and of combinatorial nature as opposed to Theorem 6.9 for instance), but they only apply to binomials. Next, [Fan22] investigates polynomial inequalities in integrals of function variables and their derivatives. This setup can be compared with our framework on moment polynomial inequalities in random variables in light of Proposition 2.4. The setup in [Fan22] is more general because it considers moments of not only function variables, but also of their derivatives. However, [Fan22] derives only sufficiency conditions for polynomial integral inequalities, which are comparable to the reverse implication in Theorem A, but does not establish their necessity. Lastly, let us refer to [IKR14, IKKM23] on infinite-dimensional moment problems. Often, moment problems are duals of sum-of-squares positivity certificates, and our paper provides the latter for moment polynomials, which are symmetric algebras on an infinite-dimensional space (spanned by the formal moments). However, this point of view is not compatible with the scope of [IKR14, IKKM23]. Namely, these papers investigate moment problems for algebras generated by nuclear spaces [SW99] (e.g., infinitely differentiable functions on $\mathbb{R}^n$ with compact support form a nuclear space), and the topological assumption of nuclearity plays a crucial role in their moment problem solvability criteria. In contrast, our positivity certificates are based on algebraic manipulations of formal moment symbols without substantial topological considerations, and leverage the real-algebraic nature of $\mathbb{R}[\underline{x}]$, which leads to a computationally viable framework. Therefore our positivity certificates do not aspire to a moment problem, nor do moment problems associated to nuclear spaces carry implications for moment polynomial positivity.

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 3.7). 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 terms of the form g^2 and $\mathtt{m}(g^2)$ for $g \in \mathscr{M}[\underline{x}]$.

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. The relation between the matrix H_d and moments of squares in $\mathscr{M}[\underline{x}]$ is captured by the following statement.

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

Such homomorphism is called a pseudo-moment evaluation.

Proof. The implication (i)$\Rightarrow$(ii) is clear. Before proving (ii)$\Rightarrow$(iii)$\Rightarrow$(i), notice that since the (u,v)-entry of H_d equals uv, the (u,v)-entry of $(\phi \circ \mathtt{m})(H_d)$ equals $\phi(\mathtt{m}(uv))$. To see

(ii)$\Rightarrow$(iii), suppose that (ii) holds, and let $c = (c_u)_{u \in [\underline{x}]_d}$ be a real column vector. Then

$$c^\top \left((\phi \circ \mathtt{m})(H_d) \right) c = \sum_{u,v \in [\underline{x}]_d} c_u c_v \phi(\mathtt{m}(uv)) = \phi \left(\mathtt{m} \left(\sum_{u,v \in [\underline{x}]_d} c_u c_v uv \right) \right)$$

$$= \phi \left(\mathtt{m} \left(\left(\sum_{u \in [\underline{x}]_d} c_u u \right)^2 \right) \right) \geq 0$$

by (ii). Since c was arbitrary, it follows that $(\phi \circ \mathtt{m})(H_d)$ is a positive semidefinite matrix, and (iii) holds. To see (iii)$\Rightarrow$(i), suppose that (iii) holds. Let $f \in \mathscr{M}[\underline{x}]$; since $\mathscr{M}[\underline{x}]$ is the tensor product of $\mathscr{M}$ and $\mathbb{R}[\underline{x}]$, there is $d \in \mathbb{N}$ such that $f = \sum_{u \in [\underline{x}]_d} q_u u$ for some $q_u \in \mathscr{M}$. Then

$$\phi \left(\mathtt{m}(f^2) \right) = \phi \left(\mathtt{m} \left(\sum_{u,v \in [\underline{x}]_d} q_u q_v uv \right) \right) = \sum_{u,v \in [\underline{x}]_d} \phi(q_u) \phi(q_v) \phi \left(\mathtt{m}(uv) \right) = c^\top \left((\phi \circ \mathtt{m})(H_d) \right) c,$$

where $c := (\phi(q_u))_{u \in [\underline{x}]_d}$ is a real column vector. Since $(\phi \circ \mathtt{m})(H_d)$ is positive semidefinite, we have $\phi \left(\mathtt{m}(f^2) \right) \geq 0$, so (i) holds. $\qquad\square$

Remark 3.2. A pseudo-moment evaluation $\phi : \mathscr{M}[\underline{x}] \to \mathbb{R}$ is uniquely determined by $\phi(x_j) \in \mathbb{R}$ for $j = 1, \ldots, n$, 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 3.3. There is a certain subtlety in Lemma 3.1. Namely, the implication (ii)$\Rightarrow$(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 $\mathtt{m}_2 - \mathtt{m}_1^2 = \det \mathtt{m}(H_2) = \mathtt{m}((x_1 - \mathtt{m}_1)^2) \in \mathrm{qm}(\emptyset, \emptyset)$. This example stems from the fact that the quadratic module $\mathrm{qm}(\emptyset, \emptyset)$ in $\mathscr{M}$, which is generated by $\mathtt{m}(f^2)$ for $f \in \mathscr{M}[\underline{x}]$, cannot be generated by a finite set. Consequently, one cannot apply Tarski's tranfer principle (see e.g. [Mar08, Theorem 11.2.1]) to translate the validity of (ii)$\Rightarrow$(i) for homomorphisms into $\mathbb{R}$ to homomorphisms into general real closed fields.

Pseudo-moment evaluations form a strictly larger class than moment evaluations. For example, 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. This is shown in [BRS$^+$22, Example 5.1] (see Example 3.8 below for an alternative argument) where, more generally, *binomial* moment and pseudo-moment inequalities are characterized using tropicalization techniques. Theorem 3.7 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 6.9.

In preparation for Theorem 3.7, the next two lemmas address extensions and strictly positive approximations of positive functionals on $\mathbb{R}[\underline{x}]_d$, which are relevant to pseudo-moment evaluations in light of Remark 3.2.

Lemma 3.4. *If a linear 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 linear 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:

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

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

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

where $L(H_d), K, B$ are independent of α, and $L(H_d)$ and K are positive definite matrices. Since $L(H_d)$ is invertible, $L(H_{d+1})$ is positive definite if and only if the Schur complement $\alpha K - BL(H_d)^{-1}B^\top$ is positive definite. This is indeed the case for a sufficiently large $\alpha > 0$. For such an α, the functional L_α on $\mathbb{R}[\underline{x}]_{2d+2}$ is positive on $\mathtt{m}(p^2)$ for $p \in \mathbb{R}[\underline{x}]_{d+1} \setminus \{0\}$, and agrees with L on $\mathbb{R}[\underline{x}]_{2d}$. Continuing in this fashion by induction on d, we obtain $\widetilde{L} : \mathbb{R}[\underline{x}] \to \mathbb{R}$ that extends L and satisfies $\widetilde{L}(p^2) > 0$ for $p \in \mathbb{R}[\underline{x}] \setminus \{0\}$. $\qquad\square$

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 linear 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. $\qquad\square$

The proof of Theorem 3.7 requires some additional terminology from real algebra [Mar08]. A *preordering* P in a commutative unital ring A is a quadratic module closed under multiplication. Let Ω denote the preordering in $\mathscr{M}$ generated by $\mathrm{qm}(\emptyset, \emptyset)$, and let $\widehat{\Omega}$ denote the preordering in $\mathscr{M}[\underline{x}]$ generated by $\mathrm{QM}(\emptyset, \emptyset)$. More concretely,

$$\Omega = \left\{ \sum_i \mathtt{m}(f_{i1}^2) \cdots \mathtt{m}(f_{ik_i}^2) \colon f_{ij} \in \mathscr{M}[\underline{x}] \right\},$$

$$\widehat{\Omega} = \left\{ \sum_i f_{i0}^2 \mathtt{m}(f_{i1}^2) \cdots \mathtt{m}(f_{ik_i}^2) \colon f_{ij} \in \mathscr{M}[\underline{x}] \right\}.$$

Preorderings play a crucial role in the Krivine-Stengle Positivstellensatz [Mar08, Theorem 2.2.1], which we recall now. If $f, g_1, \ldots, g_\ell \in \mathbb{R}[y_1, \ldots, y_k]$ and f is nonnegative on $\{g_1 \geq 0, \ldots, g_\ell \geq 0\}$, then the Krivine-Stengle Positivstellensatz implies that f is a quotient of elements from the preordering in $\mathbb{R}[y_1, \ldots, y_k]$ generated by $g_1, \ldots, g_\ell$.

The following lemma relates the symbolic Hankel matrix H_d to the preordering Ω.

Lemma 3.6. *Every principal minor of $\mathtt{m}(H_d)$ is a quotient of elements in Ω.*

Proof. This is an adaptation of [KMVW23, Proposition 4.2]; let us sketch the main idea without technical details. By [KŠV18, Proposition 4.3], a symmetric matrix X of a fixed size D is positive semidefinite if and only if $\mathrm{tr}(XT_1(X)^2) \geq 0, \ldots, \mathrm{tr}(XT_\ell(X)^2) \geq 0$, where

T_j are polynomials in X and traces of its powers, dependent only on D. In particular, by the Krivine-Stengle Positivstellensatz, minors of X are quotients of elements from the preordering in $\mathbb{R}[\operatorname{tr}(X), \ldots, \operatorname{tr}(X^D)]$ generated by $\operatorname{tr}(XT_1(X)^2), \ldots, \operatorname{tr}(XT_\ell(X)^2)$. This fact is then applied to $X = \mathtt{m}(H_d)$, in which case one can furthermore verify that $\operatorname{tr}(\mathtt{m}(H_d)T_j(\mathtt{m}(H_d))^2) \in \operatorname{qm}(\emptyset, \emptyset)$ for $j = 1, \ldots, \ell$. $\qquad\qquad\square$

Theorem 3.7. *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 elements in $\widehat{\Omega}$.*

Proof. ($\Rightarrow$) Let $d = \deg f$. Let A be the polynomial ring generated by $x_1, \ldots, x_n$ and $\mathtt{m}(u)$ for $u \in [\underline{x}]_{2d}$; in particular, A is a finitely generated polynomial ring. Let P be the preordering in A generated by the principal minors of $\mathtt{m}(H_d)$. Then P is a finitely generated preordering, and contained in $\widehat{\Omega}$ by Lemma 3.6. Assume f is not a quotient of elements in $\widehat{\Omega}$. Then f is not a quotient of elements in P. By the Krivine-Stengle Positivstellensatz [Mar08, Theorem 2.2.1] there is a homomorphism $\phi : A \to \mathbb{R}$ such that $\phi(f) < 0$ and ϕ is nonnegative on the principal minors of $\mathtt{m}(H_d)$. In particular, the matrix $(\phi \circ \mathtt{m})(H_d)$ is positive semidefinite. Note that ϕ is determined by $\phi(x_j)$ for $j = 1, \ldots, n$ 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 $L(p^2) > 0$ 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 $f = \frac{g}{h}$ for some nonzero $g, h \in \widehat{\Omega}$. Suppose $\phi(f) < 0$ for a pseudo-moment evaluation $\phi : \mathscr{M}[\underline{x}] \to \mathbb{R}$. Note that $\psi(g), \psi(h) \geq 0$ for all pseudo-moment evaluations $\psi : \mathscr{M}[\underline{x}] \to \mathbb{R}$. Since $h \neq 0$, there exists a pseudo-moment evaluation $\phi' : \mathscr{M}[\underline{x}] \to \mathbb{R}$ such that $\phi'(h) > 0$ (indeed, one can even choose ϕ' arising from evaluation at a point in $\mathbb{R}^n$ and a measure in $\mathbf{P}(\mathbb{R}^n)$). For $\varepsilon \in [0, 1]$ let $\phi_\varepsilon : \mathscr{M}[\underline{x}] \to \mathbb{R}$ be the pseudo-moment evaluation determined by $\phi_\varepsilon(\mathtt{m}(p)) = (1 - \varepsilon)\phi(\mathtt{m}(p)) + \varepsilon\phi'(\mathtt{m}(p))$ for $p \in \mathbb{R}[\underline{x}]$. Then there exists a sufficiently small $\varepsilon > 0$ so that $\phi_\varepsilon(f) < 0$ and $\phi_\varepsilon(h) > 0$. Hence

$$0 > \phi_\varepsilon(h)\phi_\varepsilon(f) = \phi_\varepsilon(g) \geq 0,$$

a contradiction. $\qquad\qquad\square$

Example 3.8 ([BRS$^+$22, Example 5.1])**.** Let $f = \mathtt{m}_{4,2}\mathtt{m}_{2,4} - \mathtt{m}_{2,2}^3$. On the one hand, all moment evaluations of f are nonnegative by Hölder's inequality applied to the three polynomials $x_1^4 x_2^2, x_1^4 x_2^2, 1$ as

$$\int (x_1^4 x_2^2 \cdot x_1^4 x_2^2 \cdot 1) \, \mathrm{d}\mu \leq \left(\int x_1^4 x_2^2 \, \mathrm{d}\mu\right)^{\frac{1}{3}} \left(\int x_1^4 x_2^2 \, \mathrm{d}\mu\right)^{\frac{1}{3}} \left(\int 1 \, \mathrm{d}\mu\right)^{\frac{1}{3}}.$$

On the other hand, consider the functional $L : \mathbb{R}[\underline{x}]_6 \to \mathbb{R}$ given on the Hankel matrix H_3 as

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

Note that the right-hand side of (3.1) is positive definite, and $L(x_1^4 x_2^2)L(x_1^2 x_2^4) - L(x_1^2 x_2^2)^3 = 1 - 2^3 = -7$. 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 3.7.

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]. We start by observing how archimedianity in $\mathbb{R}[\underline{x}]$ transfers to archimedianity in $\mathscr{M}[\underline{x}]$.

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

Proof. Let $(i_1, \ldots, i_n) \in \mathbb{N}_0^n$ be arbitrary. Since $\mathrm{QM}(S)$ is archimedean, 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 $s \in S \cup \{1\}$. Hence, $N \pm \mathrm{m}_{i_1,\ldots,i_n} \in \mathrm{qm}(S, \emptyset)$. Consequently, $\mathrm{qm}(S, \emptyset)$ and $\mathrm{QM}(S, \emptyset)$ are archimedean by [Mar08, Proposition 5.2.3]. $\qquad\qquad\square$

The following theorem is the main result of this section.

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

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

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

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

Proof. We only prove the first equivalence (the proof of second one is analogous). The implication (ii)$\Rightarrow$(i) is straightforward. Now suppose (ii) is false. Since $\mathrm{QM}(S_1, S_2)$ is archimedean, the Kadison-Dubois representation theorem [Mar08, Theorem 5.4.4] states that $\mathrm{QM}(S_1, S_2)$ contains all $g \in \mathscr{M}[\underline{x}]$ which are positive under all homomorphisms

$\mathcal{M}[\underline{x}] \to \mathbb{R}$ that are nonnegative on $\mathrm{QM}(S_1, S_2)$. In particular, since $f + \varepsilon \notin \mathrm{QM}(S_1, S_2)$, there exists a homomorphism $\varphi : \mathcal{M}[\underline{x}] \to \mathbb{R}$ such that $\varphi(\mathrm{QM}(S_1, S_2)) = \mathbb{R}_{\geq 0}$ and $\varphi(f + \varepsilon) \leq 0$ (in particular, $\varphi(f) < 0$). 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 $\mathrm{QM}(S_1)$, 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 $L(p) = \int p \, d\mu$ for all $p \in \mathbb{R}[\underline{x}]$. By the construction, $\mu \in \mathbf{K}(S_1, S_2)$. Therefore, $(\mu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_1)$ and $f(\mu, \underline{X}) = \varphi(f) < 0$. $\qquad\square$

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, so one can consider positivity of state polynomials on operators subject to archimedean constraints and commutation relations, which means that evaluations are restricted to jointly commuting bounded operators. 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 equivalence (i)⇔(ii) in Theorem 4.2.

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

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, $\mathbf{K}(S_1, S_2) \times K(S_1)$ is compact, so there is $\varepsilon > 0$ such that $f - \varepsilon \geq 0$ on $\mathbf{K}(S_1, S_2) \times K(S_1)$. Then $f \in \mathrm{QM}(S_1, S_2)$ by Theorem 4.2. $\qquad\square$

Corollary 4.4 also admits the following interpretation in the absence of the archimedean assumption.

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

   (i) *$f \geq 0$ on $\mathbf{K}(S_1, S_2) \times K(S_1)$;*
   (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 $\mathrm{QM}(\widetilde{S}_1)$ is archimedean in $\mathbb{R}[\underline{x}]$ by [Mar08, Corollary 6.1.2]. The rest then follows from Theorem 4.2. $\qquad\square$

5. MOMENT POLYNOMIAL OPTIMIZATION AND EXAMPLES

Theorem 4.2 can be applied to design a converging hierarchy of semidefinite programs (SDPs) for the moment polynomial optimization problem (1.1) from the introduction. For the sake of simplicity, we first focus on pure moment polynomial objective functions (where SDPs are based on membership in $\mathrm{qm}(S_1, S_2)$), and then indicate the necessary changes for general moment polynomial objective functions (where SDPs are based on membership in $\mathrm{QM}(S_1, S_2)$). Finally, we demonstrate how this SDP hierarchy can be applied to problems in quantum information theory (Subsections 5.1 and 5.2).

Let $S_1 \subset \mathbb{R}[\underline{x}]$ and $S_2 \subset \mathscr{M}$ be finite, and $f \in \mathscr{M}$. We are interested in solving the optimization problem

$$(5.1) \qquad \text{minimize } f(\mu) \text{ subject to } \mu \in \mathbf{K}(S_1, S_2),$$

which is the specialization of (1.1) for a pure moment polynomial objective function. Let $r \in \mathbb{N}$. Recall that $\mathrm{qm}(S_1, S_2)_{2r}$ is the convex hull of

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

Membership in $\mathrm{qm}(S_1, S_2)_{2r}$ can be certified by an SDP; indeed, its members are of the form

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

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

$$(5.2) \qquad f_r = \sup\{\alpha \in \mathbb{R} \colon f - \alpha \in \mathrm{qm}(S_1, S_2)_{2r}\}.$$

Theorem 4.2 implies (under the archimedean assumption) that f_r form a sequence of lower bounds converging to the solution of (5.1), the infimum of f on $\mathbf{K}(S_1, S_2)$.

Corollary 5.1. *Let $S_1 \subseteq \mathbb{R}[\underline{x}]$, $S_2 \subseteq \mathscr{M}$, $f \in \mathscr{M}$, and suppose $\mathrm{QM}(S_1)$ is archimedean in $\mathbb{R}[\underline{x}]$. Then the sequence $\{f_r\}_{r \geq \frac{\deg f}{2}}$ arising from the SDP hierarchy (5.2) converges monotonically to $f_* := \inf_{\mu \in \mathbf{K}(S_1, S_2)} f(\mu)$ from below.*

Proof. The sequence $\{f_r\}_r$ is increasing since $\mathrm{qm}(S_1, S_2)_{2r} \subseteq \mathrm{qm}(S_1, S_2)_{2(r+1)}$. Also, $f_r \leq f_*$ because $f - \alpha \in \mathrm{qm}(S_1, S_2)_{2r}$ implies $f \geq \alpha$ on $\mathbf{K}(S_1, S_2)$. Now let $\varepsilon > 0$ be arbitrary. Since $f - f_* \geq 0$ on $\mathbf{K}(S_1, S_2)$, Theorem 4.2 implies $f - f_* + \varepsilon \in \mathrm{qm}(S_1, S_2)$. Thus, there exists $r \in \mathbb{N}$ such that

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

so $f_* - \varepsilon \leq f_r$. Therefore, $\lim_{r \to \infty} f_r = f_*$. $\qquad\square$

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 (5.2) and its dual,

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

(2) If the solution of the dual of (5.2) satisfies certain rank conditions, then the SDP hierarchy (5.2) stops, and one can extract a concrete finitely supported optimizer for f_*.

(3) While the sizes of SDPs (5.2) 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 as in (1.1),

$$(5.3) \qquad \text{minimize } f(\mu, \underline{X}) \text{ subject to } \underline{X} \in K(S_1) \text{ and } \mu \in \mathbf{K}(S_1, S_2),$$

one needs to first address the following obstacle. Recall that $QM(S_1, S_2)_{2r}$ is the convex hull of

$$f_1^2 \mathtt{m}(g^2 s), \ f_2^2 t: \quad s \in \{1\} \cup S_1, \ t \in S_1 \cup S_2, \ f_1, f_2, g \in \mathscr{M}[\underline{x}],$$
$$\deg s_1 + 2(\deg f_1 + \deg g), \deg s_2 + 2 \deg f_2 \leq 2r.$$

While membership in $QM(S_1, S_2)_{2r}$ is a feasibility linear conic program [Bar02, Section IV.6], it does not seem to be a semidefinite program. For this reason we consider a slightly larger cone. Sums of elements of the form $f^2 s$, where $f \in \mathscr{M}[\underline{x}]$ and $s \in S_1 \cup S_2$ satisfy $\deg s + 2 \deg f \leq 2r$, are precisely

$$(5.4) \qquad\qquad \sum_{s \in S_1 \cup S_2} \sum_{v_1, v_2} G^{(s)}_{v_1, v_2} \cdot v_1 v_2 s$$

where v_i are monomials in $\mathscr{M}[\underline{x}]_{r - \frac{\deg s}{2}}$, and $G^{(s)}$ are positive semidefinite matrices of dimensions $\dim \mathscr{M}[\underline{x}]_{r - \frac{\deg s}{2}}$. Furthermore, sums of elements of the form $f^2 \mathtt{m}(g^2 s)$, where $f, g \in \mathscr{M}[\underline{x}]$ and $s \in \{1\} \cup S_1$ satisfy $\deg s + 2(\deg f + \deg g) \leq 2r$, can be written as

$$(5.5) \qquad\qquad \sum_{s \in \{1\} \cup S_1} \sum_{(u_1, v_1), (u_2, v_2)} G^{(s)}_{(u_1, v_1), (u_2, v_2)} \cdot \mathtt{m}(u_1 u_2 s) v_1 v_2,$$

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 $\deg s + 2(\deg u_i + \deg v_i) \leq 2r$, and $G^{(s)}$ are positive semidefinite matrices of size $\sum_{i+j \leq r - \frac{\deg s}{2}}(\dim \mathbb{R}[\underline{x}]_i + \dim \mathscr{M}[\underline{x}]_j)$. Let $\widetilde{QM}(S_1, S_2)_{2r}$ denote the set of sums of elements of the form (5.4) and (5.5). Then the cone $\widetilde{QM}(S_1, S_2)_{2r}$ contains $QM(S_1, S_2)_{2r}$, but is typically larger than $QM(S_1, S_2)_{2r}$. For example, if $n \geq 4$ then

$$\mathtt{m}(x_1^2) x_2^2 + 2 \mathtt{m}(x_1 x_3) x_2 x_4 + \mathtt{m}(x_3^2) x_4^2$$

belongs to $\widetilde{QM}(S_1, S_2)_{2r}$ (namely, it is of the form (5.5)) but not to $QM(S_1, S_2)_{2r}$. Nevertheless, elements of $\widetilde{QM}(S_1, S_2)_{2r}$ are nonnegative on $\mathbf{K}(S_1, S_2) \times K(S_1)$, and membership in $\widetilde{QM}(S_1, S_2)_{2r}$ can be determined by a feasibility SDP. Given $f \in \mathscr{M}[\underline{x}]$, the optimization problems

$$(5.6) \qquad\qquad f_r = \sup\{\alpha : f - \alpha \in \widetilde{QM}(S_1, S_2)_{2r}\}$$

for $r \geq \frac{\deg f}{2}$ are then SDPs, and the following analog of Corollary 5.1 holds for the optimization problem (5.3).

Corollary 5.2. *Let $S_1 \subseteq \mathbb{R}[\underline{x}]$, $S_2 \subseteq \mathscr{M}$, $f \in \mathscr{M}[\underline{x}]$, and suppose $QM(S_1)$ is archimedean in $\mathbb{R}[\underline{x}]$. Then the sequence $\{f_r\}_{r \geq \frac{\deg f}{2}}$ arising from the SDP hierarchy (5.6) converges monotonically to $f_* := \inf_{(\mu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_1)} f(\mu, \underline{X})$ from below.*

Next, we demonstrate the above SDP hierarchy method on two open optimization problems arising from nonlinear Bell inequalities in quantum physics [CG22]. A Bell scenario refers to a repeated experiment, where two or more parties, not communicating between themselves, perform measurements on a source of classical randomness and record outcomes in each round. Each party has a fixed number of measurements at disposal (with a fixed set of possible outputs), and uses one of them in each round. These measurements are viewed as random variables on a probability space that governs the source measured by the parties. The conditional expectations of joint outcomes relative to measurements in such an experiment are called correlations. A Bell inequality for a given Bell scenario is an inequality in correlations that is valid regardless of the underlying probability

space and measurements that model the experiment. Bell inequalities provide universal constraints on classical probabilistic models of a Bell scenario, and are widely used in quantum physics to certify quantum nonlocality (which violates Bell inequalities). Bell scenarios that model more sophisticated quantum networks call for nonlinear polynomial Bell inequalities, which are (from the perspective of this paper) special cases of moment polynomial inequalities. Section 5.1 confirms a covariance Bell inequality proposed in [PHBB17], and Section 5.2 rectifies a bilocal Bell inequality proposed in [TGB21].

Encoding of moment polynomials and preparation of the SDPs was done in Mathematica, and numerical solutions of the SDPs were obtained with the SDP solver MOSEK [ApS24]. The final exact certificates provided below are heavily inspired by these numerical solutions. The ad-hoc Mathematica code for constructing and solving SDPs is available on GitHub: `https://github.com/magronv/mompop` (it terminates within a few minutes on a standard desktop PC), as well as the notebook for verifying the obtained exact certificates.

5.1. **Covariance Bell inequality.** For $j = 1, 2, 3$ let A_j, B_j be binary random variables (valued in $\{-1, 1\}$) on a probability space $(\mathcal{X}, \Sigma, \pi)$, and consider the expression

$$\mathrm{cov}_{3322}(\underline{A}, \underline{B}) := \mathrm{cov}(A_1, B_1) + \mathrm{cov}(A_1, B_2) + \mathrm{cov}(A_1, B_3)$$
$$+ \mathrm{cov}(A_2, B_1) + \mathrm{cov}(A_2, B_2) - \mathrm{cov}(A_2, B_3)$$
$$+ \mathrm{cov}(A_3, B_1) - \mathrm{cov}(A_3, B_2)$$

where $\mathrm{cov}(X, Y) = \int XY \, \mathrm{d}\pi - \int X \, \mathrm{d}\pi \cdot \int Y \, \mathrm{d}\pi$. In [PHBB17], the authors ask for the largest possible value of cov_{3322}. They provide concrete examples of probability spaces (on a three-element set) and binary random variables where cov_{3322} attains the value $\frac{9}{2}$. In the quest for proving that $\mathrm{cov}_{3322} \leq \frac{9}{2}$ 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, rendering this particular approach infeasible. As an alternative, they suggest maximizing cov_{3322} via classical polynomial optimization similarly as in Remark 2.7. However, the corresponding polynomial problem has too many variables for global optimization tools to work. Thus, they use numerical nonlinear optimization to look for local maxima of cov_{3322} from numerous starting points, which lends confidence to their conjecture that $\mathrm{cov}_{3322} \leq 4.5$. Below, we settle this conjecture using the methods developed in this paper. Let

$$f = \mathfrak{m}_{100100} - \mathfrak{m}_{100000}\mathfrak{m}_{000100} + \mathfrak{m}_{100010} - \mathfrak{m}_{100000}\mathfrak{m}_{000010} + \mathfrak{m}_{100001} - \mathfrak{m}_{100000}\mathfrak{m}_{000001}$$
$$+ \mathfrak{m}_{010100} - \mathfrak{m}_{010000}\mathfrak{m}_{000100} + \mathfrak{m}_{010010} - \mathfrak{m}_{010000}\mathfrak{m}_{000010} - \mathfrak{m}_{010001} + \mathfrak{m}_{010000}\mathfrak{m}_{000001}$$
$$+ \mathfrak{m}_{001100} - \mathfrak{m}_{001000}\mathfrak{m}_{000100} - \mathfrak{m}_{001010} + \mathfrak{m}_{001000}\mathfrak{m}_{000010}.$$

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

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

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

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

with $S = \{\pm(1 - x_j^2) \colon j = 1, \ldots, 6\}$. 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, $f_* = 4.5$. After carefully inspecting the output of the SDP

solver (namely, heuristically extracting exact constraints from the kernel of the numerical positive semidefinite constraint) one can obtain an exact certificate for $f_* \leq \frac{9}{2}$. Concretely, modulo the relations $x_j^2 = 1$ we have $\frac{9}{2} - f = \mathtt{m}(v^\top G v)$, where

$$G = \begin{pmatrix} \frac{4}{3} & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{32} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{3}{8} & \frac{1}{16} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{16} & \frac{1}{8} & -\frac{1}{8} & 0 & \frac{3}{64} & \frac{3}{64} \\ 0 & 0 & 0 & -\frac{1}{8} & \frac{1}{4} & -\frac{1}{8} & -\frac{1}{16} & -\frac{1}{16} \\ 0 & 0 & 0 & 0 & -\frac{1}{8} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{64} & -\frac{1}{16} & 0 & \frac{3}{64} & 0 \\ 0 & 0 & 0 & \frac{3}{64} & -\frac{1}{16} & 0 & 0 & \frac{3}{64} \end{pmatrix} \in \mathbb{R}^{8 \times 8}$$

is positive definite, and

$$v = \begin{pmatrix} x_6 \mathtt{m}_{110000} \\ (x_1 - x_2)(x_4 + x_5) \\ 1 + x_6(x_2 - x_1 + 4\mathtt{m}_{100000}) \\ 2 - (x_1 + x_2)(x_4 + x_5) - 8x_6 \mathtt{m}_{100000} \\ (x_3 + 2\mathtt{m}_{001000})(x_4 - x_5) \\ 2x_4 \mathtt{m}_{001000} - x_3 x_5 \\ 2 + (x_4 - x_5)(4\mathtt{m}_{001000} - x_1 - x_2) + \frac{8}{3}x_4 \mathtt{m}_{110000} + 8x_6 \mathtt{m}_{100000} \\ 2 + (x_4 - x_5)(4\mathtt{m}_{001000} + x_1 + x_2) + \frac{8}{3}x_5 \mathtt{m}_{110000} + 8x_6 \mathtt{m}_{100000} \end{pmatrix} \in \mathscr{M}[\underline{x}]^8.$$

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

$$(5.7) \qquad \frac{1}{3} \sum_{i \in \{1,2,3\}} \Big(\mathrm{E}(B_i C_i) - \mathrm{E}(A_i B_i) \Big) - \sum_{\{i,j,k\}=\{1,2,3\}} \mathrm{E}(A_i B_j C_k)$$

where A_j, B_j, C_j for $j = 1, 2, 3$ are binary random variables on a probability space $(\mathcal{X}, \Sigma, \pi)$ satisfying bilocality constraints

$$(5.8) \qquad \mathrm{E}(A_1^{k_1} A_2^{k_2} A_3^{k_3} C_1^{k_4} C_2^{k_5} C_3^{k_6}) = \mathrm{E}(A_1^{k_1} A_2^{k_2} A_3^{k_3}) \mathrm{E}(C_1^{k_4} C_2^{k_5} C_3^{k_6})$$

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

$$
\begin{aligned}
& \mathrm{E}(A_i) = \mathrm{E}(B_i) = \mathrm{E}(C_i) = 0 \quad \text{for } i \in \{1, 2, 3\}, \\
(5.9) \qquad & \mathrm{E}(A_i B_j) = \mathrm{E}(B_i C_j) = 0 \quad \text{for } i \neq j, \\
& \mathrm{E}(A_i B_j C_k) = 0 \quad \text{for } |\{i, j, k\}| \leq 2.
\end{aligned}
$$

Here, $\mathrm{E}(X) = \int X \, d\pi$. In [TGB21] it is shown that the largest value of (5.7) for bilocal models with the tetrahedral symmetry is 3. Furthermore, [TGB21, TPKLR22] suggest that (5.7) 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 (5.7) subject to (5.8) and (5.9) is 4.

Consider the moment polynomial optimization problem

$$\sup \frac{1}{3} \sum_{i \in \{1,2,3\}} \Big(\mathrm{m}(x_{i+3}x_{i+6}) - \mathrm{m}(x_i x_{i+3}) \Big) - \sum_{\{i,j,k\}=\{1,2,3\}} \mathrm{m}(x_i x_{j+3} x_{k+6})$$

subject to

(5.10)
$$\mathrm{m}(x_1^{k_1} x_2^{k_2} x_3^{k_3} x_7^{k_4} x_8^{k_5} x_9^{k_6}) = \mathrm{m}(x_1^{k_1} x_2^{k_2} x_3^{k_3}) \mathrm{m}(x_7^{k_4} x_8^{k_5} x_9^{k_6}) \quad \text{for } k_i \in \{0,1\},$$
$$x_j^2 = 1 \text{ and } \mathrm{m}(x_j) = 0 \quad \text{for } j \in \{0,\dots,9\},$$
$$\mathrm{m}(x_i x_{j+3}) = \mathrm{m}(x_{i+3} x_{j+6}) = 0 \quad \text{for } i,j \in \{1,2,3\},\ i \neq j,$$
$$\mathrm{m}(x_i x_{j+3} x_{k+6}) = 0 \quad \text{for } i,j,k \in \{1,2,3\},\ |\{i,j,k\}| \leq 2.$$

Corollary 5.1 provides a converging sequence of upper bounds for the solution of (5.10). For $r = 3$, one obtains the upper bound 4.0 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. The value of (5.7) subject to (5.8) and (5.9) is therefore at most 4.0. Next, we show that the value 4 is attained; incidentally, the below construction was inspired by the numerical output of the dual SDP. Denote

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

and let $e_i \in \mathbb{R}^4$ be the ith 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 (5.8) 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 (5.9) follow by direct calculation. Finally, (5.7) evaluates to 4 for this ensemble of binary random variables.

Lastly, let us provide an analytic proof that the value of (5.7) subject to (5.8) and (5.9) is at most 4. The sub-cases in the proof were identified by looking at the numerical output of the aforementioned dual SDP for $r = 3$.

Proposition 5.3. *If binary random variables* A_i, B_i, C_i *for* $i = 1,2,3$ *satisfy constraints* (5.8) *and* (5.9), *then*

$$\frac{1}{3} \sum_{i \in \{1,2,3\}} \Big(\mathrm{E}(B_i C_i) - \mathrm{E}(A_i B_i) \Big) - \sum_{\{i,j,k\}=\{1,2,3\}} \mathrm{E}(A_i B_j C_k) \leq 4.$$

Proof. It suffices to see that

(5.11)
$$\mathrm{E}(B_i C_i) - \mathrm{E}(A_i B_i) \leq 1$$

for $i = 1,2,3$, and

(5.12)
$$-\mathrm{E}(A_i B_j C_k) - \mathrm{E}(A_k B_j C_i) \leq 1$$

for $j = 1,2,3$ and $\{i,k\} = \{1,2,3\} \setminus \{j\}$. First, (5.11) follows from

$$1 - \mathrm{E}(B_i C_i) + \mathrm{E}(A_i B_i) = \mathrm{E}\Big((1 - B_i C_i)(1 + A_i B_i) \Big) \geq 0;$$

here, the equality holds by $B_i^2 = 1$ and $\mathrm{E}(A_iC_i) = \mathrm{E}(A_i)\mathrm{E}(C_i) = 0$, and the inequality holds since $1 - B_iC_i$ and $1 + A_iB_i$ are nonnegative random variables. Second, (5.12) follows from the equality

$$1 + \mathrm{E}(A_iB_jC_k) + \mathrm{E}(A_kB_jC_i) = \mathrm{E}\left(\left(\tfrac{1}{2}\big((A_iA_k + C_iC_k)B_j + A_iC_i + A_kC_k\big) - \mathrm{E}(A_iA_k)B_j\right)^2\right),$$

which is verified by expanding the right-hand side, and applying the relations (5.8) and (5.9). $\qquad\square$

6. Lasserre's perturbative Positivstellensatz for moment polynomials

This section addresses moment polynomial positivity on general moment semialgebraic sets (without the archimedean assumption from Section 4). We show that moment polynomials nonnegative on $\mathbf{K}(S_1, S_2) \times K(S_1)$ belong to the quadratic module $\mathrm{QM}(S_1, S_2)$ *up to an arbitrarily small perturbation* of their coefficients (Theorem 6.9). This is achieved through the analysis of a sequence of conic optimization problems and their duals, and a sufficient condition for partially estimating a positive functional on $\mathscr{M}[\underline{x}]$ with a moment evaluation. Finally, a corollary for polynomial positivity on semialgebraic sets is given (Corollary 6.14).

The aforementioned perturbations arise from polynomials

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

and pure moment polynomials

$$\Psi_r = \sum_{j=1}^{n} \sum_{\substack{k,\ell \in \mathbb{N}, \\ k\ell \leq r}} \frac{\mathtt{m}(x_j^{2k})^{\ell}}{(k!)^{\ell}\ell!}$$

for $r \in \mathbb{N}$. We start with a crude estimate pertaining to the sequence $(\Psi_r)_r$.

Lemma 6.1. *For all $y \geq 0$,*

$$\sum_{k,\ell \in \mathbb{N}} \frac{y^{k\ell}}{(k!)^{\ell}\ell!} < \infty.$$

In particular, $\sum_{k,\ell \in \mathbb{N}} \frac{1}{(k!)^{\ell}\ell!} \leq e^2$.

Proof. Since $\exp(t) - 1 \leq t\exp(t)$ for $t \geq 0$,

$$(6.1) \qquad \sum_{k,\ell \in \mathbb{N}} \frac{y^{k\ell}}{(k!)^{\ell}\ell!} = \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{1}{\ell!}\left(\frac{y^k}{k!}\right)^{\ell} = \sum_{k=1}^{\infty}\left(\exp\left(\frac{y^k}{k!}\right) - 1\right) \leq \sum_{k=1}^{\infty} \frac{y^k}{k!}\exp\left(\frac{y^k}{k!}\right).$$

Since $\lim_{k\to\infty} \sqrt[k]{k!} = \infty$, there exists $k_0 \in \mathbb{N}$ such that $\frac{y^k}{k!} \leq 1$ for all $k \geq k_0$ (in particular, if $y = 1$ one can take $k_0 = 1$). Then we estimate (6.1) as

$$\sum_{k=1}^{\infty} \frac{y^k}{k!}\exp\left(\frac{y^k}{k!}\right) = \sum_{k=1}^{k_0-1} \frac{y^k}{k!}\exp\left(\frac{y^k}{k!}\right) + \sum_{k=k_0}^{\infty} \frac{y^k}{k!}\exp\left(\frac{y^k}{k!}\right)$$

$$\leq \sum_{k=1}^{k_0-1} \frac{y^k}{k!}\exp\left(\frac{y^k}{k!}\right) + \sum_{k=k_0}^{\infty} \frac{y^k}{k!} \cdot e$$

$$\leq \sum_{k=1}^{k_0-1} \frac{y^k}{k!}\exp\left(\frac{y^k}{k!}\right) + e \cdot \exp(y) < \infty.$$

If $y = 1$, the first sum above is empty, so (6.1) is bounded by e^2. □

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

$$0 \leq \Psi_r(\nu) = \sum_{j=1}^n \sum_{\substack{k,\ell \in \mathbb{N}, \\ k\ell \leq r}} \frac{\left(\sum_{i=1}^m \alpha_i X_{ij}^{2k}\right)^\ell}{(k!)^\ell \ell!} \leq \sum_{j=1}^n \sum_{\substack{k,\ell \in \mathbb{N}, \\ k\ell \leq r}} \frac{y_j^{k\ell}}{(k!)^\ell \ell!} \leq \sum_{j=1}^n \sum_{k,\ell \in \mathbb{N}} \frac{y_j^{k\ell}}{(k!)^\ell \ell!} < \infty$$

by Lemma 6.1, so $\Psi_r(\nu)$ is uniformly bounded for all $r \in \mathbb{N}$.

The next two subsections prepare the technical groundwork for the proof of the main result of this section, Theorem 6.9.

6.1. **Two linear conic problems.** Let $S_1 \subseteq \mathbb{R}[\underline{x}]$, $S_2 \subseteq \mathcal{M}$ and $f \in \mathcal{M}[\underline{x}]$. In the proof of Theorem 6.9, the following pair of optimization problems for every $r \geq \frac{\deg f}{2}$ and $M > 0$ plays an important role:

$$(6.2) \qquad Q_{r,M} : \begin{cases} \sup_{z \in \mathbb{R}} & z \\ s.t. & f - z \in \mathrm{QM}(S_1, S_2)_{2r} + \mathbb{R}_{\geq 0} \cdot \left(M - \Phi_r - \Psi_r\right); \end{cases}$$

$$(6.3) \qquad Q_{r,M}^\vee : \begin{cases} \inf_{L \in \mathcal{M}[\underline{x}]_{2r}^\vee} & L(f) \\ s.t. & L(1) = 1, \\ & L\left(M - \Phi_r - \Psi_r\right) \geq 0, \\ & L(g) \geq 0 \quad \text{for all } g \in \mathrm{QM}(S_1, S_2)_{2r}. \end{cases}$$

Here, $\mathbb{R}_{\geq 0} \cdot g$ for $g \in \mathcal{M}[\underline{x}]$ denotes the set of all nonnegative multiples of g. In the following two lemmas and their proofs we abbreviate $\mathcal{C}_{r,M} = \mathrm{QM}(S_1, S_2)_{2r} + \mathbb{R}_{\geq 0} \cdot (M - \Phi_r - \Psi_r)$. The next lemma provides information on the closure of $\mathcal{C}_{r,M}$.

Lemma 6.2. *For all $r \in \mathbb{N}$ and $M > 0$, the closure of the cone $\mathcal{M}[\underline{x}]_r \cap \mathcal{C}_{2r,M}$ in $\mathcal{M}[\underline{x}]_r$ is contained in*

$$(6.4) \qquad \{g \in \mathcal{M}[\underline{x}]_r : g + \varepsilon \in \mathcal{C}_{2r,M} \text{ for all } \varepsilon > 0\}.$$

Proof. Let us fix $r \in \mathbb{N}$, $M > 0$ and let us endow the finite-dimensional space $\mathcal{M}[\underline{x}]_r$ with some norm $\|\cdot\|$. First we claim that for every monomial $u \in \mathcal{M}[\underline{x}]_r$ there exists $A_u > 0$ such that

$$\pm u + A_u \in \mathrm{QM}(\emptyset, \emptyset)_{2r} + \mathbb{R}_{\geq 0} \cdot \left(M - \Phi_{2r} - \Psi_{2r}\right).$$

Due to the definition of Φ_{2r} and Ψ_{2r}, it suffices to see that

$$(6.5) \qquad \pm u \in \mathrm{QM}(\emptyset, \emptyset)_{2r} - \sum_{k=1}^{2^r} \mathbb{R}_{\geq 0} \cdot x_j^{2k} - \sum_{k=1}^{2^r} \mathbb{R}_{\geq 0} \cdot \mathtt{m}\left(x_j^{2k}\right).$$

We see this by recursively using the following identities on the factors of u,

$$\pm a^k b = \frac{1}{2}(a^k \pm b)^2 - \frac{1}{2}a^{2k} - \frac{1}{2}b^2,$$

$$(6.6) \qquad \pm \mathtt{m}(a^k b) = \frac{1}{2}\mathtt{m}\left((a^k \pm b)^2\right) - \frac{1}{2}\mathtt{m}\left(a^{2k}\right) - \frac{1}{2}\mathtt{m}\left(b^2\right),$$

$$-\mathtt{m}_{i_1,\ldots,i_n}^{2k} = \left(\mathtt{m}_{2ki_1,\ldots,2ki_n} - \mathtt{m}_{i_1,\ldots,i_n}^{2k}\right) - \mathtt{m}_{2ki_1,\ldots,2ki_n}$$

for all $a, b \in \mathscr{M}[\underline{x}]$; note that $\mathrm{m}_{2ki_1,\ldots,2ki_n} - \mathrm{m}_{i_1,\ldots,i_n}^{2k} \in \mathrm{qm}(\emptyset, \emptyset)$ by Lemma 2.2(2). For example, using the first, the third and the second identity in (6.6) yields

$$
\begin{aligned}
\pm x_1 \mathrm{m}(x_1 x_2) &= \frac{1}{2} \left(x_1 \pm \mathrm{m}(x_1 x_2) \right)^2 - \frac{1}{2} x_1^2 - \frac{1}{2} \mathrm{m}(x_1 x_2)^2 \\
&= \frac{1}{2} \left(x_1 \pm \mathrm{m}(x_1 x_2) \right)^2 + \frac{1}{2} \left(\mathrm{m}(x_1^2 x_2^2) - \mathrm{m}(x_1 x_2)^2 \right) - \frac{1}{2} x_1^2 - \frac{1}{2} \mathrm{m}(x_1^2 x_2^2) \\
&= \frac{1}{2} \left(x_1 \pm \mathrm{m}(x_1 x_2) \right)^2 + \frac{1}{2} \left(\mathrm{m}(x_1^2 x_2^2) - \mathrm{m}(x_1 x_2)^2 \right) + \frac{1}{4} \mathrm{m} \left(\left(x_1^2 - x_2^2 \right)^2 \right) \\
&\quad - \frac{1}{2} x_1^2 - \frac{1}{4} \mathrm{m}(x_1^4) - \frac{1}{4} \mathrm{m}(x_2^4),
\end{aligned}
$$

confirming (6.5) for $u = x_1 \mathrm{m}(x_1 x_2)$. The (crude) degree bound 2^r in (6.5) is valid because at most $\deg u \leq r$ applications of (6.6) are required, and every application of (6.6) doubles the degree on its right-hand side.

Since $\mathscr{M}[\underline{x}]_r$ is finite-dimensional, by the preceding paragraph there exists $A > 0$ (dependent on r and M) such that

$$(6.7) \qquad\qquad g + A\|g\| \in \mathrm{QM}(\emptyset, \emptyset)_{2^r} + \mathbb{R}_{\geq 0} \cdot \left(M - \Phi_{2^r} - \Psi_{2^r} \right)$$

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

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

This function satisfies the following properties:

(a) $F(g) \geq -A\|g\|$ for all $g \in \mathscr{M}[\underline{x}]_r$;
(b) $F(g_1 + g_2) \geq F(g_1) + F(g_2)$ for all $g_i \in \mathscr{M}[\underline{x}]_r$;
(c) $F(g) \geq 0$ if and only if g belongs to (6.4).

Here, (a) follows from (6.7) because $\mathrm{QM}(\emptyset, \emptyset)_{2^r} + \mathbb{R}_{\geq 0} \cdot \left(M - \Phi_{2^r} - \Psi_{2^r} \right)$ is contained in $\mathcal{C}_{2^r, M}$; (b) holds since $g_1 - z_1, g_2 - z_2 \in \mathcal{C}_{2^r, M}$ implies $g_1 + g_2 - (z_1 + z_2) \in \mathcal{C}_{2^r, M}$; and (c) follows directly by definitions of F and (6.4). 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

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

for all i, and so $F(g) \geq 0$. Therefore, g belongs to (6.4). $\qquad\square$

Remark 6.3. 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 $x_1 \notin \mathcal{C}_{1,1}$.

One particular aspect of the proof of Lemma 6.2 is reused later, so we record it in detail next.

Lemma 6.4. *For all* $i_1, \ldots, i_n \in \mathbb{N}_0$,

$$\sum_{j=1}^{n-1} \frac{1}{2^j} \mathrm{m}(x_j^{2^{j+1} i_j}) + \frac{1}{2^{n-1}} \mathrm{m}(x_n^{2^n i_n}) - \mathrm{m}_{i_1,\ldots,i_n}^2 \in \mathrm{qm}(\emptyset, \emptyset).$$

Proof. By Lemma 2.2(2) it suffices to see that

$$(6.8) \qquad \sum_{j=1}^{n-1} \frac{1}{2^j} \mathrm{m}(x_j^{2^{j+1} i_j}) + \frac{1}{2^{n-1}} \mathrm{m}(x_n^{2^n i_n}) - \mathrm{m}_{2i_1,\ldots,2i_n} \in \mathrm{qm}(\emptyset, \emptyset),$$

which we prove by induction on n. If $n = 1$, the left-hand side of (6.8) is 0. Assume (6.8) holds for $n - 1$; applying it with variables $x_2, \ldots, x_n$ and exponents $2i_2, \ldots, 2i_n$ yields

$$(6.9) \qquad \sum_{j=1}^{n-2} \frac{1}{2^j} \mathtt{m}(x_{j+1}^{2^{j+1} \cdot 2i_{j+1}}) + \frac{1}{2^{n-2}} \mathtt{m}(x_n^{2^{n-1} \cdot 2i_n}) - \mathtt{m}_{0,4i_2,\ldots,4i_n} \in \mathrm{qm}(\emptyset, \emptyset).$$

Now consider the left-hand side of (6.8) for n, and isolate x_1 in its first and third term:

$$(6.10) \qquad \begin{aligned} \sum_{j=1}^{n-1} \frac{1}{2^j} \mathtt{m}(x_j^{2^{j+1}i_j}) &= \frac{1}{2}\mathtt{m}(x_1^{4i_1}) + \frac{1}{2}\sum_{j=1}^{n-2} \frac{1}{2^j}\mathtt{m}(x_{j+1}^{2^{j+1}\cdot 2i_{j+1}}), \\ -\mathtt{m}_{2i_1,\ldots,2i_n} &= \frac{1}{2}\mathtt{m}\left(\left(x_1^{2i_1} - x_2^{2i_2} \cdots x_n^{2i_n}\right)^2\right) - \frac{1}{2}\mathtt{m}(x_1^{4i_1}) - \frac{1}{2}\mathtt{m}_{0,4i_2,\ldots,4i_n}. \end{aligned}$$

Adding the right-hand sides in (6.10) implies

$$\sum_{j=1}^{n-1} \frac{1}{2^j}\mathtt{m}(x_j^{2^{j+1}i_j}) + \frac{1}{2^{n-1}}\mathtt{m}(x_n^{2^n i_n}) - \mathtt{m}_{2i_1,\ldots,2i_n}$$

$$= \frac{1}{2}\left(\sum_{j=1}^{n-2} \frac{1}{2^j}\mathtt{m}(x_{j+1}^{2^{j+1}\cdot 2i_{j+1}}) + \frac{1}{2^{n-2}}\mathtt{m}(x_n^{2^{n-1}\cdot 2i_n}) - \mathtt{m}_{0,4i_2,\ldots,4i_n}\right) + \frac{1}{2}\mathtt{m}\left(\left(x_1^{2i_1} - x_2^{2i_2} \cdots x_n^{2i_n}\right)^2\right),$$

which belongs to $\mathrm{qm}(\emptyset, \emptyset)$ by the induction hypothesis (6.9). $\qquad\square$

We can now show that the solutions of (6.2) and (6.3) converge to the same value.

Lemma 6.5. *For all r and M, the optimization problem (6.2) is a linear conic problem, and (6.3) is its dual. Sequences $(\sup Q_{r,M})_r$ and $(\inf Q_{r,M}^\vee)_r$ are increasing. If S_2 is finite, $\mathbf{K}(S_1, S_2) \neq \emptyset$ and M is large enough, then*

$$(6.11) \qquad \limsup_{r\to\infty} Q_{r,M} = \liminf_{r\to\infty} Q_{r,M}^\vee.$$

Proof. The optimization problems (6.2) and (6.3) have linear objective functions and constraints that consist of linear equations and cone memberships, so (6.2) and (6.3) are linear conic problems. Furthermore, they are dual to each other, by the definition of duality [Bar02, Section IV.6]. 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 $Q_{r,M}^\vee$, then its restriction is feasible for $Q_{r',M}^\vee$ if $r' < r$; hence the sequence $(\inf Q_{r,M}^\vee)_{r\geq d}$ is increasing (the set of constraints on L increases with r, and consequently so does the infimum). By weak duality [Bar02, Theorem IV.6.2] we have $\sup Q_{r,M} \leq \inf Q_{r,M}^\vee$. The rest of the proof is dedicated to showing that $\inf Q_{r,M}^\vee \leq \sup Q_{2r,M}$; then (6.11) follows immediately from $\sup Q_{r,M} \leq \inf Q_{r,M}^\vee \leq \sup Q_{2r,M}$. Since $\mathbf{K}(S_1, S_2) \neq \emptyset$, there exists $(\nu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_1)$. By Proposition 2.6, we can further assume that ν is supported on a finite set. Let $M \geq \sup_r(\Phi_r(\underline{X}) + \Psi_r(\nu))$ (note that the right hand side is finite). Then $L \in \mathscr{M}[\underline{x}]_{2r}^\vee$ defined by $L(p) = p(\nu, \underline{X})$ is clearly feasible for (6.3), whence $\inf Q_{r,M}^\vee < \infty$ for all $r \geq \frac{\deg f}{2}$. Note that $L(f - \inf Q_{r,M}^\vee) \geq 0$ for all $L \in \mathcal{C}_{r,M}^\vee$. This implies that $f - \inf Q_{r,M}^\vee$ is in $\mathcal{C}_{r,M}^{\vee\vee}$, which is the closure of $\mathcal{C}_{r,M}$. Therefore, $f - \inf Q_{r,M}^\vee + \varepsilon \in \mathcal{C}_{2r,M}$ for all $\varepsilon > 0$ by Lemma 6.2, so $\inf Q_{r,M}^\vee \leq \sup Q_{2r,M}$. $\qquad\square$

6.2. **Growth of positive functionals on moment polynomials.** Next, we turn to a partial estimate of positive functionals on moment polynomials with moment evaluations (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. This classical result states that if the marginals of a positive linear

functional L on $\mathbb{R}[\underline{x}]$ do not grow too fast (Carleman's condition, which is in particular satisfied if there is $M > 0$ such that $L(x_j^{2k}) \leq M^k k!$ for all k), then L equals integration with respect to some measure on $\mathbb{R}^n$.

Proposition 6.6. *Let* $S_1 \subseteq \mathbb{R}[\underline{x}]$, $S_2 \subseteq \mathscr{M}$, $M > 0$, *and* $f \in \mathscr{M}[\underline{x}]$. *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) $L(x_j^{2k}) \leq k!M$ *and* $L(\mathtt{m}(x_j^{2k})^\ell) \leq \ell!(k!)^\ell M$ *for all* $j \in \{1, \dots, n\}$ *and* $k, \ell \in \mathbb{N}$.

Then there exists $(\mu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_1)$ *such that* $L(f) \geq f(\mu, \underline{X})$.

Before diving into the proof, two remarks are in order. Firstly, Proposition 6.6 does not provide a full estimate of a functional with a moment evaluation; given a functional L with properties (a), (b), (c) as above, it does not provide a point $\underline{X}$ and a measure μ such that $L(g) \geq g(\mu, \underline{X})$ for *all* $g \in \mathscr{M}[\underline{x}]$. Rather, it provides $\underline{X}$ and μ such that $L(f) \geq f(\mu, \underline{X})$ for a *single* $f \in \mathscr{M}[\underline{x}]$ given in advance. Secondly, while the proof of Proposition 6.6 below relies on classical ideas from functional analysis and moment problems, it is rather intricate. At its core is a version of the Gelfand-Naimark-Segal (GNS) construction producing $\underline{X}$ and μ. However, handling unboundedness requires special care (hence analyticity of vectors and strong commutation of unbounded operators are invoked), and obtaining $\underline{X}$ and μ is a three-step process, as follows. We start by applying the GNS construction to the functional L, in order to construct a representation of $\mathscr{M}[\underline{x}]$ as strongly commuting unbounded operators on a Hilbert space (whose inner product is induced by L). Next we use the direct integral decomposition of this representation to isolate a one-dimensional sub-representation that bounds L from below at f (this sub-representation already determines $\underline{X}$). Lastly, using Carleman's condition, we argue that the restriction of this sub-representation to $\mathtt{m}(\mathbb{R}[\underline{x}])$ arises from a measure μ.

For the theory of unbounded operators, which is extraneous to the rest of the paper outside of this proof, we refer the reader to the books [Sch90, Sch12] (see also footnotes for short explanations).

Proof of Proposition 6.6. We divide the proof into three steps.

Step 1: unbounded GNS construction. On $\mathscr{M}[\underline{x}]$ we define a semi-inner product $\langle p, q \rangle = L(pq)$. Let $\mathcal{N} = \{p \in \mathscr{M}[\underline{x}]: L(p^2) = 0\}$. By the Cauchy-Schwarz inequality for semi-inner products, $\mathcal{N}$ is an subspace of $\mathscr{M}[\underline{x}]$: if $p, q \in \mathcal{N}$ then $L(pq)^2 \leq L(p^2)L(q^2) = 0$ implies that $L((p + q)^2) = L(p^2) + 2L(pq) + L(q^2) = 0$, so $p + q \in \mathcal{N}$. Furthermore, if $p \in \mathcal{N}$ then $p^2 \in \mathcal{N}$. Indeed, if $L(p^2) = 0$, then

$$0 \leq L\left((p \pm \varepsilon p^3)^2\right) = L(p^2) \pm 2\varepsilon L(p^4) + \varepsilon^2 L(p^6) = \varepsilon\left(\pm 2L(p^4) + \varepsilon L(p^6)\right)$$

and hence $\pm 2L(p^4) + \varepsilon L(p^6) \geq 0$ for every $\varepsilon > 0$, implying $L(p^4) = 0$. Consequently, $\mathcal{N}$ is an ideal of $\mathscr{M}[\underline{x}]$: if $p \in \mathcal{N}$ and $q \in \mathscr{M}[\underline{x}]$ then $L((pq)^2)^2 \leq L(p^4)L(q^4) = 0$ by the Cauchy-Schwarz inequality, and so $pq \in \mathcal{N}$. 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[1] 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 Cauchy-Schwarz inequality and the condition (c) on L imply that for all

---

[1] An operator $S: D \to \mathcal{H}$ on a dense subspace $D \subseteq \mathcal{H}$ is *symmetric* if $\langle Su, v \rangle = \langle u, Sv \rangle$ for all $u, v \in D$.

$j \in \{1, \ldots, n\}$ and $k, \ell \in \mathbb{N}$, the inequalities

$$\|X_j^k p\|^2 = L(x_j^{2k} p^2) \leq \sqrt{L(x_j^{4k})} \sqrt{L(p^4)} \leq \sqrt{(2k)!} \sqrt{M} \|p^2\|$$

$$\|Y_{ke_j}^\ell p\|^2 = L(\mathtt{m}(x_j^k)^{2\ell} p^2) \leq \sqrt{L(\mathtt{m}(x_j^{2k})^{4\ell}} \sqrt{L(p^4)} \leq \sqrt{(4\ell)!} \sqrt{(k!)^{4\ell} M} \|p^2\|$$

hold for all $p \in \mathscr{M}[\underline{x}]/\mathcal{N}$ and $k, \ell \in \mathbb{N}$ (here, $Y_{ke_j} = Y_{\ldots,0,k,0\ldots}$ with k at the jth position is the operator given by multiplication with $\mathtt{m}(x_j^k)$). By induction on k it is easy to see that $\sqrt[m]{(mk)!} \leq m^k k!$ for all $k, m \in \mathbb{N}$. Therefore,

$$\|X_j^k p\| \leq \sqrt{k!} \sqrt{2}^k \sqrt[4]{M} \sqrt{\|p^2\|},$$

$$\|Y_{ke_j}^\ell p\| \leq \ell! (4 \cdot k!)^\ell \sqrt[4]{M} \sqrt{\|p^2\|}$$

for all k, ℓ, j and $p \in \langle \underline{x} \rangle / \mathcal{N}$. In terms of [Sch12, Definition 7.1], the elements of $\mathscr{M}[\underline{x}]/\mathcal{N}$ are analytic[2] vectors for X_j and Y_{ke_j}, for every $j \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$. By [Sch12, Theorem 7.18], the closures[3] $\overline{X}_j$ and $\overline{Y}_{ke_j}$ are pairwise strongly commuting self-adjoint operators[4]. In particular, $\overline{Y}_{2ke_j}$ are positive self-adjoint operators. By Lemma 6.4 and the Cauchy-Schwarz inequality we have

$$
\begin{aligned}
(6.12) \quad \|Y_{i_1,\ldots,i_n} p\|^2 &= L(\mathtt{m}_{i_1,\ldots,i_n}^2 p^2) \\
&\leq \sum_{j=1}^{n-1} \frac{1}{2^j} L\left(\mathtt{m}(x_j^{2^{j+1} i_j}) p^2\right) + \frac{1}{2^{n-1}} L\left(\mathtt{m}(x_n^{2^n i_n}) p^2\right) \\
&\leq \sum_{j=1}^{n-1} \frac{1}{2^j} \left\|\left(\overline{Y}_{(2^{j+1} i_j)e_j} + I\right) p\right\|^2 + \frac{1}{2^{n-1}} \left\|\left(\overline{Y}_{(2^n i_n)e_n} + I\right) p\right\|^2
\end{aligned}
$$

for all $p \in \mathscr{M}[\underline{x}]/\mathcal{N}$. That is, the operators $Y_{i_1,\ldots,i_n}$ are majorized in the sense of (6.12) by self-adjoint operators that pairwise strongly commute with all $\overline{X}_j$ and $\overline{Y}_{ke_j}$. Thus, for every collection $i_1, \ldots, i_n, i_1', \ldots, i_n', j$ there exists a self-adjoint operator S on $\mathcal{H}$ that commutes with $X_j, Y_{i_1,\ldots,i_n}, Y_{i_1',\ldots,i_n'}$ on $\mathscr{M}[\underline{x}]/\mathcal{N}$, and satisfies

$$\|X_j p\|, \|Y_{i_1,\ldots,i_n} p\|, \|Y_{i_1',\ldots,i_n'} p\| \leq \|S p\|$$

for all $p \in \mathscr{M}[\underline{x}]/\mathcal{N}$. Then the closures $\overline{X}_j, \overline{Y}_{i_1,\ldots,i_n}, \overline{Y}_{i_1',\ldots,i_n'}$ are pairwise strongly commuting self-adjoint operators by [Sch88, Proposition 2]. Since real polynomials in strongly commuting self-adjoint operators are again self-adjoint,

$$\varphi(x_j) = \overline{X}_j, \quad \varphi(\mathtt{m}_{i_1,\ldots,i_n}) = \overline{Y}_{i_1,\ldots,i_n}$$

defines an integrable representation[5] φ of $\mathscr{M}[\underline{x}]$ on $\mathcal{H}$ according to [Sch90, Definition 9.1.1]. By the construction, $\varphi(\mathrm{QM}(S_1, S_2))$ consists of positive self-adjoint operators.

---

[2]A vector $v \in D$ is *analytic* for $S : D \to D$ if there is $M > 0$ such that $\|S^k v\| \leq M^k k!$ for all k.

[3]The *closure* of an operator T on $\mathcal{H}$ is the operator on $\mathcal{H}$ whose graph in $\mathcal{H} \oplus \mathcal{H}$ is the closure of the graph of T. Every densely defined symmetric operator admits the closure.

[4]An operator is *self-adjoint* if it is equal to its own adjoint on $\mathcal{H}$ (this notion is stronger than being symmetric). *Spectral projections* of a self-adjoint operator S are projections on $\mathcal{H}$ corresponding to measurable subsets of the spectrum of S via the spectral theorem. Two self-adjoint operators *strongly commute* if all their spectral projections commute.

[5]An *integrable representation* of a commutative algebra (with trivial involution) is a homomorphism into a subalgebra of strongly commuting self-adjoint operators on a Hilbert space.

Step 2: integral decomposition of φ. Irreducible integral representations of $\mathscr{M}[\underline{x}]$ are one-dimensional by the unbounded analog of Schur's lemma[6] [Sch90, Corollary 9.1.11]. Since $\mathscr{M}[\underline{x}]$ is countably generated, by [Sch90, Theorem 12.3.5 and subsequent Remark 2] there exist a compact metric space T with a Borel probability measure λ and a measurable family of integral representations φ_t of $\mathscr{M}[\underline{x}]$ acting on Hilbert spaces $\mathcal{H}_t$ for $t \in T$, such that φ is unitarily equivalent to the direct integral[7]

$$\int_T^\oplus \varphi_t \, \mathrm{d}\lambda(t)$$

on the Hilbert space $\mathcal{H}' = \int_T^\oplus \mathcal{H}_t \, \mathrm{d}\lambda(t)$, and almost every (relative to λ) $\mathcal{H}_t$ is one-dimensional. Let $e_t \in \mathcal{H}_t$ be such that that the vector $\int_T^\oplus e_t \, \mathrm{d}\lambda(t) \in \mathcal{H}'$ corresponds to the unit vector $1 \in \mathcal{H}$ under the aforementioned unitary equivalence. Since $1 \in \mathcal{H}$ is a cyclic[8] vector for $\varphi(\mathscr{M}[\underline{x}])$, we have $e_t \neq 0$ for almost all t. Define $L_t : \mathscr{M}[\underline{x}] \to \mathbb{R}$ as $L_t(p) = \langle \varphi_t(p)e_t, e_t \rangle$; then

$$(6.13) \qquad\qquad L(p) = \int_T L_t(p) \, \mathrm{d}\lambda(t)$$

for all $p \in \mathscr{M}[\underline{x}]$. Since $\varphi(\mathrm{QM}(S_1, S_2))$ consists of positive self-adjoint operators, it follows by [Sch90, Proposition 12.2.3(ii)] that $\varphi_t(\mathrm{QM}(S_1, S_2))$ consists of positive selfadjoint operators, for almost every t. In particular, $L_t(\mathrm{QM}(S_1, S_2)) = \mathbb{R}_{\geq 0}$ for almost every t. The subset

$$Z' = \{t \in T \colon L_t(f) \leq L(f)L_t(1)\}$$

of T is measurable. Note that $\lambda(Z') = 0$ would by (6.13) give rise to the contradiction

$$L(f) = \int_T L_t(f) \, \mathrm{d}\lambda(t) > \int_T L(f)L_t(1) \, \mathrm{d}\lambda(t) = L(f)L(1) = L(f),$$

so $m := \lambda(Z') > 0$.

Let $N > 0$ be sufficiently large so that $\sum_{k=1}^\infty \frac{nM}{N^k} < m$, and consider the sets

$$Z_{j,k} = \left\{t \in T \colon L_t\big(\mathtt{m}(x_j^{2k})\big) \geq N^k k!\right\}$$

for $j \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$. Then $Z_{j,k}$ is measurable, and

$$k!M \geq L\big(\mathtt{m}(x_j^{2k})\big) = \int_T L_t\big(\mathtt{m}(x_j^{2k})\big) \, \mathrm{d}\lambda(t) \geq \int_{Z_{j,k}} L_t\big(\mathtt{m}(x_j^{2k})\big) \, \mathrm{d}\lambda(t) \geq N^k k!\lambda(Z_{j,k}),$$

so $\lambda(Z_{j,k}) \leq \frac{M}{N^k}$. By the choice of N,

$$\lambda\left(\bigcup_{j=1}^n \bigcup_{k \in \mathbb{N}} Z_{j,k}\right) \leq \sum_{j=1}^n \sum_{k=1}^\infty \lambda(Z_{j,k}) \leq n \sum_{k=1}^\infty \frac{M}{N^k} < m,$$

and so

$$\lambda\left(Z' \setminus \bigcup_{j,k} Z_{j,k}\right) > 0.$$

---

[6]The Schur-Dixmier lemma implies that irreducible representations of an abelian $\mathbb{C}$-algebra of countable dimension are one-dimensional.

[7]*Direct integrals* of Hilbert spaces, operators and representations generalize direct sums of these objects. While infinite-dimensional integral representations do not necessarily admit a direct sum decomposition into irreducible ones, they do admit a direct integral decomposition into irreducible ones.

[8]A vector $v \in \mathcal{H}$ is *cyclic* for an algebra of operators $\mathcal{A}$ on $\mathcal{H}$ if $\mathcal{A} \cdot v$ is dense in $\mathcal{H}$.

In particular, there exists $t \in T$ such that φ_t is one-dimensional, $e_t \neq 0$, and

$$L_t(\mathrm{QM}(S_1, S_2)) = \mathbb{R}_{\geq 0}, \quad L_t(f) \leq L(f) L_t(1), \quad L_t\big(\mathtt{m}(x_j^{2k})\big) \leq N^k k! \text{ for } k \in \mathbb{N}.$$

Since $\varphi_t : \mathscr{M}[\underline{x}] \to \mathbb{R}$ is a nonzero homomorphism, it is unital, and so $L_t(p) = \varphi_t(p) L_t(1)$ for all $p \in \mathscr{M}[\underline{x}]$. Since $L_t(1) = \|e_t\|^2 > 0$,

$$\varphi_t(\mathrm{QM}(S_1, S_2)) = \mathbb{R}_{\geq 0}, \quad \varphi_t(1) = 1, \quad \varphi_t(f) \leq L(f), \quad \varphi_t\big(\mathtt{m}(x_j^{2k})\big) \leq \frac{N^k k!}{L_t(1)} \text{ for } k \in \mathbb{N}.$$

Step 3: conclusion. Consider the functional $\widetilde{L} : \mathbb{R}[\underline{x}] \to \mathbb{R}$ given by $\widetilde{L}(p) = \varphi_t(\mathtt{m}(p))$. Then $\widetilde{L}(\mathrm{QM}(S_1)) = \mathbb{R}_{\geq 0}$, and $\widetilde{L}$ satisfies the Carleman condition since $\widetilde{L}(x_j^{2k}) \leq \frac{1}{L_t(1)} N^k k!$ for all $j = 1, \ldots, n$ and $k \in \mathbb{N}$. By a refined version of Nussbaum's theorem [Sch17, Theorem 14.25] applied to $\widetilde{L}$, there is $\mu \in \mathbf{K}(S_1, S_2)$ such that $\varphi_t(\mathtt{m}_{i_1, \ldots, i_n}) = \widetilde{L}(x_1^{i_1} \cdots x_n^{i_n}) = \int x_1^{i_1} \cdots x_n^{i_n} \, d\mu$ for all $(i_1, \ldots, i_n) \in \mathbb{N}_0^n$. Let $\underline{X} = (\varphi_t(x_1), \ldots, \varphi_t(x_n))$; then $\underline{X} \in K(S_1)$ because $\varphi_t(\mathrm{QM}(S_1, S_2)) = \mathbb{R}_{\geq 0}$. Lastly, $f(\mu, \underline{X}) = \varphi_t(f) \leq L(f)$. $\qquad\square$

The following auxiliary lemma bounds the growth of a positive linear functional on moment polynomials in terms of its marginal values (on x_j^{2k} and $\mathtt{m}(x_j^{2k})$).

Lemma 6.7. *Let $r \in \mathbb{N}$, $M \geq 1$ and $L \in \mathscr{M}[\underline{x}]_{2r}^\vee$. Suppose $L(\mathrm{QM}(\emptyset, \emptyset)_{2r}) = \mathbb{R}_{\geq 0}$ and*

$$(6.14) \qquad\qquad L\left(x_j^{2k}\right), L\left(\mathtt{m}(x_j^{2k})\right) \leq k! \, M$$

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

$$(6.15) \qquad\qquad |L(w)| \leq \sqrt{(\deg w)! \, M}$$

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 $i_1 + \cdots + i_n \leq r$, one obtains

$$(6.16) \qquad\qquad L(\mathtt{m}_{2i_1, \ldots, 2i_n}) \leq (i_1 + \cdots + i_n)! \, M$$

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

$$(6.17) \qquad\qquad L(\mathtt{m}_{i_1, \ldots, i_n}^{2k}) \leq L(\mathtt{m}_{2ki_1, \ldots, 2ki_n})$$

for $k(i_1 + \cdots + i_n) \leq r$ since $\mathtt{m}_{2ki_1, \ldots, 2ki_n} - \mathtt{m}_{i_1, \ldots, i_n}^{2k} \in \mathrm{qm}(\emptyset, \emptyset)$ by Lemma 2.2(2). By (6.16), (6.17) and [Las06, Lemma 6.2],

$$L(w^2) \leq (\deg w)! \, M$$

for all monomials $w \in \mathscr{M}[\underline{x}]_r$. Finally, (6.15) follows from $L(w^2) - L(w)^2 = L((w - L(w))^2) \geq 0$ for $\deg w \leq r$. $\qquad\square$

In the final preparation for Theorem 6.9, we connect the solution of the optimization problem (6.3) with the infimum of f on $\mathbf{K}(S_1, S_2) \times K(S_1)$.

Lemma 6.8. *Suppose S_2 is finite, $\mathbf{K}(S_1, S_2) \neq \emptyset$ and $f \in \mathscr{M}[\underline{x}]$ is bounded below on $\mathbf{K}(S_1, S_2) \times K(S_1)$; denote $f_* := \inf_{(\mu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_2)} f(\mu, \underline{X}) > -\infty$. For large enough $M > 0$, (6.3) is feasible for $2r \geq \deg f$, and $\inf Q_{r,M}^\vee \nearrow f_M$ as $r \to \infty$ for some $f_M \geq f_*$.*

Proof. Feasibility of (6.3) follow by the same argument as in the proof of Lemma 6.5. Let L be feasible for $Q_{r,M}^\vee$. Then $L(M - \Phi_r - \Psi_r) \geq 0$ implies

$$(6.18) \qquad L(x_j^{2k}) \leq k! \, M \quad \text{for } k \leq r, \qquad L(\mathtt{m}(x_j^{2k})^\ell) \leq \ell! \, (k!)^\ell M \quad \text{for } k\ell \leq r.$$

Let $d \in \mathbb{N}$, and $r \geq d$. By Lemma 6.7 and (6.18) (for $\ell = 1$),

$$(6.19) \qquad\qquad |L(w)| \leq \sqrt{d!\, M} =: c_d$$

for all monomials w in $\mathscr{M}[\underline{x}]_d$. In particular, $L(f)$ is uniformly bounded for large enough r. Therefore, $(\inf Q^{\vee}_{r,M})_r$ 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 $f_M \geq f_*$.

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 (6.3), and let $s_r \in \ell^{\infty}$ be given as $s_r(w) = \frac{1}{c_{\deg w}} L^{(r)}(w)$ for monomials w with $\deg w \leq 2r$, 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], the closed unit ball in ℓ^{∞} is compact in the weak-* topology. In particular, the sequence $(s_r)_r$ has an accumulation point in ℓ^{∞} with respect to the weak-* topology. Hence, there is $s \in \ell^{\infty}$ and a subsequence $(s_{r_m})_m$ converging to s. Define

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

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

$$L(x_j^{2k}) \leq k!\, M \quad \text{for } k \leq r_m, \qquad L\big(\mathtt{m}(x_j^{2k})^{\ell}\big) \leq \ell!\, (k!)^{\ell}\, M \quad \text{for } k\ell \leq r_m$$

by (6.18) since $L^{(r_m)}$ is feasible for $Q^{\vee}_{r_m,M}$. Consequently,

$$L(x_j^{2k}) \leq k!\, M \quad \text{for } k \in \mathbb{N}, \qquad L\big(\mathtt{m}(x_j^{2k})^{\ell}\big) \leq \ell!\, (k!)^{\ell}\, M \quad \text{for } k, \ell \in \mathbb{N}.$$

Therefore, $L(f) \geq f_*$ by Proposition 6.6. $\qquad\qquad\qquad\qquad\qquad\square$

6.3. **Positivstellensatz.** We are now ready to prove a moment polynomial analog of Lasserre's Positivstellensatz [Las06, LN07]. Its proof relies on the interplay of solutions of optimization problems (6.2) and (6.3) with the infimum of f on $\mathbf{K}(S_1, S_2) \times K(S_1)$.

Theorem 6.9 (Perturbative Positivstellensatz)**.** *Let $S_1 \subseteq \mathbb{R}[\underline{x}]$ and $S_2 \subseteq \mathscr{M}$, and suppose S_2 is finite. The following statements are equivalent for $f \in \mathscr{M}[\underline{x}]$:*

 (i) *$f \geq 0$ on $\mathbf{K}(S_1, S_2) \times K(S_1)$;*
 (ii) *for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that $f + \varepsilon(\Phi_r + \Psi_r) \in \mathrm{QM}(S_1, S_2)$.*

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

 (i') *$f \geq 0$ on $\mathbf{K}(S_1, S_2)$;*
 (ii') *for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that $f + \varepsilon(1 + \Psi_r) \in \mathrm{qm}(S_1, S_2)$.*

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

$$\Phi_r(\underline{X}) + \Psi_r(\nu) \leq M$$

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 $\mathbf{K}(S_1, S_2)$ interpolate any measure in $\mathbf{K}(S_1, S_2)$ up to moments of any fixed order by Proposition 2.6, it follows that $f(\mu, \underline{X}) \geq 0$ for all $(\mu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_1)$.

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

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

First suppose $f_* > 0$. By Proposition 2.6, there exists $(\nu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_1)$, with a finitely supported ν. Denote $M_0 := \sup_r (\Phi_r(\underline{X}) + \Psi_r(\nu)) < \infty$, and let $M > \max\{\frac{1}{f_*}, M_0\}$ be arbitrary. By Lemmas 6.5 and 6.8, there exists $r_M > 0$ such that $\sup Q_{r_M, M} > f_* - \frac{1}{M}$. That is, there are $z_M \geq f_* - \frac{1}{M}$, $\lambda_M \in \mathbb{R}_{\geq 0}$ and $q_M \in \mathrm{QM}(S_1, S_2)_{2r_M}$ such that

$$(6.20) \qquad f - z_M = q_M + \lambda_M \left(M - \Phi_{r_M} - \Psi_{r_M} \right).$$

Evaluating (6.20) at $(\nu, \underline{X}) \in \mathbf{K}(S_1, S_2) \times K(S_1)$ gives

$$f(\nu, \underline{X}) - f_* + \frac{1}{M} \geq f(\nu, \underline{X}) - z_M$$
$$= q_M(\nu, \underline{X}) + \lambda_M \left(M - \Phi_{r_M}(\underline{X}) - \Psi_{r_M}(\nu) \right)$$
$$\geq \lambda_M (M - M_0),$$

and therefore

$$(6.21) \qquad \lambda_M \leq \frac{f(\nu, \underline{X}) - f_* + \frac{1}{M}}{M - M_0}.$$

The right-hand side of (6.21) goes to 0 as $M \to \infty$. By (6.20),

$$f + \lambda_M \left(\Phi_{r_M} + \Psi_{r_M} \right) = z_M + q_M + \lambda_M M \in \mathrm{QM}(S_1, S_2)_{2r_M},$$

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

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

Case (b): assume $\mathbf{K}(S_1, S_2) = \emptyset$, and let $f \in \mathbb{R}[\underline{x}]$ and $\varepsilon > 0$ be arbitrary. Let x_{n+1} be an auxiliary variable, and consider $S_1' = x_{n+1} \cdot (\{1\} \cup S_1) \subset \mathbb{R}[x_1, \ldots, x_{n+1}]$ and $S_2' = \mathtt{m}_{1,0,\ldots,0}^2 \cdot S_2 \subset \mathcal{M}$. Then $K(S_1')$ contains $\mathbb{R}^n \times \{0\}$, $\mathbf{K}(S_1', S_2')$ contains all $\mu \in \mathbf{P}(\mathbb{R}^n \times \{0\})$ such that $\int x_1 \, d\mu = 0$ (and is in particular nonempty), and $x_{n+1}f \geq 0$ on $\mathbf{K}(S_1', S_2')$. By the case (a) of the proof above, there exists $r \in \mathbb{N}$ such that

$$(6.22) \qquad x_{n+1}f + \frac{\varepsilon}{n + e + e^2} \left(\Phi_r + \Psi_r + \sum_{k=0}^{r} \frac{x_{n+1}^{2k}}{k!} + \sum_{\substack{k, \ell \in \mathbb{N}, \\ k\ell \leq r}} \frac{\mathtt{m}(x_{n+1}^{2k})^\ell}{(k!)^\ell \ell!} \right) \in \mathrm{QM}(S_1', S_2').$$

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

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

Note that ξ intertwines with $\mathtt{m}$. Applying ξ to (6.22) then gives

$$f + \frac{\varepsilon}{n + e + e^2} \left(\Phi_r + \Psi_r + \sum_{k=0}^{r} \frac{1}{k!} + \sum_{\substack{k, \ell \in \mathbb{N}, \\ k\ell \leq r}} \frac{1}{(k!)^\ell \ell!} \right) \in \mathrm{QM}(S_1, S_2') \subseteq \mathrm{QM}(S_1, S_2),$$

and therefore $f + \varepsilon(\Phi_r + \Psi_r) \in \mathrm{QM}(S_1, S_2)$ because $\sum_{k=0}^{r} \frac{1}{k!} \leq e$ and $\sum_{k\ell \leq r} \frac{1}{(k!)^\ell \ell!} \leq e^2$ by Lemma 6.1.

(i')$\Leftrightarrow$(ii') The proof is analogous to (i)$\Leftrightarrow$(ii), and utilizes the straightforward counterparts of Lemmas 6.5, 6.7, 6.8 and Proposition 6.6 for $\mathrm{qm}(S_1, S_2)$. $\square$

Remark 6.10. In Theorem 6.9, one can replace the cone $\mathrm{QM}(S_1, S_2)$ with the larger cone $\widetilde{\mathrm{QM}}(S_1, S_2)$ from Section 5. While the resulting statement is slightly weaker than Theorem 6.9, it has the advantage that it is more amenable for computations. Namely, for every $r \in \mathbb{N}$, the smallest $\varepsilon_r \geq 0$ such that $f + \varepsilon_r(\Phi_r + \Psi_r) \in \widetilde{\mathrm{QM}}(S_1, S_2)_{2r}$ can be computed via SDP. Theorem 6.9 then implies that $\lim_{r \to \infty} \varepsilon_r = 0$ if $f \geq 0$ on $\mathbf{K}(S_1, S_2) \times K(S_1)$.

Example 6.11. The implications (ii)$\Rightarrow$(i) and (ii')$\Rightarrow$(i') of Theorem 6.9 fail in general when S_2 is not finite. Let $n = 1$, $f = -1$ and

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

Since the $2i^{\text{th}}$ moment of $\mu = e^{-\sqrt{|t|}} \, dt$ is $(4i + 1)!$, we have $\mathbf{K}(\emptyset, S_2) \neq \emptyset$ 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

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

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

Example 6.12. 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 $\mathbf{P}(\mathbb{R}^2)$; however, f does not belong to $\mathrm{qm}(\emptyset, \emptyset)$. For a fixed $r \in \mathbb{N}$, searching for the smallest $\varepsilon_r \geq 0$ such that $f + \varepsilon_r(1 + \Psi_r) \in \mathrm{qm}(\emptyset, \emptyset)$ can be formulated as an SDP. For small values of r one obtains $\varepsilon_2 = 0.33333$, $\varepsilon_3 = 0.06330$, $\varepsilon_4 = 0.01416$.

Example 6.13. Let us return to $f = \mathtt{m}_{4,2}\mathtt{m}_{2,4} - \mathtt{m}_{2,2}^3$ from Example 3.8. Then f is nonnegative on $\mathbf{P}(\mathbb{R}^2)$, but $\rho(f) < 0$ for some pseudo-moment evaluation $\rho : \mathscr{M} \to \mathbb{R}$, so f does not admit a sum-of-squares certificate with denominators in the sense of Hilbert's 17th problem by Theorem 3.7. Nevertheless, for every $\varepsilon > 0$, Theorem 6.9 guarantees an $r \in \mathbb{N}$ such that $f + \varepsilon(1 + \Psi_r) \in \mathrm{qm}(\emptyset, \emptyset)$. Alternatively, since f is homogeneous with respect to the degree on $\mathscr{M}$, its nonnegativity on $\mathbf{P}(\mathbb{R}^2)$ is equivalent to nonnegativity on $\mathbf{P}([-1, 1]^2)$. By Theorem 4.2, $f + \varepsilon \in \mathrm{qm}(\{1 - x_1, 1 + x_1, 1 - x_2, 1 + x_2\}, \emptyset)$ for every $\varepsilon > 0$.

6.4. **Polynomial positivity on arbitrary semialgebraic sets.** Theorem 6.9 also carries implications for classical (non-moment) polynomials (see Remark 6.15 for comparison with earlier results).

Corollary 6.14. *Let $S \subseteq \mathbb{R}[\underline{x}]$. Then the following statements 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 $f + \varepsilon\Phi_r \in \mathrm{QM}(S)$.*

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 $\zeta(\mathtt{m}_{i_1,\ldots,i_n}) = 0$ (for i_j not all zero) maps $\mathrm{QM}(S, \emptyset)$ into $\mathrm{QM}(S)$. Applying ζ to the conclusions of Theorem 4.2 for f and $\mathrm{QM}(S, \emptyset)$ gives the desired statement. $\qquad\square$

Remark 6.15. The special case of Corollary 6.14 for $S = \emptyset$ is given in [Las06, Theorem 4.1]. Furthermore, the conclusion of Corollary 6.14 has been established in [LN07, Corollary 3.7] under additional assumptions, namely that S is finite and has the strong moment property[9], and the interior of $K(S)$ is nonempty; furthermore, [LN07, Corollary

---

[9]S has the *strong moment property* if every functional on $\mathbb{R}[\underline{x}]$ that is nonnegative on the preordering generated by S, equals integration with respect to a measure supported on $K(S)$. This is a rather restrictive property; e.g., $S = \{x_1, x_2\}$ does not have it.

3.7] requires preorderings instead of quadratic modules. Corollary 6.14 disposes of all these assumptions. Nevertheless, [Las06] and [LN07] provide stronger certificates in two special cases: when nonnegativity on $K(S) \cap [-1, 1]^n$ is considered (and S is finite with the strong moment property, and $K(S)$ has nonempty interior), a simpler perturbation can be used in place of Φ_r [LN07, Corollary 3.6]; when S consists of concave polynomials and f is convex, only sums of squares and conic combinations of S are needed in place of $\mathrm{QM}(S)$ [Las06, Corollary 4.3].

Remark 6.16. Corollary 6.14 characterizes polynomial positivity on arbitrary basic closed semialgebraic sets (and even more general sets, since infinitely many constraints are allowed), 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 6.17. Let us record one of the simplest cases to which [LN07, Corollary 3.7] does not apply. Clearly, $x_1 x_2 \geq 0$ on $K(\{x_1, x_2\})$. By Corollary 6.14, for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that $x_1 x_2 + \varepsilon \Phi_r \in \mathrm{QM}(\{x_1, x_2\})$, even though the set $\{x_1, x_2\}$ does not have the strong moment property.

For a fixed $r \in \mathbb{N}$, one can find the smallest $\varepsilon_r \geq 0$ such that $x_1 x_2 + \varepsilon_r \mathtt{m}(\Phi_r) \in \mathrm{qm}(\emptyset, \emptyset)$ by solving an SDP. For $r = 2, \ldots, 8$ the values of ε_r are

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

Remark 6.18. As is evident from the proof of Theorem 6.9, the sequence of polynomials Φ_r can be replaced by

$$(6.23) \qquad \sum_{j=1}^{n} \sum_{k=0}^{r} \frac{x_j^{2k}}{c_k} \qquad \text{for } r \in \mathbb{N},$$

where $c_k > 0$ are such that $(c_k)_k$ has super-exponential growth (to ensure point-wise convergence of (6.23), which is used for (ii)$\Rightarrow$(i) of Theorem 6.9 and for feasibility of (6.2) and (6.3)) and $(k^{-k} c_k)_k$ has at most exponential growth (which is needed for applying Proposition 6.6). One might further contemplate whether only the constant term and the leading terms of Φ_r are essential in Corollary 6.14; 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 $S = \{-1 + \frac{x_1^{2k}}{k!-1} : k \geq 2\}$. Then $K(S) = \emptyset$ and $f \geq 0$ on $K(S)$. We claim that $-2 + (1 + \frac{x_1^{2r}}{r!}) \notin \mathrm{QM}(S)$ for every $r \in \mathbb{N}$. Indeed, suppose

$$(6.24) \qquad -1 + \frac{x_1^{2r}}{r!} = \sigma_1 + \sum_{k=2}^{\ell} \sigma_k \cdot \left(-1 + \frac{x_1^{2k}}{k!-1} \right)$$

where $\ell \geq 2$ and $\sigma_k \in \mathbb{R}[x]$ are sums of squares. Note that $\ell \leq r$. Let $X = \sqrt[2\ell]{\ell!-1}$. Then the right-hand side of (6.24) is nonnegative at X, while the left-hand side of (6.24) is negative at X, a contradiction.

Nevertheless, if $\varepsilon > 0$ is arbitrary and $r \geq \frac{2}{\varepsilon} - 1$, then

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

as anticipated by Corollary 6.14.

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

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

Corollary 6.19. *Let* S, f, ε *be as above. Then* $(f_r^{(\varepsilon)})_r$ *is an increasing sequence, and*

$$(6.25) \qquad f_* \leq \lim_{r \to \infty} f_r^{(\varepsilon)} \leq \inf_{\underline{X} \in K(S)} \left(f(\underline{X}) + \varepsilon(\exp(X_1^2) + \cdots + \exp(X_n^2)) \right).$$

In particular,

$$(6.26) \qquad \lim_{\varepsilon \downarrow 0} \lim_{r \to \infty} f_r^{(\varepsilon)} = f_*.$$

Proof. The first inequality in (6.25) holds by Corollary 6.14, and the second inequality in (6.25) is straightforward. Lastly, (6.26) follows from $\lim_{\varepsilon \downarrow 0} \inf_K (f + \varepsilon g) = \inf_K f$ for any nonnegative function g on $\mathbb{R}^n$. $\qquad \square$

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. 25

[ApS24] MOSEK ApS. *Introducing the MOSEK Optimization Suite 10.2.3*, 2024. 19

[Are46] Richard Arens. The space L^ω and convex topological rings. *Bull. Amer. Math. Soc.*, 52:931–935, 1946. 7

[Bar02] Alexander Barvinok. *A course in convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002. 18, 25, 30

[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. 2

[BKP16] Sabine Burgdorf, Igor Klep, and Janez Povh. *Optimization of polynomials in non-commuting variables*. SpringerBriefs in Mathematics. Springer, 2016. 3

[BP05] Dimitris Bertsimas and Ioana Popescu. Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal on Optimization*, 15(3):780–804, 2005. 2

[BRS+22] Grigoriy Blekherman, Felipe Rincón, Rainer Sinn, Cynthia Vinzant, and Josephine Yu. Moments, sums of squares, and tropicalization. *arXiv:2203.06291*, 2022. 3, 5, 11, 12, 14

[CG22] Davide Castelvecchi and Elizabeth Gibney. 'Spooky' quantum-entanglement experiments win physics Nobel. *Nature*, 610:241–242, 2022. 18

[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. 2

[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. 3

[Fan22] Giovanni Fantuzzi. Verification of some functional inequalities via polynomial optimization. *IFAC-PapersOnLine*, 55(16):166–171, 2022. 18th IFAC Workshop on Control Applications of Optimization CAO 2022. 2, 11

[HIKV23] Didier Henrion, Maria Infusino, Salma Kuhlmann, and Victor Vinnikov. Infinite-dimensional moment-SOS hierarchy for nonlinear partial differential equations. *arXiv:2305.18768*, 2023. 2, 3

[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. 3, 9

[HM04] J. William Helton and Scott A. McCullough. A Positivstellensatz for non-commutative polynomials. *Trans. Amer. Math. Soc.*, 356(9):3721–3737, 2004. 3

[IKKM23]   Maria Infusino, Salma Kuhlmann, Tobias Kuna, and Patrick Michalski. Moment problem for algebras generated by a nuclear space. *arXiv:2301.12949*, 2023. 3, 11

[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. 3, 11, 25

[KMV22]   Igor Klep, Victor Magron, and Jurij Volčič. Optimization over trace polynomials. *Ann. Henri Poincaré*, 23(1):67–100, 2022. 3, 10

[KMVW23]   Igor Klep, Victor Magron, Jurij Volčič, and Jie Wang. State polynomials: positivity, optimization and nonlinear bell inequalities. *Math. Program.*, 2023. 3, 10, 13, 16, 17

[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. 2

[KPV21]   Igor Klep, James E. Pascoe, and Jurij Volčič. Positive univariate trace polynomials. *J. Algebra*, 579:303–317, 2021. 12

[KŠV18]   Igor Klep, Špela Špenko, and Jurij Volčič. Positive trace polynomials and the universal Procesi–Schacher conjecture. *Proc. Lond. Math. Soc.*, 117(6):1101–1134, 2018. 13

[KSV22]   Igor Klep, Claus Scheiderer, and Jurij Volčič. Globally trace-positive noncommutative polynomials and the unbounded tracial moment problem. *Math. Ann.*, 387:1403–1433, 2022. 33

[Las01]   Jean-Bernard Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2000/01. 2, 4

[Las06]   Jean-Bernard Lasserre. A sum of squares approximation of nonnegative polynomials. *SIAM Journal on Optimization*, 16(3):751–765, 2006. 5, 29, 30, 32, 33

[Las09]   Jean-Bernard Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press optimization series. Imperial College Press, 2009. 3, 9

[LJ10]   Yuan Liao and Wenxin Jiang. Bayesian analysis in moment inequality models. *The Annals of Statistics*, 38(1):275–316, 2010. 2

[LN07]   Jean-Bernard Lasserre and Tim Netzer. Sos approximations of nonnegative polynomials via simple high degree perturbations. *Math. Z.*, 256:1432–1823, 2007. 5, 30, 32, 33

[Mar08]   Murray Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008. 2, 3, 5, 12, 13, 14, 15, 16, 33

[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. 2

[MW23]   Victor Magron and Jie Wang. *Sparse polynomial optimization: theory and practice*, volume 5 of *Series on Optimization and Its Applications*. World Scientific Press, 2023. 17

[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. 3

[Par05]   Kalyanapuram Rangachari Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Soc., Providence, RI, 2005. 7, 16

[PHBB17]   Victor Pozsgay, Flavien Hirsch, Cyril Branciard, and Nicolas Brunner. Covariance Bell inequalities. *Phys. Rev. A*, 96(6):062128, 13, 2017. 2, 3, 4, 19

[PPHI15]   Ariel Pakes, Jack Porter, Kate Ho, and Joy Ishii. Moment inequalities and their application. *Econometrica*, 83(1):315–334, 2015. 2

[Pro76]   Claudio Procesi. The invariant theory of $n \times n$ matrices. *Adv. Math.*, 19(3):306–381, 1976. 10

[PS01]   Victoria Powers and Claus Scheiderer. The moment problem for non-compact semialgebraic sets. *Adv. Geom.*, 1(1):71–88, 2001. 24

[Put93]   Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993. 2, 4

[Put97]   Mihai Putinar. A note on Tchakaloff's theorem. *Proc. Amer. Math. Soc.*, 125(8):2409–2414, 1997. 9

[Rud87]   Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Company, 1987. 3rd edition. 6

[Sch88] Konrad Schmüdgen. Strongly commuting selfadjoint operators and commutants of un-
         bounded operator algebras. *Proc. Amer. Math. Soc.*, 102:365–372, 1988. 27

[Sch90] Konrad Schmüdgen. *Unbounded Operator Algebras and Representation Theory*, volume 37
         of *Operator Theory: Advances and Applications*. Birkhäuser Basel, 1990. 26, 27, 28

[Sch12] Konrad Schmüdgen. *Unbounded self-adjoint operators on Hilbert space*, volume 265 of *Grad-
         uate Texts in Mathematics*. Springer Dordrecht, 2012. 10, 16, 26, 27

[Sch17] Konrad Schmüdgen. *The moment problem*, volume 277 of *Graduate Texts in Mathematics*.
         Springer Cham, 2017. 2, 9, 16, 25, 29

[SW99]   Helmut H. Schaefer and Manfred P. Wolff. *Topological vector spaces*, volume 3 of *Graduate
         Texts in Mathematics*. Springer-Verlag, New York, second edition, 1999. 11

[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. 2,
         4, 19, 20

[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. 2, 3, 4, 20

Index of recurring symbols

IGOR KLEP: FACULTY OF MATHEMATICS AND PHYSICS, DEPARTMENT OF MATHEMATICS, UNIVER-
SITY OF LJUBLJANA & FAMNIT, UNIVERSITY OF PRIMORSKA, KOPER & 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, UNIVERSITY OF AUCKLAND, NEW ZEALAND
    Email address: `jurij.volcic@auckland.ac.nz`